ISSN 2286 - 9131

Transactions

on Computer and Information Technology

Volume 10, No. 2, November 2016

ECTI-CIT

Invited Article

Research Article

Manuscript Submission Guideline

ECTI TRANSACTIONS ON COMPUTER AND INFORMATION TECHNOLOGY
VOLUME 10 NO. 2, November 2016

ECTI Transactions on Computer and Information Technology
ECTI Association
Sirindhon International Institute of Technology
131 Moo 5 Tiwanon Rd.,Bangkadi
Muang, Pathumthani 12000,Thailand

Tel: 02-5012578
Email: editorcit@ecti.or.th
http://www.ecti-thailand.org

Color Systems and Color Image Enhancement Methods

Akira Taguchi[1], Non-member

ABSTRACT

There are many color systems. Some systems are correspond to the human visual system, such as the Munsell color system. Other systems are formulated to ease data processing in machines, such as RGB color space. At first, Munsell color system is introduced in this paper. Next, RGB color system and hue-saturation-intensity (HSI) color system which is derived from RGB color systems are reviewed. HSI color system is important, because HSI color system is closely related to Munsell color system. We introduce the advantage and drawbacks of the conventional HSI color space. Furthermore, the improved HSI color system is introduced. The second half of this paper, we introduce a lot of color image enhancement methods based on the histogram equalization or the differential histogram equalization. Since hue preserving is necessary for color image processing, intensity processing methods by using both intensity and saturation in HSI color space are reviewed. Finally, hue preserving color image enhancement methods in RGB color system are explained.

Keywords: Munsell Color System, RGB Color System, HSI Color System, Histogram Equalization, Differential Histogram Equalization, Color Image Enhancement, Hue-preserving Processing

1. INTRODUCTION

Color signal can be expressed in many different ways, each with its advantage and drawbacks. Some representations are formulated to help humans select colors, such as the Munsell color system. This kind of color systems is called the color appearance system. And others are formulated to ease data making and processing in machines, such as RGB color system. This kind of color systems is called the additive color system.

In this paper, we first explain about the Munsell color system in order to clarify the prosperity of the human visual system.

Secondary, we refer the RGB color system (CIERGB) which was the utmost additive color system which created by the International Commission on Illumination (CIE). We can derive various color systems which are belonged to the additive color system by the color system conversion such as CIEXYZ and CIELAB. The hue-saturation-intensity (HSI) color system [1] and the hue- saturation-value (HSV) color system [2] are included in such color systems. HSI and HSV color systems are important, since the HSI or HSV color system is close to human being's perception to color and is closely related to Munsell color system. The HSI color system is explained and the problem of the color system is pointed out in this paper. This problem is called the gamut problem. Then, we introduce the improved HSI color system whose gamut is same of RGB color system [3]. This means that the gamut problem is solved.

The second half of this paper, we introduce a lot of color image enhancement methods based on the histogram equalization (HE) [1] or the differential histogram equalization (DHE) [4]. Since hue preserving is necessary for color image processing, intensity processing methods in HSI color system are reviewed [5],[6]. Next we show an intensity processing method while preserving hue in RGB color system which is proposed by Naik and Murthy [7]. In this method, intensity value can be controlled. However saturation value is decided depend on the processing of intensity. Saturation is always decreased by this method. Therefore the colorfulness of the output image is decreased compared to the input image. Next we show a new processing scheme in RGB color system which can preserve hue [8]. The method also only can control intensity value and saturation value cannot be controlled. Saturation value of the output is increased by this processing which is different from Naik's method. High contrast and high colorfulness image is can be obtained by this method.

2. MUNSELL COLOR SYSTEM -THE HUMAN VISUAL PERCEPTION OF COLOR

Albert Henry Munsell who was an American painter, teacher of art was the first to separate hue, value, and chroma into perceptually uniform and independent dimensions, and he was the first to systematically illustrate the colors in three-dimensional system which is depicted in the Munsell color tree (see in Fig.1). Three attributes of the Munsell color system is illustrated in Fig.2.

Munsell's system is based on rigorous measurements of human subjects' visual responses to color. Though several replacements for the Munsell system

Manuscript received on October 25, 2016.
Final manuscript received on October 25, 2016.
[1] The author is with Department of Computer Science, Tokyo City University, Japan, E-mail: ataguchi@tcu.ac.jp

have been invented, building on Munsell's foundational ideas including the International Commission on Illumination's CIELAB (L*a*b*) , the Munsell system has outlasted its contemporary color models.

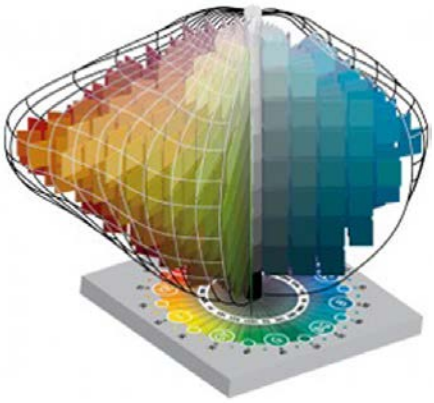


Fig.1: *Munsell color tree [9].*

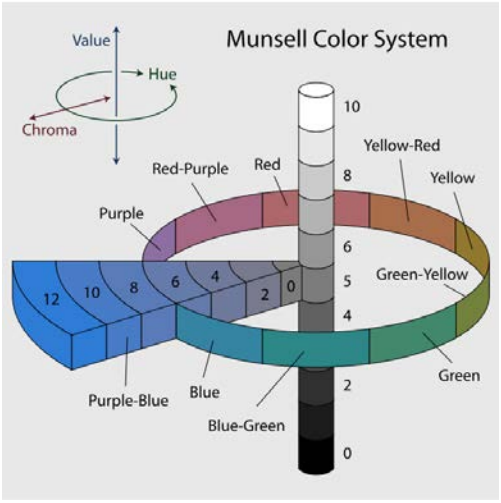


Fig.2: *Munsell color system [10].*

2.1 Munsell hue

Hue is that attribute of a color by which we distinguish red from green, blue from yellow, etc. Munsell hue circle is shown in Fig.3. Each horizontal circle Munsell divided into five principal hues: Red, Yellow, Green, Blue, and Purple, along with 5 intermediate hues (e.g., YR) halfway between adjacent principal hues. Each of these 10 steps, with the named hue given number 5, is then broken into 10 sub-steps, so that 100 hues are given integer values (see in Fig.3).

Two colors of equal Munsell value and Munsell chroma, on opposite sides of a hue circle, are complementary colors, and mix additively to the neutral gray of the same value.

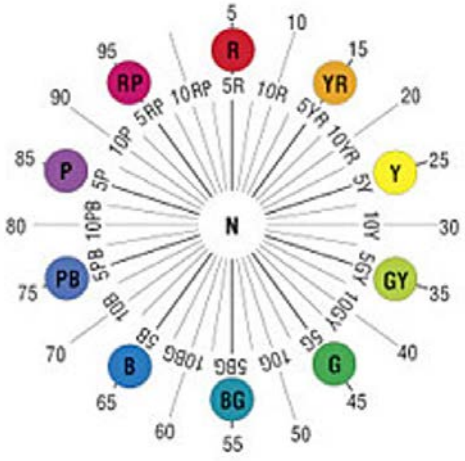


Fig.3: *Munsell hue [11].*

2.2 Munsell value

Munsell value varies vertically along the color solid, from black (value 0) at the bottom, to white (value 10) at the top and neutral grays lie along the vertical axis between black and white shown as Fig.2.

2.3 Munsell chroma

Munsell chroma shows the quality of a color's purity or saturation. For example, a gray color is an extreme low chroma. On the other hand, Fire-engine red may be a high-chroma red.

Note that there is no intrinsic upper limit to chroma. Different areas of the color system have different maximal chroma coordinates. For example, maximal chroma is 4 in the case of hue 5Y and value 3. On the other hand maximal chroma is 12 in the case of hue 5PB and value 6. This property makes the unique shape of the Munsell tree.

2.4 Munsell color system limit

The Munsell color system is the appeared color system. Therefore the color of color materials is described precisely in the Munsell color system. However, we cannot process the color image in the Munsell color system. Furthermore, the color cannot be evaluated in the Munsell color system.

In order to process the color image take the human visual system into account, the HSI system is used as a substitute for Munsell color system. The HSI color system will be explained later. On the other hand CIELAB is used as a substitute for Munsell color system in ordered to evaluate the color.

3. RGB COLOR SYSTEM AND COLOR SYSTEMS TRANSFORMED BY RGB COLOR SYSTEM

3.1 RGB color system -the additive color model

The RGB color model is an additive color model in which red, green and blue light are added together in various ways to reproduce a broad array of colors. The name of the model comes from the initials of the three additive primary colors, red, green and blue.

The main purpose of the RGB color model is for the sensing, representation and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Before the electronic age, the RGB color model already had a solid theory behind it, based in human sensing system by three cone cells of the human retina.

The values of RGB are assumed to be in the range of [0,1] or in some cases in the range of [0-255] (i.e., 8bit). This way black is represented as (0, 0, 0), white is represented as (1, 1, 1) or (255, 255, 255). These black and the white colors are represented by 2 of the opposite corner of the cube that can be defined by the R, G, B axes of the Cartesian coordinate systems shown in Fig.4. Other corners of the cube represent the red, green, blue, cyan, magenta and yellow colors. Grayscale colors are represented with identical R, G, B components. Because the RGB color system is widely used in monitors, digital cameras, it is the most important color system in image processing.

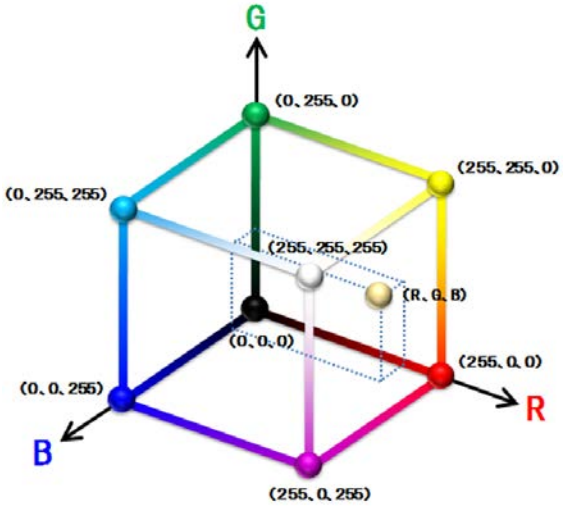


Fig.4: *RGB color system [12].*

represents every color with three components: hue (H), saturation (S), intensity (I). Figure 5 shows the ideal shape of the HSI color system.

The HSI color system is closely related to the Munsell color system. Hue, saturation and intensity correspond to Munsell hue, Munsell chroma and Munsell value, respectively. Three attributes of the HSI color system are same as those of the Munsell color system. However, the HSI color system is not a perceptually uniform color system unlike the Munsell color system.

The hue circle is shown in Fig.6. Three primary colors (i.e., red, green and blue) are located 120° intervals. The complementary color is also defined in the HSI hue. The complementary colors of red, green and blue are cyan, magenta and yellow, respectively. The location of three primary colors of HSI hue is different from the location of those of Munsell hue (see in Fig.3).

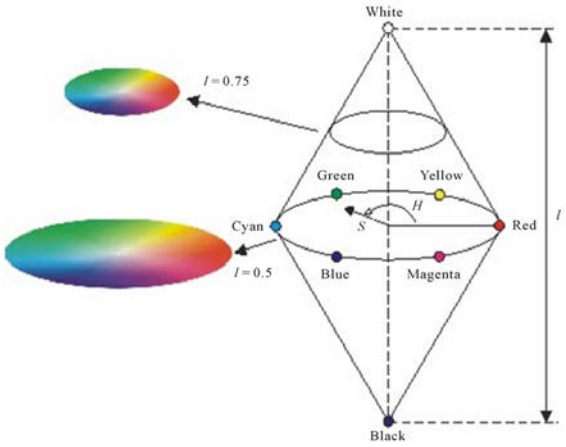


Fig.5: *Ideal HSI color system [1].*

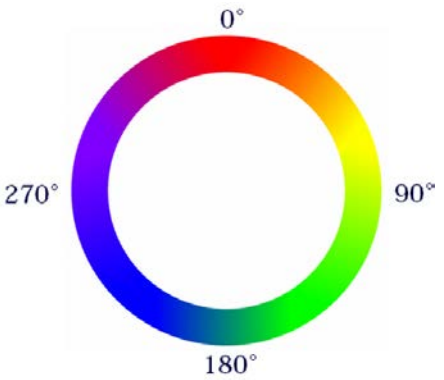


Fig.6: *HSI hue circle.*

3.2 Ideal HSI color system

The HSI color system is very important and attractive color system for image processing and its applications because it represents colors similarly how the human eye senses colors. The HSI color model

3.3 HSI color system

Real HSI color system will be introduced. A lot of HSI color systems have been proposed [1], [2]. In this paper, we adopt the HSI color system which is defined in Ref. [1], since the definition of hue and intensity is

appropriate. Furthermore, this system is regarded as the most common HSI color system which is employed in many researches [13]-[16]. We show the conversion formula between the RGB color system and the HSI color system.

(1) From RGB color system to HSI color system

It is assumed that the RGB values have been normalized to the range [0,1]. Intensity I is defined as

$$I = \frac{1}{3}(R+G+B) \tag{1}$$

The range of the intensity is [0,1]. The hue component is obtained as

$$H = \begin{cases} \theta & if\ B \le G \\ 360-\theta & if\ B > G \end{cases} \tag{2}$$

where

$$\theta = \cos^{-1}\left[\frac{\{(R-G)+(R-B)\}/2}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right]$$

The saturation component is given by

$$S = 1 - \frac{3}{(R+G+B)}\{\min(R,G,B)\} \tag{3}$$

The range of saturation is also from 0 to 1. In the case of (R,G,B)=(0,0,0), the denominator of second term of Eq.(3) becomes 0. Thus, we cannot calculate the equation. We define the saturation S as 0 in this case.

(2) From HSI color system to RGB color system

C_1, C_2 and C_3 are defined as shown in Table 1, then, the conversion formula from HSI color system to RGB color system is given as follows:

$$C_1 = I(1-S) \tag{4}$$

$$C_2 = I\left[1 + \frac{S\cos H}{\cos(60^\circ - H)}\right] \tag{5}$$

$$C_3 = 3I - (C_1 + C_2) \tag{6}$$

Table 1: *Conversion from HSI color system to RGB color system.*

H	0°-120°	120°-240°	240°-360°
	$H = H$	$H = H - 120$	$H = H - 240$
C_1	B	R	G
C_2	R	G	B
C_3	G	B	R

3.4 Relationship between the RGB color system and the HSI color system

The RGB and HSI color systems are shown in Fig.7. The diagonal which links point (1,1,1) to point (0,0,0) of the RGB color system is equivalent to the intensity (I) axis of the HSI color system. Seeing from point (1,1,1) (i.e., white) to (0,0,0) (i.e., black) the RGB color system, the RGB color system is shown as the regular hexagon such as Fig.8. The center point of the regular hexagon corresponds to the diagonal line from (1,1,1) to (0,0,0). The hue is defined as Fig.8. That is, red corresponds to 0° and green and blue correspond to 120° and 240°, respectively.

The SI planes of $H = 0°, 30°$ and 60° are illustrated in Fig.9. The SI plane of H= 0° is also drawn in Fig.7. A black triangle in Fig.9 shows a cross-section with the RGB color system. From Fig.9(a), we understood that the gamut of SI color plane of H= 0° is larger than that of RGB color system. Only in the condition of $I \le 1/3$, the HSI gamut and the RGB gamut is same each other. In $I > 1/3$, only R (i.e., C_2) component is larger than 1. Next, we study in the case of H=30°. Figure 9(b) shows the relationship between HSI and RGB color systems when H=30°. The common area of two color systems with H=30° get larger compared to the case of H=0°. In the condition of $I \le 1/2$, the gamut of HSI and RGB color systems is same each other. In $I > 1/2$, only R (i.e., C_2) component is larger than 1 same as the case of H=0°. Figure 9(c) shows the relationship between HSI and RGB color systems when H=60°. In this case, the common area of two color systems becomes maximum. In the condition of $I \le 2/3$, the gamut of two color systems is same each other. In the condition $I > 2/3$, both R (i.e., C_2) and G (i.e., C_3) are larger than 1.

The shape of common region of two color systems when H=90° is same as that when H=30°. Furthermore, the shape of common region of two color systems when H=120° is same as that when H=0°.

Next, we would like to make study about the volume of two color systems. The RGB color system is the cube whose length of a side is 1. Thus, the volume of the RGB color system is 1. On the other hand, the HSI color system is the triangular pyramid. The area of the bottom face is $9\sqrt{3}/2$ and the height is $\sqrt{3}$. The volume of the HSI color system is 9/2 and 4.5-times of the volume of the RGB color system.

From the study mentioned above, the width of common gamut of two color systems depends on hue value. The boundary condition that two color systems become common for intensity is shown in Fig.10

4. HSI COLOR SYSTEM WITH SAME GAMUT OF RGB COLOR SYSTEM [3]

We studied the relationship between RGB gamut and HSI gamut. The shape of HSI color system is a triangular pyramid and the HSI gamut is larger

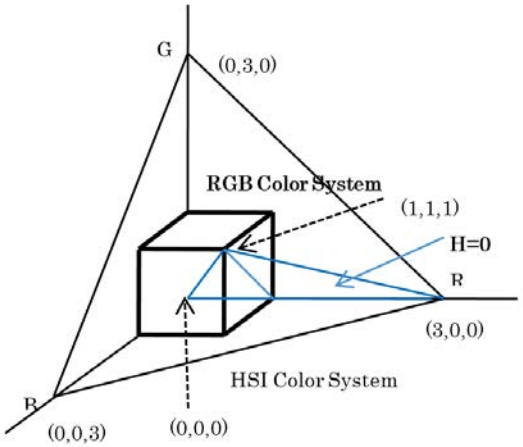


Fig.7: *RGB color system and HSI color system.*

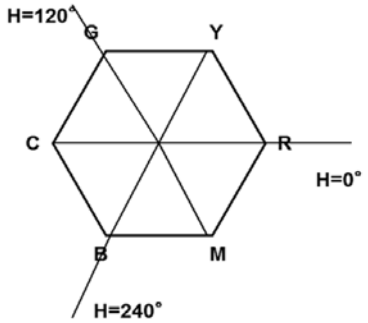


Fig.8: *Seeing the RGB color system from (1,1,1) to (0,0,0).*

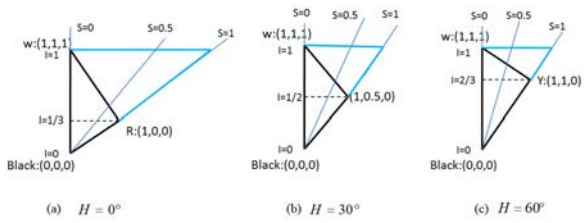


Fig.9: *SI planes of HSI color system with $H = 0°$, $30°$ and $60°$.*

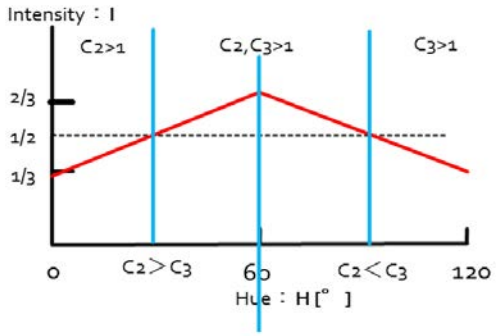


Fig.10: *The condition that two color systems become common.*

than the RGB gamut were made clear. Furthermore, we showed that the relationship between the gamut of two color systems is changed depend on the hue value. Then, the condition of intensity value that two color systems become common was also illustrated in Fig.10. In this chapter, a novel HSI system with same gamut of RGB color system is shown.

In order to introduce the novel conversion method, we utilize the CMY color system. We can also clarify the relations of the CMY color system and the HSI color system. The relationship between the CMY color system and the HSI color system is shown in Fig.11.

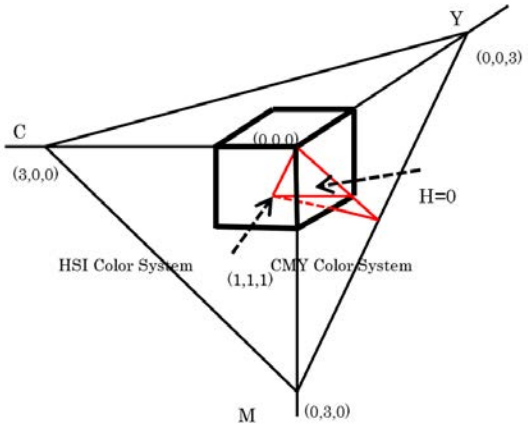


Fig.11: *CMY color system and HSI color system.*

The SI plane with $H = 0°$ which is converted from CMY color system is drawn by red line in Fig.11 and Fig.12. A blue triangle in Fig.12 is the SI plane which is converted from RGB color system. If we use the conversion rule between RGB color system and HSI color system in the condition $0 \leq I \leq 1/3$ and use the conversion rule between CMY color system and HSI color system in the condition $1/3 < I \leq 1$, HSI gamut become same as the RGB gamut. The switching value of intensity is changed depend on the hue value. The red line which is described in Fig.10 shows the switching value.

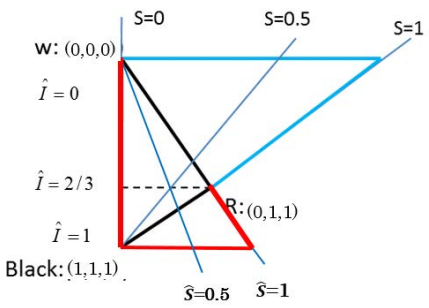


Fig.12: *SI planes of HSI color system with H=0° and cross-sections RGB and CMY systems (CMY coordinate system).*

A new conversion formula between RGB color system and HSI color system with same gamut of RGB color system is shown as follows.

(1) Conversion color systems from RGB to HSI

Given an image in RGB color format, the intensity component is given by Eq. (1) same as the conventional conversion formula.

We divide intensity range into two parts. First part is from 0 to the value on red line in Fig.10, and second part is from the value on red line to 1.

$$\text{(A)} \begin{array}{l} 0 \le I \le \left(\frac{\theta}{180} + \frac{1}{3}\right) : 0^\circ \le \theta < 60^\circ \\ 0 \le I \le \left(-\frac{\theta}{180} + 1\right) : 60^\circ \le \theta < 120^\circ \end{array}$$

The hue and saturation components are obtained by Eq.(2) and (3), respectively. These conversion rules are same as the conventional rules.

$$\text{(B)} \begin{array}{l} \left(\frac{\theta}{180} + \frac{1}{3}\right) < I \le 1 : 0^\circ \le \theta < 60^\circ \\ \left(-\frac{\theta}{180} + 1\right) < I \le 1 : 60^\circ \le \theta < 120^\circ \end{array}$$

The hue component is same as the case (A), thus, given by Eq. (2).

The saturation component is derived by using CMY components ($C = 1-R, M = 1-G, Y = 1-B$) as follow:

$$S = \hat{S} \tag{7}$$

where

$$\hat{S} = 1 - \frac{3}{(C + M + Y)}\{\min(C, M, Y)\}$$

The range of saturation is also from 0 to 1. In the case of (C,M,Y)=(0,0,0) (i.e., (R,G,B)=(1,1,1)), the denominator of second term of the above equation becomes 0. Thus, we define the saturation $\hat{S}$ as 0 in this case.

In our method, out-of-gamut problem is resolved since we give a new conversion method in the intensity condition (B).

(2) Conversion color systems from HSI to RGB

We show the conversion formula from the proposed HSI color system to RGB color system.

$$\text{(A)} \begin{array}{l} 0 \le I \le \left(\frac{\theta}{180} + \frac{1}{3}\right) : 0^\circ \le \theta < 60^\circ \\ 0 \le I \le \left(-\frac{\theta}{180} + 1\right) : 60^\circ \le \theta < 120^\circ \end{array}$$

In this case, the RGB components are given by the equations which are same as the conventional HSI color system (i.e., Eq. (4)-(6)).

$$\text{(B)} \begin{array}{l} \left(\frac{\theta}{180} + \frac{1}{3}\right) < I \le 1 : 0^\circ \le \theta < 60^\circ \\ \left(-\frac{\theta}{180} + 1\right) < I \le 1 : 60^\circ \le \theta < 120^\circ \end{array}$$

In the case of (B), we consider CMY color system instead of RGB color system. Thus, the hue is shifting 180° as follow:

$$\hat{H} = \begin{cases} H + 180^\circ & if\ 0 \le \theta < 180^\circ \\ H - 180^\circ & if\ 180 \le \theta < 360^\circ \end{cases} \tag{8}$$

We define the $\hat{S}$ and $\hat{I}$ as follows:

$$\hat{S} = S \tag{9}$$

$$\hat{I} = 1 - I \tag{10}$$

The relationship between RGB and CMY is shown as: $C = 1-R, M = 1-G, Y = 1-B$. C_1, C_2 and C_3 are derived by using $\hat{H}, \hat{S}$ and $\hat{I}$ as follows:

$$C_1 = 1 - \hat{I}(1 - \hat{S}) \tag{11}$$

$$C_2 = 1 - \hat{I}\left[1 + \frac{\hat{S}\cos\hat{H}}{\cos(60^\circ - \hat{H})}\right] \tag{12}$$

$$C_3 = 3I - (C_1 + C_2) \tag{13}$$

where the definition of C_1, C_2 and C_3 are same as Table 1.

5. COLOR IMAGE ENHANCEMENT IN THE HSI COLOR SYSTEM

Hue, saturation and intensity are the attributes of color. Hue is that attribute of a color which decides what kind color it is, i.e., a red or an orange. Saturation shows the vividness of color and has effect on the impression of the image. For the purpose of enhancing a color images, it is to be seen that hue value should not change for any pixel. If hue value is changed, thereby distorting the image. For general hue-preserving color image processing, the original RGB image is usually transformed to another color system such as HSI or HSV, then, the intensity /saturation components is processed, but the hue component remains unchanged. We would like to introduce color image enhancement methods in HSI color system in this chapter.

5.1 Saturation Correction Algorithm for solving the gamut problem

In 3.4, we showed the RGB gamut and the HSI gamut is different. Therefore, there is a risk that the processed pixel is located outside of the gamut of RGB color system. According to the relationship between two color systems discussed 3.4, we can see that the processed pixel is located whether outside or inside of RGB color system without converting from HSI color system to RGB color system. New conversion formula from HSI color system to RGB color system with preserving hue and intensity is introduced.

In the processed pixel is located outside of RGB color system, saturation correction is carried out in order to move the pixel inside RGB color system.

5.1.1 New conversion formula from HSI color system to RGB color system [17]

We show the conversion procedure from HSI color system to RGB color system with the saturation correction algorithm as follows:
[Case 1] $0° \leq H < 60°$
C_2 is calculated by Eq. (5).
(1-a) $C_2 \leq 1$
C_1 and C_3 are also smaller than 1. Thus, C_1 is given by Eq.(4) and C_3 is given by Eq.(6).
(1-b) $C_2 > 1$
In this case, the pixel is located outside RGB color system and saturation correction is necessary. Saturation value is revised to make the value of $C_2 = 1$. From Eq.(5), the corrected saturation is given by

$$\hat{S} = \frac{(1-I)\cdot\cos(60° - H)}{I\cdot\cos H} \tag{14}$$

C_1 and C_3 are derived by using the corrected saturation value.

$$C_1 = I(1-\hat{S}) \tag{15}$$

$$C_3 = I\left[1 + \frac{\hat{S}(\cos(60° - H) - \cos H)}{\cos(60° - H)}\right] \tag{16}$$

[Case 2] $60° \leq H < 120°$
C_3 is calculated by Eq. (6).
(2-a) $C_3 \leq 1$
C_1 and C_2 are also smaller than 1. Thus, C_1 is given by Eq.(4) and C_2 is given by Eq.(5).
(2-b) $C_3 > 1$
In this case, saturation correction is necessary. Saturation value is revised to make the value of $C_3 = 1$. From Eq.(6), the corrected saturation is given by

$$\hat{S} = \frac{(1-I)\cdot\cos(60° - H)}{I\cdot\cos(60° - H) - \cos H} \tag{17}$$

C_1 and C_2 are derived by using the corrected saturation value.

$$C_1 = I(1-\hat{S}) \tag{18}$$

$$C_2 = I\left[1 + \frac{\hat{S}\cos H}{\cos(60° - H)}\right] \tag{19}$$

5.1.2 Experimental Results

We apply the HE to intensity component and saturation component independently for color image enhancement. We compare the transformation from HSI color space to RGB color space with saturation correction to without saturation correction. In the case of the transformation without saturation correction, the clipping operation is applied to transformed pixels which are located outside of the RGB gamut. The clipping causes the hue shifting.

Two images are prepared for enhancement and these images are shown in Fig.13. These images are 8-bit for each component.

Enhancement results without saturation correction are shown in Fig.14. On the other hand, enhancement results with saturation correction are shown in Fig.15.

First we give a qualitative analysis on enhancement results of "old town". We can easily distinguish that the most different section among these images is the house located at right side. Obviously, there is a great hue distortion in Fig.14(b). On contrast, Fig.15(b) is the result of proposed algorithm which has a perfect hue matching the original image (i.e., Fig.13(b)) and the intensity and saturation is properly enhanced. From Fig.14(a), we understand that the hue of the skin becomes red. The proposed saturation correction method is effective for color image processing in HSI color space.

Next, we perform the numerical evaluation. There are 24% pixels of "couple" is necessary for saturation correction. There are 27% pixels of "old town" need to be modified. In the case of color image enhancement in HSI color space, the gamut problem is often occurred and the correction algorithm is necessary for preserve hue components.

(a) couple (b) old town

Fig.13: *Test images (1).*

5.2 Differential histogram equalization (DHE) for color images

5.2.1 DHE for gray-scale images

Histogram equalization is the most famous image enhancement method. On the other hand DHE which is a modified method of HE is not popular. Therefore the DHE is explained briefly as follow.

Consider an input image $f(i,j)$, which is a the

(a) couple (b) old town

Fig.14: *Enhancement results without saturation correction.*

(a) couple (b) old town

Fig.15: *Enhancement results with saturation correction.*

total number of N pixels with gray-levels in the range $[0, L-1]$. We calculate the differential gray-levels of the input image as follows:

$$d(i,j) = \text{int}\left\{\sqrt{d_H(i,j)^2 + d_V(i,j)^2}\right\} \tag{20}$$

where

$$d_H(i,j) = \{f(i+1,j+1)+2\cdot f(i+1,j)+f(i+1,j-1)\} - \{f(i-1,j+1)+2\cdot f(i-1,j)+f(i-1,j-1)\}$$

$$d_V(i,j) = \{f(i+1,j+1)+2\cdot f(i,j+1)+f(i-1,j+1)\} - \{f(i+1,j-1)+2\cdot f(i,j-1)+f(i-1,j-1)\}$$

int{} in Eq.(1) represents the integer transform processing.

The differential gray-level histogram (DH) $h_d(r)$ is given by

$$h_d(r) = \sum_{(i,j)\in D_r} d(i,j) \tag{21}$$

where D_r is a region composed of pixels whose value is r. Thus, the horizontal axis of DH is gray-level r and the vertical axis is the total differential gray-levels of (i,j) points which meet the condition $f(i,j) = r$.

The DHE will map an input gray-level r into an output gray level s using the following transformation function $T(r)$.

$$s = T(r) = (L-1)\cdot c(r) \tag{22}$$

where

$$c(r) = \frac{\sum_{k=0}^{r} h_d(k)}{\sum_{k=0}^{L-1} h_d(k)}$$

5.2.2 Intensity DHE (IDHE) for color images

First, the input image which is described RGB color system transforms to HSI color system. $I(i,j)$ is assumed to be the L-level intensity component of the input color image. The output of IDHE $O_I(I(i,j))$ is given from Eq.(20) to Eq.(22) by replacing $f(i,j)$ with $I(i,j)$.

5.2.3 Saturation DHE (SDHE) for color images

In color images another differential histogram which is the saturation differential histogram (SDH) can be defined. $S(i,j)$ is assumed the saturation component of input color image. $S(i,j)$ is expanded to L-level.

The differential of saturation $d_s(i,j)$ is derived by using Eq.(20) by replacing $f(i,j)$ with $S(i,j)$. The SDH is given by

$$h_d^{S_{(i,j)}}(r) = \sum_{(i,j)\in D_r^{I_{i,j}}} \{|\rho(i,j)|\cdot d_s(i,j)\} \tag{23}$$

where

$$\rho(i,j) = \frac{\sum_{k=-2}^{2}\sum_{l=2}^{2}[I\{(i+kj+l)-\bar{I}(i,j)\}\{S(i+kj+l)-\bar{S}(i,j)\}]}{\sqrt{\sum_{k=-2}^{2}\sum_{l=2}^{2}\{I(i+kj+l)-\bar{I}(i,j)\}^2}\sqrt{\sum_{k=2}^{2}\sum_{l=2}^{2}\{S(i+kj+l)-\bar{S}(i,j)\}^2}}$$

D_r is a region composed of pixels whose intensity value is r (i.e., $I(i,j) = r$). In the case of SDHE, we also emphasize the intensity component of color images by using saturation information. The saturation information is generally not appropriate at locations where the component images do not have same basic image structure. Thus, we introduce $\rho(i,j)$ which is local correlation of intensity and saturation data, in order to measure the structural similarity. The magnitude of $\rho(i,j)$ determines how appropriate the saturation data is at a given location. In uncorrelated

regions, since $|\rho(i,j)| \cdot d_s(i,j)$ shows small value, it does not contribute to the Eq.(23). $\bar{I}(i,j)$ and $\bar{S}(i,j)$ are local mean of intensity and local mean of saturation, respectively. We set 5×5 window for calculating the local mean in this paper.

The output of SDHE $O_S(I(i,j))$ is given by Eq.(22) by replacing $h_d(r)$ with $h_d^{S_{i,j}}(r)$.

5.2.4 The combination of IDHE and SDHE

The output $O(I(i,j))$ of the combination of two DHEs is derived easily as follow.

$$O(I(i,j)) = \alpha \cdot O_s(I(i,j)) + (1-\alpha) \cdot O_I(I(i,j)) \quad (24)$$

where $0 \leq \alpha \leq 1$. The output of the combination method with $\alpha = 0$ and $\alpha = 1$ are equivalent to IDHE and SDHE, respectively.

5.2.5 Experimental Results

We prepare two test images with RGB 24-bit shown in Fig.16. We demonstrate how the results of the combination method changed, when parameter α is changed. Enhancement results with $\alpha = 0$, 0.2, 0.4, 0.6 0.8 and 1 of two test images are shown in Fig.17 and Fig.18.

(a) bridge (b) cherry blossom

Fig.16: *Test images (2).*

Subjective evaluation is performed by 20 men and women in its twenties. Subjects choose their favorite one enhancement image from 11 enhancement images with from 0 to 1 at 0.1 intervals. The result is shown in Table 2. In the case of "bridge", the average value of is 0.435. In the other hand, the average of is 0.525 in "cherry blossom". The average values of two images are different each other. It is necessary to give a decision guideline of depend on the original image which is agree in human sense.

Table 2: *Subjective evaluation (Number of persons).*

α	0.3	0.4	0.5	0.6	0.7
bridge	2	10	7	1	0
cherry	0	3	10	6	1

In order to evaluate the contrast and colorfulness of images, we introduce two variance-based indexes in CIELAB color system. The contrast of color images corresponds to the width of $L*$ axial color signal distribution. Thus, the contrast is quantified by the variance value of $L*$ component of color signals which is the first index of the color image quality evaluation. The colorfulness can be evaluated by the value based on signal variance on $a*b*$ plane. M1 [17] is introduced as the second index to evaluate the colorfulness of the color image as follow.

$$M1 = \sigma_{ab} + 0.37\mu_{ab} \quad (25)$$

where σ_{ab} and μ_{ab} are the standard deviation and average of the color image on $a*b*$ plane, respectively.

We plot of two indexes vs. α in Fig.19. Two straight lines in each graph show the index value of the original test image. In the case of "bridge", $M1$ of all enhancement results is larger than that of original "bridge" image. On the other hand, L* variance of enhancement results with α from 0 to 0.5 is larger than that of original image. It is preferable for two indexes of enhancement result to be larger than those of the original image. From these points of view, it is preferable for α to set smaller than 0.5. We study about "cherry blossom" in the same way, it is preferable for α to set larger than 0.3.

From Table 2, 95% of subjects chose α smaller than 0.5 for "bridge" and 100% of subjects chose α larger than 0.3 for "cherry blossom".

We come to see that parameter can be decided by the following guideline.

(1) Deciding α to meet the condition that two indexes of the enhancement result are larger than those of original image.

(2) It is preferable to decide around 0.5, in order to use two DHEs effectively.

6. COLOR IMAGE ENHANCEMENT IN THE RGB COLOR SYSTEM

It is necessary to preserve the hue before/after the color image enhancement. The color image is captured by RGB color system. Therefore, it is preferable to process the color image in RGB color system. However, if thought simply, it is considered that hue preserving processing is difficult in RGB color system. Nevertheless two processing schemes in the RGB color system which can be preserved hue are shown.

6.1 Hue-preserving processing scheme proposed by Naik and Murthy.

Naik and Murthy proposed a scheme to generalize any gray-scale contrast enhancement techniques to color images with hue preserving in RGB color system. Therefore this scheme frees from the gamut problem.

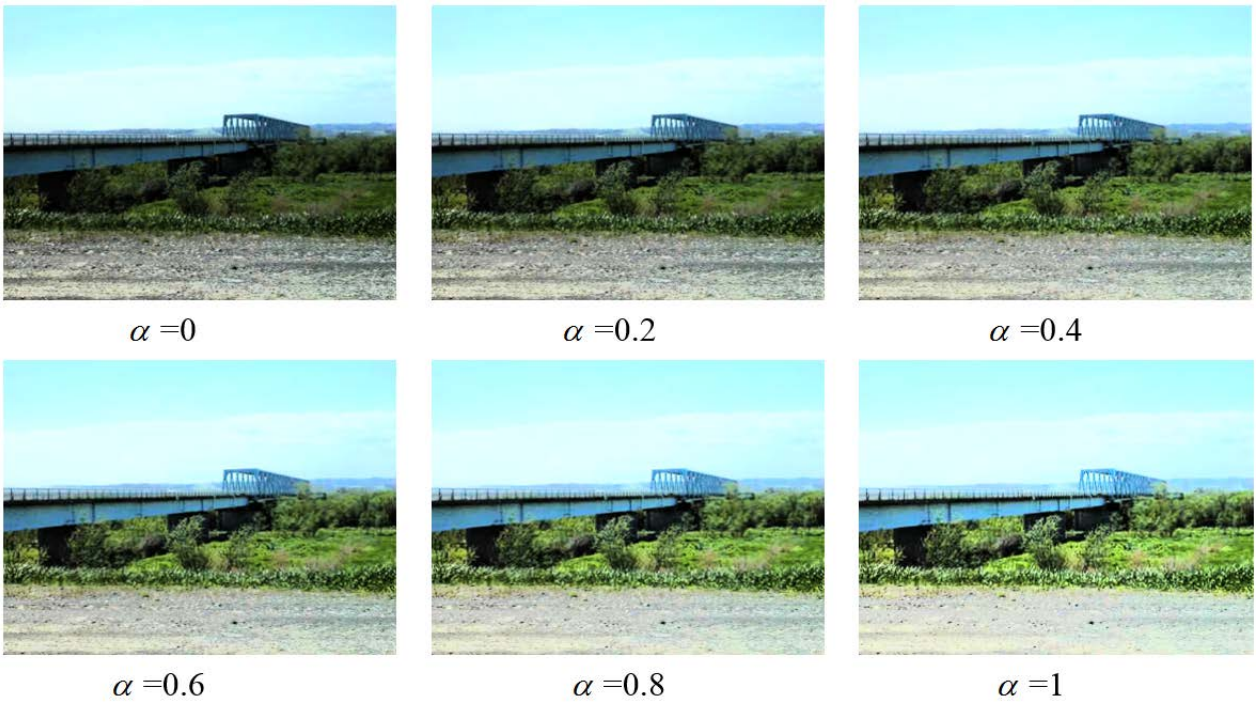
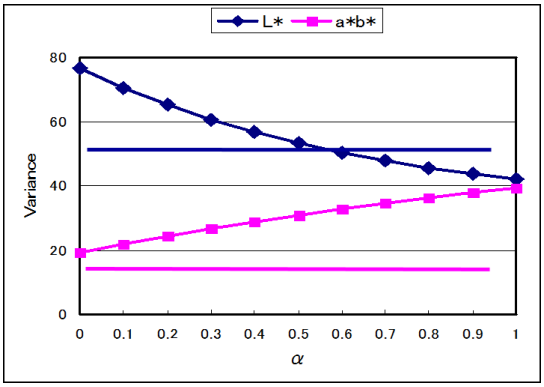
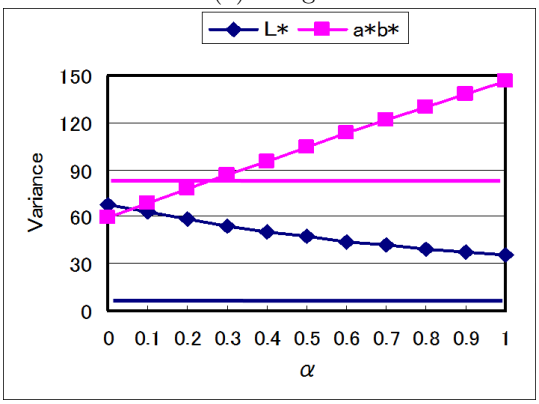


Fig.17: *Enhancement results (bridge).*

Fig.18: *Enhancement results (cherry blossom).*



(a) bridge



(b) cherry blossom

Fig.19: *Two indexes variances vs.α (L∗: L∗ variance, a ∗ b∗ : M1).*

Let the RGB element of the original input image be (R_{in}, G_{in}, B_{in}) and the RGB element of enhancement image be $(R_{out}, G_{out}, B_{out})$. The RGB element is assumed to be normalized. And we define l_{in} as

$$l_{in} = (R_{in} + G_{in} + B_{in})/3 \tag{26}$$

Let "f" denote the arbitrary gray-scale transfer function. This will be deal with in two separate cases.
[Case 1] $\alpha(l_{in}) = f(l_{in})/l_{in} \leq 1$

$$R_{out}{=}\alpha(l_{in}) \cdot R_{in},\ G_{out}{=}\alpha(l_{in}) \cdot G_{in},\ B_{out}{=}\alpha(l_{in}) \cdot B_{in} \tag{27}$$

[Case 2] $\alpha(l_{in}) > 1$

In this case, first transform the RGB color element (R_{in}, G_{in}, B_{in}) to CMY color element (C_{in}, M_{in}, Y_{in})

$$C_{in} = 1 - R_{in}, M_{in} = 1 - G_{in}, Y_{in} = 1 - B_{in} \tag{28}$$

We define l_{in}^{CMY} as

$$l_{in}^{CMY} = (C_{in} + M_{in} + Y_{in})/3 = 1 - l_{in} \tag{29}$$

The output is given by

$$C_{out}{=}\beta(l_{in}^{CMY}){\cdot}C_{in}, M_{out}{=}\beta(l_{IN}^{CMY}){\cdot}M_{in} Y_{out}{=}\beta(l_{in}^{CMY}) \cdot Y_{in} \tag{30}$$

where

$$\beta(l_{in}^{CMY}) = g(l_{in}^{CMY})/l_{in}^{CMY}$$

$$g(l_{in}^{CMY}) = 1 - f(l_{in})$$

Finally, we transform the output signal from CMY color system to RGB color system.

In this scheme, when the condition $\alpha(l_{in}) > 1$ is satisfied, the color signals of RGB space is transformed to CMY space. Therefore, each element of resulting images (i.e., $R_{out}, G_{out}, B_{out}$) should not exceed 1. However, the saturation value of enhancement images is always decreased.

6.2 Hue-Preserving Processing Scheme with rich saturation

We propose a novel scheme to generalize any gray-scale image processing to color images, which is the extended version of the Naik's scheme.
[Case 1]

$$(1{-}a)\ \alpha(l_{in}) \leq 1\ and\ \beta(l_{in}^{CMY}) \leq 1/\max\{C_{in}, M_{in}, Y_{in}\}$$

Enhancement processing is carried out in CMY color system.

$$C_{out}{=}\beta(l_{in}^{CMY}) \cdot C_{in},\ M_{out}{=}\beta(l_{in}^{CMY}) \cdot M_{in},\ \ Y_{out}{=}\beta(l_{in}^{CMY}) \cdot Y_{in} \tag{31}$$

Next CMY elements are converted into RGB elements.

$$(1{-}b)\ \alpha(l_{in}) \leq 1\ and\ \beta(l_{in}^{CMY}) > 1/\max\{C_{in}, M_{in}, Y_{in}\}$$

In this case, the output signal is derived by two steps.
Step 1)

$$\hat{C}_{in} = \gamma(l_{in}^{CMY}) \cdot C_{in},\ \ \hat{M}_{in} = \gamma(l_{in}^{CMY}) \cdot M_{in},\ \ \hat{Y}_{in} = \gamma(l_{in}^{CMY}) \cdot Y_{in} \tag{32}$$

where

$$\gamma(l_{in}^{CMY}) = 1/\max\{C_{in}, M_{in}, Y_{in}\}$$

$$\hat{R}_{in} = 1 - \hat{C}_{in},\ \ \hat{G}_{in} = 1 - \hat{M}_{in},\ \ \hat{B}_{in} = 1 - \hat{Y}_{in}$$

Step 2)

$$\hat{l}_{in} = (\hat{R}_{in} + \hat{G}_{in} + \hat{B}_{in})/3$$

$$R_{out} = \hat{\alpha}(l_{in}) \cdot \hat{R}_{in}, G_{out} = \hat{\alpha}(l_{in}) \cdot \hat{G}_{in}, B_{out} = \hat{\alpha}(l_{in}) \cdot \hat{B}_{in} \quad (33)$$

where

$$\hat{\alpha}(l_{in}) = f(l_{in})/\hat{l}_{in}$$

[Case 2]

$(2-a)$ $\alpha(l_{in}) > 1$ *and* $\alpha(l_{in}) \leq 1/\max\{R_{in}, G_{in}, B_{in}\}$

$$R_{out} = \alpha(l_{in}) \cdot R_{in},\ G_{out} = \alpha(l_{in}) \cdot G_{in},\ B_{out} = \alpha(l_{in}) \cdot B_{in} \quad (34)$$

$(2-b)$ $\alpha(l_{in}) > 1$ *and* $\alpha(l_{in}) > 1/\max\{R_{in}, G_{in}, B_{in}\}$

Step 1)

$$\hat{R}_{in} = \delta(l_{in}) \cdot R_{in},\ \hat{G}_{in} = \delta(l_{in}) \cdot G_{in},\ \hat{B}_{in} = \delta(l_{in}) \cdot B_{in} \quad (35)$$

where

$$\delta(l_{in}) = 1/\max\{R_{in}, G_{in}, B_{in}\}$$

Step 2)

$$\hat{C}_{in} = 1 - \hat{R}_{in},\quad \hat{M}_{in} = 1 - \hat{G}_{in},\quad \hat{Y}_{in} = 1 - \hat{B}_{in}$$

$$\hat{l}_{in}^{CMY} = (\hat{C}_{in} + \hat{M}_{in} + \hat{Y}_{in})/3$$

$$C_{out} = \hat{\beta}(l_{in}^{CMY}) \cdot \hat{C}_{in},\quad M_{out} = \hat{\beta}(l_{in}^{CNY}) \cdot \hat{M}_{in},\quad Y_{out} = \hat{\beta}(l_{in}^{CMY}) \cdot \hat{Y}_{in} \quad (36)$$

where

$$\hat{\beta}(l_{in}^{CMY}) = g(l_{in}^{CMY})/\hat{l}_{in}^{CMY}$$

The output signal transformed into RGB color system.

$$R_{out} = 1 - C_{out},\quad G_{out} = 1 - M_{out},\quad B_{out} = 1 - Y_{out}$$

We illustrate the principle of the proposed scheme based on comparisons with the scheme which is proposed by Naik and Murthy. In order to explain clearly, we take a SI plane (i.e., equal hue plane) for example.

We show an example of [Case 2](2-b). The line which is connected (0,0,0) and (1,1,1) is intensity axis. Therefore, the saturation gets larger as the distance from intensity axis. Points "A" shown in Fig.20 is the input signal. Points "B" and "C" shown in Fig.20 mean the result of Naik's scheme and the proposed scheme, respectively. In the scheme of Naik and Murthy, the output signal is located on the line which connected input signal and (1,1,1). Thus, the saturation of the output signal is decreased. On the other hand, in our scheme, first, the input signal is translated to the boundary of the RGB color system, next, translated on the line which connected the boundary point signal and (1,1,1). The saturation of output signal is always higher than that of the output of Naik's scheme. Furthermore, the output signal of our scheme is also always located inside of the RGB system as is the case of the Naik's scheme.

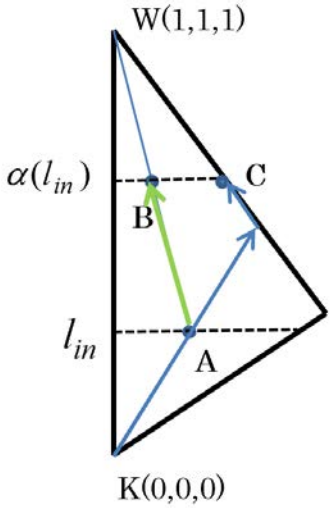


Fig.20: *Our scheme compared to Naik's scheme in the case of [Case 2](2-b) (A: Input B: Naik's result C: Our result)).*

6.3 Experimental Results

Two color images, "balloon" and "airplane" are used for simulation. HE is used as the transformation function $f()$. We compare our method to Naik's method.

Fig.21 and Fig.22 show the original image and the output images of 2 methods on "balloon" and "airplane", respectively. Two methods can preserve the hue component. Since HE is applied the contrast is improved efficiently by both methods. The intensity of both output images is same. However, the colorfulness of two outputs is different. Because of the saturation of the proposed method is larger than that of Naik's method for all pixels. The enhancement results of the proposed scheme show more colorful than those of Naik's method.

7. CONCLUSIONS

First half of this paper color systems are reviewed. Color systems are divided into two categories. One is the color appearance system and the other is the

(a) Original (b) Naik's method (c) Our method

Fig.21: *Enhancement results (balloon).*

(a) Original (b) Naik's method (c) Our method

Fig.22: *Enhancement results (airplane).*

additive color system. Munsell color system which is belonged to the color appearance system and RGB and HSI color systems which are belonged to the additive color system are reviewed. Munsell color system is constructed based on the human visual system. It is important point that HSI color system is closely related to Munsell color system. However the conventional HSI color system has the gamut problem. Therefore HSI color system with same gamut of RGB color system is introduced.

The second part of this paper color image enhancement methods are reviewed. It is necessary to preserve the hue before/after the color image enhancement. Enhancement in HSI color system is effective to preserve hue. We introduce high quality enhancement methods in HSI color system. Finally we show two processing schemes in RGB color system which can preserve the hue. These schemes to generalize any gray-scale contrast enhancement techniques to color images. These methods cannot control the saturation value, however, free from the gamut problem.

References

[1] R.C. Gonzalez and R.E. Woods, *Digital Image Processing* 3rd ed., Pearson Education, Inc., 2008.

[2] https://en.wikipedia.org/wiki/HSL_and_HSV

[3] M. Kamiyama and A. Taguchi, "HSI color space with same gamut of RGB color space," *IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences*, Vol.E100-A, No.1, 2007 (to be appeared).

[4] F. Saitoh, "Image contrast improvement based on differential gray-levels histogram," *IEEJ Trans. EIS*, vol.126, no.2, pp.228-236, Feb. 2006 (in Japanese).

[5] K. Nakai, Y.Hoshi and A. Taguchi, "Color image contrast enhacement method based on differential intensity/saturation gray-levels histograms," *Proc. of 2013 International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS)*, pp. 445 - 449, Nov. 2013.

[6] N. Nakajima and, A. Taguchi, "Intensity/Saturation enhancement based on differential gray-levels histogram equalization," *Proc. 2014 IEEE Asia Pacific Conference on Circuits and Systems*, pp. 527-530, 2014.

[7] S.K. Naik and C.A. Murthy, "Hue-preserving color image enhancement without gamut problem," *IEEE Trans. on Image processing*, vol.12, no.12, pp.1591-1598, December 2003.

[8] K. Murahira and A. Taguchi "Hue-preserving color image enhancement in RGB color space

with rich saturation," *Proc. of 2012 International Symposium on Intelligent Signal Processing and Communication Systems*, pp.266-269 (2012).
[9] https://notanothergardeningblog.com/tag/munsell-colour-system/
[10] https://en.wikipedia.org/wiki/Munsell_color_system#/media/File:Munsell-system.svg
[11] https://jp.pinterest.com/ythomasetyparis/the-munsell-color-system/
[12] http://imagingsolution.blog107.fc2.com/blog-entry-247.html
[13] L. Gao and X. Ling, "Color edge detection based on mathematical morphology in HSI space," *Proc. of 2010 International Conference on Computer and Information Application (ICCIA)*, pp.5-8, 2010.
[14] L.J. Ubong, S.T. Chee, and W.N. Giap, "A comparison of RGB and HSI color segmentation in real - time video images: A preliminary study on road sign detection," *Proc. of International Symposium on Information Technology (ITSim 2008)*, vol.4, pp.1-6, 2008.
[15] J. Zengwei, C. Jiazhong and Z. Jingli, "Image segmentation based on the HSI color space and an improved mean shift," *Proc. of International Conference on Information and Communications Technologies (IETICT 2013)*, pp.135-140, 2013.
[16] D. Yu, L-H. Ma and H-Q. Lu, "Normalized SI correction for hue-preserving color image enhancement," *Proc. of the Sixth International Conference on Machine Learning and Cybernetics*, pp.1498-1503, August 2007.
[17] A. Taguchi and Y. Hoshi, "Color image enhancement in HSI color space without gamut problem," *IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences*, vol.E98-A, no.2, pp.792-795, 2015.
[18] D. Hasler and S. Susstrunk, "Measuring colourfulness in natural images," *Proceedings of SPIE*, Vol 5007, pp.87-95, jun. 2003.

Akira Taguchi received the B.E., M.E. and Dr. Eng. Degrees in electrical engineering from Keio University, Yokohama, in 1984, 1986, 1989 respectively. In 1989, he joined Musashi Institute Technology (the old name of Tokyo City University). He is currently a Professor at the Department of Computer Science, Tokyo City University. From 1993 to 1994 he spent one year at the Signal Processing Laboratory, Tampere University of Technology, Finland as a visiting Researcher. His research interests are in the areas of digital image processing including nonlinear signal processing and color image processing. He served General Chair for IEEE ISPACS 2013. He is currently a Vice Chair of International Steering Committee of ISPACS. He is a senior member of IEEE and a fellow of IEICE.

A Copyright- and Privacy-Protected Diabetic Retinopathy Diagnosis Network

Wannida Sae-Tang[1], Non-member

ABSTRACT

This paper proposes a copyright- and privacy-protected diabetic retinopathy (DR) diagnosis network. In the network, DR lesions are automatically detected from a fundus image by firstly estimating non-uniform illumination of the image, and then the lesions are detected from the balanced image by using level-set evolution without re-initialization. The lesions are subsequently marked by using contours. The lesion-marked fundus image is subsequently shared for intra or inter hospital network diagnosis with copyright and privacy protection. Watermarking technique is used for image copyright protection, and visual encryption is used for privacy protection. Sign scrambling of two dimensional (2D) discrete cosine transform (DCT) and one dimensional (1D) DCT is proposed for lesion-marked fundus image encryption. The proposed encryption methods are compared with other transform-based encryption methods, i.e., discrete Fourier transform (DFT) amplitude-only images (AOIs), DCT AOIs, and JPEG 2000-based discrete wavelet transform (DWT) sign scrambling which were proposed for image trading system. Since the encryption is done after DR diagnosis, contours used for DR marking must also be visually encrypted. The proposed encryption methods are effective for strong-edge images that are suitable for lesion-marked fundus images, while random sign-based JPEG 2000, DFT AOIs, and DCT AOIs encrypt the images imperfectly. Moreover, the proposed methods are better in terms of image quality. In addition, watermarking performance and compression performance are confirmed by experiments.

Keywords: Diabetic Retinopathy, Fundus Image, Copyright Protection, Privacy Protection, Network Diagnosis

1. INTRODUCTION

DR is a major cause of vision loss. Early detection of DR helps reduce the risk of blindness [1]. However, it is not an easy task to detect the DR precisely. The earliest sign of DR is broken blood vessels. Exudates, a sign of DR, is lipids leaked form broken blood vessels. The characteristics of exudates are bright and yellowish with arbitrary shapes and sizes. For hard exudates, the boundaries are well-defined. On contrary, for soft exudates or cotton wool spots, boundaries are not clear. In some cases, intra or inter hospital network diagnosis is required for diabetic retinopathy diagnosis. In those networks, diagnosis results are needed to be confirmed, or cooperated diagnosis is required. Security of retinal image storage and transmission is then a task of this paper. Reversible watermarking for fundus images was proposed for image copyright protection in [2]. In [3], watermark is encrypted before embedding to medical images for reliable and robust transmission. Anyway, image visual encryption was not considered for fundus images and lesion-marked fundus images.

In copyright-and privacy-protected image trading system [4], the AOI which is the inversely transformed spectra of an image was used as image visual encryption. In the system, an image is transformed using DFT. Then, amplitude components are extracted. Finally, inverse transform is applied to amplitude components to obtain the AOI. This method is quite complex, since intensity range (IR) reduction of the AOI was required [4, 5]. In [6], JPEG 2000-based DWT sign scrambling was proposed in order to encrypt-then-compress the image. However, JPEG 2000 is still not widely used, and the visual encryption performance of the method is not good enough for strong edge-images especially for lesion-marked fundus images.

This paper proposes a copyright- and privacy-protected DR diagnosis network. In the network, DR lesions are automatically detected from a fundus image. Then, the lesion-marked image is shared for intra or inter hospital network diagnosis with copyright and privacy protection. Watermarking technique is used for image copyright protection, and sign scrambling of 2D DCT and 1D DCT is proposed for lesion-marked fundus image encryption for privacy protection. The proposed encryption methods are compared with other transform-based encryption methods, i.e., DFT AOIs, DCT AOIs, and JPEG 2000-based DWT sign scrambling. Visual encryption performance, encrypted image quality, security, complexity, watermarked image quality, watermark extracting performance, and image compression performance are considered in this paper.

Manuscript received on December 8, 2015 ; revised on June 1, 2016.

Final manuscript received on June 14, 2016.

[1] The author is with The Sirindthorn International Thai-German Graduate School of Engineering, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand. E-mail: wannida.s.sse@tggs-bangkok.org

The rest of this paper is organized as follows. Section 2 introduces DR, fundus images, and requirements for copyright- and privacy-protected DR diagnosis network. Section 3 describes the proposed network and the proposed encryption methods. Experimental results and discussions are given in Section 4. Finally, Section 5 concludes this paper.

2. PRELIMINARIES

This section gives basic knowledge of DR and fundus images. In addition, requirements for copyright- and privacy-protected DR diagnosis network is described.

2.1 Diabetic Retinopathy and Fundus Images

The number of diabetic patients has grown rapidly in the last few years and is expected to raise in the future, since people have the longevity up. The longer a person has diabetes, the more likely they will develop DR. If left untreated, DR may cause human vision loss [7]. Therefore, the early detection of DR is very important. Signs of the early stage of DR are capillary aneurysms, micro-aneurysms, exudates, hemorrhages, and cotton wool spots. In the next stage, abnormal blood vessels are found.

Figure 1 shows anatomical parts of human eye [8]. Retina which captures scenes and sends the images to the brain is located at the back of the eyeball. Fundus images which are human retinal images are generally used for DR diagnosis. A fundus image normally contains optic disk, fovea, and blood vessels. In diabetic eyes, abnormal things such as abnormal blood vessels, cotton wool spot, and exudates appear. Figure 2 shows an example of diabetic fundus images containing exudates.

There is much research on DR diagnosis from retinal images. In [9], early detection of diabetic eyes using fundus images was proposed. Moreover, fundus images were registered for long-term DR analysis in [10, 11].

Anyway, fundus images normally have non-uniform illumination due to the retina response, image acquisitions, instrumental limitations, and environments. This problem makes the diagnosis of DR difficult. Exudate detection from non-dilated retinal images was proposed in [12]. In addition, non-uniform illumination of retinal images was considered in [13, 14]. It was estimated/corrected before DR diagnosis.

2.2 Requirements for Copyright- and Privacy-Protected Diabetic Retinopathy Diagnosis Network

The following are requirements for copyright- and privacy-protected DR diagnosis network.

(A) DR Diagnosis Performance: DR lesions are detected, and the results are evaluated by sensitivity, specificity, positive predictive value (PPV), accuracy, and misclassified proportion (MP) as described in Section 3.1

(B) Visual encryption performance: It is expected that the image in transmission process is visually encrypted instead of using only secure transmission channels for privacy protection of the diagnosed image. Only the person who has a decryption key, for example, the ophthalmologist can view the image while other staffs cannot. Therefore, the encryption method used in the network should be robust against attacks.

(C) Image quality: The quality of the image which the receiver obtains is expected to be identical with the original diagnosed image. The processes that may degrade the image quality, i.e., watermarking and encryption, should be considered.

(D) Watermarking performance: Watermarking performance is evaluated by watermarked image quality and correct watermark extracting rate. It is expected that both are high, however, there is a trade off between both things.

(E) Compression performance: Compression performance could be evaluated by the compression ratio with the quality of the compressed image.

(F) Complexity: Complexity is desired to be as low as possible. However, it is not the most important requirement, if the task is an offline task.

3. PROPOSED COPYRIGHT- AND PRIVACY-PROTECTED DIABETIC RETINOPATHY DIAGNOSIS NETWORK

Figure 3 shows the proposed network on an assumption of two hospital network diagnosis. In practical, there may be more than two hospitals in the network. In addition, it can also be applied to intra hospital network diagnosis. The result of DR is sent from one user to another/other user (s) for checking (by computer or human) or helping diagnose since there are many types of DR signs. In this paper, exudates detection is performed by the first user. Before sending the diagnosed results to another user, the image is encrypted, watermarked, and compressed, respectively. DR diagnosis is presented in Section 3.1 . The proposed encryption method is described in Section 3.2 Security analysis and complexity analysis are also provided.

3.1 Diabetic Retinopathy Diagnosis

In this paper, exudate detection method proposed in [15] is used. However, the DR diagnosis method is flexible for the proposed network. Figure 4 illustrates the illumination estimation method for color RGB fundus images proposed in [15]. The illumination is estimated from R, G, and B components independently on the assumption that the significant details of retinal structures and lesions may be on any color band of images. Weighted surface fitting is used

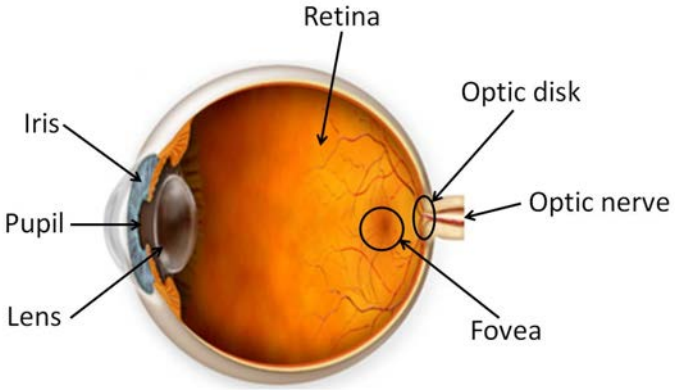


Fig.1: *Anatomical parts of human eye [8].*

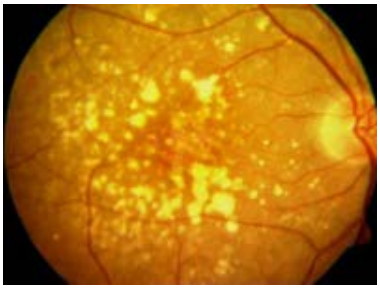


Fig.2: *Diabetic fundus image.*

for illumination estimation. On the assumption that the illumination is smooth, the polynomial surface is applied as described by Eq. 1 and Eq. 2.

$$\vec{E} = S\vec{P}, \tag{1}$$

where $\vec{E}$ is the estimated background vector, S is the surface matrix, and $\vec{P}$ is the parametric surface vector.

$$\begin{bmatrix} E(1,1) \\ \vdots \\ E(x,y) \\ \vdots \\ E(X,Y) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ x^A & x^{A-1}y & \dots & y & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ X^A & X^{A-1}Y & \dots & Y & 1 \end{bmatrix} \begin{bmatrix} P_1 \\ \vdots \\ P_N \end{bmatrix}, \tag{2}$$

where A is the order of the polynomial, N is the number of terms, $X \times Y$ is the image size, and (x, y) is the pixel coordinate. $\vec{P}$ can be calculated by Eq. 3.

$$\vec{P} = (S^T DS)^{-1}(S^T D)\vec{I}, \tag{3}$$

where $\vec{I}$ is the intensity vector generated by transforming the original image to a vector, and D is the diagonal weight matrix defined as

$$D = \begin{bmatrix} D(1,1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & D(X,Y) \end{bmatrix} \tag{4}$$

This matrix defines whether a pixel will be used in surface fitting process or not. The elements in the diagonal of the matrix correspond to pixels in the original image by order. The elements that correspond to the wanted pixels are set to 1, while the elements that correspond to the unwanted pixels are set to 0.

Unwanted pixels for surface fitting include optic disk, fovea, blood vessles, and lesions. To determine the unwanted pixels, upper bound and lower bound are calculated. Each single band image is smoothed by low pass filtering. Then, a low frequency image is subtracted from the original image to obtain a high frequency image. The upper bound is then calculated by adding the low frequency image, $I_L(x, y)$, by the standard deviation of the high frequency image, $I_H(x, y)$, as described by Eq. 5.

$$UB(x,y) = I_L(x,y) + \sqrt{\frac{1}{mn}\sum_{\forall(x,y)}(I_H(x,y) - \bar{I}_H)^2}, \tag{5}$$

where $UB(x, y)$ denotes the upper bound at pixel (x, y). I_H denotes the statistic mean of $I_H(x, y)$. In the same way, the lower bound is calculated by subtracting the low frequency image by the standard deviation of the high frequency image as described by Eq. 6.

$$LB(x,y) = I_L(x,y) - \sqrt{\frac{1}{mn}\sum_{\forall(x,y)}(I_H(x,y) - \bar{I}_H)^2}, \tag{6}$$

where $LB(x, y)$ denotes the lower bound. The upper bound and the lower bound are used to determine the diagonal weight matrix as Eq. 7.

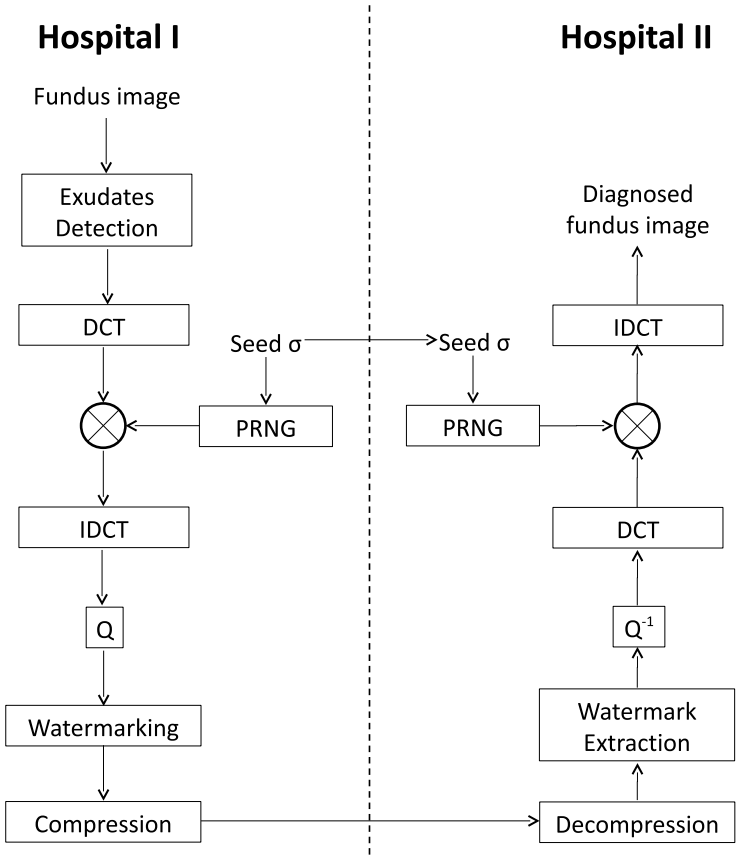


Fig.3: *Proposed copyright- and privacy-protected diabetic retinopathy diagnosis network (DCT: discrete cosine transform, PRNG: pseudo random number generator, IDCT: inverse discrete cosine transform, Q: quantization, Q^{-1}: inverse quantization).*

$$D(x,y) = \begin{cases} 0, UB(x,y) < I(x,y) \\ 0, I(x,y) < LB(x,y) \\ 1, LB(x,y) \leq I(x,y) \leq UB(x,y) \end{cases}. \quad (7)$$

Figure 5 shows an example of unwanted pixels detected by using upper/lower bounds. Figure 6 shows the result of illumination estimation. By using this method, there are two parameters needed to be controlled: filter mask and order of polynomial [15].

After illumination estimation process, exudates and other lesions are detected using level-set evolution without re-initialization [16]. Sensitivity, specificity, PPV, accuracy, and MP described by Eqs. (8), (9), (10), (11), and (12), respectively, are used to evaluate the effectiveness of the DR diagnosis method.

$$Sensitivity = TP/(TP + FN); \quad (8)$$

$$Specificity = TN/(TN + FP); \quad (9)$$

$$PPV = TP/(TP + FP); \quad (10)$$

$$Accuracy = (TP+TN)/(TP+TN+FP+FN); \quad (11)$$

$$MP = FP/(TP + TN + FP + FN), \quad (12)$$

where true positive (TP) is the number of lesion pixels correctly detected, true negative (TN) is the number of non-lesion pixels correctly identified as non-lesion pixels, false positive (FP) is the number of non-lesion pixels wrongly detected as lesion pixels, and false negative (FN) is the number of lesion pixels that cannot be detected. Sensitivity is the probability that the proposed method can classify the pixels of lesions as lesions. Specificity is the probability that the proposed method can classify the pixels of non-lesions as non-lesions. PPV is the probability that the pixels classified as lesions are really lesions. Specificity and accuracy are not very meaningful because the true negative value is always very high. Specificity and accuracy are always close to 100% regardless of the detection method.

3.2 Proposed Encryption Method

Each component of the color fundus image is processed independently. From now, only one color band of the image is considered for easy explanation. The algorithm for image encryption is as follows:

(1) Transform the image $f(x, y)$ by using DCT.
(2) Generate a random sign matrix R which consists of $+1$ and -1 randomly with the same size as the

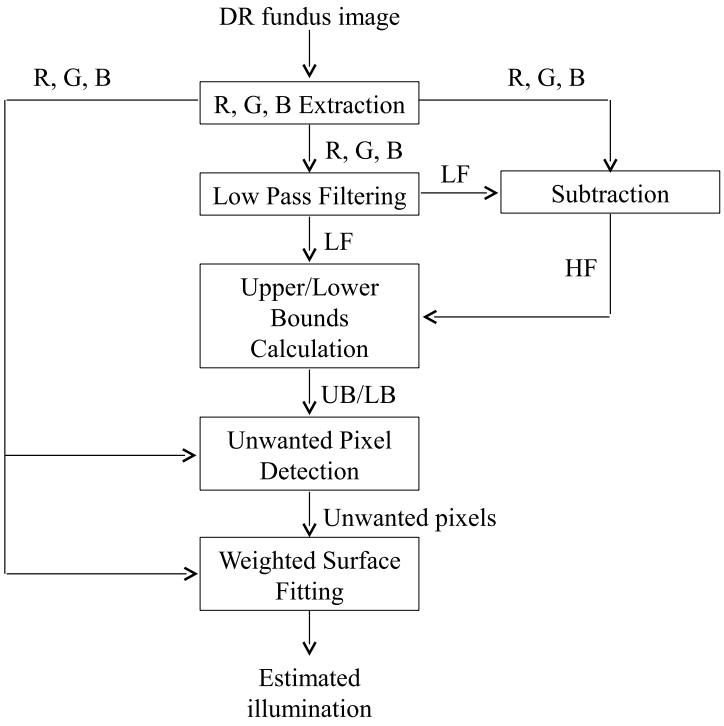


Fig.4: *Block diagram for illumination estimation [15] (DR: diabetic retinopathy, LF: low frequency, HF: high frequency, UB: upper bound, LB: lower bound).*

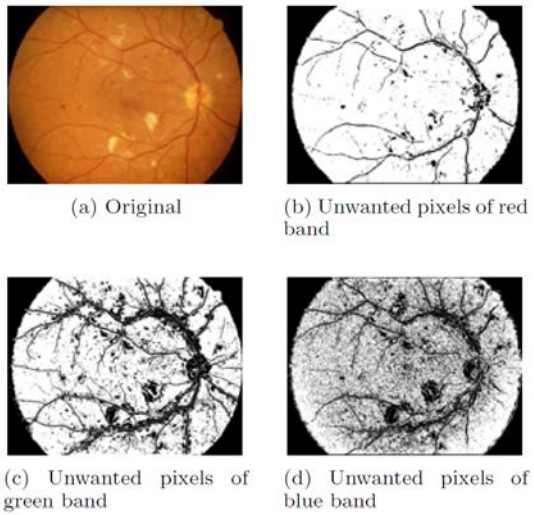


Fig.5: *Unwanted pixels detected by using upper/lower bounds*

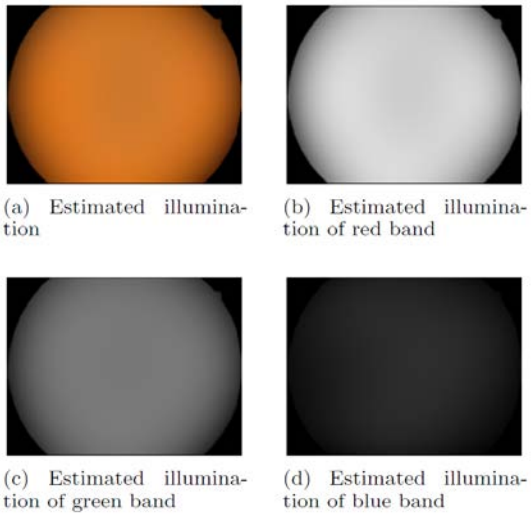


Fig.6: *Estimated illumination*

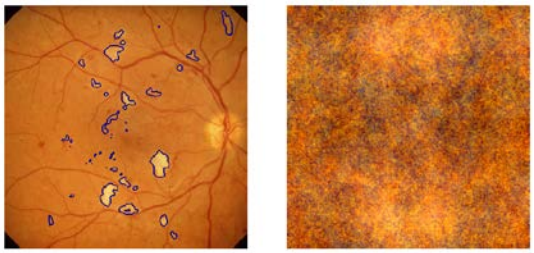
image $f(x, y)$. A seed σ for pseudo random number generator (PRNG) is input for image encryption and is sent to another user as the decryption key.

(3) Multiply DCT coefficients $F(u, v)$ of image $f(x, y)$ with the random matrix R pixel-wise.

(4) Apply inverse discrete cosine transform (IDCT) to the sign-scrambled coefficients $F'(u, v)$.

In this framework, 2D DCT and 1D DCT sign scrambling are proposed for image encryption.

For 2D DCT, the diagnosed fundus image is transformed to a frequency domain by using 2D DCT. Let $F(u, v)$ be the $X \times Y$-sized discrete cosine transformed (DCTed) coefficients of $X \times Y$-sized diagnosed fundus image $f(x, y)$, where $x = 0, 1, \ldots, X - 1$, $y = 0, 1, \ldots, Y - 1$, $u = 0, 1, \ldots, X - 1$, and $v = 0, 1, \ldots, Y - 1$;



(a) Diagnosed (b) Encrypted

Fig.7: *Diagnosed fundus image and encrypted diagnosed fundus image using 2D DCT sign scrambling.*

$$F(u,v) = \sum_{x=0}^{X-1}\sum_{y=0}^{Y-1} f(x,y)\cos\left(\frac{\pi}{X}\left(x+\frac{1}{2}\right)u\right)\dots \cos\left(\frac{\pi}{Y}\left(y+\frac{1}{2}\right)v\right). \quad (13)$$

DCTed coefficients $F(u,v)$ can also be expressed in the polar form as

$$F(u,v) = |F(u,v)|S_{DCT}(u,v), \quad (14)$$

where $|F(u,v)|$ and $S_{DCT}(u,v)$ denote the amplitude and sign components of $F(u,v)$, respectively.

By applying sign scrambling encryption, the encrypted coefficients can be formulated as

$$F'(u,v) = |F(u,v)|S_{DCT}(u,v)R(u,v), \quad (15)$$

where $R(u,v)$ denotes a random sign matrix with $X \times Y$ in size which consists of $+1$ and -1. Apply IDCT to $F'(u,v)$ to obtain:

$$f'(x,y) = \sum_{u=0}^{X-1}\sum_{v=0}^{Y-1} \alpha(u)\alpha(v)F'(u,v)\dots \cos\left(\frac{\pi}{X}\left(x+\frac{1}{2}\right)u\right)\cos\left(\frac{\pi}{Y}\left(y+\frac{1}{2}\right)v\right), \quad (16)$$

where

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{X}}, u = 0 \\ \sqrt{\frac{2}{X}}, u = 1,2,\dots,X-1, \end{cases} \quad (17)$$

$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{Y}}, v = 0 \\ \sqrt{\frac{2}{Y}}, v = 1,2,\dots,Y-1, \end{cases} \quad (18)$$

and $f'(x,y)$ denotes the visually encrypted diagnosed fundus image. Figures 7 (a) and (b) show a diagnosed fundus image and the related encrypted image, respectively.

It is necessary to check the IR of the encrypted image. Eq. (19) defines the IR.

$$\text{IR} = \max\left(f'(x,y)\right) - \min\left(f'(x,y)\right). \quad (19)$$

To store the images into 8 bits per pixel (bpp) images, a simple linear quantization (LQ) is used to quantize the image $f'(x,y)$ by taking into account the entire IR. $f'(x,y)$ is quantized as

$$f'_{id}(x,y) = \left(\frac{f'(x,y) - \min(f'(x,y))}{s}\right), \quad (20)$$

$$s = \frac{\max(f'(x,y)) - \min(f'(x,y))}{2^n - 1}. \quad (21)$$

$f'_{id}(x,y)$ is the index image which ranges in $[0, 2^n - 1]$, i.e., the number of quantization levels is equal to 2^n. To obtain quantized image $f'_{LQ}(x,y)$, the inverse quantization with constant step size s and bias constant $\min(f'(x,y))$ are applied to $f'_{id}(x,y)$ as

$$f'_{LQ}(x,y) = sf'_{id}(x,y) + \min(f'(x,y)). \quad (22)$$

For image decryption, firstly the quantized image $f'_{LQ}(x,y)$ is transformed using 2D DCT as shown in Eq. (23).

$$F'_{LQ}(u,v) = \sum_{x=0}^{X-1}\sum_{y=0}^{Y-1} f'_{LQ}(x,y)\cos\left(\frac{\pi}{X}\left(x+\frac{1}{2}\right)u\right)\dots \cos\left(\frac{\pi}{Y}\left(y+\frac{1}{2}\right)v\right), \quad (23)$$

where $F'_{LQ}(u,v)$ denotes the encrypted coefficients with quantization effects. Then, the random sign matrix R is multiplied pixel-wise with the coefficients $F'_{LQ}(u,v)$ as

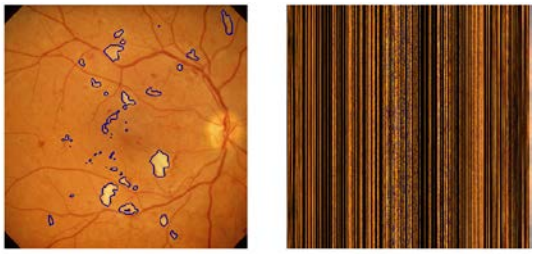
$$F_{LQ}(u,v) = F'_{LQ}(u,v)R(u,v), \quad (24)$$

where $F_{LQ}(u,v)$ denotes the decrypted coefficients with quantization effects. Finally, 2D IDCT is applied to the coefficients $F_{LQ}(u,v)$ as

$$f_{LQ}(x,y) = \sum_{u=0}^{X-1}\sum_{v=0}^{Y-1} \alpha(u)\alpha(v)F_{LQ}(u,v)\dots \cos\left(\frac{\pi}{X}\left(x+\frac{1}{2}\right)u\right)\cos\left(\frac{\pi}{Y}\left(y+\frac{1}{2}\right)v\right), \quad (25)$$

where $f_{LQ}(x,y)$ denotes the decrypted diagnosed fundus image with quantization effects. It is noted that if the quantization is not applied, $f_{LQ}(x,y) = f(x,y)$. Anyway, other quantization methods can be used in order to minimize quantization errors.

For the last method using 1D DCT, Let $F_c(x,v)$ be the $X \times Y$-sized column-wise 1D discrete cosine transformed (DCTed) coefficients of the diagnosed



(a) Diagnosed (b) Encrypted

Fig.8: *Diagnosed fundus image and encrypted diagnosed fundus image using 1D DCT sign scrambling.*

fundus image $f(x, y)$, where $x = 0, 1, \ldots, X-1$, $y = 0, 1, \ldots, Y-1$, and $v = 0, 1, \ldots, Y-1$;

$$F_c(x, v) = \sum_{y=0}^{Y-1} f(x, y) \cos\left(\frac{\pi}{Y}\left(y + \frac{1}{2}\right) v\right). \quad (26)$$

DCT coefficients $F_c(x, v)$ can also be expressed in the polar form as

$$F_c(x, v) = |F_c(x, v)| S_c(x, v), \quad (27)$$

where $|F_c(x, v)|$ and $S_c(x, v)$ denote the amplitude and sign components of $F_c(x, v)$, respectively.

To encrypt the image $f(x, y)$, the coefficients, $|F_c(x, v)|$ is multiplied with $R(x, v)$ as

$$F'_c(x, v) = |F_c(x, v)| S_c(x, v) R(x, v), \quad (28)$$

Apply 1D IDCT to $F'_c(x, v)$ to obtain

$$f'_c(x, y) = \sum_{v=0}^{Y-1} \alpha(v) F'_c(x, v) \cos\left(\frac{\pi}{Y}\left(y + \frac{1}{2}\right) v\right), \quad (29)$$

where $f'_c(x, y)$ denotes the visually encrypted diagnosed fundus image, and

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{Y}}, v = 0 \\ \sqrt{\frac{2}{Y}}, v = 1, 2, \ldots, Y-1. \end{cases} \quad (30)$$

Figures 8 (a) and (b) show a diagnosed fundus image and the related encrypted image, respectively. It is seen that the visual encryption performance of 1D DCT-based method is as good as that of 2D DCT-based method.

It is noted that for 1D DCT, quantization and the similar decryption method are also applied.

3.3 Security Analysis

Comparing security of encryption algorithms is not an easy task. It depends on skills and levels of attackers. One important aspect of security analysis of encryption is key space size. For both 2D DCT and 1D DCT sign scrambling, assume that the attacker applies brute-force attack. To guess the random sign matrix R, the number of possible ways in generating the matrix called encryption key space size, S_k, is:

$$S_k = 2^{XY}. \quad (31)$$

It is clearly seen that the key space size depends on the image size. Since medical images normally have quite big sizes, this key space size is big enough to this kind of attack. For example, suppose that the fundus image size $XY = 1024 \times 1024$ pixels. In this case, the key space $S_k = 2^{1024 \times 1024}$ which is much bigger than that of an AES-like cipher with a 256-bit key ($S_k = 2^{256}$).

From an experiment (see Figs. 9 and 10), 1%, 5%, and 10% wrong bits for the decryption make decrypted images have lower quality with respect to original images. In security aspect, with these numbers, it seems that the proposed encryption methods are still sensitive. However, for 30% and 40% wrong bits for the decryption, the results show that the decrypted images are unrecognizable. It would need about 60% to 70% correct decryption bits to obtain recognizable images. For an 1024×1024-pixel image with the key space $2^{1024 \times 1024}$, to get 60% to 70% correct decryption bits from guessing, it is still not an easy task.

3.4 Complexity Analysis

Even though low complexity is not a major requirement of the proposed network, this section compares the complexity of both encryption methods. The difference between methods is the dimension of image transformation. The complexity of 2D DCT for an $X \times Y$ sized image is $O((XY)^2)$ or $O((XY \log_2 (XY))$ with fast algorithm, whereas that of 1D DCT is $O((X^2)Y)$ or $O((X \log_2 X)Y)$ with fast algorithm. It is clearly seen that 1D DCT outperforms 2D DCT in terms of complexity.

4. EXPERIMENTAL RESULTS

The experiment is divided into two parts: DR diagnosis performance and copyright and privacy protection performance. Eight 32-bit color fundus images from REVIEW databases [17] and eighty nine 24-bit color fundus images with 1500×1152 pixels in size from DIARETDB1 database [18] are used in the experiments.

4.1 DR Diagnosis Performance

In this part, eighty nine 24-bit color fundus images with 1500×1152 pixels in size from DIARETDB1 database are used to evaluate the exudate detection performance. Forty seven images contain exudates, but the rest images do not. The database includes

Table 1: *Exudate detection results from forty seven images that contain hard exudates.*

Method	Average sensitivity (%)	Average specificity (%)	Average PPV (%)	Average MP (%)
Sopharak et al. [19]	43.48	99.31	25.48	0.68
Ravishankar et al. [9]	58.21	98.09	13.37	1.90
Welter et al. [20]	66.00	98.64	19.45	1.34
Welfer et al. [21]	70.48	98.84	21.32	1.10
Kande et al. [22]	86.00	98.00	-	-
Sae-Tang et al. [15]	89.38	99.08	80.74	0.81

Table 2: *Exudate detection results from forty two images that do not contain hard exudates.*

Method	Average specificity (%)	Average MP (%)
Sopharak et al. [19]	99.28	0.71
Ravishankar et al. [9]	97.53	2.47
Welter et al. [20]	99.22	0.77
Welfer et al. [21]	98.74	1.20
Kande et al. [22]	-	-
Sae-Tang et al. [15]	99.48	0.52

manually labeled images that are used as a ground truth in the experiment. Table 1 and 2 show the experimental results. The method used in this paper is compared with the methods proposed by Soparak et al. [19], Ravishankar et al. [9], Walter et al. [20], Welfer et al. [21], and Kande et al. [22]. All compared methods are based on mathematical morphology and were validated using DIARETDB1 database. Welfer et al. [21] used the contrast enhancement before detecting exudates. The method improved the sensitivity of exudate detection, but it also increases FP causing low PPV. Kande et al. [22] used local contrast enhancement, but it introduces noises to images. For images containing hard exudates, the method used in this paper achieves higher average PPV than those of other methods. For images without hard exudates, the method achieves the highest average specificity and the lowest average MP compared to other methods. It shows that the method significantly reduces FP in detecting exudates and simultaneously keeps high sensitivity in detecting exudates. This is because of the effective illumination estimation.

4.2 Copyright and Privacy Protection Performance

The proposed encryption methods are compared with JPEG 2000-based DWT sign scrambling [6], 2D DFT AOI [4], 1D DFT AOI [5], 2D DCT AOI [5], and 1D DCT AOI [5] which were proposed for copyright- and privacy-protected image trading systems in which a trusted third party (TTP) can process images using visual scrambled images. Even though the application is different from the proposed application, these encryption methods could also be applied to the proposed network. Figures 11, 12, and 13 show the proposed copyright- and privacy-protected DR diagnosis network using the compared encryption methods.

4.2.1 Visual Encryption Performance

The original fundus images and the lesion-marked fundus images are encrypted by several methods as shown in Figs. 14, 15, 16, and 17. The results show that the proposed 1D DCT sign scrambling and 2D DCT sign scrambling methods perfectly visually encrypt the original fundus images and the lesion-marked fundus images, while other methods encrypt the images imperfectly especially for the lesion-marked fundus images. Among several methods, JPEG 2000-based DWT sign scrambling [6] is the worst encryption method. 2D DFT AOI and 2D DCT AOI are better than 1D DFT AOI and 1D DCT AOI.

4.2.2 Image Quality

Tables 3 and 4 compare IRs of encrypted diagnosed fundus images "HRIS001" and "HRIS003," respectively. Between two proposed methods, 2D DCT sign scrambling gives narrower IRs than those of 1D DCT sign scrambling for every color band of images. From this fact, by applying LQ to the encrypted images, 2D DCT sign scrambling method has a possibility to achieve better image quality. For JPEG 2000-based DWT sign scrambling method, the IRs are not defined, because the scrambled DWT coefficients are encoded, and by using the proprietary JPEG 2000 decoder, the original images are recovered for lossless compression, viz., quantization is not required. Therefore, infinite PSNRs are obtained. However, this method will not be considered anymore in the rest experiments because of its bad visual encryption performance. For AOI methods, 2D DFT AOI and 2D DCT AOI methods give wide IRs, while 1D DFT AOI and 1D DCT AOI methods give much lower IRs than those of 2D transform methods. However, they are worse than the proposed methods. Besides IRs of images, the distribution of pixel intensities are also an important factor of image quality. Consequently, the image quality is evaluated by peak signal-to-noise ratio (PSNR). The PSNRs of decrypted images with respect to the original diagnosed fundus images are calculated. From Tables 5 and 6, the proposed methods give the highest PSNRs than the conventional methods. 2D DCT sign scrambling method gives higher PSNRs than those of 1D DCT sign scrambling. 1D DFT AOI method and 1D DCT AOI method give higher PSNRs than those of 2D DFT AOI method and 2D DCT AOI method. Figures 18 and 19 show the decrypted images. Even though 2D DCT sign scrambling gives higher PSNRs than those of 1D DCT sign scrambling, the image visual quality for 1D DCT sign scrambling method is also good. Therefore, 1D DCT sign scrambling method can be considered in case that the complexity of the system is concerned.

Table 3: *Intensity range of encrypted diagnosed fundus image "HRIS001."*

	Red	Green	Blue
2D DCT sign scrambling	[49.4, 351.2]	[−2.5, 219.3]	[−55.4, 110.2]
1D DCT sign scrambling	[−414.1, 420.3]	[−245.6, 264.2]	[−176.0, 160.7]
JPEG 2000-based DWT sign scrambling [6]	N/A	N/A	N/A
2D DFT AOI [4]	[130.9, 7688.7]	[53.4, 5068.4]	[4.4, 3651.8]
1D DFT AOI [5]	[82.9, 1203.7]	[22.1, 746.7]	[−13.7, 538.9]
2D DCT AOI [5]	[95.0, 8434.1]	[31.7, 5725.7]	[−6.1, 4046.6]
1D DCT AOI [5]	[75.8, 1186.5]	[15.3, 740.4]	[−25, 532.4]

Table 4: *Intensity range of encrypted diagnosed fundus image "HRIS003."*

	Red	Green	Blue
2D DCT sign scrambling	[−112.5, 369.3]	[−75.0, 174.3]	[−147.6, 160.0]
1D DCT sign scrambling	[−407.1, 449.6]	[−215.8, 217.6]	[−260.6, 237.5]
JPEG 2000-based DWT sign scrambling [6]	N/A	N/A	N/A
2D DFT AOI [4]	[24.1, 1242.2]	[0.4, 6070.3]	[−7.3, 8260.3]
1D DFT AOI [5]	[−66.5, 1387.8]	[−11.0, 715.2]	[−62.0, 896.5]
2D DCT AOI [5]	[−45.3, 1325.3]	[−38.3, 6621.9]	[−70.7, 8890.2]
1D DCT AOI [5]	[−66.7, 1386.7]	[−41.8, 739.1]	[−70.1, 913.6]

Table 5: *Peak signal-to-noise ratio of decrypted diagnosed fundus image "HRIS001" [dB].*

	Red	Green	Blue
2D DCT sign scrambling	62.2152	64.9031	67.4253
1D DCT sign scrambling	53.3933	57.6723	61.2826
JPEG 2000-based DWT sign scrambling [6]	Inf	Inf	Inf
2D DFT AOI [4]	34.1981	37.8061	40.5690
1D DFT AOI [5]	50.6424	54.1384	56.1575
2D DCT AOI [5]	33.3508	36.6959	39.6654
1D DCT AOI [5]	50.9035	54.6350	56.8893

Table 6: *Peak signal-to-noise ratio of decrypted diagnosed fundus image "HRIS003" [dB].*

	Red	Green	Blue
2D DCT sign scrambling	58.1709	63.9040	62.0651
1D DCT sign scrambling	53.1801	59.0853	57.8542
JPEG 2000-based DWT sign scrambling [6]	Inf	Inf	Inf
2D DFT AOI [4]	30.3778	36.2587	34.1194
1D DFT AOI [5]	48.7052	54.3634	51.9720
2D DCT AOI [5]	29.3981	35.2937	32.7385
1D DCT AOI [5]	48.6078	53.9613	51.9457

4.2.3 Watermarking Performance

Almost any arbitrary digital watermarking technique can be applied. This paper uses a simple non-blind additive digital watermarking technique in the discrete wavelet transformed (DWTed) domain, which is based on the essence of [23].

(A) Watermarking: The process of watermarking is as follows:
  (1) $X \times Y$-sized DWTed coefficients are divided into $X_B \times Y_B$-sized coefficient blocks where $X_B = Y_B = 8$ in the experiments. That is, the DWTed coefficients are divided into $K = XY/X_BY_B$ blocks.
  (2) A binary watermark sequence, $b = [b_1b_2b_3...b_K]$, is provided, where $b_k \in \{0, 1\}$, and $k = 1, 2, 3, ..., K$.
  (3) Each single bit of the watermark sequence, b_k, is represented by an M-sequence with the length L, and the length of M-sequences used in the experiments is the smallest, i.e., $L = 3$, where $L = B^P - 1$, $B = 2$, and $P = 2$, even longer sequences serve better correct watermark extracting rate. The M-sequences differ from each other for different single bits of the watermark. Therefore, two different M-sequences are prepared for $b_k \in \{0, 1\}$.
  (4) Each M-sequence is subsequently embedded to a divided block of DWTed coefficients by linearly scaling and adding it to three bottom-right coefficients of the block as

$$C'_{k,l} = \mathrm{sgn}\left(C_{k,l}\right)\left(|C_{k,l}| + \delta W_{b_k,l}\right), \quad (32)$$

  where $C'_{k,l}$ denotes the l-th watermarked DWTed coefficient in the k-th block, $C_{k,l}$ denotes the l-th original DWTed coefficient in the k-th block, δ is a scaling factor, and $W_{b_k,l} \in \{0, 1\}$ denotes the l-th element of the M-sequence for bit b_k, where $l = 1, 2, \ldots, L$. It is noted that all variables in Eq. (32) are integers.

(B) Watermark Extraction: The watermark is extracted from the DWTed coefficients by using a correlation-based detector as

$$b'_k = \arg \max_{b_k \in \{0,1\}} \sum_{l=1}^{L} \left(W_{b_k,l} - 0.5\right) ... \left(\frac{\left|C''_{k,l}\right| - |C_{k,l}|}{\delta} - 0.5\right), \quad (33)$$

where b'_k is an extracted watermark bit, and $\left|C''_{k,l}\right|$ denotes the l-th watermarked DWTed coefficient in the block. It is noted that $\left|C''_{k,l}\right|$ may be different from $\left|C'_{k,l}\right|$ because of clipping in JPEG 2000 decoding.

Table 7 shows PSNRs of watermarked diagnosed fundus images averaged by 97 images without compression effects, when δ is varied from 1 to 8 with stepping by 1. All methods are comparable in terms of watermarked image quality. However, these slight differences in PSNRs of watermarked images between methods are not significant, because the final image quality significantly depends on the encryption method, and from Table 5 and 6, it is clearly seen that the proposed methods give much higher PSNRs than those of other methods. For watermark extracting performance, all methods have achieved 100% correct watermark extracting rates using a correlation-based detector. It is noted that if the compression is lossless, the images can be recovered, because the watermark is extracted 100% correctly. It means that this watermarking method is very effective as reversible watermarking. Due to the fact that watermarking performance depends on the watermarking technique, optimization or other watermarking techniques such as [23–26] could be applied.

4.2.4 Compression Performance

Kakadu [27], a software to encode and decode JPEG 2000 images, is used in the experiments. It is noted that any image compression standard could be used in the proposed network instead of JPEG 2000.

Usually, compression and encryption are combined by firstly compressing the content and encrypting it subsequently. However, the encryption is performed independently from the compression process in the proposed network for flexibility in choosing the encryption method and the compression method. Therefore, it is necessary to confirm the compression performance when the encryption is done first. In this part, the compression performance is compared with and without the proposed encryption method and other encryption methods without considering watermarking process, where the diagnosed fundus image "HRIS001" with 446 kbytes is used in the experiment. The encrypted diagnosed fundus images are linearly quantized and then compressed using JPEG 2000. The compression bit rate is varied by 0.25, 0.5, 1, 2, 3, 4, and 5 bpp for lossy compression. Table 8 compares the compressed file sizes of the quantized encrypted diagnosed fundus images, while the file sizes of the quantized encrypted diagnosed fundus images are 446 kbytes. PSNRs of the compressed images are calculated with respect to the quantized images as shown in Table 9.

From the results, all encryption methods degrade the compression performance. Comparing the proposed methods to other methods, 1D DFT AOI and 1D DCT AOI are better than the proposed methods in terms of compression performance. However, these methods are worse in terms of visual encryption performance (see Section 4.2.1) which is more important for privacy protected network. 2D DFT AOI and 2D DCT AOI are better than the proposed methods in terms of compression performance, when the compression is lossy. However, they are worse than the proposed methods in terms of visual encryption performance (see Section 4.2.1) and much worse in terms of image quality (see Section 4.2.2). In addition, for lossless compression, the proposed methods are comparable with them. It is expected that the compression is lossless in practical terms, since high quality images are required for DR diagnosis. Therefore, the proposed methods are the better choices.

Table 10 concludes the effectiveness of the proposed encryption methods and the conventional encryption methods. Symbols "×," "Δ," "◯," and "◎" mean "bad," "fair," "good," and "excellent," respectively.

5. CONCLUSIONS

The paper has proposed a copyright- and privacy-protected DR diagnosis network. In the network, DR lesions are automatically detected from a fundus image by firstly estimating the non-uniform illumination of the image, and then the lesions are detected from the balanced image by using level-set evolution without re-initialization. The lesions are subsequently marked by using contours. The lesion-marked fundus image is subsequently shared for intra or inter hospital network diagnosis with copyright and privacy protection. Watermarking technique is used for image copyright protection, and visual encryption is used for privacy protection. 2D DCT sign scrambling and 1D DCT sign scrambling have been proposed for privacy protection. The proposed encryption methods are effective for strong-edge images that is suitable for lesion-marked fundus images, while random sign-based JPEG 2000, DFT AOIs, and DCT AOIs encrypt the images imperfectly. Moreover, the proposed methods are better in terms of image quality. Eventhough 2D DCT sign scrambling method is the best in terms of image quality, the proposed 1D DCT sign scrambling could be used instead of 2D DCT sign scrambling in cases that the complexity is concerned. In addition, the proposed encryption methods could be applied for general images and other applications, and other encryption methods that encrypt images containing strong edges perfectly could be applied to the proposed network. All methods are comparable in terms of watermarked image quality. For watermark extracting performance, all methods have achieved 100% correct watermark extracting rates using a correlation-based detector.

References

[1] G. G. Yen and W. F. Leong, "A sorting system for hierarchical grading of diabetic fundus images: A preliminary study," *IEEE Trans. Inf.*

Table 7: *Averaged peak signal-to-noise ratios of watermarked diagnosed fundus images [dB].*

δ	1	2	3	4	5	6	7	8
2D DCT sign scrambling	64.3396	60.9186	58.5993	56.8373	55.1256	53.7174	52.5801	51.5593
1D DCT sign scrambling	64.3865	61.1529	58.8164	57.0382	55.3445	53.9888	52.8151	51.8253
2D DFT AOI [4]	65.1744	61.8779	59.4572	57.6995	56.0046	54.6690	53.5241	52.5157
1D DFT AOI [5]	65.5210	62.4001	59.9945	58.2029	56.5058	55.1782	54.0052	52.9781
2D DCT AOI [5]	63.8137	60.8935	58.4202	56.7237	54.9686	53.6312	52.4312	51.4703
1D DCT AOI [5]	63.9598	60.8524	58.5301	56.7642	55.0644	53.7950	52.6293	51.6443

Table 8: *Compressed file sizes of encrypted diagnosed fundus images for "HRIS001" [kbytes].*

Bit rate [bit per pixel]	0.25	0.5	1	2	3	4	5	Lossless
original	4.69	9.36	18.5	37.2	55.7	75.2	92.6	211
2D DCT sign scrambling	4.65	9.07	18.5	37.1	55.7	74.2	92.8	317
1D DCT sign scrambling	4.61	9.36	18.6	37.0	55.7	74.2	92.7	282
2D DFT AOI [4]	4.71	9.25	18.6	37.1	55.7	74.3	92.8	299
1D DFT AOI [5]	4.68	9.25	18.4	37.1	55.2	74.2	92.8	231
2D DCT AOI [5]	4.56	9.20	18.6	36.8	55.7	74.3	92.8	316
1D DCT AOI [5]	4.71	9.35	18.5	37.1	55.7	74.3	92.8	251

Table 9: *Peak signal-to-noise ratios of compressed encrypted diagnosed fundus images for "HRIS001" [dB].*

Bit rate [bit per pixel]	0.25	0.5	1	2	3	4	5	Lossless
original	25.5465	27.4390	29.4362	34.7269	38.6917	41.6652	43.7407	Inf
2D DCT sign scrambling	23.8068	24.5345	25.7157	27.5950	29.0822	29.9125	31.1130	Inf
1D DCT sign scrambling	16.1688	19.2890	22.4141	25.4751	30.0771	32.3156	34.2390	Inf
2D DFT AOI [4]	28.5696	29.2500	30.1917	31.3830	32.5048	33.5065	34.3004	Inf
1D DFT AOI [5]	30.6858	31.7375	33.8367	35.8697	38.2836	40.0649	41.5370	Inf
2D DCT AOI [5]	26.8727	27.6097	28.5678	29.7366	31.1661	31.9076	32.7252	Inf
1D DCT AOI [5]	28.6237	29.5979	31.6235	33.9189	35.8974	37.8213	39.5469	Inf

Table 10: *Conclusions of effectiveness of considered methods (P_1: 2D DCT sign scrambling (Proposed 1), P_2: 1D DCT sign scrambling (Proposed 2), C_1: JPEG 2000-based DWT sign scrambling [6] (Conventional 1), C_2: 2D DFT AOI [4] (Conventional 2), C_3: 1D DFT AOI [5] (Conventional 3), C_4:2D DCT AOI [5] (Conventional 4), C_5: 1D DCT AOI [5] (Conventional 5)).*

Requirement	P_1	P_2	C_1	C_2	C_3	C_4	C_5
Visual encryption performance	◎	◎	×	○	△	○	△
Image quality	◎	◎	◎	△	○	△	○
Watermarking performance	◎	◎	-	◎	◎	◎	◎
Compression performance	○	○	-	○	◎	○	◎
Complexity	○	◎	-	○	◎	○	◎

Technol. Biomed., vol. 12, no. 1, pp. 118–130, Jan. 2008.

[2] J. Nayak, P. S. Bhat, R. A. U., and M. S. Kumar, "Efficient storage and transmission of digital fundus images with patient information using reversible watermarking technique and error control codes," *Journal of medical systems*, vol. 33, issue 3, pp. 163–171, Jun. 2009.

[3] M. Nergui, U. S. Acharya, U. R. Acharya, W. Yu, "Reliable and robust transmission and storage techniques for medical images with patient information," *Journal of Medical systems* vol. 34, issue 6, pp. 1129–1139, 2010.

[4] W. Sae-Tang, M. Fujiyoshi, and H. Kiya, "A Generation Method of Amplitude-Only Images with Low Intensity Ranges," *IEICE Trans. Fundamentals*, vol. E96-A, no. 6, pp. 1323–1330, Jun. 2013.

[5] W. Sae-Tang, S. LIU, M. FUJIYOSHI, and H. KIYA, "1D Frequency Transformation-Based Amplitude-Only Images for Copyright- and Privacy-Protection in Image Trading Systems," *ECTI-CIT*, vol. 8, no. 2, Nov. 2014.

[6] W. Sae-Tang, S. Liu, M. Fujiyoshi, and H. Kiya, "A copyright- and privacy-protected image trading system using fingerprinting in discrete wavelet domain with JPEG 2000," *IEICE Trans. Fundamentals*, vol. E97-A, no. 11, Nov. 2014.

[7] W. Sae-Tang, W. Chiracharit, and W. Kumwilaisak, "Exudates Detection in Fundus Images Using Non-uniform Illumination Background Subtraction," *Proc. IEEE TENCON2010*, Fukuoka, Japan, pp. 204–209, Nov. 21-24, 2010.

[8] [Online]. Available: http://www.mastereyeassociates.com/presbyopia.

[9] S. Ravishankar, A. Jain, and A. Mittal, "Automated feature extraction for early detection of diabetic retinopathy in fundus images," *Proc. IEEE CVPR*, Miami, FL, 2009.

[10] H. Narasimha-Iyer, A. Can, B. Roysam, Charles V. Stewart, Howard L. Tanenbaum, A. Majerovics, and H. Singh, "Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 6, pp. 1084–1098, Jun. 2006.

[11] H. Narasimha-Iyer, A. Can, B. Roysam, Howard L. Tanenbaum, and A. Majerovics, "Integrated analysis of vascular and nonvascular changes

from color retinal fundus image sequences," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 8, pp. 1436–1445, Aug. 2007.

[12] A. Sopharak, B. Uyyanonvara, and S. Barman, "Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy c-means clustering," *Sensors*, pp. 2148–2161, 2009.

[13] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Medical Image Analysis*, vol. 3, no. 9, pp. 179–190, 2005.

[14] E. Grisan, A. Giani, E. Ceseracciu, and A. Ruggeri, "Model-based Illumination Correction in Retinal Images," *IEEE Int. Sym. on Biomed. Imag.*, pp. 984–987, Apr. 2006.

[15] W. Sae-Tang, W. Chiracharit, S. Kiattisin, and W. Kumwilaisak, "Non-Uniform Illumination Estimation in Fundus Images Using Bounded Surface Fitting," *The international journal on applied biomedical engineering (IJABME)*, vol. 5, no. 1, pp. 37–45, 2012.

[16] C. Li, C. Xu, C. Gui, and M. D. Fox, "Level-set evolution without re-initialization: A new variational formulation," *IEEE CVPR'05*, 2005.

[17] University of LINCOLN, Department of Computing and Informatics (MHAC MC3201), Retinal Vessel Image set for Estimation of Widths (REVIEW) databases, [Online]. Available: http://aldiri.info.

[18] T. Kauppi, V. Kalesnykiene, J-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, et al. DIARETDB1: diabetic retinopathy database and evaluation protocol. In: Medical image understanding and analysis (MIUA), pp. 61–65, 2007.

[19] A. Sopharak, B. Uyyanonvara, S. Barman, S. Vongkittirux, and N. Wongkamchang, "Fine exudate detection using morphological reconstruction enhancement," *The international journal on applied biomedical engineering (IJABME)*, vol. 1, no. 1, pp. 45–50, 2010.

[20] T. Walter, J. C. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1236–1243, Oct. 2002.

[21] D. Welfer, J. Scharcanski, and D. R. Marinho, "A coarse-to-fine strategy for automatically detecting exudates in color eye fundus images," *Computerized Medical Imaging and Graphics*, pp. 228-235, 2010.

[22] G. Kande, P. Subbaiah, and T. Savithri, "Segmentation of exudates and optic disk in retinal images," *IEEE Sixth Indian Conference on Computer Vision, Graphic & Image Processing*, 2008.

[23] S. Pereira, S. Voloshynoskiy, and T. Pun, "Optimal transform domain watermark embedding via linear programing," *Signal Processing 81*, pp. 1251–1260, 2001.

[24] H. Inoue, A. Miyazaki, and T. katsura, "An image watermarking method based on the wavelet transform," *Proc. IEEE ICIP*, pp. 296–300, 1999.

[25] Z. Zhang, Q. Sun, and W. Wong, "A novel lossy-to-lossless watermarking scheme for JPEG2000 images," *Proc. IEEE ICIP*, pp. 573–576, 2004.

[26] Y. Chen and H. Huang, "A progressive image watermarking scheme for JPEG2000," *Proc. IEEE IIH-MSP*, pp. 230–233, 2012.

[27] D. Taubman, "Kakadu software-a comprehensive framework for JPEG2000," 2005.

Wannida Sae-Tang received her B.Eng. degree in Electronic and Telecommunication Engineering with the first class honors and her M.Eng. degree in Electrical Engineering from King Mongkut's University of Technology Thonburi, Thailand in 2007 and 2011, respectively, and her Ph.D. degree in Information and Communication Systems from Tokyo Metropolitan University, Japan in 2014 with Tokyo metropolitan governmental Asian human resources fund. From 2007 to 2009, she was an IC packaging design engineer of New product design and research and development team at United Test and Assembly Center Thai Ltd. She is currently a lecturer of The Sirindthorn International Thai-German Graduate School of Engineering, King Mongkut's University of Technology North Bangkok, Thailand. Her research interests include image processing and multimedia communication. She received the Best Paper Award of the IEICE/ITE/KSBE IWAIT in 2014.

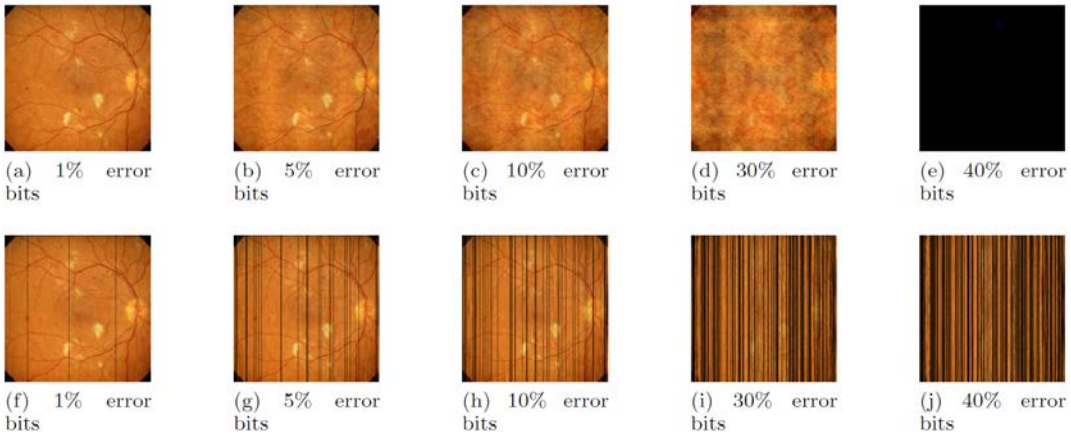


Fig.9: *Decrypted images of "HRIS001" for several percentages of error decryption bits. The upper row is for the proposed 2D DCT sign scrambling, and the lower row is for the proposed 1D DCT sign scrambling.*

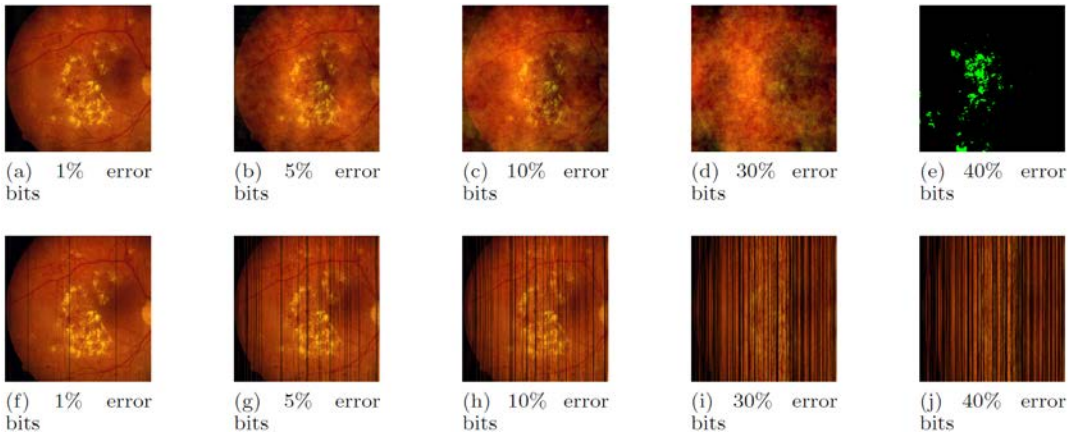


Fig.10: *Decrypted images of "HRIS003" for several percentages of error decryption bits. The upper row is for the proposed 2D DCT sign scrambling, and the lower row is for the proposed 1D DCT sign scrambling.*

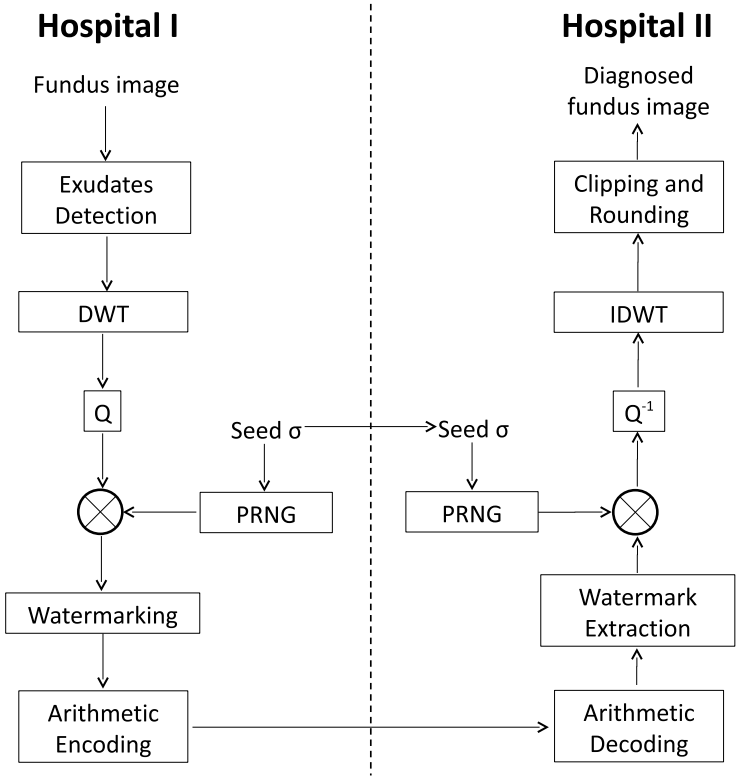


Fig.11: *Copyright- and privacy-protected diabetic retinopathy diagnosis network using JPEG 2000-based DWT sign scrambling (DWT: discrete wavelet transform, Q: quantization, PRNG: pseudo random number generator, Q^{-1}: inverse quantization, IDWT: inverse discrete wavelet transform) [6].*

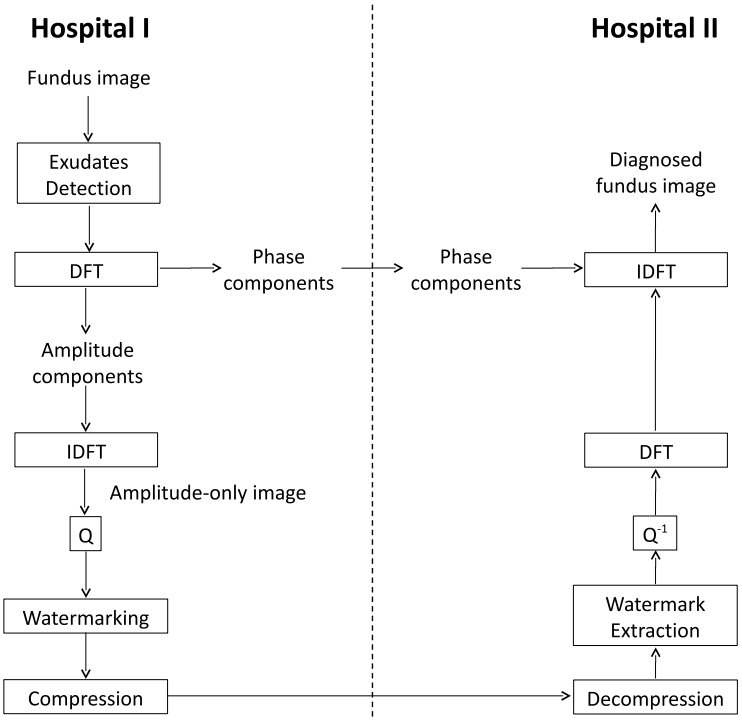


Fig.12: *Copyright- and privacy-protected diabetic retinopathy diagnosis network using DFT AOI (DFT: discrete Fourier transform, IDFT: inverse discrete Fourier transform, Q: quantization, Q^{-1}: inverse quantization) [4, 5].*

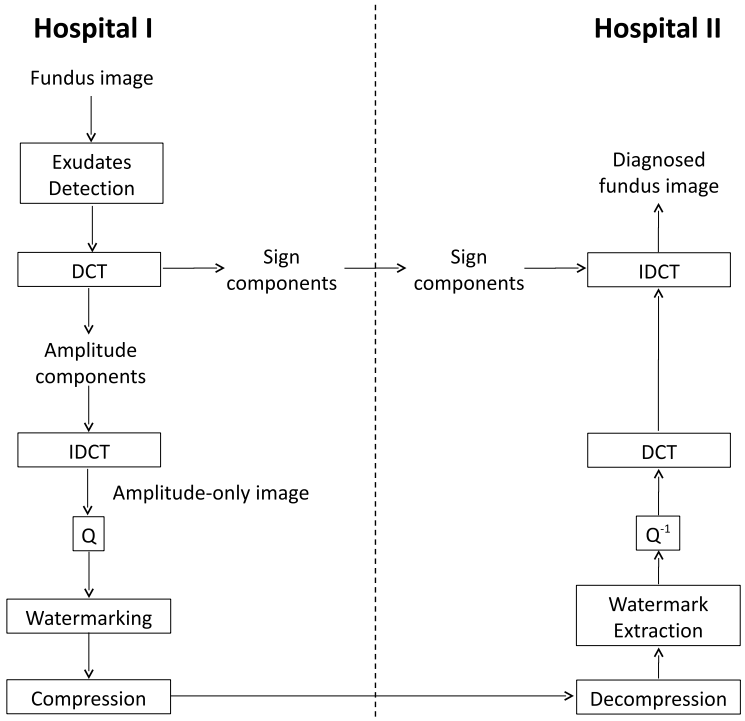
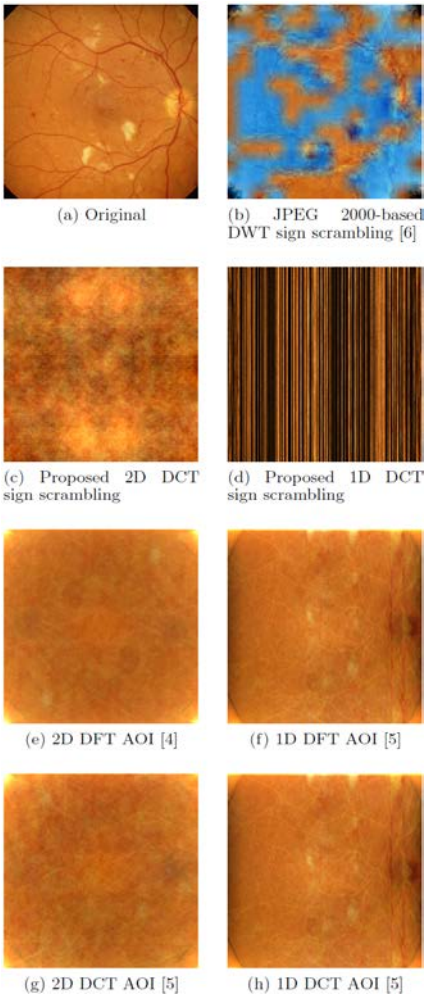


Fig.13: *Copyright- and privacy-protected diabetic retinopathy diagnosis network using DCT AOI (DCT: discrete cosine transform, IDCT: inverse discrete cosine transform, Q: quantization, Q^{-1}: inverse quantization) [5].*



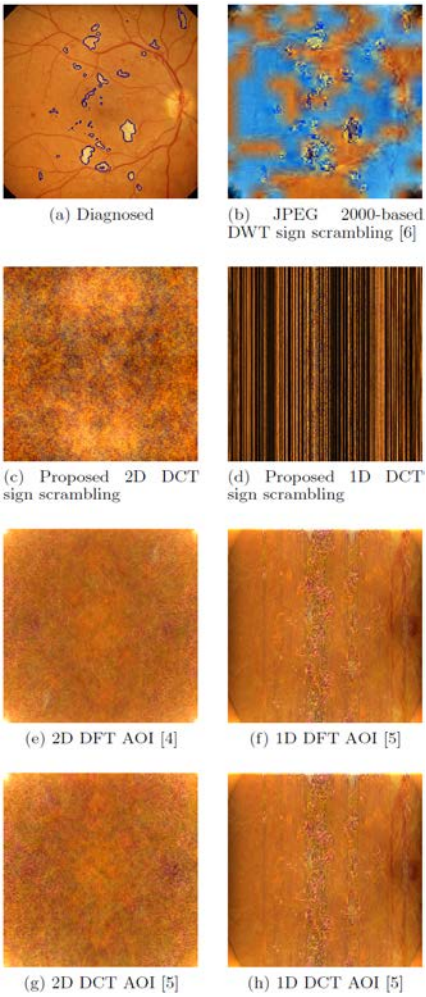
(a) Original (b) JPEG 2000-based DWT sign scrambling [6]

(c) Proposed 2D DCT sign scrambling (d) Proposed 1D DCT sign scrambling

(e) 2D DFT AOI [4] (f) 1D DFT AOI [5]

(g) 2D DCT AOI [5] (h) 1D DCT AOI [5]

Fig 14. Visually encrypted fundus image "HRIS001."



(a) Diagnosed (b) JPEG 2000-based DWT sign scrambling [6]

(c) Proposed 2D DCT sign scrambling (d) Proposed 1D DCT sign scrambling

(e) 2D DFT AOI [4] (f) 1D DFT AOI [5]

(g) 2D DCT AOI [5] (h) 1D DCT AOI [5]

Fig 15. Visually encrypted diagnosed fundus image "HRIS001."

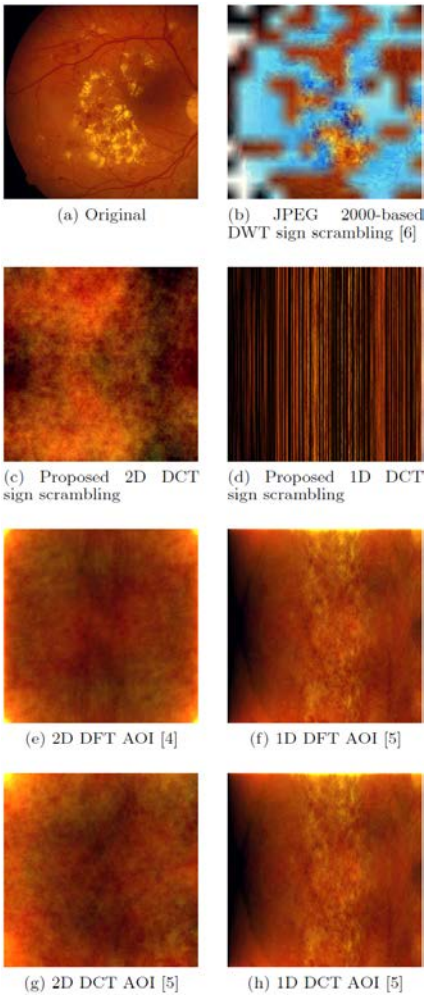


Fig 16. Visually encrypted fundus image "HRIS003."

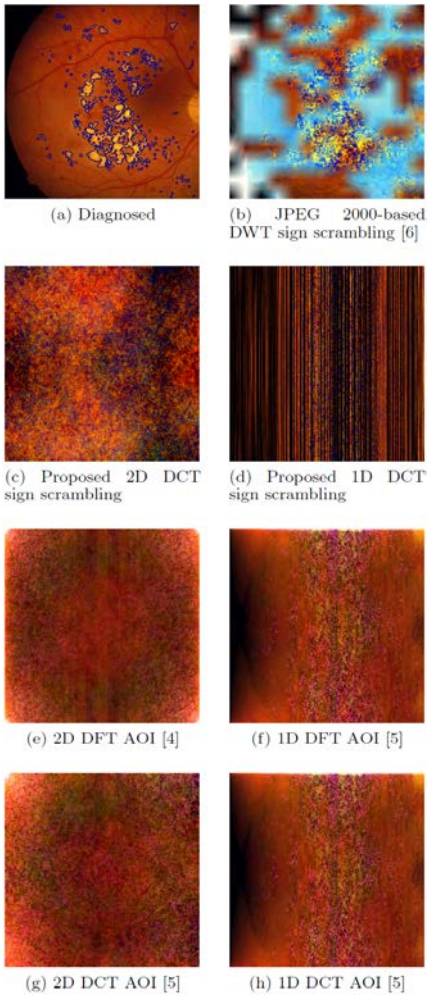


Fig 17. Visually encrypted diagnosed fundus image "HRIS003."

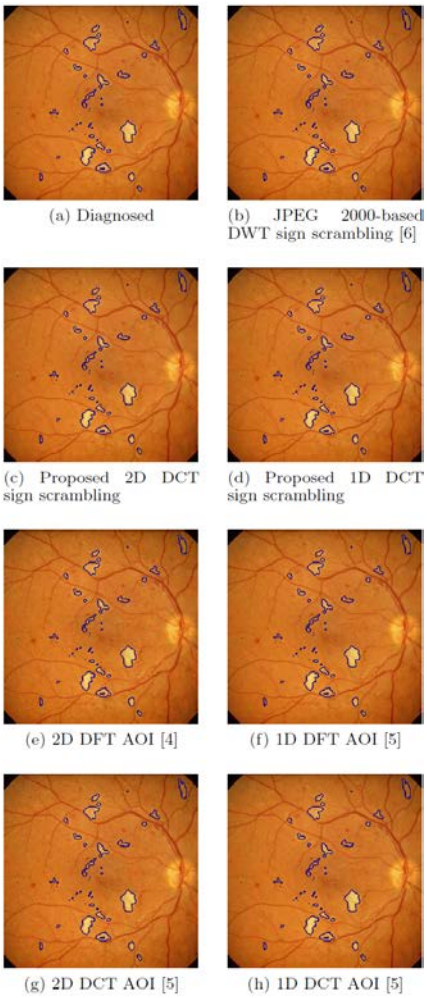


Fig 18. Decrypted diagnosed fundus image "HRIS001."

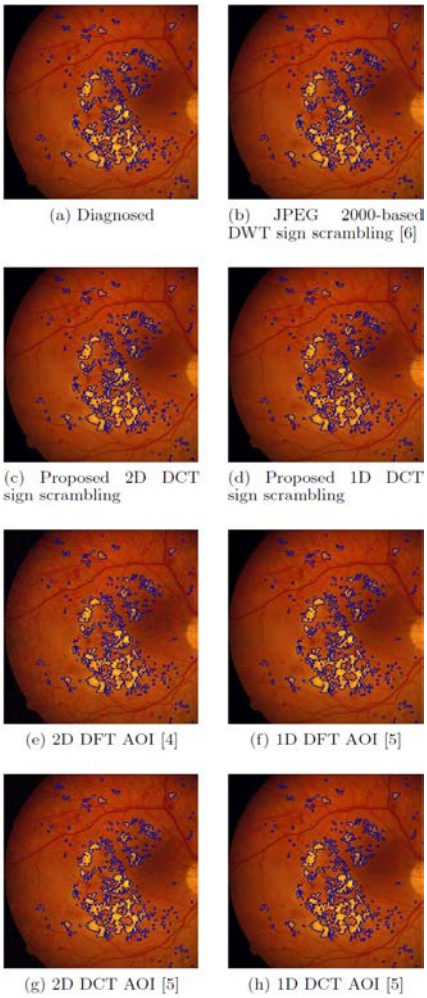


Fig 19. Decrypted diagnosed fundus image "HRIS003."

In-mole-fraction of InGaAs Insertion Layers Effects on the Structural and Optical Properties of GaSb Quantum Dots Grown on (001) GaAs Substrate

Suwit Kiravittaya[1], **Kamonchanok Khoklang**[2], **Supachok Thainoi**[3], **Somsak Panyakeow**[4], and **Somchai Ratanathammaphan**[5], Non-members

ABSTRACT

GaSb quantum dots (QDs) have been grown by solid-source molecular beam epitaxy on a 4-monolayer (ML) $In_xGa_{1-x}As$ (x = 0.07, 0.15, 0.20, and 0.25) to investigate the effects of In-mole-fraction of InGaAs insertion layers on the structural and optical properties of the GaSb QDs. The density of grown GaSb QDs is approximately 1.2-2.8×10^9 cm^{-2} on InGaAs insertion layers which depends on the In-mole-fraction. Dot shape and size change substantially when In-mole-fraction of InGaAs insertion layers is varied. The uniformity of GaSb QDs improves when the indium content increases. The change in free-standing QD morphology is likely due to the modified strain at different values of indium compositions in InGaAs insertion layers. The effects of In-mole-fraction of InGaAs insertion layer on optical properties of the QDs are studied by photoluminescence (PL). PL results show the blueshift of the emission when the indium content in InGaAs insertion layer increases.

Keywords: InGaAs Insertion Layers, GaSb Quantum Dots, Molecular-beam Epitaxy System, GaAs, Stranski-Krastanov

1. INTRODUCTION

Semiconductor quantum dots (QDs) have gained considerable interests due to their promising properties for novel device applications in the past few decades. Many electronic and optoelectronic devices based on IIIV QDs have been realized and investigated [1-6]. They include single-electron transistor, QD laser, light-emitting diode (LED), photodectectors and memory devices. Among these devices, self-assembled In(Ga)As/GaAs QDs grown in Stranski-Krastanov growth mode are major investigated. However, spontaneous formation of QDs on substrate surface allows us to fabricate not only In(Ga)As/GaAs QDs but also QDs in many different strained material systems. GaSb/GaAs QDs, which have a type-II band alignment, might be useful and more suitable for some applications. For example, the GaSb/GaAs QDs have been proposed to use in memory device application. Due to a very large valence band offset of ∼0.6 eV, long storage time ($> 10^{-5}$ s) is expected [5]. Longer storage time might be archived by using GaSb QDs grown on different materials such as AlAs.

Challenges in QD technology are related to the control of spontaneous QD formation process [7]. In order to utilize QD technology, detail understandings of natural growth process must be first investigated. For most of QD applications, these self-assembled QDs must be capped. The segregation and intermixing effects for each material system e.g. GaSb/GaAs must be understood. The performance of QD devices depends on the electronic properties. Probing them by structural and optical investigations is necessary.

For optoelectronic devices, GaSb/GaAs QDs have been realized and some unique properties on different devices, such as wide tunable wavelength for LED and higher operation temperature for infrared photodetectors due to the large hole confinement and long carrier lifetime, have been demonstrated [8]. Device performance strongly depends on the QD properties. It is therefore of interest to investigate structural properties of obtained QDs under various growth conditions in order to fine-tune them for any specific applications. Nevertheless, the studies on the structural and optical properties of GaSb/InGaAs/GaAs QDs are still lacking.

In the past, there are reports on the growth as well as structural and optical properties of InAs QDs grown in InGaAs matrix [9,10]. This structure is known as dot-ina- well structure. The obtained InAs QDs have significant different properties. Due to the increasing number of adjustable parameters, perfor-

Manuscript received on August 14, 2015 ; revised on March 7, 2016.
Final manuscript received on September 3, 2016.
[1] The author is with Advanced Optical Technology (AOT) Laboratory, Department of Electrical and Computer Engineering, Naresuan University, Phitsanulok, Thailand, E-mail: suwitki@gmail.com
[2,3,4,5] The authors are with The Semiconductor Device Research Laboratory (SDRL), Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand, E-mail: kamonchanok.r19@gmail.com, supachok.t@chula.ac.th, s_panyakeow@yahoo.com and Somchai.R@chula.ac.th

mance of specific devices based on this structure can be optimized. In this work, we are interested in tuning the GaSb QD properties by adjusting the indium content of InGaAs layer below GaSb QDs.

In this paper, we present a study on the structural and optical properties of GaSb/InGaAs type-II QD material system with $In_xGa_{1-x}As$ insertion layers with $x = 0.07$, 0.15, 0.20, and 0.25 are introduced. All samples are fabricated by molecular beam epitaxy (MBE) and the QD structures are investigated by atomic force microscope (AFM). The photoluminescence (PL) measurements are performed to examine the optical properties of the samples at 20 K by lock-in method using a 514.5-nm Ar ion laser as an excitation source and a liquid-nitrogen cooled InGaAs photodetector.

2. EXPERIMENTS

Samples studied here were grown on (001) GaAs substrates in a solid-source molecular beam epitaxy system (Riber Compact 21) equipped with a reflection high-energy electron diffraction (RHEED) observation system and an Sb valved cracker cell. Other cells (Ga, In and As_4) are standard effusion cells. Valved cracker cell is used in this work because a relatively fast switching time between As-rich and Sb-rich environment is needed. After the desorption of surface oxide on GaAs substrate at 580°C under As4 beam in As-rich condition, a 300-nm thick GaAs buffer layer was grown at 600°C with the growth rate of 0.5 monolayer per second (ML/s). During the growth, RHEED pattern was observed and well-prepared GaAs buffer layer (flat (001) GaAs surface) showed a clear (2×4) surface reconstruction. In order to grow InGaAs insertion layer, the substrate temperature was ramped down to 500°C. After the substrate

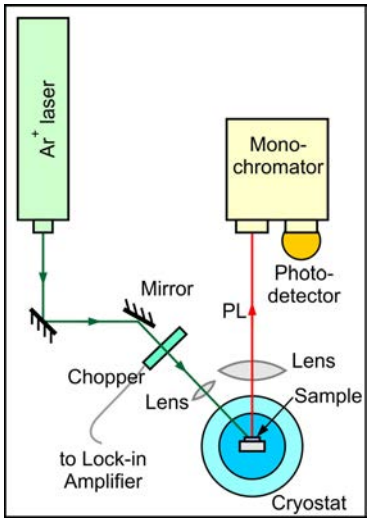


Fig.1: *Schematic of the PL setup for optical investigation of GaSb QD samples.*

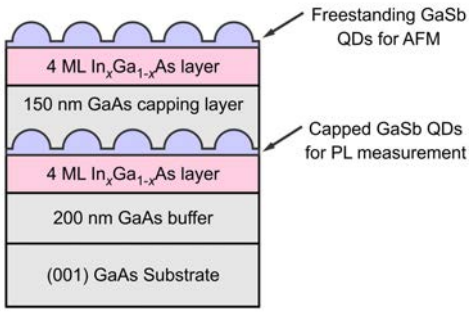


Fig.2: *Schematic structure of 3-ML GaSb QDs with 4-ML InGaAs insertion layer on undoped (001) GaAs substrate.*

temperature stabilized, 4-ML $In_xGa_{1-x}As$ was grown. Indium and gallium growth rates were changed in order to obtain the desired content of indium in InGaAs layer. For $x = 0.07$, 0.15, 0.20, and 0.25, the indium (gallium) rate are 0.025, 0.025, 0.025, and 0.033 (0.33, 0.14, 0.10, and 0.10) ML/s, respectively. After the 4-ML InGaAs insertion layer growth, the sample surface was soaked by Sb flux for 60 s. Self-assembled GaSb QDs are obtained by depositing 3-ML GaSb at the gallium growth rate of 0.1 ML/s. The V/III flux ratio (Sb/Ga) is kept constant at 4. For optical investigation, the QDs were capped with a 150-nm-thick GaAs and the 4-ML InGaAs insertion layer and GaSb QDs were grown again on top surface without capping. The top GaSb QDs are for surface morphology examination. After the growth, the sample was cooled down immediately. The surface morphology was characterized by AFM (Seiko SPA-400) in dynamic force mode in air. For PL tests, a 514.5-nm Ar-ion laser was used an excitation source, and data were obtained by a monochromator (HORIBA Jobin Yvon, iHR320) and InGaAs photodetector (Hamamatsu G9494). During the PL measurement, the sample is cooled down to 20 K by using a closed circuit helium cryostat. The schematic of the PL setup is shown in Fig. 2 and the schematic structure of the investigated samples is displayed in Fig. 2. For this work, only the indium content of InGaAs insertion layer is varied.

3. RESULTS AND DISCUSSION

Figure 3 shows AFM images of GaSb QDs on $In_xGa_{1-x}As$ (Figs. 3(a)-(d)) surfaces from the central regions of each sample. On the left column, AFM images with the colorscale corresponding to the height scale are shown while the same images with the colorscale corresponding to the slope scale is shown in the right column. The slope scale is obtained by calculating the amplitude of the surface gradient, i.e., if we define the AFM height data as $h(x, y)$, the image with the colorscale corresponding to the surface slope is

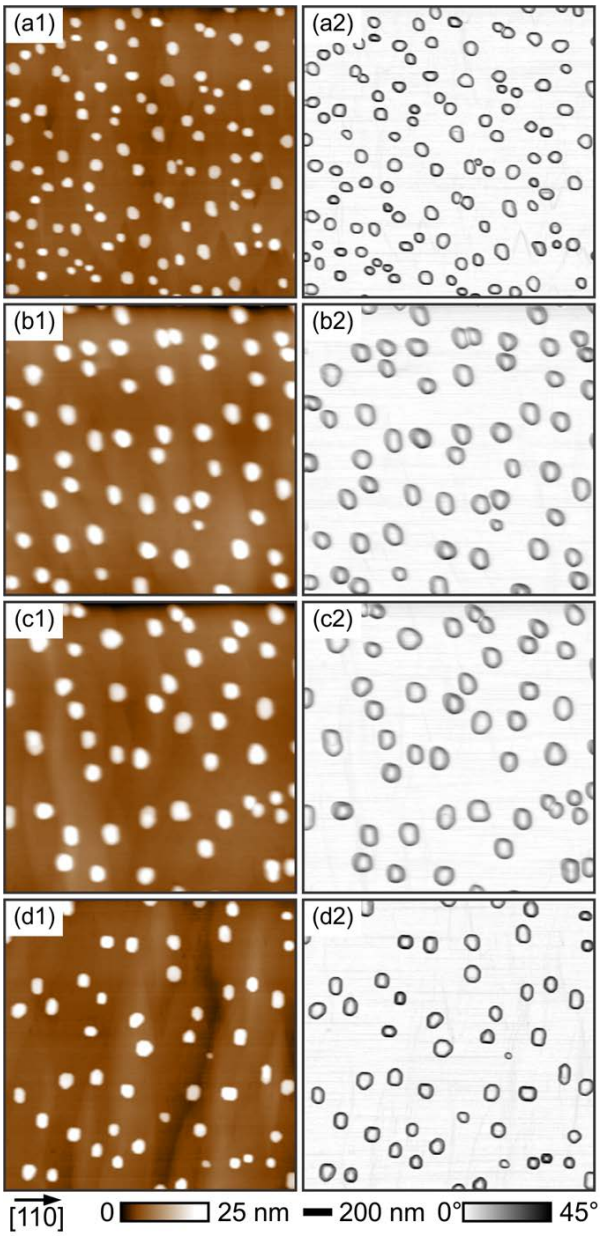


Fig.3: *(a) - (d) 2×2 μm^2 AFM images of GaSb QDs on 4-ML $In_xGa_{1-x}As$ with x = 0.07, 0.15, 0.20, and 0.25, respectively. Left column shows the images in height scale and right column shows the images in slope scale.*

$$|\nabla h| = \sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}. \qquad (1)$$

From all AFM images, we can clearly indicate the GaSb QDs on the surface. Wetting layer surface has a natural pattern of surface steps aligned along [1-10] direction. GaSb QDs are typically located on the steps and well-distributed over the whole surface. This result is consistent with the reports of the InAs/GaAs QD growth [11].

Distinct surface morphology is observed when In-

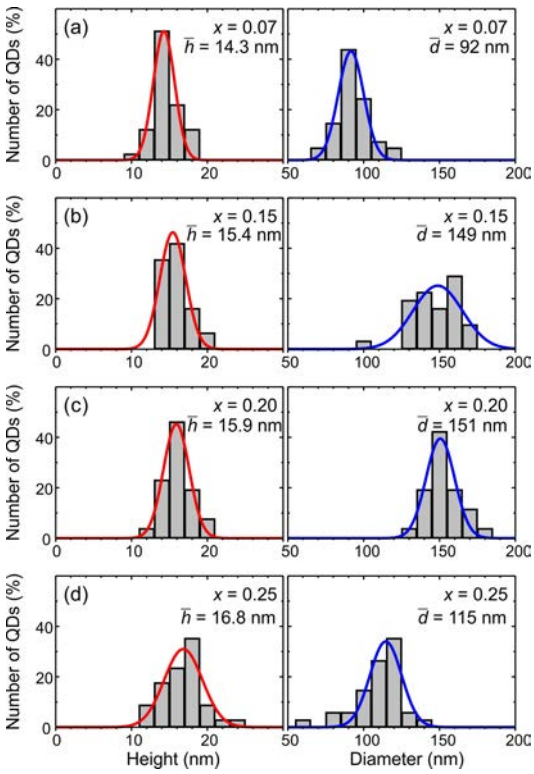


Fig.4: *(a) - (d) histograms of QD height and diameter distribution of GaSb QDs on 4 ML $In_xGa_{1-x}As$ with x = 0.07, 0.15, 0.20, and 0.25, respectively. Solid lines are Gaussian fits. Average height $\bar{h}$ and average diameter $\bar{d}$ are indicated in the figures.*

GaAs layer is introduced. The dot densities are approximately 1.2-2.8×10^9 cm^{-2}. QD density substantially decreases and QD size increases when the indium content increases from x = 0.07 (Fig. 3(a)) to 0.25 (Fig. 3(d)). Gradual morphology changes are still observed. The elongation direction of the base slightly changes from [110] to [1-10] when InGaAs insertion layer is introduced. This is due to the presence of different amount of indium in the InGaAs insertion layer. The uniformity of GaSb QDs improves when the indium content increases. Increasing more indium content into InGaAs insertion layer might induce self-assembled InGaAs QD formation prior to the growth of GaSb QDs.

From each AFM image of GaSb QDs shown in Fig. 3, analysis of height and diameter of each free-standing GaSb QD is performed. QD height is defined as the difference between surface level of QD apex and its base while the diameter is calculated from the largest closedloop contour line of individual QD. Area is converted to QD diameter d by using circular approximation, i.e.,

$$d = \sqrt{4A_c/\pi}, . \qquad (2)$$

where A_c is area of the largest closed-loop contour line. Height and diameter distribution of GaSb QDs

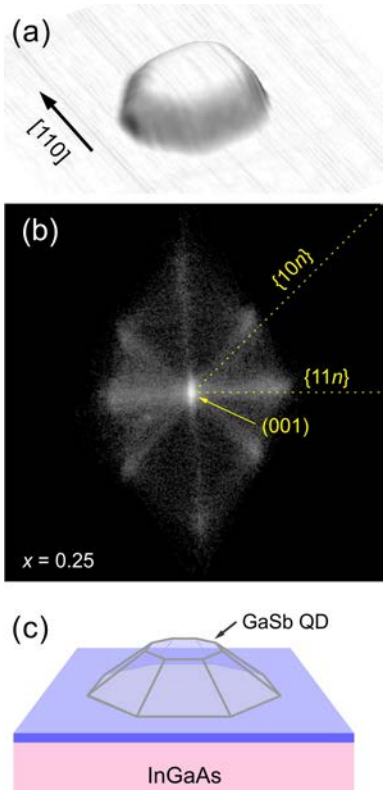


Fig.5: *(a) a magnified AFM image of single GaSb QD grown on 4-ML $In_{0.25}Ga_{0.75}As$, (b) facet plot of the free-standing QD surface and (c) schematic of the GaSb QD shape grown on InGaAs.*

on different InGaAs insertion layers is displayed as histograms in Fig. 4. All height and diameter data can well be fitted with Gaussian (normal) distribution functions. From this result, one can clearly see that both QD height and diameter increase when the diameter increase, except only at the indium content of $x = 0.25$ where we observe the reduction of diameter. This non-monotonic diameter change might be due to the excessively high indium content on the surface. Some indium atoms may diffuse and incorporate into the GaSb QDs. The presence of indium in GaSb increases the mismatch strain and therefore modifies the QD shape. These observations are due to the change of initial surface energy, interface energy as well as the strain energy of the QD system [12].

Closer look on individual QDs, one can observe distinct QD shape [13]. For GaSb grown on InGaAs insertion layer, the free-standing QDs have a flat (001) surface on the top. Figure 5(a) shows a magnified AFM image under perspective view. The GaSb QD is grown on $In_{0.25}Ga_{0.75}As$ surface. In order to quantify the orientation of QD surface, we perform a statistical analysis of QD surface measured by AFM (Fig. 3). The surface slopes *on QD surface* in two dimensions, i.e., $\partial h/\partial x$ and $\partial h/\partial y$) are collected and plotted in a form of two-dimensional histogram. The result for the GaSb QDs grown on $In_{0.25}Ga_{0.75}$ As is shown in Fig. 5(b). From this result, we see clearly the spot at the origin. This confirms a flat (001) surface of QD. Moreover, preferential facetted orientations of QD surface can also be extracted. They are $\{10n\}$ and $\{11n\}$, where n is a positive integer. Slightly misaligned of the indicated dotted lines for $\{10n\}$ and $\{11n\}$ is due to the slightly misalignment of the sample occurred during the AFM scan.

According to the statistical analysis shown above, we can draw a schematic of obtained GaSb/InGaAs QD shape as shown in Fig. 5(c). The QD consists of multi-facetted surface with a flat (001) on the top. This modelled QD shape might be important for further analysis of QD, i.e., for developing an overgrowth model [14] or a comparison with the results from other measurement techniques such as cross-sectional scanning tunnelling microscopy [12].

In order to utilize QDs in optoelectronic applications, QDs must be buried into a matrix. We therefore investigate the optical properties of capped GaSb QDs by using PL spectroscopy. We have interested in tuning the QD electronic properties by changing the indium content in InGaAs insertion layer. Both QD size and shape and PL peak energy are our measures. Consistency to our results from free-standing QD array is also observed from the PL measurement of capped samples. The capping process induces shape change and arsenic-intermixing into the capped QDs. PL reveals the optical characteristics of capped QDs. Figure 6 shows the PL spectra of capped GaSb/InGaAs QDs. It clearly shows the difference in the 20 K PL spectra under In-mole fraction of InGaAs insertion layer. PL of the capped samples shows the emission of QD peak at 1.2-1.3 eV. This corresponds to the emission from bulk $GaAs_{1-y}Sb_y$

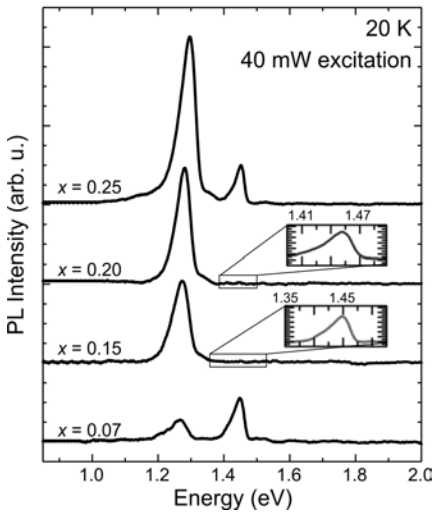


Fig.6: *PL measurement of capped GaSb QDs with 4-ML $In_xGa_{1-x}As$ insertion layer. Insets show the 10× magnification of the peaks around 1.45 eV.*

with the Sb content of only 11-17% [15]. However, this is an overestimated intermixing value because the quantum confinement and complicated band structure have not been taken into account. More realistic estimation of the Sb content in capped GaSb QDs needs more experimental evidence and/or sophisticated calculation techniques.

Concerning the line width of the QD peaks emitted at 1.2-1.3 eV, the line width shows *no* systematic change when the indium content of InGaAs insertion layer is varied. They are in the range of 55-63 meV. This non-systematic change and un-correlation with the height and diameter distribution of free-standing QDs (Fig. 4) might be due to the complication of the type-II band alignment and the intermixing process occurred during the capping. Overgrowth process is a subject for future experimental investigation.

According to the PL result (Fig. 6), the blueshift of the PL emission is observed when the indium content in InGaAs insertion layer increases. This is initially counter-intuitive because higher indium content should decrease the bandgap of the barrier layer. However, we can explain the result by the fact that the indium in the insertion layer does also alter the growth as well as the strain in the capped QD. The observed blueshift of PL emission from GaSb QDs is possibly due to the reduction of strain in the dots. This has been observed and reported without any clear explanation in Ref. [16]. This strain reduction results in the shift of the heavy-hole band and the quantized hole energy level [15]. Interestingly, relative PL intensity increases with the increasing indium content at the same excitation power (40 mW) and temperature (20 K). This indicates that the confinement of carriers in InGaAs insertion layer can possibly enhance the PL intensity, which relates to the carrier recombination in this nanostructure.

The unique characteristic of a type-II band alignment is seen when the laser power is varied [17]. As the density of photo-generated carriers in the dots increases, so does the strength of the Coulomb attraction to the electrons. Thus, as the excitation density is increased, the PL peak shifts to higher energy. This has been observed in GaSb/GaAs QDs. The expected behaviour is that the emission energy shows the third-root-of-excitation-intensity dependence [17, 18]. Figures 7(a) and 7(b) show comparative schematic of band diagram for type-II GaSb QDs when the carrier density is low (Fig. 7(a)) and high (Fig. 7(b)). Figure 7(c) shows an example of the normalized PL spectra of GaSb QDs grown on 4-ML $In_{0.25}Ga_{0.75}As$ insertion layer at 20 K with excitation intensities of 5 and 40 mW. The peaks at A and B are at 1.256 and 1.443 eV, respectively. The peaks at A' and B' are 1.28 and 1.448 eV, respectively. The peaks at A and A' are attributed to GaSb QD emission because of the exchange between antimony and arsenic on the dot surfaces during GaAs overgrowth or smaller

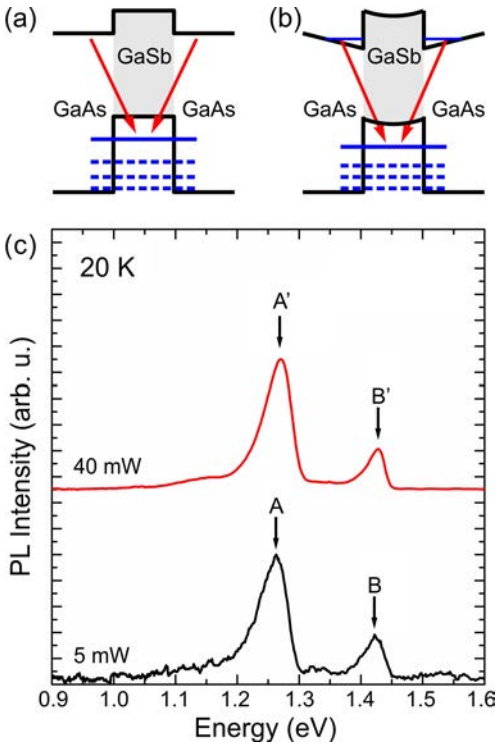


Fig.7: *Schematics of the recombination in type-II QD nanostructure when carrier density is (a) low and (b) high. Band bending is due to Coulomb attraction between photo-generated electrons and holes. (c) The normalized PL spectra of GaSb QDs grown on 4-ML $In_{0.25}Ga_{0.75}As$ insertion layer at 20 K with an excitation intensity of 5 and 40 mW. Slightly blueshift of both peaks is observed.*

effective dot size after overgrowth [19, 20]. Slightly blueshift of the QD peak is observed when the excitation power increases. This observation is consistent with the theoretical characteristics of type-II QDs. We attribute the peaks B and B' at around 1.44 eV, which has been reported to occur in Si (As-site) with bulk GaAs, to other system-dependent impurities [21]. In this case (peak B-B'), the blueshift of the emission energy is also seen with increasing excitation intensity.

4. CONCLUSION

We have systematically investigated the effects of In-mole-fraction of InGaAs insertion layers on the structural and optical properties of the GaSb QDs grown on (001) GaAs substrate. Self-assembled growth process is used to realize GaSb QDs. The density of obtained GaSb QDs on InGaAs insertion layers is approximately 1.2-2.8$\times 10^9$ cm^{-2}. The uniformity of free-standing GaSb QDs is improved when the indium content increases. Observation of QDs reveals some preferential surface orientation of facetted QD with flat-top (001) surface. The PL peak energy and intensity can be altered by changing the indium

composition of the InGaAs insertion layer. This indicates that the electron confinement in InGaAs insertion layer can modify the recombination in this structure. This work enhances the basic understanding of the structural and optical properties of GaSb QDs grown on InGaAs surface.

ACKNOWLEDGEMENT

This work is supported by Naresuan University, Research Chair Grant, the National Science and Technology Development Agency (NSTDA), Thailand, Asian Office of Aerospace Research and Development (AOARD) (Grant No. FA 2386-16-1-4003), Office of Naval Research Global (ONRG), Thailand Research Fund (TRF), Nanotechnology Center of Thailand (NANOTEC), Ratchadapisek Sompoch Endowment Fund (2015), Chulalongkorn University (CU-58-051-EN), the Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission (EN264A) and Ratchadaphisek Sompoch Endowment Fund of Chulalongkorn University (RES560530079-EN). We also wish to thank Dr. Maetee Kunrugsa, Ms. Patchareewan Prongjit, Ms. Nattapa Prapasawad, and Mr. Samatcha Vorathamrong for their assistances.

References

[1] V. M. Ustinov, A. E. Zhukov, N. A. Maleev, A. R. Kovsh, S. S. Mikhrin, B. V. Volovik, Y. G. Musikhin, Y. M. Shernyakov, M. V. Maximov, A. F. Tsatsul'nikov, N. N. Ledentsov, Zh. I. Alferov, J. A. Lott, and D. Bimberg, "1.3 μm InAs/GaAs quantum dot lasers and VCSELs grown by molecular beam epitaxy," *J. Cryst. Growth*, Vol. 227-228, pp.1155-1161, 2001.

[2] J. Sabarinathan, P. Bhattacharya, P-C. Yu, and S. Krishna, "An electrically injected InAs/GaAs quantum-dot photonic crystal microcavity light-emitting diode," *Appl. Phys. Lett.*, Vol. 81, pp. 3876-3878, 2002.

[3] K. Hirakawa, S.-W. Lee, Ph. Lelong, S. Fujimoto, K. Hirotani, and H. Sakaki, "High-sensitivity modulation-doped quantum dot infrared photodetectors," *Microelec. Eng.*, Vol. 63, pp. 185-192, 2002.

[4] K. Tanabe, D. Guimard, D. Bordel, and Y. Arakawa, "High-efficiency InAs/GaAs quantum dot solar cells by metalorganic chemical vapor deposition," *Appl. Phys. Lett.*, Vol. 100, pp. 193905, 2012.

[5] M. Hayne, R. J. Young, E. P. Smakman, T. Nowozin, P. Hoidgson, J. K. Garleff, P. Rambabu, P. M. Koenraad, A. Marent, L. Bonato, A. Schliwa, and D. Bimberg, "The structural, electronic and optical properties of GaSb/GaAs nanostructures for charge-based memory," *J. Phys. D: Appl. Phys.*, Vol. 46, pp. 264001, 2013.

[6] C.-C. Tseng, W.-H. Lin, S.-Y. Wu, S.-H. Chen, and S.-Y. Lin, "The transition mechanisms of type-II GaSb/GaAs quantum-dot infrared light-emitting diodes," *J. Cryst. Growth*, Vol. 323, pp. 466-469, 2011.

[7] S. Kiravittaya, A. Rastelli, and O. G. Schmidt, "Advanced quantum dot configurations," *Rep. Prog. Phys.*, Vol. 72, pp. 046502, 2009. and references therein.

[8] W.-H. Lin, C.-C. Tseng, S.-Y. Wu, W.-H. Wu, S.-Y. Lin, and M-C. Wu, "The influence of background As on GaSb/GaAs quantum dots and its application in infrared photodetectors." *Phys. Status Solidi C*, Vol. 9, No.2, pp. 314-317, 2012.

[9] A. Stinz, G. T. Liu, H. Li, L. F. Lester, and K. J. Malloy, "Low-threshold current density 1.3-m InAs quantum-dot lasers with the dots-in-a-well (DWELL) structure," *IEEE Photon. Technol. Lett.*, Vol. 12, pp. 591-593, 2000.

[10] R. S. Attaluri, S. Annamalai, K. T. Posani, A. Stinz, and S. Krishna, "Effects of Si doping on normal incidence InAs/In0.15Ga0.85As dots-in-well quantum dot infrared photodetectors," *J. Appl. Phys.*, Vol. 99, pp. 083105, 2006.

[11] E. Placidi, F. Arciprete, V. Sessi, M. Fanfoni, F. Patella, and A. Balzarotti, "Step erosion during nucleation of InAs/GaAs(001) quantum dots," *Appl. Phys. Lett.*, Vol. 86, pp. 241913, 2005.

[12] H. Eisele and M. Dähne, "Critical thickness of the 2-dimensional to 3-dimensional transition in GaSb/GaAs(001) quantum dot growth," *J. Cryst. Growth*, Vol. 338, pp.103-106, 2012.

[13] K. Khoklang, S. Kiravittaya, M. Kunrugsa, P. Prongjit, S. Thainoi, S. Ratanathammaphan, and S. Panyakeow, "Molecular beam epitaxial growth of GaSb quantum dots on (0 0 1) GaAs substrate with InGaAs insertion layer," *J. Cryst. Growth*, Vol. 425, pp.291-294, 2015.

[14] R. Songmuang, S. Kiravittaya, and O. G. Schmidt, "Shape evolution of InAs quantum dots during overgrowth," *J. Cryst. Growth*, Vol. 249, pp. 416-421, 2003.

[15] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys," *J. Appl. Phys.*, Vol. 89, pp.5815-5875, 2001.

[16] J. Tatebayashi, A. Khoshakhlagh, S. H. Huang, G. Balakrishnan, L. R. Dawson, D. L. Huffaker, D. A. Bussian, H. Htoon, and V. Klimov, "Lasing characteristics of GaSb/GaAs self-assembled quantum dots embedded in an InGaAs quantum well," Appl. Phys. Lett., Vol. 90, pp. 261115, 2007.

[17] N. N. Ledentsov, J. Böhrer, M. Beer, F. Heinrichsdorff, M. Grundmann, D. Bimberg, S. V. Ivanov, B. Ya Meltser, S. V. Shaposhnikov, I. N. Yassievich, N. N. Faleev, P. S. Kop'ev, and Zh. I. Alferov, "Radiative states in type-II GaSb/GaAs

quantum wells," *Phys. Rev. B*, vol. 52, pp. 14058-14066, 1995.

[18] I. Farrer, M. J. Murphy, D. A. Ritchie, and A. J. Shields, "Room temperature 1.3μm emission from self-assembled GaSb/GaAs quantum dots," *J. Cryst. Growth*, Vol. 251, No.1-4, pp. 771-776, 2003.

[19] G. Balakrishnan, J. Tatebayashi, A. Khoshakhlangh, S. H. Huang, A. Jallipalli, L. R Dawson, and D.L. Huffaker, "III/V ratio based selectivity between strain Stranki-Krastanov and strain-free GaSb quantum dots on GaAs," *Appl. Phys. Lett.*, Vol. 89, pp. 161104, 2006.

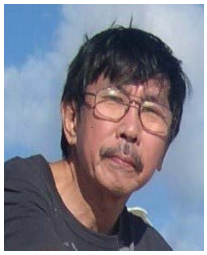
[20] C.-K. Sun, G. Wang, J. E. Bowers, B. Brar, H.-R. Blank, H. Kroemer, and M. H. Pikunhn, "Optical investigation of the dynamic behavior of GaSb/GaAs quantum dots," *Appl. Phys. Lett.*, Vol. 68, pp.1543-1545, 1996.

[21] E. H. C. Parker, *The Technology and Physics of Molecular Beam Epitaxy*, Sir John Cass School of Physical Science and Technology city of London Polytechnic London, England pp. 107-109, 1985.

Suwit Kiravittaya was born in Bangkok, Thailand, in 1976. He received the B.S. and Ph.D. degrees in Electrical Engineering from Chulalongkorn University in 1998 and 2003, respectively. After his Ph.D., he worked as a post-doc and became a scientist in Max-Planck Institute for Solid State Research, Stuttgart, Germany. In 2009, he became a scientist in the Institute for Integrative Nanoscience, IFW Dresden. Germany. Between May 2010 - October 2011, he led a research group named "Strained Nanomembranes for Optical, Electronic, and Fluidic Systems" in Institute for Integrative Nanoscience. Since 2012, he became a lecturer in the Department of Electrical and Computer Engineering, Faculty of Engineering, Naresuan University, Phitsanulok, Thailand. He is currently an assistant professor at Naresuan University. His current research interests are topics related to several advanced optical technologies including optical system for quantum information processing, novel opto-electronic devices. In 2015, he found a research unit so-called Advanced Optical Technology Laboratory.

Kamonchanok Khoklang was born in Phetchabun, Thailand, in 1990. She received the B.S. degrees in Advanced Materials and Nanotechnology Engineering from Silpakorn University in 2012 and M.S. degree in Electrical Engineering from Chulalongkorn University in 2015. From 2013 to 2014, she was with Osaka University as a research student by a supporting from JASSO, Japan. She current research interests and topics related to several advanced materials nanotechnology and optical technologies with semiconductor device.

Supachok Thainoi was born in Bangkok, Thailand, on October 12, 1956. He received the M.S. Degree in Nuclear Technology from Chulalongkorn University, Bangkok, Thailand in 1999. He is a Research Assistant in the Semiconductor Device Research Laboratory, Electrical Engineering Department, Faculty of Engineering, Chulalongkorn University. His current interests owe in the field of III-V Compound semiconductor and radiation detectors



Somsak Panyakeow received the B. Eng., M. Eng. and D. Eng. degrees, all in Electrical Engineering from Osaka University, Japan in 1969, 1971 and 1974 respectively. In 1974, he began to work at the Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University as a lecturer. He was a pioneer to set up the Semiconductor Device Research Laboratory (SDRL) at Chulalongkorn University in 1975. He was an Assistant Professor and an Associate Professor in 1977 and 1980 respectively. He was appointed to be a Full Professor of the department in 1982. He was promoted to be Professor C11 (the highest rank) in 1992. He has been engaged in research on photovoltaic devices and systems since 1975. He has been appointed to be Senior Research scholar by Thailand Research Fund (TRF) since 2005. In 2008, he was highly honored receiving prestigious "Dussadeemala" medal from his Majesty the King on his life-long research contribution. His recent research work is in the area of Molecular Beam Epitaxy for quantum devices and nanoelectronics & nanophotonics as well as for the development of high performance nanostructure solar cells.

Somchai Ratanathammaphan received the B.Eng, M.Eng and D.Eng. degree in Electrical Engineering from Chulalongkorn University, Thailand in 1984, 1988, and 1993, respectively. From 1986 to 1987, he was with Waseda University as a research student. From 1989 to 1990, he was with Tokyo Institute of Technology as a research student by a supporting from MEXT, Japan. During his doctoral work, he developed the combined technique of super-saturation and two-phase solution LPE growth of GaAs-GaAlAs single quantum well lasers by using a horizontal sliding boat with lateral sliding mechanism. In 1993, he joined the Department of Electrical Engineering, Chulalongkorn University as a lecturer. He became an assistant Professor and an Associate Professor in 2000 and 2003 respectively. Recently, he also has been engaged in the research of fabrication technologies of ultra-fine structure for nanoelectronics and photonics by molecular-beam epitaxy and the application of semiconductor nanowires on artificial photosynthesis and nanogenerators.

Low computational cost matching pursuit algorithm for LCX-based Intruder Detection System

Tomonori Sato[1], **Ziji Ma**[2], **Takeshi Higashino**[3], and **Minoru Okada**[4], Non-members

ABSTRACT

The leaky coaxial cable (LCX) based intruder detection system detects intruders by making efficient use of time-variation of the channel impulse response among the cables. Recently, a compressed sensing (CS)-assisted intruder detection system has been proposed for improving the accuracy of the intruder detection. Although the CS-assisted system can improve the false detection probability performance efficiently, it is difficult to implement it because of its huge computational cost requirement. This paper proposes a modified matching pursuit (MMP) algorithm for reducing the computational cost of the CS-assisted intruder detection scheme. MMP can reduce the computational cost by limiting the search range in the vicinity of impulse response peaks estimated at the last measurement. Computer simulation results show the proposed system can reduce the computational cost for the conventional CS-based algorithm without degradation in the false detection performance.

Keywords: LCX, Compressed Sensing, Matching Pursuit, Low Complexity Algorithm

1. INTRODUCTION

There has been a great demand for intruder detection systems for managing the securities of the key facilities such as the airports, rail tracks, military bases, nuclear power plants, schools, and so on. Many kinds of sensor devices, including infrared sensors, cameras and microwave radars, etc. are widely used for intruder detection. However, these sensors are not suitable for wide area surveillance because of limitation in the sensing range. To realize reliable intruder detection in wide areas, LCX-based intruder detection systems have been studied[1–3].

A typical LCX-based intruder detection system is composed of a pair of LCX cables and an impulse response measuring equipment. The wide-band RF (Radio Frequency) signal is transmitted through either one of the LCX cables. The signal is received by the other side of the cables. The signal is propagated among a pair of LCX cables. The received signal is applied to the impulse response estimator, where the cross-correlation between the transmitted and received signals is calculated to derive the impulse response of the propagation channel. When an intruder enters the region between the two cables, the intruder scatters the transmitted signal, and it induces a change in the impulse response. We can detect the intruder by this variation in the impulse response.

The LCX-based intruder detection system has several advantages over the conventional intruder detection systems. The LCX-based intruder detection system allows us to detect the intruders in a narrow and long range even with hundreds of meters. Furthermore, it can localize the position of intruders along the LCX cables.

Although the LCX based detection system has the attractive advantages, it has problems in the accuracy of the intruder detection. Especially, the false alarm rate becomes unacceptably high when the disturbances such as the vibration of the LCX cables occurs. Since the LCX-based intruder detection system detects an intruder by sensing the time variation of the channel impulse response, the vibration of LCX cables could induce a false alarm. Therefore a new detection algorithm is required to solve this problem.

The conventional schemes employ the Least Square (LS) algorithm to estimate the impulse response. For the LS method, additive noise gives considerable effect on the received signal, so it is difficult to improve false error rate performance. To solve the problem, we have proposed a detection method based on Compressed Sensing (CS) algorithm[4]. CS method can estimate the impulse response vector accurately in a noisy environment, only it the vector in sparse.

However, the proposed CS-based intruder detection system requires huge computational cost because it repeatedly performs large size matrix calculations. To reduce the computational cost, we propose a modified matching pursuit algorithm that can largely decrease the computational complexity for solving CS. Since the motion of the targets is relatively slow comparing with the measuring interval, the difference in target positions between the current and last estima-

Manuscript received on December 15, 2012 ; revised on May 29, 2014.

Final manuscript received on October 17, 2016.

[1,2,3,4] The authors were with Graduate School of Information Science, Nara Institute of Science and Technology(NAIST), Ikoma-shi,630-0192 Japan, E-mail: tomonori-sa@is.naist.jp, zijima@hnu.edu.cn, higa@is.naist.jp and mokada@is.naist.jp

Fig.1: *the structure of LCX-based intruder detection system*

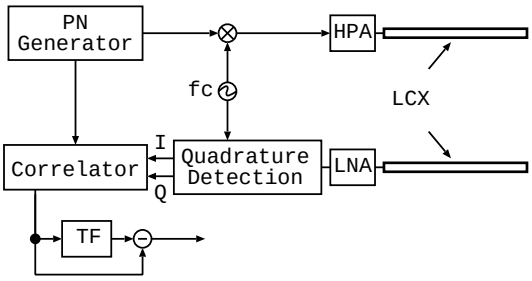


Fig.2: *System Model*

tion period is small. Firstly, the system estimates the largest positions using the conventional CS algorithm. After that, the search for matching pursuit is conducted only around the vicinity of the target positions estimated at the previous estimation. The search is repeatedly performed at the each estimation period. In this algorithm, we can reduce the computational cost thanks to the limited search range. Computer simulation result shows that our method can reduce the processing time for estimating the target positions.

2. INTRUDER DETECTION SYSTEM USING LCX

2.1 System Model

Figure 1 illustrates the overview of intruder detection system using LCX cables. In the system, a pair of LCX cables is implemented around vital areas as shown in Fig. 1. The sounding signal is transmitted from the transmitter-side LCX cable(LCX-Tx). The transmitted signal is received by the receiver-side cable(LCX-Rx). When an intruder enters the area between two cables, the sounding signal is scattered by the intruder. The signal processing unit measures the impulse response by calculating the cross-correlation between transmitted and received signals. We can detect the intruder by observing the change in the impulse response.

The block diagram of the transmitter and receiver are shown in Figure 2. The RF (Radio Frequency) carrier signal of the frequency, f_c, is modulated by the PN (Pseudo-random Noise) code generated by the PN generator. The modulated signal is amplified by the high power amplifier (HPA) and applied the transmitter-side LCX cable. The receiver(Rx) is connected to receiver-side LCX cable. The received signal is first amplified by the low noise amplifier (LNA) and then applied to the quadrature detector for frequency down-conversion. The channel impulse response is estimated by calculating the cross-correlation among the frequency down-converted received signal and the PN code. The estimated channel impulse response is stored in the time-frame memory (TF). Finally, the time-variation in the estimated channel impulse response is calculated by subtracting the stored response from the current one. We can detect the obstacles between LCX cables by observing the time-variation [5].

Let us go into the further details of the detection system. The transmitted signal in equivalent low-pass expression[6] $x(t)$ is given by

$$x(t) = \sum_{k=0}^{N-1} x_k f(t - kT_c), \tag{1}$$

where $x_k \in \pm 1$ is a PN sequence of length, N, T_c is the chip duration, and

$$f(t) = \begin{cases} \frac{1}{\sqrt{T_c}} & |t| \leq \frac{T_c}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

is the pulse waveform.

The generated signal, $x(t)$, is frequency up-converted, amplified and emitted from the the transmitter-side LCX cable. The transmitted signal propagates through the gap between the transmitter and receiver-side LCX cables. The Received signal from the receiver-side LCX cable is amplified, and frequency down-converted. The received signal in the equivalent low-pass expression is given by

$$y(t) = \int_{\tau=0}^{\infty} g(\tau; t)x(t - \tau)d\tau + z(t), \tag{3}$$

where $g(\tau; t)$ is an channel impulse response, and $z(t)$ is an Additive White Gaussian Noise (AWGN) component.

The channel impulse response can be divided into two components, namely,

$$g(\tau; t) = g^{normal}(\tau) + g^{scat}(\tau; t), \tag{4}$$

where $g^{normal}(\tau)$ and $g^{scat}(\tau; t)$ are the direct paths and the scattered signal components, respectively. The direct path component is considered as the impulse response when no obstacles between the paired

LCX cables exist. Therefore, $g^{normal}(\tau)$ is given by

$$g^{normal}(\tau) = \sum_{m=0}^{N_{txslot}} \sum_{n=0}^{N_{rxslot}} a_{m,n}\delta(\tau - \tau_{m,n})e^{-j\theta_{m,n}}, \tag{5}$$

where N_{txslot} and N_{rxslot} are the numbers of slot in transmitter and receiver LCXs respectively. Also, $a_{m,n}$ and $\tau_{m,n}$ denote the amplitude and delay time among the m-th slot of the transmitter-side LCX n-th slot of the receiver-side one.

The impulse response corresponding to scattered signal component is given by

$$g^{scat}(\tau;t) = \tag{6}$$
$$\sum_{m=0}^{N_{txslot}} \sum_{n=0}^{N_{rxslot}} \frac{\lambda^2\sigma_0}{(4\pi)^3 d_1^2 d_2^2}\delta(\tau - \tau'_{m,n})e^{-j(\theta'_{m,n}+2\pi f_{obst}t)},$$

where λ is the free-space wavelength. d_1 and d_2 are the distance between transmitter-side cable and the obstacle, and the distance between he obstacle and the receiver-side cable, respectively. f_{obst} is the Doppler frequency caused by the motion of the obstacle, and σ_0 is the effective cross-section due to the scattering obstacle.

The cross correlation between $y(t)$ and $x(t)$ is given by

$$R(\xi) = \int_0^{NT_c} y(t+\xi)x^*(t)dt. \tag{7}$$

Substituting Eq (3) into Eq (7), we can rewrite the correlation function as:

$$R(\xi) = \int_0^{\infty} \int_0^{NT_c} g(\tau;t)x(t+\xi-\tau)x^*(t)dtd\tau. \tag{8}$$

Let us assume that $g(\tau;t)$ is time-invariant between $t=0$ and $t=NT_c$, or the time variation in the channel impulse response can be neglected during the PN code duration. This implies that $g(\tau;t)$ is approximately written as:

$$g(\tau;t) = \tilde{g}(\tau). \tag{9}$$

Therefore, Eq (8) is now given by

$$R(\xi) = \int_0^{\infty} \tilde{g}(\tau) \int_0^{NT_c} x(t+\xi-\tau)x^*(t)dtd\tau. \tag{10}$$

According to the auto-correlation property of the PN code, the correlation function of the transmitted signal is derived as

$$\int_0^{NT_c} x(t)x^*(t-\tau)dt = \begin{cases} N & (\tau = 0) \\ \cong 0 & (\tau \neq 0) \end{cases}. \tag{11}$$

Substituting Eq (11) into Eq (10), we can obtain the channel impulse responses corresponding as:

$$R(\xi) \quad = \quad Ng(\tau)\mid_{\xi=\tau}. \tag{12}$$

The scattered signal component caused by an obstacle between the paired LCX cables, $g^{scat}(\tau)$, can be obtained by $g^{obst}(\tau) - g^{normal}(\tau)$ and finally the location of the intruding obstacle can be estimated.

2.2 Least Square Method

Let us define the vector form representation of the received signal as: $\mathbf{y} = [y_0, y_1, \ldots, y_{M-1}]^T$, where $y_k = y(kT_c)$ and $M = 2N-1$. Then, Eq (3) can be expressed as [7]

$$\mathbf{y} = \mathbf{X}\mathbf{g} + \mathbf{z}, \tag{13}$$

where

$$\mathbf{X} = \begin{bmatrix} x_0 & 0 & \cdots & 0 \\ x_1 & x_0 & \cdots & 0 \\ \vdots & x_1 & \ddots & 0 \\ x_{N-1} & \vdots & \ddots & x_0 \\ 0 & x_{N-1} & \ddots & x_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{N-1} \end{bmatrix} \tag{14}$$

is the convolution matrix, and $\mathbf{z} = [z_0, z_1, \ldots, z_{M-1}]$ is the AWGN vector.

According to Eq (11), the matrix $\mathbf{X}$ satisfies the following relationship:

$$\mathbf{X}^H\mathbf{X} \backeq N\mathbf{I}_N, \tag{15}$$

where $\mathbf{I}_N$ is the $N \times N$ identify matrix. The channel impulse response vector, $\mathbf{g}$, is given by

$$\mathbf{g}^{LS} = \arg\min_{\mathbf{g}} ||\mathbf{y} - \mathbf{X}\mathbf{g}||_2, \tag{16}$$

where $||\mathbf{a}||_2 = \sqrt{\sum |a_i|^2}$ is the l_2 norm of $\mathbf{a}$, and $\arg\min_{\mathbf{x}} f(\mathbf{x})$ gives $\mathbf{x}_0$ which minimizes $f(\mathbf{x})$.

Therefore, impulse response vector $\mathbf{g}^{LS}$ is represented by following equation.

$$\mathbf{g}^{LS} = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{y}. \tag{17}$$

According to Eq (15), we can get

$$\mathbf{g}^{LS} = \mathbf{X}^H\mathbf{y}. \tag{18}$$

2.3 Compressed Sensing Method

If the unknown vector is sparse, we can estimate the vector by making efficient use of compressed sensing (CS) method[8][9]. In the LCX based intruder detection system, the impulse response corresponding to the scattered signal has non-zero value only at the location where the intruder exists. Therefore,

the impulse response vector is sparse. We can apply CS to the channel response estimation of the intruder detection system.

Now, let us define that the impulse response vector corresponding to the scattered signal is $\mathbf{g}^{CS}$. Then, CS is formulated as [10, 11]:

$$\begin{aligned} \mathbf{g}^{CS} &= \arg\min_{\tilde{\mathbf{g}}} ||\tilde{\mathbf{g}}||_1 \\ s.t. &\quad \left|\left|\mathbf{X}^H(\mathbf{y}-\mathbf{X}\tilde{\mathbf{g}})\right|\right|_1 \le \lambda_c \end{aligned} \tag{19}$$

where $||\mathbf{a}||_1 = \sum_i |a_i|$ is the l_1 norm of $\mathbf{a}$, and

$$\lambda_c = \sqrt{2N\sigma(1+a)\log N} \tag{20}$$

is the noise margin, where σ is the standard deviation of the noise, and $a \ge 0$ is the relaxation factor. Eq. (20) can be solved by using the linear programming. However CS using the linear programming requires huge computational cost. Therefore, in the following, we propose an another algorithm for solving CS.

3. COMPUTATIONAL COST USING MODIFIED MATCHING PURSUIT

As we mentioned in the previous section, CS-based channel impulse response estimation requires huge computational cost. To reduce the computational cost, we propose a modification of matching pursuit algorithm. In the following, we will briefly explain the principle of Matching Pursuit[12]. Then we propose a modified algorithm for reducing the cost.

3.1 Matching Pursuit

Suppose that $\mathbf{g}^{mp}$ is an arbitrary m-sparse signal. The M-dimensional observe signal vector $\mathbf{y}$ is given by

$$\mathbf{y} = \mathbf{X}\mathbf{g}^{mp}, \tag{21}$$

where $M \times N$ matrix $\mathbf{X}$ whose *rows* are the measurement vectors. We refer to $\mathbf{X}$ as the measurement matrix and denote its columns by $[\mathbf{x}_0, ..., \mathbf{x}_{N-1}]$.

To identify the ideal signal $\mathbf{g}^{mp}$, we need to determine which columns of $\mathbf{X}$ participate in the measurement vector $\mathbf{y}$. The idea behind the algorithm is to pick columns in a greedy fashion. At each iteration, we choose the column of $\mathbf{X}$ that is most strongly correlated with the remaining part of $\mathbf{y}$. Then we subtract off its contribution to $\mathbf{s}$ and iterate on the residual. One hopes that, after m iterations, the algorithm will have identified the correct set of columns.

The procedure of the recovering signals is represented by the following steps.

1. Initialize the residual $\mathbf{r}_0 = \mathbf{y}$, the index set $\Lambda_0 = 0$, and the iteration counter $i = 1$.
2. Find the index λ_i that solves the easy optimization problem

$$\lambda_i(t) = \arg \max_{k=0,...,N-1} |\langle \mathbf{r}_{i-1}, \mathbf{x}_k \rangle|. \tag{22}$$

3. Augment the index set and the matrix of chosen atoms: $\Lambda_i = \Lambda_{i-1} \cup \{\lambda_i\}$ and $\mathbf{X}_i = [\mathbf{X}_{i-1}\mathbf{x}_{\lambda_i}]$. We use the convention that X_0 is empty matrix.
4. Solve a least square problem to obtain a new signal estimate:

$$\mathbf{g}_i = \arg\min_{\mathbf{g}} ||\mathbf{y} - \mathbf{X}_i\mathbf{g}||_2 . \tag{23}$$

5. Calculate the new approximation of the data and the new residual

$$\mathbf{r}_i = \mathbf{y} - \mathbf{X}_i\mathbf{g}_i. \tag{24}$$

6. Increment i, and return to Step 2 if $i < m$ or average power of $||\mathbf{r}_i||_2 > \eta_r$, where η_r is threshold.
7. The estimate $\mathbf{g}^{mp}$ for the ideal signal has nonzero indices at the components listed in Λ_m. The value of the estimate $\mathbf{g}^{mp}$ in component λ_k equals the kth component of $\mathbf{g}_i$.

3.2 Modified Matching Pursuit

Although the matching pursuit algorithm mentioned above is capable of detecting the intruders, it is difficult to carry out the calculation within the practical period due to huge computational cost. To reduce the computational cost, we propose a modified matching pursuit algorithm. The motion of the intruder was relatively slow in compared with the observation interval, NT_c. This implies that the channel impulse response at the current observation interval does not drastically change at the next interval. The proposed modified matching pursuit algorithm makes efficient use of the slow variation of the channel impulse response. In the proposed algorithm, the estimation is once carried out by using the original matching pursuit algorithm. After the first estimation is completed, the algorithm stores the positions where the peak value that is larger than the threshold. In the following estimation, the peak search operation original described in step 2 is carried out only in the areas around the vicinity of the peaks detected in the previous estimation.

3.2.1 Derivation Scheme for Peak Position Vector

Now, we explain the further detail of the proposed algorithm in Figure 3. The proposed intruder detection system obtains the sparse impulse response vector $\mathbf{g}^{CS} = [g_0^{CS}, g_1^{CS}, ..., g_{N-1}^{CS}]$ using CS process at t. Then, the system keeps the indices of elements whose values are larger than threshold η_p. Let us define the peak position vector at the initial stage as $\mathbf{v}_{init}(t) = [v_0^{init}, v_1^{init}, ..., v_{N-1}^{init}]$. As shown in the graph on top of Figure fig:mmp, the initial position

vector is given by

$$v_k^{init} = \begin{cases} 1 & |g_k^{CS}| > 0 \\ 0 & |g_k^{CS}| = 0 \end{cases}. \quad (25)$$

We could use the initial position vector as a search region if the peak position would not change. In practice, the position varies depending on the motion of the intruders. Therefore, we need to expand the peak position vector.

At Step 1 of Figure 3, we modify the channel impulse response vector $\mathbf{g}^{CS}$ as follows. The modified channel impulse response is given by

$$g_k'^{CS} = \frac{1}{2}|g_k^{CS}| + \frac{1}{4}|g_{k-1}^{CS}| + \frac{1}{4}|g_{k+1}^{CS}|. \quad (26)$$

Then, as shown in Step 2 of Figure 3, the threshold η_p is given by

$$\eta_p = p\left|\left|g_k'^{CS}\right|\right|_2, \quad (27)$$

where p is a coefficient for determining the threshold. As shown in Step 3, $\mathbf{g}''^{CS}$ is calculated as:

$$g_k''^{CS} = \begin{cases} g_k'^{CS} & |g_k'^{CS}| \geq \eta_p \\ 0 & |g_k'^{CS}| < \eta_p \end{cases}. \quad (28)$$

Finally, the modified peak position vector $\mathbf{v}(t) = [v_0, v_1(t), ..., v_{N-1}(t)]$ is obtained by the following equation:

$$v_k = \begin{cases} 1 & |g_k''^{CS}| > 0 \\ 0 & |g_k''^{CS}| = 0 \end{cases}. \quad (29)$$

In this method, it is important how to decide the threshold η_p. We discuss it in Section 4.1

3.2.2 Detection Scheme using Modified Matching Pursuit

In this section, we show the algorithm of Modified Matching Pursuit using the peak position vector $\mathbf{v}(t-1)$ at certain time $t-1$. In Step 2 in Section 3.1 the inner product operation is carried out for all N vectors as in Eq. (22). The proposed algorithm calculate the system explores by the following algorithm, the inner product and peak search is carried out only for the positions that the elements of the peak position vector obtained in the previous section are '1.' The peak search procedure is

$$\lambda_i(t) = \arg \max_{\{k|v_k(t-1)=1\}} |\langle \mathbf{r}_{i-1}, \phi_k \rangle|. \quad (30)$$

The intruder could exist where the positions corresponding to the non-zero element in $\mathbf{v}(t-1)$.

The procedures of the proposed modified matching pursuit algorithm are the same as the original matching pursuit algorithm described in 3.1except for the peak search in Eq.(30). The peak search is repeated

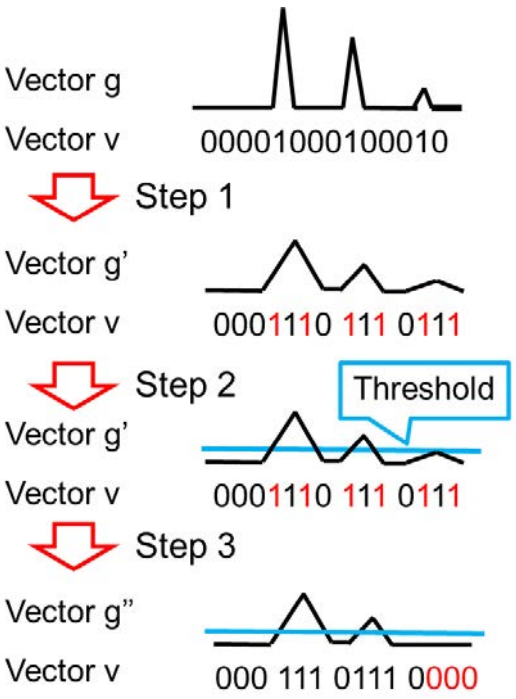


Fig.3: *Flow Chart of Modified Matching Pursuit*

unless the length of the residual vector is larger than the threshold, η_r^{mmp}.

The threshold is determined by the following procedure. Firstly, we calculate the expected received signal $||\mathbf{y}_{t-1}^{mmp}||_2$ when the channel impulse response vector is $\mathbf{g}_{t-1}^{CS}$. The expected received signal is given by

$$\mathbf{y}_t^{mmp} = \mathbf{X}\mathbf{g}_t^{CS}, \quad (31)$$

We employs

$$\eta_r^{mmp} = ||\mathbf{y}_t - \mathbf{y}_{t-1}^{mmp}||_2. \quad (32)$$

The number of iteration is at most the number of non-zero elements of $\mathbf{v}(t-1)$. When the l_2 norm the residual vector $\mathbf{r}_i$ is still larger than η_r^{mmp} after the number of iteration reaches the number of non-zero elements, it is probable that the peak arises at a position other than the non-zero positions of $\mathbf{v}(t-1)$. In this case, the original matching pursuit algorithm is carried out for obtaining the proper result.

4. NUMERICAL RESULTS

In this section, we show some numerical results to confirm the performance of the proposed method. Table 1 is the system configuration to run the simulation.

Figure 4 shows a detection area model. We use two 50-meter LCX cables placed in parallel. The transmitter is connected to the upper left endpoint of the LCX cable, and the receiver is connected to the bot-

Table 1: *System Configuration*

Parameters	Values
LCX length	50 [m]
Width between LCXs	10 [m]
Time symbol(N_t)	32
Frequency	300 [MHz]
bandwidth	60 [MHz]
ε_{lcx}	2.4
PN code size	63
Target speed	1.2[m/s]

Table 2: *Computer Simulation Environment*

CPU	Intel Core2 Quad Q9650 3.00GHz
RAM	4.00GB
OS	Linux 3.2.0
Programming Language	C++

tom left side.

In the simulation, we assume that one target obstacle exists between two LCX cables and it moves at the speed of 1.2 m/s as in Figure 4.

The simulation is carried out at the simulation environment shown in Table 2. We repeated simulation runs N_t times for evaluating the performance.

4.1 Verification of coefficient using proposal method

As described in Section 3.2.1, both the false error rate and the processing time vary depending on the peak threshold η_p using Modified Matching Pursuit. Figure 5 and 6 respectively show the false negative error rate and processing time against the coefficient p when CNR is 20dB, 22dB and 25dB, where CNR is the received signal to thermal noise power ratio. Figure 5 shows that the false negative error rate grows with increase in the coefficient and we can get the acceptable false negative error rate if the coefficient p is less than 1. In contrary, the processing time can be reduced with increase in p.

Figure 7 shows the relationship between the false negative error rate and the processing time. The required processing time for obtaining the acceptable error rate performance is about 0.11[s]-0.12[s]. According to Figures 5 and 6, it corresponds that p is around 1. Therefore, we set $p = 1$ in the following analyses.

4.2 Performance evaluations

We evaluate the computational cost of the proposed modified matching pursuit algorithm based intruder detection system. Figure 8 and 9 respectively show iteration count and processing time against CNR using the conventional and the proposed methods. The proposed method can reduce the number of

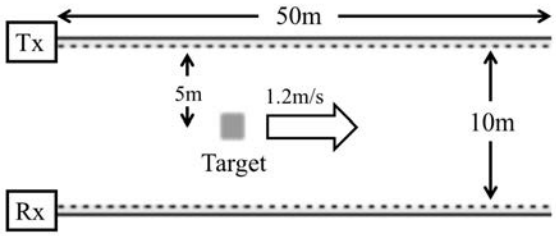


Fig.4: *Propagation Area Model*

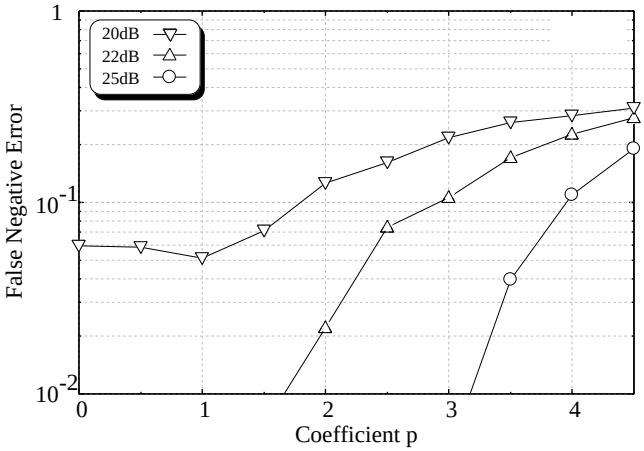


Fig.5: *False negative error rate when coefficient p changes.*

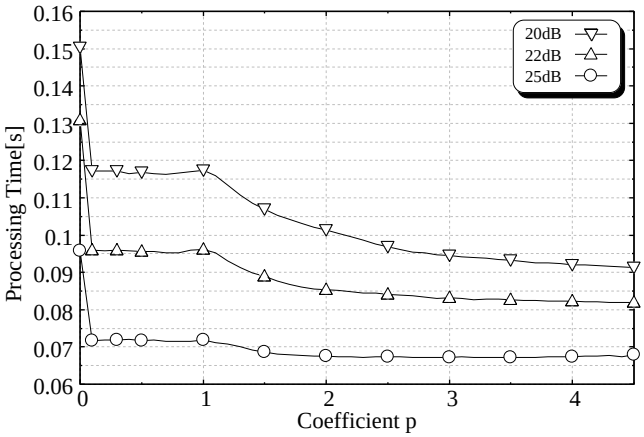


Fig.6: *Processing time when coefficient p changes.*

iterations by half. Furthermore, the processing time for the proposed method is about 30% of the conventional one.

The false positive rate (FPR) and false negative rate (FNR) performance against CNR are respectively shown in Figure 10 and 11. In Figure 10, the proposed scheme gives lower false positive rate than the conventional one in high CNR region. Also, the proposed scheme can slightly improve the false negative rate performance.

According to these results, we can conclude that

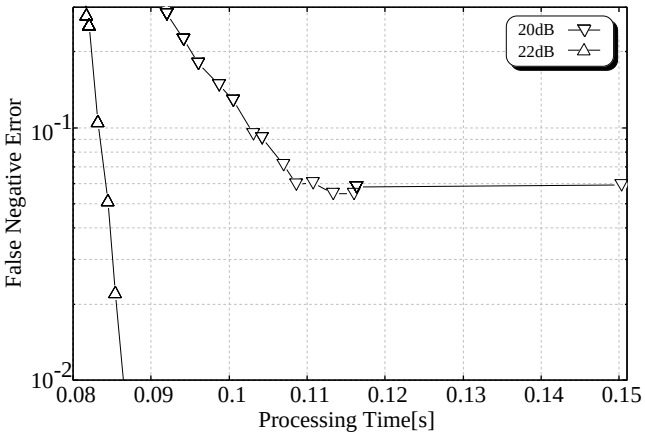


Fig.7: *False negative error rate when processing time changes.*

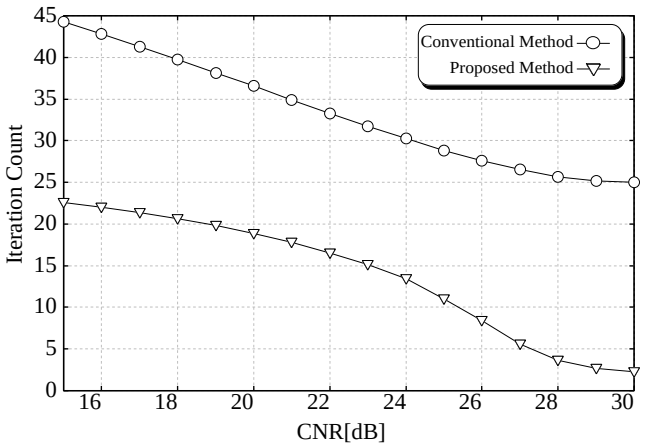


Fig.8: *Iteration Count*

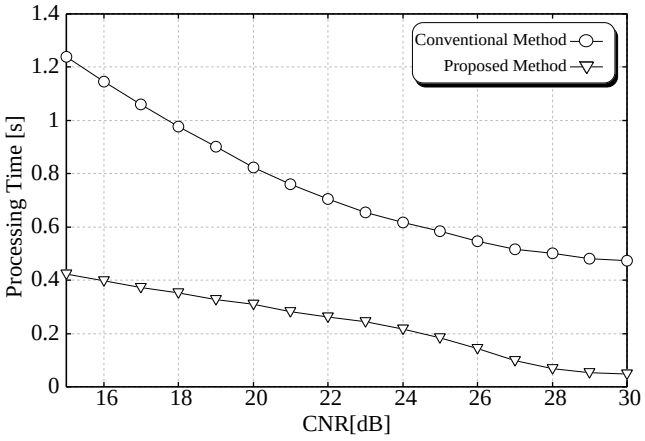


Fig.9: *Processing Time*

the proposed scheme is capable of reducing the computational cost while maintaining the error rate performance of the conventional CS-based intruder detection algorithm.

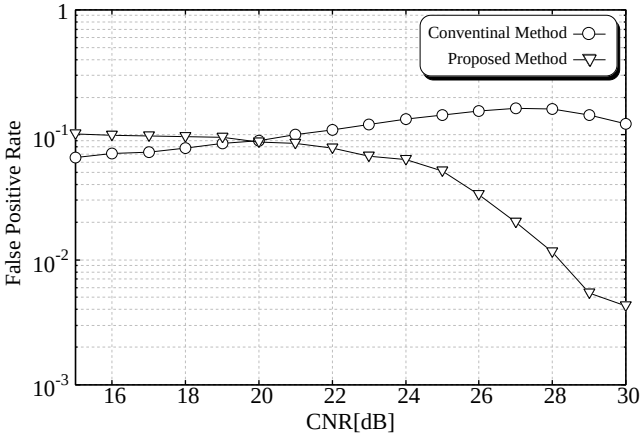


Fig.10: *False positive error rate when CNR changes.*

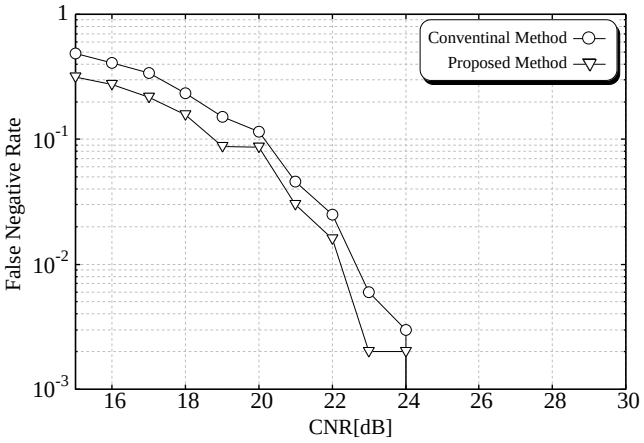


Fig.11: *False negative error rate when CNR changes.*

5. CONCLUSION

In this paper, we have proposed a reduced complexity intruder detection system based on the modified matching pursuit-based compressed sensing method. We verified the optimal coefficient value of threshold using proposed method and evaluated the iteration count, the procedure time and the false error rate performance by computer simulation.

The optimal coefficient value is 1 in the supposed intruder detection system configuration. The result shows that the proposed method can reduce the computational cost while maintaining the error rate performance.

References

[1] K. Inomata and T. Hirai, "Microwave back-projection radar for wide-area surveillance system," 34th European Microwave Conference, vol. 3, no. 11, pp. 1425–1428, Oct 2004.

[2] K. Nishikawa, T. Higashino, K. Tukamoto, and S. Komaki, "TDOA based wireless positioning method using leaky coaxial cable [in japanese],"

The transactions of the Institute of Electronics, Information and Communication Engineers. B, vol. J92-B, no. 1, pp. 320–327, Jan 2009.

[3] K. Inomata and T. Hirai, "An analysis of propagation loss characteristics in indoor short range wireless communications using ray-tracing technique [in japanese]," Technical report of IEICE, vol. SANE2005, no. 101, pp. 29–34, Feb 2006.

[4] T. S. Okada, "A study on detection accuracy for intruder detection system using leaky coaxial cable," 2012 International Conference on Embedded Systems and Intelligent Technology, pp. 167–170, Jan 2012.

[5] G. Marubayashi, M. Nakagawa, and R. Kohno, *Spread Spectrum Communication and Its Applications*. The Institute of Electronics, Information and Communication Engineers, 1998.

[6] S. Stein and J.Jay.Jones, *Modern Communication Principles With Application to Degital Signaling*. McGRAW HILL, 1970.

[7] A. Goldsmith, *Wireless Communications*. New York: Cambridge University Press, 2005.

[8] D. L. Donoho, "Compressed sensing," Information Theory, IEEE transaction on, vol. 52, no. 4, pp. 1289–1306, Apr 2006.

[9] T. Tanaka, "Mathematics of compressed sensing," IEICE Fundamentals Review, vol. 4, no. 1, pp. 39–47, June 2010.

[10] W. U.Bajwa, J. Haupt, G. Raz, and R. Nowak, "Compressed channel sensing," Conf. Information Sciences and Systems, pp. 5–10, Mar 2008.

[11] W. U.Bajwa, J. Haupt, A. M.Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," Proceedings of the IEEE, vol. 98, no. 6, pp. 1058–1076, June 2010.

[12] J. A. C.Gilvert, "Signal recovery from random measurement via orthogonal matching pursuit," IEEE Transactions on Information Theory, vol. 53, no. 12, pp. 4655–4666, Dec 2007.

Tomonori Sato received his B.E. degree in communications engineering from Osaka University Osaka Japan in 2008, and his M.E. and Ph.D. degrees from Graduate School of Information Science, Nara Institute of Science and Technology(NAIST) in 2010 and 2013, respectively. His research interest includes the radio sensing and researching intruder detection system using LCX in the school with Okada. His research interests include LCX, signal processing, wireless communication. Now, he works at Hitachi Metals, Ltd.

Ziji Ma received the B.E. degree in electronics engineering and the M.S. degree in electronics science and technology from Hunan University, China, in 2001 and 2007, respectively. He received the Ph.D in information systems from Nara Institute of Science and Technology in 2012. Now he is working at Hunan University, China. His major research interests include wireless communications, digital signal processing and measurement systems.

Takeshi Higashino received the B.E., M.E. and Ph.D. degrees in communications engineering from Osaka University, in 2001, 2002 and 2005 respectively. From 2005 to 2012, he was an assistant professor in the Division of Electrical, Electronic and Information Engineering at Osaka University. In 2012, he joined the Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan, as an Associate Professor. From 2014 to 2015, he was a visiting researcher at Georgia Institute of Technology, Atlanta, U.S. His research interest is on radio over fiber technique, wireless communication including OFDM, MIMO, digital broadcasting, wireless position location and wireless power transfer. Dr. Higashino is a member of IEEE and IEICE.

Minoru Okada received the B.E. degree from the University of Electro-Communications, Tokyo, Japan, in 1990. The M.E and Ph.D. degrees was received from Osaka University, Osaka, Japan, in 1992 and 1998, respectively, all in communications engineering. In 2000, he joined the Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan, as an Associate Professor and became a Professor. From 1993 to 2000, he was a Research Associate with Osaka University. From 1999 to 2000, he was Visiting Research Fellow at University of Southampton, Southampton, U.K.. His research interest is wireless communications, including WLAN, WiMAX, CDMA, OFDM, and satellite communications. Dr. Okada is a member of the Institute of Image, Information, and Television Engineers of Japan (ITEJ), the Institute of Television Engineers of Japan, the Institute of Electrical, Information, and Communication Engineers of Japan (IEICE), and the Information Processing Society of Japan (IPSJ). He received the Young Engineer Award from IEICE in 1999.

Dynamic Codebook for Foreground Segmentation in a Video

Worapan Kusakunniran [1], Non-member and **Rawitas Krungkaew** [2], Member

ABSTRACT

The foreground segmentation in a video is a way to extract changes in image sequences. It is a key task in an early stage of many applications in the computer vision area. The information of changes in the scene must be segmented before any further analysis could be taken place. However, it remains with difficulties caused by several real-world challenges such as cluttered backgrounds, changes of the illumination, shadows, and long-term scene changes. This paper proposes a novel method, namely a dynamic codebook (DCB), to address such challenges of variations in the background scene. It relies on a dynamic modeling of the background scene. Initially, a codebook is constructed to represent the background information of each pixel over a period of time. Then, a dynamic boundary of the codebook will be made to support variations of the background. The revised codebook will always be adaptive to the new background's environments. This makes the foreground segmentation more robust to the changes of background scene. The proposed method has been evaluated by using the changedetection.net (CDnet) benchmark which is a well-known video dataset for testing change-detection algorithms. The experimental results and comprehensive comparisons have shown a very promising performance of the proposed method.

Keywords: Foreground, Segmentation, Video, Dynamic, Codebook

1. INTRODUCTION

The foreground segmentation aims to detect moving objects in a video sequence. It is a very important step in many computer vision based applications [1][2][3][4]. For example, in [1], to analyze the performance of a tennis match, key events such as the tennis ball being hit and the tennis ball bouncing on the ground must be detected beforehand. In order to detect such events, the foreground of the tennis ball must be extracted. This is where the foreground segmentation will be employed. It is similar to the work in [2] related to the application domain of the traffic surveillance, the foreground of the vehicle must be segmented before any further analysis such as the count of vehicles, the calculation of average speed of individual vehicle will be taken place. In [3][4], the foreground or the movement of human must be extracted before being used in recognizing human gaits and human actions respectively.

In addition, there are many techniques [5][6][7][8][9][10][11][12][13][14][15][16][17][18][19][20] proposed to address the problem of the foreground segmentation in a video. For example, the method [5], namely frame differencing, is one of the most conventional techniques for the foreground segmentation. It works well under the static background scene especially in the indoor environment. The core concept is to detect any movement in a scene by subtracting the current frame from the previous frame. However, this technique cannot extract a full silhouette of a moving object. It can detect only a moving part of an object. Also, it has a difficulty in detecting an object that moves with a very slow speed such as a movement of a few pixels per dozen frames [6].

In [16], the performance of the foreground segmentation is boosted by using the artificial neural network to learn variations in the background scene. The learning rate must be calibrated to match the nature of each video scene. It is automatically chosen by using the fuzzy technique and the coherence-based self-organizing background subtraction algorithm. In [17], the genetic programming is used to combine the outputs of multiple foreground-detection algorithms. The proposed technique will match the right algorithm for the right video. As a result, it can achieve the high performance based on the CDnet benchmark. However, several challenges of the foreground segmentation still remain unsolved.

In [11], a Gaussian mixture model (GMM) is used to model the background information in each pixel of the video frame. This is performed under the assumption that the GMM can model a variation of background's intensity values in each particular pixel by using a multimodal background model. Then, the likelihood of a pixel's value being a part of backgrounds or foregrounds will be determined based on a simple heuristic of the constructed GMM of that particular pixel. The one that does not fit to the constructed GMM is identified as a foreground pixel.

Manuscript received on July 31, 2016 ; revised on December 13, 2016.

Manuscript received on January 15, 2017.

[1] The author is with Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand 73170, Phone(+662)4410909, E-mail: worapan.kunmahidol.ac.th

[2] The author is with Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand 73170, Phone(+662)4410909, E-mail: rawitagmail.com

Otherwise, it will be identified as a background pixel.

The key advantage of this technique is that it can deal with repetitive motions in background scenes. The multimodal background model of the GMM can be used to represent multiple surfaces appeared in each particular pixel such as water surfaces. However, it cannot easily and efficiently model fast variations in background scenes with just a few Gaussians. Moreover, this technique also has to deal with the trade-off issue on adjusting the learning rate. For the low learning rate, the background model will be gradually adjusted to the recent background scene. In this case, it will have a difficulty in adapting the background model to the sudden change in the background scene. In contrast, for the high learning rate, the background model will be rapidly adjusted to the recent background scene. It will make the slow-moving object to be eventually absorbed into the background model and faultily classified as a part of the background [12].

To overcome such limitation of the GMM-based method, the codebook-based technique was adopted for the foreground segmentation in a video by not making use of any parametric parameters [12]. One codebook will be constructed to model the background information of each pixel. It will consist of one or more codewords. One codeword will represent one variation of the background model in the scene. For the background scene with the high variations, the large number of codewords will be created for each codebook during the training phase.

Then, in the testing phase, if a pixel's value does not belong to any codewords of the corresponding codebook, then that pixel will be identified as the foreground. The advantage of this technique is that it can efficiently deal with a dynamic background scene. However, it is not flexible enough to deal with a very high dynamic background environment. Also, it can suffer from the false negative foreground detection due to its quantization criterion [13]. In [13][14][15], the codebook-based method was improved by combining with other techniques.

Based on our literature review, the challenge of variations in the background scene is remaining unsolved. It can be caused by the moving of the background scene itself such as waving trees and water surfaces, or caused by the unstable camera. In this paper, the DCB-based method is proposed to address such challenge. The boundary of the codeword will be adjusted and adapted to the content of the recent background scene. In addition, the new codeword can be created to represent a new background information. Also, the existing codeword can be deleted if it represents the out-of-date background information

It has been known that the evaluation is the crucial process to identify the strength and weakness of each method [19]. In this paper, the proposed method has been evaluated based on the datasets and evaluation approaches as described in the CDnet [18]. The experimental results and comprehensive comparisons are carried out.

It can be concluded that the proposed method outperforms the existing techniques in the literature review, for the case of the baseline category containing the combination of mild challenges on the foreground segmentation in a video, which are a subtle background motion, isolated shadows, an abandoned object, and stopping pedestrians for a short period. Particularly, the proposed method is also shown to be very promising on solving the challenges of the unstable camera in the camera jitter category and the strong background motion in the dynamic background category.

The rest of this paper is organized as follows. Section 2 describes the details of the proposed method. Section 3 shows the experimental results and the relevant discussions. Then, the conclusion is drawn in section 4.

2. THE PROPOSED DYNAMIC CODEBOOK (DCB)

The proposed DCB is explained in this section. The DCB itself is for the background modeling. Later, its models can be used for the foreground segmentation in a video, which will be also explained in the section 2.2 It is developed to overcome the limitation of the conventional codebook-based method [21] in order to address the challenge of the dynamic background scene. The conventional codebook-based method performs under the assumption that the cylindrical codebook-model in RGB color space can sufficiently model the pixel intensity distortions. However, for the high-dynamic background scene, such variations of the background information cannot be coped by using the cylindrical shape. To address this problem, the dynamic adjustment of the codeword's range/shape is embedded in the development process of DCB.

2.1 Codebook

The framework of the codebook-based method is shown in the algorithm 1. One codebook is constructed as the background model of each pixel. It contains at least one codeword. Each codeword is represented by a range of pixel's intensity values of the background of that pixel. The number of codewords for each codebook depends on a degree of variations in the background scene of the training phase. The high variation in the background pixel will result in the high number of codewords for the corresponding codebook. Thus, the codebook is created by observing changes of pixel's intensity in video frames.

In this paper, to simplify the explanation, it is assumed that the codebook is constructed for a particular pixel i (CB_i). Each codeword j ($CW_{i,j}$) of CB_i is modeled by using the rgb vector $v_{i,j} = (r_{i,j}, g_{i,j}, b_{i,j})$

Algorithm 1 Construction of the codebook (CB_i) in the training phase

Input: The given video sequence of the background scene (V)
Output: The codebook (CB_i)

```
Initialize CB_i to be empty (i.e. contains no codeword)
for (each frame t in V )
{
    Initialize flag = codeword not found
    for (each codeword CW_{i,j})
    {
        if(δ(I_{i,t}, v_{i,j}) < ε and I_min < I_{i,t} < I_max)
        {
            flag = codeword found
        {
    }
    if (flag = codeword found)
    {
        v_{i,j} = ((f×r_{i,j}+r)/(f+1), (f×g_{i,j}+g)/(f+1), (f×b_{i,j}+b)/(f+1))
        h_{i,j} = (min(I_{i,t}, I_min), max(I_{i,t}, I_max), f + 1, neg, p, t)
    {
    else
    {
        The new codeword for CB_i is constructed
    {
}
return CB_i
```

and the hex-tuple $h_{i,j} = (I_{min}, I_{max}, f, neg, p, q)$, where I_{min} is the minimum brightness value assigned to $CW_{i,j}$, I_{max} is the maximum brightness value assigned to $CW_{i,j}$, f is the frequency or the number of times that $CW_{i,j}$ occurs, neg is the maximum negative run length (MNRL) which is the longest period that $CW_{i,j}$ is not accessed, and p and q are the first and the last access times of $CW_{i,j}$ respectively. In addition, the color space does not have to always be rgb. It can be other color spaces such as $l\alpha\beta$ as used in this paper.

In the training phase, a video sequence of the background scene is used to construct codebooks as the background models. If the training data sufficiently covers most of possible background variations, then the constructed codebooks will be robust to such variations during the process of the foreground segmentation.

The training process of constructing the codebook (CB_i) begins with checking the pixel's intensity value $I_{i,t} = (r, g, b)$ from each background frame at time t against each existing $CW_{i,j}$ based on the following two criteria. First, as shown in the equation (1), the color distortion (δ) between $I_{i,t}$ and $CW_{i,j}$ is less than the detection threshold (ϵ).

$$\delta(I_{i,t}, v_{i,j}) < \epsilon \tag{1}$$

The color distortion is calculated as shown in the equation (2).

$$\delta(I_{i,t}, v_{i,j}) = \sqrt{||I_{i,t}||^2(1 - cos^2\theta)} \tag{2}$$

where $||I_{i,t}||$ is the $l2$-norm of $I_{i,t}$ as shown in the equation (3) and $cos\theta$ is the cosine similarity between $I_{i,t}$ and $v_{i,j}$ as shown in the equation (4). As in the calculation, $cos\theta$ is used to measure the similarity between $I_{i,t}$ and $v_{i,j}$. Thus, 1 - $cos^2\theta$ is used to measure the difference or the distortion between $I_{i,t}$ and $v_{i,j}$. The $||I_{i,t}||^2$ is used as the multiplication factor in the equation (2), in order to geometrically normalize the codeword's vector ($v_{i,j}$) to the brightness of the input pixel ($I_{i,t}$) [21].

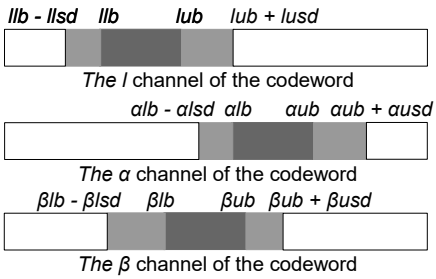


Fig.1: *The structure of each codeword in DCB.*

$$||I_{i,t}|| = \sqrt{r^2 + g^2 + b^2} \tag{3}$$

$$cos\theta = \frac{r \times r_{i,j} + g \times g_{i,j} + b \times b_{i,j}}{\sqrt{r^2 + g^2 + b^2}\sqrt{r_{i,j}^2 + g_{i,j}^2 + b_{i,j}^2}} \tag{4}$$

Second, the brightness of the input pixel is between the minimum and maximum brightness values. If there is at least one $CW_{i,j}$ that satisfies these two criteria, then the corresponding $v_{i,j}$ and $h_{i,j}$ will be updated as shown in the equations (5) and (6). Otherwise, the new codeword will be created for this CB_i by using the information of the input pixel.

In the testing phase, these two criteria are used for segmenting the foreground from the background. That is, if there is at least one $CW_{i,j}$ that satisfies the two criteria, the pixel will be classified as the background. Also, the corresponding $v_{i,j}$ and $h_{i,j}$ will be updated as shown in the equations (5) and (6). Otherwise, the pixel will be classified as the foreground.

$$v_{i,j} = (\frac{f \times r_{i,j} + r}{f + 1}, \frac{f \times g_{i,j} + g}{f + 1}, \frac{f \times b_{i,j} + b}{f + 1}) \tag{5}$$

$$h_{i,j} = (min(I_{i,t}, I_{min}), max(I_{i,t}, I_{max}), f+1, neg, p, t) \tag{6}$$

2.2 Dynamic codebook (DCB)

In this paper, the DCB is developed under the $l\alpha\beta$ color space. The l channel represents the illumination information, while α and β channels represent the color information. This color space was shown to be an effective choice for the process of foreground segmentation in a video [22]. This is because, in the $l\alpha\beta$ color space, the chromatic component (i.e. α and β) is almost completely separated from the achromatic component (i.e. l) [23].

Table 1: *The symbols used in the structure of each codeword in DCB.*

Symbols	Descriptions
l	the pixel's intensity value in the l channel
α	the pixel's intensity value in the α channel
β	the pixel's intensity value in the β channel
llb	the lower bound of a codeword in the l channel
αlb	the lower bound of a codeword in the α channel
βlb	the lower bound of a codeword in the β channel
lub	the upper bound of a codeword in the l channel
αub	the upper bound of a codeword in the α channel
βub	the upper bound of a codeword in the β channel
$llsd$	the standard deviation of $\|llb - l\|$ over a training period
αlsd	the standard deviation of $\|\alpha lb - \alpha\|$ over a training period
βlsd	the standard deviation of $\|\beta lb - \beta\|$ over a training period
$lusd$	the standard deviation of $\|lub - l\|$ over a training period
αusd	the standard deviation of $\|\alpha ub - \alpha\|$ over a training period
βusd	the standard deviation of $\|\beta ub - \beta\|$ over a training period

Moreover, in the proposed DCB, the l channel of the codeword can be modeled to be more dynamic than the α and β channels of the codeword. This will make the DCB to be more robust to changes of illumination in the background scene, when compared with the rgb color space.

As mentioned, DCB has the key feature of the dynamic decision boundary of each codeword. Thus, the range and/or shape of the codeword in the $l\alpha\beta$ color space will be dynamically adjusted to the recent context of the background environment. To simplify the explanation, Fig. 1 illustrates the structure of each codeword and Table 1 describes the meaning of symbols used in DCB as shown in Fig. 1.

The structure of each codeword consists of three channels according to the used color space i.e. $l\alpha\beta$. In each channel of the codeword, there are two layers of the boundary in order to cope with variations of the background in each pixel. The first layer is ranged by using the lower bound and upper bound. The second layer is extended from the first layer by using the standard deviation of the variations of the background.

Table 2: *The construction of a new codeword in DCB. Note: δ is a constant value for a new codeword construction.*

Components	Values
llb	$l - \delta$
lub	$l + \delta$
$llsd$	δ
$lusd$	δ
αlb	$\alpha - \delta$
αub	$\alpha + \delta$
αlsd	δ
αusd	δ
βlb	$\beta - \delta$
βub	$\beta + \delta$
βlsd	δ
βusd	δ

Table 3: *The seven measurements for the performance evaluation.*

Measurements	Calculations
Re	$\frac{TP}{TP+FN}$
Sp	$\frac{TN}{TN+FP}$
FPR	$\frac{FP}{FP+TN}$
FNR	$\frac{FN}{TP+FN}$
PWC	$\frac{100(FN+FP)}{TP+FN+FP+TN}$
Pr	$\frac{TP}{TP+FP}$
F1	$\frac{2(Pr)(Re)}{Pr+Re}$

This proposed DCB-based method has two phases which are training and testing phases. They are described in the following paragraphs.

2.2.1 The training phase

It needs a video sequence of the background scene for constructing the codebooks CB_i. For DCB, the pixel' values $I_{i,t} = (l, \alpha, \beta)$ from each background frame at time t will be checked against each existing $CW_{i,j}$ based on the following three conditions.

$$(llb - llsd) \leq l \leq (lub + lusd) \tag{7}$$

Algorithm 2 The updates of the matched codeword $CW_{i,j}$ in DCB

Input: The pixel' values $I_{i,t} = (l, \alpha, \beta)$ from each background frame at time t
Output: The updated codeword $CW_{i,j}$

```
if (llb − llsd) ≤ l ≤ (lub + lusd)
{
    if (l < llb)
    {
        llb = llb − llsd
        llsd = ((llsd × f) + k(llb − l)) / (f + 1)
    }
    if (l > lub)
    {
        lub = lub + lusd
        lusd = ((lusd × f) + k(l − lub)) / (f + 1)
    }
}
if (αlb − αlsd) ≤ α ≤ (αub + αusd)
{
    if (α < αlb)
    {
        αlb = αlb − αlsd
        αlsd = ((αlsd × f) + k(αlb − α)) / (f + 1)
    }
    if (α > αub)
    {
        αub = αub + αusd
        αusd = ((αusd × f) + k(α − αub)) / (f + 1)
    }
}
if (βlb − βlsd) ≤ β ≤ (βub + βusd)
{
    if (β < βlb)
    {
        βlb = βlb − βlsd
        βlsd = ((βlsd × f) + k(βlb − β)) / (f + 1)
    }
    if (β > βub)
    {
        βub = βub + βusd
        βusd = ((βusd × f) + k(β − βub)) / (f + 1)
    }
}
return CW_{i,j}
```

$$(\alpha lb - \alpha lsd) \leq \alpha \leq (\alpha ub + \alpha usd) \tag{8}$$

$$(\beta lb - \beta lsd) \leq \beta \leq (\beta ub + \beta usd) \tag{9}$$

If there is no codeword that satisfies with these three conditions, a new codeword of CB_i will be created as shown in Table 2. Otherwise, the matched codeword will be updated as shown in the algorithm 2. In the Table 2, the constant value (δ) for the construction of a new codeword is empirically adjusted. In this paper, it is set to be 5.

2.2.2 The testing phase

In the testing phase, for each pixel i in a current frame, if it belongs to any codeword of CB_i by satisfying the three conditions in the equations (7), (8) and (9), then it will be classified as the background pixel. Otherwise, it will be classified as the foreground pixel. If the pixel is classified as the background, it will be used to update the matched codeword as shown in the algorithm 2.

In the algorithm 2, k is the learning factor for updating the value of the standard deviation of the pixel's intensity value away from the corresponding codeword. In this paper, k is empirically set to be 1. Therefore, the codewords and codebooks modeling the background will be dynamically updated based on the changes of the scene. That means, the DCB has a property to deal with variations of the background in a cluttered environment.

In addition, the out-of-date codeword can be deleted from the corresponding codebook. In this case, the out-of-date codeword means the codeword in which there is no any pixels belong to it for a period of time which can be empirically adjusted. In this paper, this period of time is set to be 50 frames. However, the out-of-date codeword will not deleted if it is the only one codeword left in the corresponding codebook.

3. EXPERIMENTAL RESULTS

In this paper, the CDnet benchmark [18] is used to evaluate the performance of the proposed method. The results of the foreground segmentation by using the proposed method are compared with the ground truth provided by the CDnet. The performance of the proposed method is also compared to the performances of the other existing methods in the literature review, using the same CDnet dataset. The CDnet dataset provides realistic, camera capture videos (with no CGI) from both indoor and outdoor environments with various camera ranging. The dataset of the six categories provides about 70,000 frames in 31 videos. The six categories are baseline, dynamic background, camera jitter, shadows, intermittent object motion and thermal.

The baseline category contains a combination of mild challenges on the foreground segmentation in a video, as can be seen in Fig. 2. The challenges include a subtle background motion, isolated shadows, an abandoned object, and stopping pedestrians for a short period. The dynamic background category contains a strong background motion such as a shimmering water and a shaking tree. The camera jitter category contains videos recorded by unstable cameras. The intermittent object motion category contains scenes of topping pedestrians for a short period or suddenly start moving pedestrians. The shadows category contains strong shadows and the thermal category contains videos recorded by far-infrared cameras.

This paper focuses on the baseline dataset. This is because the proposed method has been developed for the normal scenes and the scenes with mild to medium variations caused by several challenges as mentioned above for the baseline category. While the other five categories focus on the strong varia-

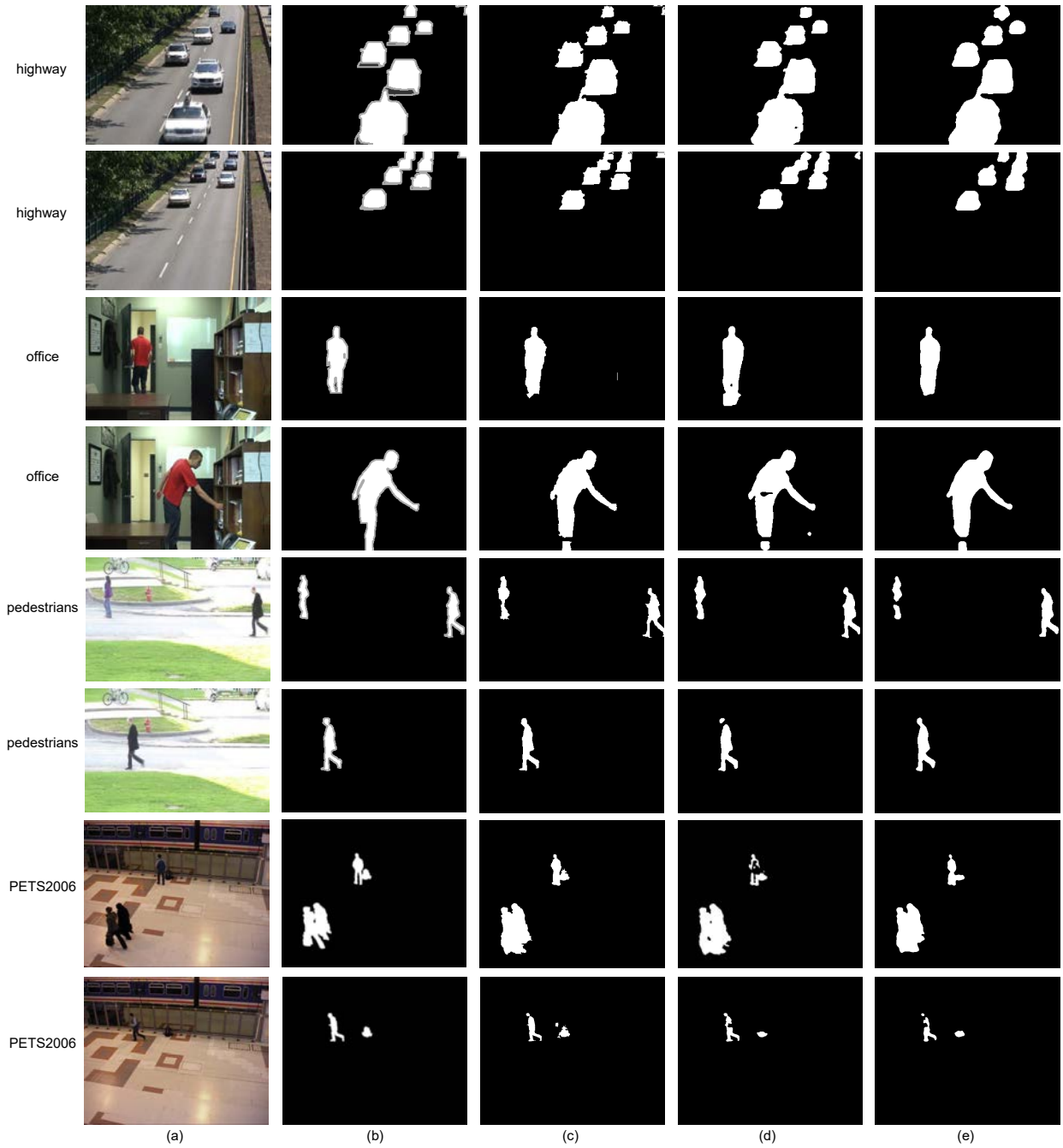


Fig.2: *The sample segmented foreground images by using the proposed method, the IUTIS-1 [17] and the multimode background subtraction [24], compared with the ground truth images. (a) Original images. (b) Ground truth images. (c) Segmented images by using the proposed method. (d) Segmented images by using the IUTIS-1 [17]. (e) Segmented images by using the multimode background subtraction [24].*

tion of its only type. However, the additional experiments and comparisons are also carried out under these five categories of the dataset, in order to preliminarily analyze the results of applying the proposed method on the strong variations. Moreover, the proposed method will be further developed in the future work to cope with such strong variations.

The seven measurements are used to calculate the average ranking and performance across all methods in the CDnet. As described in Table 3, they are Recall (Re), Specificity (Sp), False Positive Rate (FPR), False Negative Rate (FNR), Percentage of Wrong Classifications (PWC), Precision (Pr), and F-Measure (F1). These seven measurements are calculated based on the four measurements of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

The TP, TN, FP and FN are calculated by comparing the segmentation result based on the proposed method with the ground truth. TP is calculated based on the foreground pixels that are correctly identified as the foreground. TN is calculated based on

Table 4: *The performance comparisons between the proposed method and the other existing methods in the literature review, based on the baseline video category from the CDnet benchmark.*

Method	**Average Re**	**Average Sp**	**Average FPR**	**Average FNR**	**Average PWC**	**Average F1**	**Average Pr**
Multimode Background Subtraction [24]	0.9158	0.9979	0.0021	0.0842	0.4361	0.9287	0.9431
IUTIS-1 [17]	0.9214	0.9979	0.0021	0.0786	0.4538	0.9298	0.9391
SOBS-CF [25]	0.9347	0.9978	0.0022	0.0653	0.3912	0.9299	0.9254
FTSG [26]	0.9513	0.9975	0.0025	0.0487	0.4766	0.9330	0.9170
Spectral-360 [27]	0.9616	0.9968	0.0032	0.0384	0.4265	0.9330	0.9065
M4CD Version 1.5 [28]	0.9521	0.9975	0.0025	0.0479	0.4402	0.9272	0.9057
CwisarDRP [29]	0.8580	0.9981	0.0019	0.1420	0.8778	0.8880	0.9347
CwisarDH [30]	0.8972	0.9980	0.0020	0.1028	0.5679	0.9145	0.9337
AAPSA [31]	0.9092	0.9979	0.0021	0.0908	0.5826	0.9183	0.9286
M4CD Version 1.0 [28]	0.9382	0.9976	0.0024	0.0618	0.4934	0.9204	0.9063
C-EFIC [32]	0.9455	0.9970	0.0030	0.0545	0.5201	0.9309	0.9170
KDE-ElGammal [11]	0.8969	0.9977	0.0023	0.1031	0.5499	0.9092	0.9223
EFIC [33]	0.9349	0.9971	0.0029	0.0651	0.5223	0.9172	0.9023
Mahalanobis distance [34]	0.3154	0.9991	0.0009	0.6846	2.8698	0.4642	0.9270
KNN [35]	0.7934	0.9979	0.0021	0.2066	1.2840	0.8411	0.9245
CP3-online [36]	0.8501	0.9972	0.0028	0.1499	0.7725	0.8856	0.9252
RMoG [37]	0.7082	0.9981	0.0019	0.2918	1.5935	0.7848	0.9125
AMBER [38]	0.8784	0.9973	0.0027	0.1216	0.9233	0.8813	0.8980
IUTIS-2 [17]	0.7452	0.9978	0.0022	0.2548	1.5115	0.7913	0.9100
Euclidean distance [34]	0.8385	0.9955	0.0045	0.1615	1.0260	0.8720	0.9114
GMM (Zivkovic) [39]	0.8085	0.9972	0.0028	0.1915	1.3298	0.8382	0.8993
Multiscale Spatio-Temporal BG Mode [40]	0.8137	0.9970	0.0030	0.1863	1.1478	0.8450	0.8870
GMM (Stauffer and Grimson) [41]	0.8180	0.9948	0.0052	0.1820	1.5325	0.8245	0.8461
GraphCutDiff [42]	0.7028	0.9960	0.0040	0.2972	1.9757	0.7147	0.8093
The proposed method (DCB)	0.9456	0.9975	0.0025	0.0544	0.4366	0.9359	0.9272

the background pixels that are correctly identified as the background. FP is calculated based on the foreground pixels that are incorrectly identified as the background. FN is calculated based on the background pixels that are incorrectly identified as the foreground.

The sample segmented foreground images by using the proposed method are shown in Fig. 2. They are from four different videos including highway, office, pedestrians, and PETS2006 in the CDnet benchmark. The results shown in Fig. 2 demonstrate the efficiency and robustness of the proposed method under various movements and conditions of the background scenes.

Table 5: *The precision comparisons between the proposed method and the other existing methods in the literature review, based on the other five categories from the CDnet benchmark including 1) dynamic background (DB), 2) camera jitter (CJ), 3) shadows (SD), 4) intermittent object motion (IO) and 5) thermal (TM). The average precision from the five categories (Avg1) and the average precision from the two categories of dynamic background and camera jitter (Avg2) are shown in the last two columns respectively.*

Method	**DB**	**CJ**	**IO**	**SD**	**TM**	**Avg1**	**Avg2**
Multimode Background Subtraction [24]	0.8651	0.8443	0.7827	0.3481	0.8268	0.7334	0.8547
IUTIS-1 [17]	0.3305	0.5299	0.5485	0.6032	0.9245	0.5873	0.4302
SOBS-CF [25]	0.5953	0.6405	0.5464	0.5899	0.8715	0.6487	0.6179
FTSG [26]	0.9129	0.7645	0.8512	0.5005	0.9088	0.7875	0.8387
Spectral-360 [27]	0.8456	0.8387	0.7374	0.5815	0.9114	0.7829	0.8422
CwisarDRP [29]	0.8723	0.8713	0.8543	0.5773	0.9116	0.8174	0.8718
CwisarDH [30]	0.8499	0.8516	0.7417	0.5547	0.8786	0.7753	0.8508
AAPSA [31]	0.7336	0.8021	0.7139	0.5877	0.8795	0.7434	0.7679
M4CD Version 1.0 [28]	0.6806	0.7901	0.8000	0.5590	0.9452	0.7550	0.7354
C-EFIC [32]	0.6993	0.8157	0.5823	0.4791	0.8690	0.6891	0.7575
KDE-ElGammal [11]	0.5732	0.4862	0.4609	0.6217	0.8974	0.6079	0.5297
EFIC [33]	0.6849	0.6389	0.5634	0.4846	0.849	0.6442	0.6619
Mahalanobis distance [34]	0.7451	0.8564	0.5098	0.8726	0.9932	0.7954	0.8008
KNN [35]	0.6931	0.7018	0.7121	0.3979	0.9186	0.6847	0.6975
CP3-online [36]	0.6122	0.4562	0.5631	0.5914	0.7663	0.5978	0.5342
RMoG [37]	0.7288	0.7605	0.8026	0.3097	0.9365	0.7076	0.7447
AMBER [38]	0.7990	0.8493	0.7530	0.4658	0.8514	0.7437	0.8242
IUTIS-2 [17]	0.5564	0.7184	0.8374	0.4480	0.9395	0.6999	0.6374
Euclidean distance [34]	0.4487	0.3753	0.4995	0.5763	0.8877	0.5575	0.4120
GMM (Zivkovic) [39]	0.6213	0.4872	0.6458	0.5428	0.8706	0.6335	0.5543
Multiscale Spatio-Temporal BG Mode [40]	0.5515	0.3979	0.6016	0.5282	0.8403	0.5839	0.4747
GMM (Stauffer and Grimson) [41]	0.5989	0.5126	0.6688	0.5352	0.8652	0.6361	0.5558
GraphCutDiff [42]	0.5357	0.5918	0.8315	0.4260	0.9111	0.6592	0.5638
The proposed method (DCB)	0.7632	0.9107	0.5291	0.3706	0.8502	0.6848	0.8370

As shown in Fig. 2, in the highway video, the segmentation performances of the three methods are equally good. However, the proposed method is shown to be better than the IUTIS-1 [17] and the multimode background subtraction [24] methods, in the office, pedestrians and PETS2006 videos. The proposed method can provide good quality segmented silhouettes, when compared with the results of the two methods [17][24].

In the row 3 of Fig. 2 (d), for the office video, the IUTIS-1 [17] incorrectly segments the shadow as the foreground. In the row 4 of Fig. 2 (d), for the office video, the IUTIS-1 [17] generates the incomplete silhouette.

In the row 5 of Fig. 2 (e), for the pedestrians video, the multimode background subtraction [24] generates

the incomplete silhouette. The human silhouette is cut into two parts. Similarly, in the row 6 of Fig. 2 (d), for the pedestrians video, the IUTIS-1 [17] also cuts the human silhouette into two parts.

In the rows 7 and 8 of Fig. 2 (d) and (e), it can be seen that the IUTIS-1 [17] and the multimode background subtraction [24] cannot provide the complete silhouettes.

Based on the 'highway' video, the proposed method is shown to be robust to the shadow of the trees, the small movement of the background (i.e. the moving trees), the fast moving objects (i.e. the moving cars), the small and far moving objects, and the outdoor environment. Based on the 'office' video, the proposed method is shown to be robust to the slow moving objects, and the indoor environment. Based on the 'pedestrians' video, the proposed method is shown to be robust to the variation of the daylight. It is also shown to be able to provide the clean silhouettes of the human body. Based on the 'PETS2006' video, the proposed method is shown to be able to work well under the conditions of the top camera-view. Also, it can segment the group of people and the unattended object.

In addition, the comprehensive comparisons between the performance of proposed method and the performance of the existing methods in the literature review are shown in Table 4. This is done based on the baseline video category provided by the CDnet. In Table 4, it can be seen that the proposed DCB-based method outperforms the other existing methods, in average (i.e. reported by the CDnet). This average ranking is calculated by averaging the ranks of the method in Re, Sp, FPR, FNR, PWC, Pr and F1. The proposed DCB achieves very high recall, specificity and precision of 94.56%, 99.75% and 92.72% respectively.

Table 5 shows the precision comparisons between the proposed method and the other existing methods in the literature, based on the other five categories from the CDnet benchmark including 1) dynamic background (DB), 2) camera jitter (CJ), 3) shadows (SD), 4) intermittent object motion (IO), and 5) thermal (TM). For the camera jitter category, the proposed method achieves the highest precision, when compared with the other existing methods in the literature. Its performance is also promising for the categories of dynamic background and thermal. This is because the proposed DCB can adapt quickly to the changes of the background. Multiple codewords are created to model variations of the background in each pixel of the scene, which can be caused by the camera jitter and/or the dynamic background itself. The dynamic of the codewords can make it quickly handle the variations of the background.

However, the proposed method cannot achieve the good performance for the shadow category. This is because the proposed DCB is explicitly developed to address the challenge of the variations in the background scene, not particularly for the shadow. In the future work, to handle the case of the shadow, the DCB can be modified especially in the luminance channel of the codewords, or the additional technique of the shadow removal can be applied to the segmentation results.

In Table 5, the average precision from the five categories (Avg1) and the average precision from the two categories of dynamic background and camera jitter (Avg2) are displayed in the last two columns respectively. The Avg2 is used to validate the performance of the techniques for the challenge of the variations in the background scene, which is the focus of this paper.

As shown in Table 5, based on the Avg1 and Avg2, the proposed method is shown to outperform many methods in the literature. In addition, when compared with [17][28][31][32][35], the proposed method achieves the lower precision in average from the five categories, but it can achieve the significantly higher precision in average from the two categories. When compared with [34][37][38], the proposed method can achieve the slightly higher precision in average from the two categories, however, it can perform significantly better in the case of the baseline category (as shown in Table 4). When compared with [24][26][27][29][30], the proposed method achieves the lower precisions in both cases of Avg1 and Avg2, however, it can perform better in the case of the baseline category (as shown in Table 4).

4. CONCLUSION

In this paper, the DCB-based foreground segmentation method in a video is proposed. It is developed to overcome the limitation of the conventional codebook-based method. The variations in background scenes are coped by using the dynamic boundary of each codeword in the DCB. Thus, the proposed method can be more robust under the cluttered environments in the real scenes. It can cope with the variations in the background scene. The CDnet benchmark is used to evaluated the performance of the proposed method. It has been shown that the proposed method achieves a very promising performance by outperforming other existing methods in the literature, for the case of the baseline category which contains the mild challenges of a subtle background motion, isolated shadows, an abandoned object, and stopping pedestrians for a short period. The proposed method is also shown to be very robust for the case of strong variations in the background scene caused by the unstable camera and the dynamic background in the camera jitter and dynamic background categories respectively.

References

[1] F. Yan, W. Christmas, and J. Kittler, "A tennis ball tracking algorithm for automatic annotation of tennis match," in *British machine vision conference*, vol. 2, 2005, pp. 619–628.

[2] R. A. Hadi, G. Sulong, and L. E. George, "Vehicle detection and tracking techniques: a concise review," *arXiv preprint arXiv:1410.5894*, 2014.

[3] W. Kusakunniran, "Recognizing gaits on spatio-temporal feature domain," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 9, pp. 1416–1423, 2014.

[4] R. Poppe, "A survey on vision-based human action recognition," *Image and vision computing*, vol. 28, no. 6, pp. 976–990, 2010.

[5] N. Prabhakar, V. Vaithiyanathan, A. P. Sharma, A. Singh, and P. Singhal, "Object tracking using frame differencing and template matching," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 4, no. 24, pp. 5497–5501, 2012.

[6] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russell, "Towards robust automatic traffic scene analysis in real-time," in *Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on*, vol. 1. IEEE, 1994, pp. 126–131.

[7] J. N. Kapur, P. K. Sahoo, and A. K. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Computer vision, graphics, and image processing*, vol. 29, no. 3, pp. 273–285, 1985.

[8] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[9] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, "Pfinder: Real-time tracking of the human body," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 780–785, 1997.

[10] T. Horprasert, D. Harwood, and L. S. Davis, "A statistical approach for real-time robust background subtraction and shadow detection," in *Ieee iccv*, vol. 99, 1999, pp. 1–19.

[11] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," in *European conference on computer vision*. Springer, 2000, pp. 751–767.

[12] A. Ilyas, M. Scuturici, and S. Miguet, "Real time foreground-background segmentation using a modified codebook model," in *Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference on*. IEEE, 2009, pp. 454–459.

[13] Y. Li, F. Chen, W. Xu, and Y. Du, "Gaussian-based codebook model for video background subtraction," in *International Conference on Natural Computation*. Springer, 2006, pp. 762–765.

[14] S. ITing, S.-C. Hsu, and C.-L. Huang, "Hybrid codebook model for foreground object segmentation and shadow/highlight removal," *Journal of Information Science and Engineering*, vol. 30, pp. 1965–1984, 2014.

[15] M. A. Mousse, E. C. Ezin, and C. Motamed, "Foreground-background segmentation based on codebook and edge detector," in *Signal-Image Technology and Internet-Based Systems (SITIS), 2014 Tenth International Conference on*. IEEE, 2014, pp. 119–124.

[16] L. Maddalena and A. Petrosino, "A self-organizing approach to background subtraction for visual surveillance applications," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1168–1177, 2008.

[17] S. Bianco, G. Ciocca, and R. Schettini, "How far can you get by combining change detection algorithms?" *arXiv preprint arXiv:1505.02921*, 2015.

[18] Y. Wang, P.-M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, "Cdnet 2014: an expanded change detection benchmark dataset," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 387–394.

[19] S. Brutzer, B. Höferlin, and G. Heidemann, "Evaluation of background subtraction techniques for video surveillance," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1937–1944.

[20] H. Sajid and S.-C. S. Cheung, "Background subtraction for static & moving camera," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 4530–4534.

[21] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis, "Real-time foreground–background segmentation using codebook model," *Real-time imaging*, vol. 11, no. 3, pp. 172–185, 2005.

[22] P. Noriega, B. Bascle, and O. Bernier, "Local kernel color histograms for background subtraction." in *VISAPP (1)*, 2006, pp. 213–219.

[23] F. J. López-Rubio, E. López-Rubio, R. M. Luque-Baena, E. Dominguez, and E. J. Palomo, "Color space selection for self-organizing map based foreground detection in video sequences," in *2014 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2014, pp. 3347–3354.

[24] H. Sajid and S.-C. S. Cheung, "Universal multimode background subtraction," *Submitted to IEEE Transactions on Image Processing*, 2015.

[25] L. Maddalena and A. Petrosino, "A fuzzy spatial coherence-based approach to background/foreground separation for moving object

detection," *Neural Computing and Applications*, vol. 19, no. 2, pp. 179–186, 2010.

[26] R. Wang, F. Bunyak, G. Seetharaman, and K. Palaniappan, "Static and moving object detection using flux tensor with split gaussian models," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 414–418.

[27] M. Sedky, M. Moniri, and C. C. Chibelushi, "Spectral-360: A physics-based technique for change detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 399–402.

[28] K. Wang, C. Gou, and Y. Liu, "M4cd: A robust change detection method with multimodal background modeling and multi-view foreground learning," *Submitted to IEEE Transactions on Image Processing*, 2015.

[29] M. D. Gregorio and M. Giordano, "Wisardrp for change detection in video sequences," in *Submitted to the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[30] M. De Gregorio and M. Giordano, "Change detection with weightless neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 403–407.

[31] G. Ramírez-Alonso and M. I. Chacón-Murguía, "Auto-adaptive parallel som architecture with a modular analysis for dynamic object segmentation in videos," *Neurocomputing*, vol. 175, pp. 990–1000, 2016.

[32] G. Allebosch, D. Van Hamme, F. Deboeverie, P. Veelaert, and W. Philips, "C-efic: Color and edge based foreground background segmentation with interior classification," in *International Joint Conference on Computer Vision, Imaging and Computer Graphics*. Springer, 2015, pp. 433–454.

[33] G. Allebosch, F. Deboeverie, P. Veelaert, and W. Philips, "Efic: edge based foreground background segmentation and interior classification for dynamic camera viewpoints," in *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer, 2015, pp. 130–141.

[34] Y. Benezeth, P.-M. Jodoin, B. Emile, H. Laurent, and C. Rosenberger, "Comparative study of background subtraction algorithms," *Journal of Electronic Imaging*, vol. 19, no. 3, 2010.

[35] Z. Zivkovic and F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern recognition letters*, vol. 27, no. 7, pp. 773–780, 2006.

[36] D. Liang and S. Kaneko, "Improvements and experiments of a compact statistical background model," *arXiv preprint arXiv:1405.6275*, 2014.

[37] S. Varadarajan, P. Miller, and H. Zhou, "Spatial mixture of gaussians for dynamic background modelling," in *Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on*. IEEE, 2013, pp. 63–68.

[38] B. Wang and P. Dudek, "A fast self-tuning background subtraction algorithm," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 395–398.

[39] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 2. IEEE, 2004, pp. 28–31.

[40] X. Lu, "A multiscale spatio-temporal background model for motion detection," in *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2014, pp. 3268–3271.

[41] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, vol. 2. IEEE, 1999.

[42] A. Miron and A. Badii, "Change detection based on graph cuts," in *2015 International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE, 2015, pp. 273–276.

Worapan Kusakunniran received the B.Eng. degree in computer engineering from the University of New South Wales (UNSW), Sydney, Australia, in 2008, and the Ph.D. degree in computer science and engineering from UNSW, in cooperation with the Neville Roach Laboratory, National ICT Australia, Kensington, Australia, in 2013. He is currently a Lecturer with the Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand. He is the author of several papers in top international conferences and journals. His current research interests include biometrics, pattern recognition, image processing, computer vision, multimedia, and machine learning.

Dr. Kusakunniran served as a Program Committee for the ICT International Student Project Conference (ISPC) 2016, the International Conference on Knowledge and Systems Engineering (KSE) 2015, the ACCV2014 Workshop on Human Gait and Action Analysis in the Wild: Challenges and Applications, the National Conference on Information Technology (NCIT) 2014, and the IEEE Workshop on the Applications of Computer Vision (WACV) 2013.

He has also served as a Reviewer for several international conferences and journals, such as the International Conference on Pattern Recognition (ICPR), the IEEE International Conference on Image Processing (ICIP), the IEEE International Conference on Advanced Video and Signal based Surveillance (AVSS), the Pattern Recognition (PR), the IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics (TSMCB), the IEEE Transactions on Image Processing (TIP), the IEEE Transactions on Information Forensics and Security (TIFS), the IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), the Computer Vision and Image Understanding (CVIU), the Digital Signal Processing (DSP), the EURASIP Journal on Image and Video Processing (JIVP),

the Machine Vision and Applications (MVA), the International Journal of Automation and Computing (IJAC), and the IEEE Signal Processing Letters (SPL).

He was a recipient of ICPR Best Biometric Student Paper Award in 2010, ISPC 1^{st} Prize Award of Software Innovation Contest (Mobile Applications) in 2014, and ISPC 1^{st} Prize Award of Software Innovation Contest (ICT Applications) in 2016.

Rawitas Krungkaew received the B.Sc. degree in computer science from Mahidol University, Nakhon Pathom, Thailand, in 2009. He is currently working as a senior software engineer at Agoda Company Pte. Ltd. and also pursuing the M.Sc. degree in computer science, Faculty of Information and Communication Technology, Mahidol University. His current research interests include pattern recognition, image and video processing, and computer vision.

Optimal Dispatch Strategy of Cogeneration with Thermal Energy Storage for Building Energy Management System

Kebsiri Manusilp[1] and **David Banjerdpongchai**[2], Members

ABSTRACT

This paper presents optimal dispatch strategy of cogeneration with thermal energy storage (TES) for building energy management system (BEMS). In previous research related to cogeneration as a supply system, it is observed that there is some excessive heat from cogeneration operation released to the atmosphere. In order to improve energy efficiency, we therefore incorporate TES to utilize the excessive heat from cogeneration into two objective functions, i.e., total operating cost (TOC) and total carbon dioxide emission (TCOE). In particular, we aim to minimize TOC which is referred to economic optimal operation and to minimize TCOE which is referred to environmental optimal operation. Both optimal operations are subjected to energy dispatch strategy which TES constraint is taken into account. We demonstrate the dispatch strategy with a load profile of a large shopping mall as a test system and compare the results to that of previous dispatch of cogeneration without TES. The proposed strategy of cogeneration with TES can reduce TOC of the test system up to 4.15% and 1.85% for economic and environmental optimal operations, respectively. Furthermore, TCOE can be reduced up to 5.25% and 6.25% for economic and environmental optimal operations, respectively.

Keywords: Combined Heat and Power Generation (CHP), Thermal Energy Storage, Building Energy Management System (BEMS), Energy Efficiency

1. INTRODUCTION

Energy consumption in buildings is essential for energy efficiency plan because energy consumption has been rising while energy supply is limited. Therefore, building energy management system (BEMS) has become an interesting topic which aims to obtain high efficiency of energy usage. Building's load can be normally classified in three types: electrical, cooling, and heating loads which can receive energy from power grids. The power grid generally provides high carbon dioxide emission and cost of importing power from power grid also depends on on/off peak hours. There are many researches [1-2] seeking for technology to reduce importing power from grid and reduce operating cost and CO_2 emission from electrical generation to the building. One of important technologies normally applied with BEMS is cogeneration or combined heat and power (CHP).

CHP [3-4] is simultaneous production of heat and power which uses only a single source. Heat energy can obtain from fuel combustion in the process. Several survey articles indicate that CHP technology can provide more reliable and higher energy efficiency compared to conventional generation. Moreover, Thailand Power Development Plan (PDP2015) [5] also promotes CHP as an energy source due to its efficiency. Since the buildings such as commercial buildings or shopping malls contain coincidence of onsite cooling, heating, and electrical loads, CHP is suitable for the building supply. In research report [6], implementation of integrated model of cogeneration, solar, and conventional energy source is applied to energy demand in commercial building. The problem aims to minimize cost of life-cycle and earn profit from exporting onsite-generated electrical energy to grid based on national policy.

Cogeneration system is focused on the economic operation under emission constraints which tries to achieve minimum cost and minimum pollutant emission of cogeneration [7]. Linear model is designed to respond the electrical, thermal, and cooling loads in the small industrial building. In [8], the economic operation aims to reduce investment, save more energy and time, and gain more benefit and comfort.

In [9], the modified model of cogeneration is integrated with BEMS to minimize operating cost and CO_2 emission. The research [9] has proposed design of cogeneration which is applied to a large shopping mall. Their framework provides optimal operation which aims to minimize total operating cost (TOC) and total CO_2 emission (TCOE). The major components of BEMS [9] consists of CHP, absorption chiller, and auxiliary boiler as energy supply. CHP is main component which supplies both electrical energy and heat energy to demand in building's load. It is observed that CHP releases some waste heat, i.e., exces-

Manuscript received on August 15, 2016 ; revised on November 10, 2016.
Final manuscript received on February 7, 2017.
[1,2] The authors are with the Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University, 254 Phayathai Road, Pathumwan, Bangkok, 10330 Thailand, E-mail: kemanusilp@gmail.com and bdavid@chula.ac.th
[2] Corresponding author.

sive thermal energy over the load demand. In this paper, we incorporate Thermal Energy Storage (TES) [10] to utilize excessive thermal energy from CHP system. TES is employed to store excessive thermal energy from CHP and cooperate to supply thermal energy to absorption chiller. We will give a guideline how to choose capacity of TES. Furthermore, we demonstrate that TES can significantly improve operation of auxiliary boiler which, in return, reduce TOC and TCOE. Then, we analyze energy flow of both economic and environmental optimal operation.

This paper is organized as follows. Section 2 describes the main components of BEMS. Section 3 presents problem formulation including objective functions and dispatch strategies. Effects of TES to TOC and TCOE are given in section 4. Section 5 presents numerical results of optimal dispatch strategies. Section 6 analyzes optimal energy flow. Conclusions are given in section 7.

2. SYSTEM DESCRIPTION

The proposed BEMS is composed of CHP, auxiliary boiler, absorption chiller, and thermal energy storage as internal energy supply of the building. Power grid is external energy supply which serves as power exchange unit with BEMS. CHP is the main energy supply which responds to load demand in the building. Following the research [9], we classify building load into two types for consideration: electrical load and cooling load. In Figure 1, CHP simultaneously produces electrical energy (EE) and thermal energy or heat energy (HE) which supply to the building load. CHP mainly supplies EE to respond the electrical load. Moreover, power grid is connected to BEMS serving as a backup supply. In case of power shortage, power grid takes responsibility to supply EE to electrical load. On the other hand, BEMS can export EE to power grid when EE production from CHP exceeds the building demand. Absorption chiller is a component which converts HE to be cooling energy before supplies to cooling load. Absorption chiller receives HE from CHP, auxiliary boiler and TES to respond cooling load demand in the building. Furthermore, excessive HE from CHP production which is over the load demand will be charged to TES.

3. PROBLEM FORMULATION

Optimal dispatch operation for the proposed BEMS can be defined by two objective functions, namely, economic optimal operation and environmental optimal operation. Both objective functions are subjected to electrical energy dispatch and cooling energy dispatch. TES utilization is taken into account of cooling energy dispatch. We simplify model of each component to be linear and neglect internal losses.

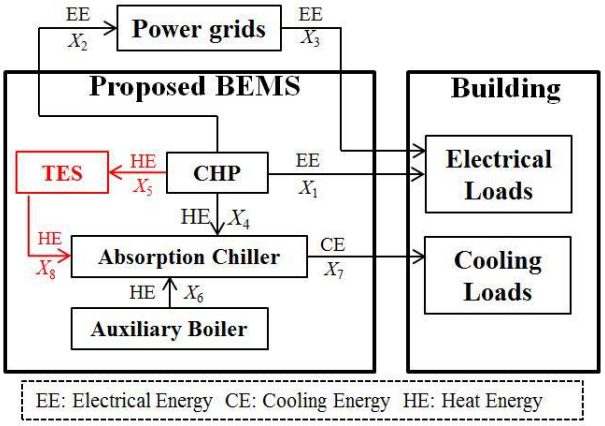


Fig.1: *Diagram of proposed BEMS.*

3.1 Objective Functions

The economic optimal operation aims to minimize total operating costs of BEMS. The objective function, total operating cost (TOC), is equal to the sum of energy cost (EC) and demand charge cost (DCC). EC consists of the sum of energy cost of CHP, selling and purchasing of power trading with power grids, and auxiliary boiler. DCC is calculated from maximum imported power from power grids. The economic objective function is stated as follows.

$$\mathrm{TOC} = \mathrm{EC} + \mathrm{DCC} \tag{1}$$

$$\mathrm{EC} = \sum_{k=1}^{nxd} c_{CHP}(x_{1,k} + x_{2,k}) - q_k x_{2,k} + p_k x_{3,k} + c_{AB} x_{6,k} \tag{2}$$

$$\mathrm{DCC} = \frac{d_{PG}}{\Delta t} \max_{h=1,\ldots,nxd} x_{3,k} \tag{3}$$

where $x_{i,k}$ is energy flow following Fig.1 in time interval of k. c_{CHP} and c_{AB} are operating cost of the CHP and the auxiliary boiler which depend on fuel price, q_k is electrical energy base price, p_k is electrical energy charge price during on-peak or off-peak time, d_{PG} is demand charge depending on maximum imported power from power grids, n is the number of time interval in one day, d is the number of days, and $\triangle t$ is time duration of each time interval. Unit of cost is baht.

The environmental optimal operation addresses issues of greenhouse gas effect. The objective function aims to minimize carbon dioxide emission which considers in terms of total carbon dioxide emission (TCOE). TCOE is equal to the sum of carbon dioxide emission from CHP, power grids, and auxiliary boiler. The environmental optimal operation is stated as follows.

$$\mathrm{TCOE} = \sum_{k=1}^{nxd} (EF_{CHP,CO_2}(x_{1,k} + x_{2,k}) + GEF x_{3,k} + \frac{EF_{AB,CO_2}}{\eta_{AB}} x_{6,k}) \tag{4}$$

where EF_{CHP,CO_2} and EF_{AB,CO_2} are CO_2 emission of the CHP operation and the auxiliary boiler which depend on fuel price, GEF is grid emission factor, and η_{AB} is efficiency of boiler. Unit of TCOE is tonCO_2.

3.2 Thermal Energy Storage Model

Thermal Energy Storage (TES) is component which stores excessive heat from CHP. Status of TES is represented by state of charge (SOC) which depends on charging and discharging of thermal energy with TES. Thermal energy charges and discharges of TES are under limitation of charge and discharge rate. Moreover, we need to take into account of the maximum and minimum states of charge of TES. Constraints of TES are given as follows.

$$\varepsilon x_{5,k} \leq R_1 \tag{5}$$

$$\frac{1}{\delta} x_{9,k} \leq R_2 \tag{6}$$

$$x_{9,k} = \text{init}(1-\mu)^k + \sum_{j=1}^{k} \left(\varepsilon x_{5,j} - \tfrac{1}{\delta} x_{8,j}\right)(1-\mu)^{k-j+1} \tag{7}$$

$$S_{min} \leq x_{9,k} \leq S_{max} \tag{8}$$

where ε and δ are TES charge and discharge efficiency, R_1 and R_2 are charge and discharge rate. The variable $x_{9,k}$ represents state of charge at time k. μ is loss coefficient of TES and init is initial heat stored in TES. Finally, S_{min} and S_{max} are minimum and maximum states of TES.

3.3 Dispatch Strategies

Dispatch strategies of optimal operation aim to efficiently supply energy to meet demand in the building. The EE dispatch concerns management of EE supplies to the load. CHP and power grid together cooperate to supply EE to meet electrical load (U_k). CHP will operate for EE production under limitation of its production, $P_{CHP,min}$ and $P_{CHP,max}$, when there is EE load. The production of CHP is usually operated with a proportion of EE and HE equal to power-to-heat ratio (P2H) of their CHP. Ramp rate (R_{CHP}), calculated from the difference between EE production at current and previous state, should be considered. EE dispatch is given as follows.

if $U_k = 0$, then

$$x_{1,k} = x_{2,k} = x_{4,k} = x_{5,k} = 0$$

else

$$P_{CHP,min}\Delta t \leq x_{1,k} + x_{2,k} \leq P_{CHP,max}\Delta t$$
$$\tfrac{x_{1,k}+x_{2,k}}{x_{4,k}+x_{5,k}} = P2H$$
$$|(x_{1,k} + x_{2,k}) - (x_{1,k-1} + x_{2,k-1})| \leq R_{CHP}\Delta t$$

end

$$x_{1,k} + x_{3,k} = U_k$$

For CE dispatch, energy supplies consist of CHP, auxiliary boiler, TES, and absorption chiller. CE dispatch manages cooling energy from these operations supply to meet cooling load (C_k) in the building which can divided in four conditions. Firstly, in case of no cooling load there is no HE supply to chiller and chiller does not operate. The excessive heat from CHP will be charged to TES under charging rate. Secondly, in case there is C_k less than minimum operation of chiller and CHP can supply enough HE to chiller, CHP and TES cooperate to supply HE to chiller. The chiller operates at its minimum cooling production ($CP_{AC,min}$) to convert HE to CE supply to C_k in the building. Auxiliary boiler is not necessary to use, the boiler shuts down. Thirdly, in case C_k is over the minimum operation of chiller, CHP and TES cooperate to supply HE to chiller. Auxiliary boiler still shuts down when CHP can supply enough HE to chiller. Lastly, auxiliary boiler will operate in case CHP cannot supply enough HE. CHP, TES, and boiler cooperate to supply HE to chiller. Chiller will operate at the C_k level but not over the maximum of cooling production of chiller ($CP_{AC,max}$). CE dispatch is given as follows.

if $C_k = 0$, then

$$x_{4,k} = x_{6,k} = x_{7,k} = x_{8,k} = 0$$
$$x_{9,k} = (x_{9,k} + x_{5,k})(1-\mu)$$
$$\varepsilon * x_{5,k} \leq R_1\Delta t$$

elsesif $C_k \leq CP_{AC,min}\Delta t$, then

$$(x_{4,k} + x_{8,k})COP_{AC} = x_{7,k}$$
$$\tfrac{1}{\delta} x_{8,k} \leq R_2\Delta t$$
$$x_{9,k} = (x_{9,k-1} - x_{8,k})(1-\mu)$$
$$x_{5,k} = x_{6,k} = 0$$
$$x_{7,k} = CP_{AC,min}\Delta t$$

elseif $C_k \leq \frac{P_{CHP,max}\Delta t}{P2H} COP_{AC}$, then

$$(x_{4,k} + x_{8,k})COP_{AC} = x_{7,k}$$
$$\tfrac{1}{\delta} x_{8,k} \leq R_2\Delta t$$
$$x_{9,k} = (x_{9,k-1} - x_{8,k})(1-\mu)$$
$$x_{5,k} = x_{6,k} = 0$$
$$x_{7,k} = \min(C_k, \tfrac{P_{CHP,max}\Delta t}{P2H} COP_{AC})$$

else

$$(x_{4,k} + x_{6,k} + x_{8,k})COP_{AC} = x_{7,k}$$
$$\tfrac{1}{\delta} x_{8,k} \leq R_2\Delta t$$
$$x_{9,k} = (x_{9,k-1} - x_{8,k})(1-\mu)$$
$$HP_{AB,min}\Delta t \leq x_{6,k} \leq HP_{AB,max}\Delta t$$
$$x5,k = 0$$

$$x_{7,k} = \min(C_k, CP_{AC,max}\left(\tfrac{P_{CHP,max}\Delta t}{P2H} + HP_{AB,max}\Delta t\right) COP_{AC})$$

end

The optimal dispatch can be formulated as optimization problem with economic and environmental optimal operation subject to operation constraints. The problems are linear program (LP) which can be efficiently solved by LP solvers.

3.4 System Parameters

The proposed BEMS requires appropriate operations to establish as a system. There are some research articles reporting CHP operations. For CHP operation, BEMS employs gas turbine as the CHP system with size and capacity based on peak electrical demand. Moreover, gas turbine can obtain high temperature steam [12]. There is report of varying CHP size from 22-25 MW and observe TOC and TCOE [9]. In this paper, we choose CHP of 24 MW which is appropriate for the load demand of test system. To select the absorption chiller, we consider heat energy input of absorption chiller and heat energy output of CHP system. The double-effect absorption chiller is selected and the coefficient of performance (COP) is 1.1 following regulation on energy usage for building recommendation [9]. We choose industrial boiler as an auxiliary boiler. Based on surveys, industrial boiler utilizes natural gas and gives approximately 75% of thermal efficiency at full load and there are various capacities [9, 13]. CO_2 emission factor from natural gas combustion (EF_{AB,CO_2}) is 0.1810 tCO_2/MWh [14].

For electricity price, there are two types of electricity price in Thailand, i.e., time-of-day (TOD) and time-of-use (TOU) but we consider TOU in this study. TOU rates depend on time of use during the day; on-peak time is on Monday to Friday 09.00-22.00 and the rest is off-peak time. Moreover, TOU tariffs consist of energy charge, demand charge, service charge, power factor charge, fuel adjustment (Ft), and VAT. In this study, we consider only energy charge and demand charge and the rest are neglected. BEMS operation pays for energy charge as 4.1283 and 2.6107 baht/kWh for on-peak and off-peak time, respectively. The demand charge for large general service with 69-kV is 74.14 baht/kW [15]. Besides, electricity price for selling EE from CHP production is referred in [16]. For grid emission factor (GEF) of Thailand, CO_2 emission is estimated when BEMS uses electricity from grid [17]. Parameters of TES and other parameters are summarized in Table. 1.

Table 1: *System parameters.*

Description	Notation	Value
CHP system		
Rated power (MW)	-	24
Electrical energy efficiency (%)	$\eta_{CHP,EE}$	33.90
Power to heat ratio	P2H	0.9244
Maximum power production (MW)	$P_{CHP,max}$	24
Minimum power production (MW)	$P_{CHP,min}$	4.8
Electrical energy ramp rate (MW)	R_{CHP}	24
CO_2 emission factor (tCO_2/MWh)	EF_{CHP,CO_2}	0.5349
Auxiliary boiler		
Rated heat power (MW)	-	13.1882
Efficiency (%)	η_{AB}	75
Maximum heat production (MW)	$HP_{AB,max}$	13.1882
Minimum heat production (MW)	$HP_{AB,min}$	2.6376
CO_2 emission factor from natural gas combustio	EF_{AB,CO_2}	0.1810
Absorption chiller		
Rated cooling power	-	42.2
Coefficient of performance	COP_{AC}	1.1
Maximum cooling production (MW)	$CP_{AC,max}$	42.2
Minimum cooling production (MW)	$CP_{AC,min}$	8.44
Thermal energy storage		
Rated heat power (MW)	-	50
Heat charge rate (MW)	R_1	15
Heat discharge rate (MW)	R_2	15
Charging efficiency	ε	0.95
Discharging efficiency	δ	0.95
TES loss coefficient	μ	0.001
Initial heat energy in TES MW)	init	5
Maximum heat storage (MW)	S_{max}	50
Minimum heat storage (MW)	S_{min}	5
Others		
Electrical energy load in time interval k (MWh)	U_k	-
Cooling energy load in time interval k (MWh)	C_k	-
Time duration of each interval (hour)	Δt	1
Number of time interval per day	n	24
Number of days in month	D	7
Average Price of Natural Gas of December 2015 (bath/MMBtu)	APNG	179.54

4. EFFECTS OF THERMAL ENERGY STORAGE

As TES is introduced to utilize excessive thermal energy from CHP, how to select a suitable capacity of TES will be considered in this section. We will compare TOC and TCOE for several TES capacities and for the case without TES. TES is varied in range from 41 to 60 MW. The comparison is based on economic and environmental optimal operation.

4.1 Economic Optimal Operation

This subsection shows TOC and TCOE under economic optimal operation. Figure 2 presents TOC comparing between BEMS with TES and BEMS without TES. It clearly shows that TOC continually decrease when TES capacity increases. It is observed that TOCs can be reduced by BEMS with TES at least 1.45% and at most 5.25%.

As Eq. (1), TOC comes from the sum of energy cost of components in BEMS and demand charge cost of imported EE from power grids. It is shown that BEMS with TES can linearly reduce energy cost or EC when capacities of TES increase as in Figure 3 (top). TOC can be reduced at least 4% and at most 6.22% when compared to that of BEMS without TES. Besides, BEMS with TES makes demand charge cost increase when compared to that BEMS without TES as shown in Figure 3 (bottom). Demand charge cost linearly decreases in range 41 to 46 MW of TES ca-

pacities and starts saturating at 47 MW of TES capacity. The reason is that CHP tries to produce EE to meet building load while HE produced from CHP cannot exceed TES capacity. Therefore, CHP cannot produce enough EE to meet the building load so EE is imported from power grids due to inexpensive cost of imported EE. As a result, BEMS with TES has higher demand charge cost.

To further investigate the effects of TES, EC is broken down by components as shown in Figure 4. Operating cost of CHP linearly increases as TES capacity ranges from 41 to 51 MW and starts saturating at 52 MW. Furthermore, the last point which makes cost of CHP operation less than the case without TES is at TES capacity 50 MW as shown in Figure 4 (top). Figure 4 (middle) shows that the main reduction of energy cost comes from auxiliary boiler which can reduce at least 29.97% and at most 44.48% when compared to the case of without TES. Figure 4 (bottom) shows the cost of importing EE from power grids linearly decreases as TES capacity ranges from 41-51 MW and starts saturating at 52 MW but the cost is a bit higher than that BEMS without TES.

TCOE consists of emission from operating components as shown in Eq. (4). In Figure 5, BEMS with TES linearly reduces TCOE as TES capacity increases. In particular, BEMS with TES can decrease TCOE at least 4.24% and at most 6.19%.

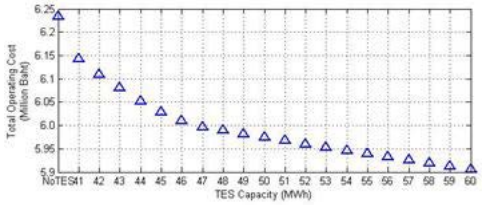


Fig.2: *Total operating cost when varying TES capacity for economic optimal operation.*

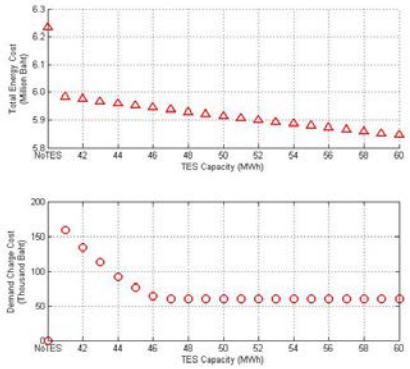


Fig.3: *Total energy cost and demand charge cost when varying TES capacity for economic optimal operation.*

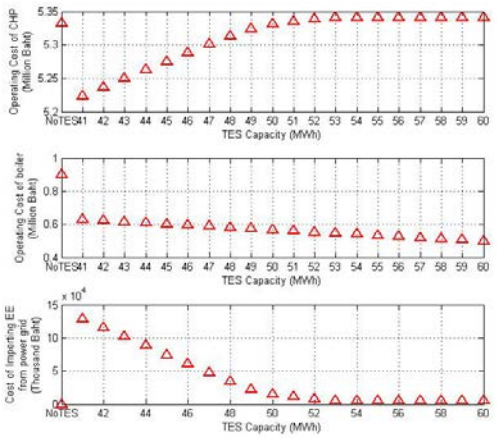


Fig.4: *Operating cost of each component when varying TES capacity for economic optimal operation.*

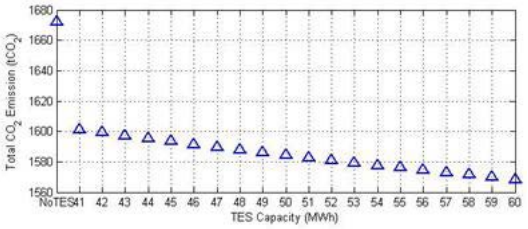


Fig.5: *Total CO_2 emission when varying TES capacity for economic optimal operation.*

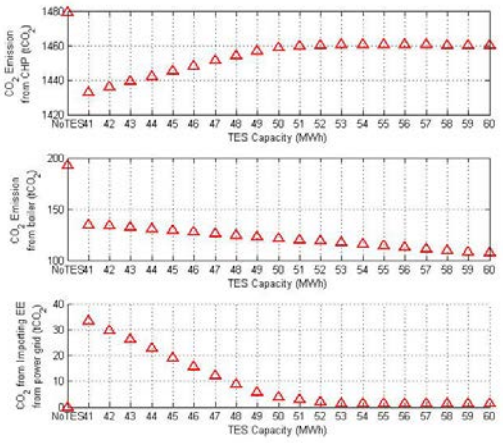


Fig.6: *CO_2 emission of each component when varying TES capacity for economic optimal operation.*

When breaking down TCOE by components, we find that the trend looks similar to TOC. Main CO_2 emission comes from auxiliary boiler operation as shown in Figure 6 (middle). CO_2 emission of CHP operation linearly increases as TES capacity ranges from 41-50 MW and starts saturating at 51 MW. Although BEMS with TES makes CO_2 emission of CHP increase, all CO_2 emission still less than that of BEMS without TES as shown in Figure 6 (top). CO_2 emission of importing EE from grids inversely varies with emission from CHP and starts saturating at 54 MW but emission is a bit higher than that BEMS without TES.

4.2 Environmental Optimal Operation

In this subsection, we show the effects of TES on TOC and TCOE under environmental optimal operation. Figure 7 shows that BEMS with TES makes TOCs continually decrease when TES capacity increases. It is observed that TES of 41-43 MW makes TOC higher than that of BEMS without TES. BEMS with TES can reduce TOC at least 0.45% and at most 3.57%. Besides, Figure 8 shows EC and DCC as Eq. (1). BEMS with TES provides EC lower than that BEMS without TES. EC linearly reduces when TES capacity increases as shown in Figure 8 (top). Demand charge cost when applied TES is still higher than the case without TES. TOC continuously decreases when TES capacity increases and starts saturating at 53 MW of TES capacity as shown in bottom subplot Figure 8.

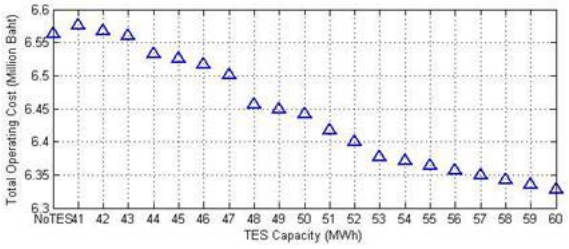


Fig.7: *Total operating cost when varying TES capacity for environmental optimal operation.*

When breaking down EC by components, the main reduction of energy cost comes from auxiliary boiler operation. BEMS with TES can linearly reduce energy cost of auxiliary boiler operation at least 15.70% and at most 27.80% as shown in Figure 9 (middle). In Figure 9 (top), operating cost of CHP linearly increases when TES capacity increases and starts saturating at 53 MW. It's observed that TES of 50 MW gives the cost of CHP operation less than that of BEMS without TES. Cost of importing EE from grids linearly decreases as TES capacity ranges from 41-50 MW and starts saturating at 53 MW. In Figure 9 (bottom), we observe the cost of importing EE from power grids with TES greater than 53 MW is slightly higher than the case without TES.

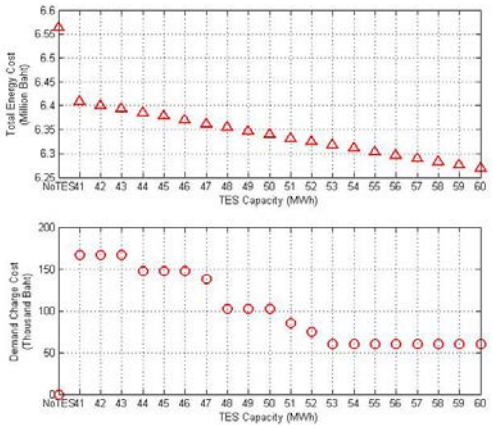


Fig.8: *Total energy cost and demand charge cost when varying TES capacity for environmental optimal operation.*

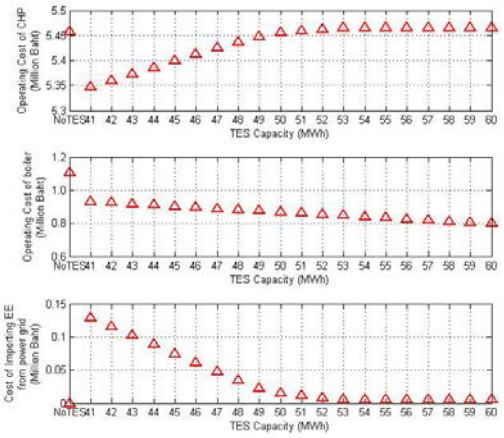


Fig.9: *Operating cost of each component when varying TES capacity for environmental optimal operation.*

Figure 10 shows TCOE linearly decreases when capacity of TES increases. BEMS with TES can reduce TCOE at least 5.25% and at most 7.25% when compared to that of BEMS without TES. As expressed in Eq. (4), TCOE is combination of emission from CHP, auxiliary boiler, and importing EE from grids. The main reduction of CO_2 emission comes from the auxiliary boiler operation as shown in Figure 11 (middle). CO_2 emission from boiler reduces at least 15.61% and at most 27.85% compared to the case without TES. In Figure 11 (top), when operating BEMS with TES, CHP system gives lower CO_2 emission compared to the case without TES. CO2 emissions from CHP start saturating at TES of 53 MW. CO_2 emissions of importing EE from grid linearly decreases when TES capacity increases and saturates at 53 MW as shown in Figure 11 (bottom). However, importing EE from

power grids when operating with TES greater than 53 MW gives slightly higher CO_2 emission when compared to the case without TES.

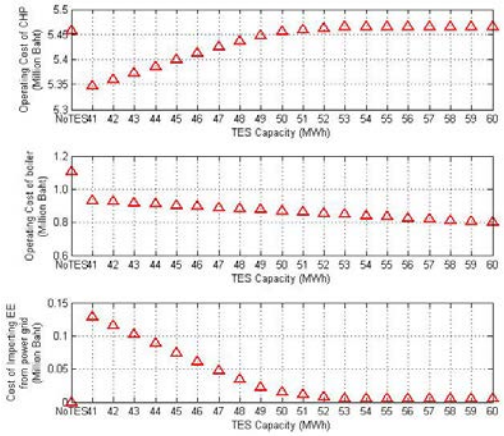


Fig.10: *Total CO_2 emission when varying TES capacity for environmental optimal operation.*

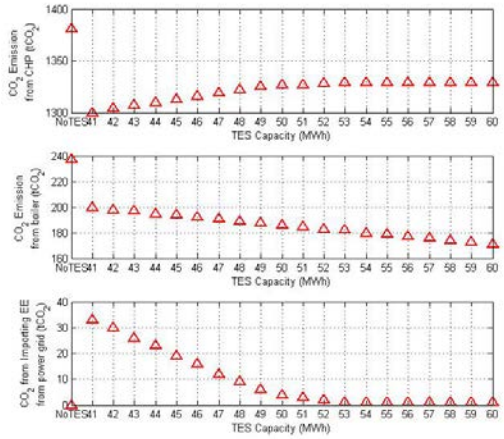


Fig.11: *CO_2 emission of each component when varying TES capacity for environmental optimal operation.*

It is observed that in both economic and environmental optimal operations, BEMS with TES can reduce more TOC and TCOE when TES capacity increases. It is essential to have some criteria for the selection of TES. When we break down EC, it is obvious that at 50 MW is the minimum capacity which can maintain cost of CHP operation lower than the cost of CHP without TES. Therefore, TES of 50 MW is chosen for BEMS.

5. NUMERICAL RESULTS

In this section, we compare TOC and TCOE between BEMS with TES and BEMS without TES. We consider electrical and cooling load profiles of a large shopping mall for 7 days. The average of natural gas price on December 2015 [11] at 179.54 baht/MMBtu is used for calculation. TES with 50 MW is applied to support CHP operation. Table 2 summarizes comparison results.

In Table 2, BEMS with TES can reduce TOC 258,504 and 121,188 baht for economic and environmental optimal operations, respectively. Moreover, when applying TES, TCOE is reduced by 88 and 101 tonCO_2 for economic and environmental optimal operation, respectively. To further investigate the benefits of TES, we note that the main reduction of both TOC and TCOE comes from auxiliary boiler operation as shown in Table 3 and Table 4. For economic operation, BEMS with TES can reduce operating cost of auxiliary boiler up to 332,475 baht and reduce CO_2 emission up to 71 tonCO_2. In part of environmental optimal operation, operating cost of auxiliary boiler is reduced up to 237,042 baht when operating TES. Moreover, CO_2 emission from boiler decreases up to 51 tonCO_2. It is obvious that TES plays a cooperative role to supply heat to cooling load of building. However, there exists some cost and CO_2 emission for importing EE from power grids when applying TES. In both economic and environmental optimal operation, we find that operating cost and CO_2 emission increases because the charge of importing EE from power grids during off-peak period is cheaper than running operation of CHP.

6. ANALYSIS OF OPTIMAL ENERGY FLOW

In this section, we demonstrate daily electrical and cooling energy flow of BEMS with TES and compare to the case without TES [9]. The analysis is divided into economic and environmental optimal operations.

Table 2: *Comparison of total operating cost and CO_2 emission.*

Objective function	BEMS without TES	BEMS with TES	% change
Economic Optimal Operation			
Total Operating Cost (TOC,baht)	6,233,275	5,974,771	-4.15
Total CO_2 Emission (TCOE,tCO_2)	1,672	1,584	-5.25
Environmental Optimal Operation			
Total Operating Cost (TOC,baht)	6,563,078	6,441,890	-1.85
Total CO_2 Emission (TCOE,tCO_2)	1,618	1,517	-6.25

6.1 Economic Optimal Operation

In Figure 12, the main EE supply to electrical load comes from CHP production. Moreover, the BEMS without TES [9] tries to utilize power grid to sup-

Table 3: *Operating cost classified by component.*

Component	BEMS without TES	BEMS with TES	% change
Economic Optimal Operation			
CHP	5,332,627	5,331,534	-0.02
Auxiliary Boiler	900,648	568,173	-36.92
Power Grids	0	15,134	N/A
Environmental Optimal Operation			
CHP	5,457,205	5,455,525	-0.03
Auxiliary Boiler	1,105,874	868,832	-21.43
Power Grids	0	15,131	N/A

Table 4: *CO_2 emission classified by component.*

Component	BEMS without TES	BEMS with TES	% change
Economic Optimal Operation			
CHP	1,479	1,459	-1.38
Auxiliary Boiler	193	122	-36.92
Power Grids	0	4	N/A
Environmental Optimal Operation			
CHP	1,381	1,327	-3.92
Auxiliary Boiler	237	186	-21.43
Power Grids	0	4	N/A

ply EE to electrical load during off-peak hour as yellow area in bottom figure. The reason is that operating cost of EE production of CHP is higher than charge of imported EE from power grids during off-peak period. The power grid takes responsibility with CHP to reduce TOC as least as possible. EE and HE production of CHP are shown in Figure 13 and 14. BEMS with TES operates quite similar to BEMS without TES. The main EE production of CHP supply to electrical load is shown as pink region in Figure 13. They try to earn profit from selling excessive EE to power grids as green area. In Figure 14, CHP also supplies main heat energy production to chiller. The excessive heat at off-peak hours is released to the atmosphere as shown in Figure 14 (top). On the other hand, BEMS with TES utilizes excessive heat energy, shown as white region in Figure 14 (bottom) by charging to TES.

Cooling energy flow are shows in Figure 15. Absorption chiller receives heat energy from supply sources, and converts to cooling energy to supply to the cooling load. CHP is the main supply source of HE to chiller in both BEMS without TES [5] and BEMS with TES. In BEMS without TES, auxiliary boiler cooperates with CHP to supply HE to chiller during on-peak hours as shown in Figure 15 (top). In contrary, BEMS with TES incorporates with auxiliary boiler and CHP to supply HE to chiller at on-peak hours as shown in Figure 15 (bottom). It is obvious that TES comes to reduce the operation of auxiliary boiler at on-peak hours which makes the main reduction of operating cost and CO_2 emission.

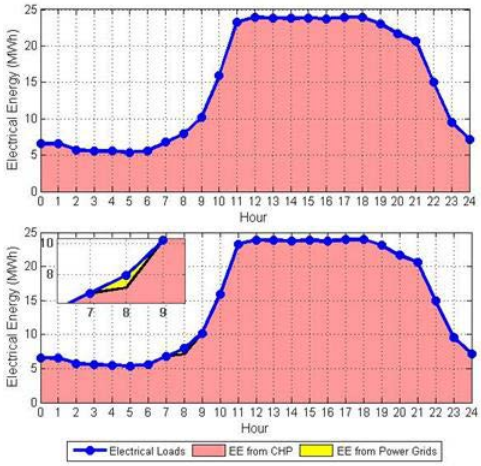


Fig.12: *Electrical energy flow to electrical load of BEMS without TES (top) and BEMS with TES (bottom) for economic optimal operation.*

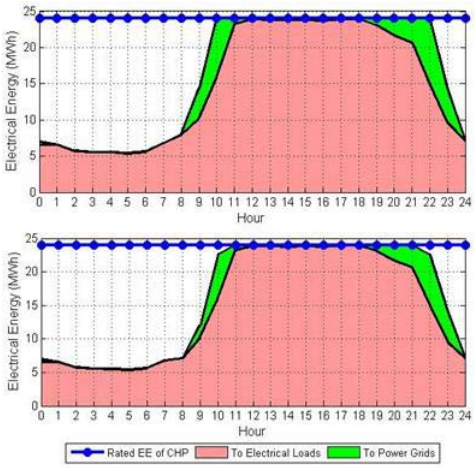


Fig.13: *Electrical energy production of CHP of BEMS without TES (top) and BEMS with TES (bottom) for economic optimal operation.*

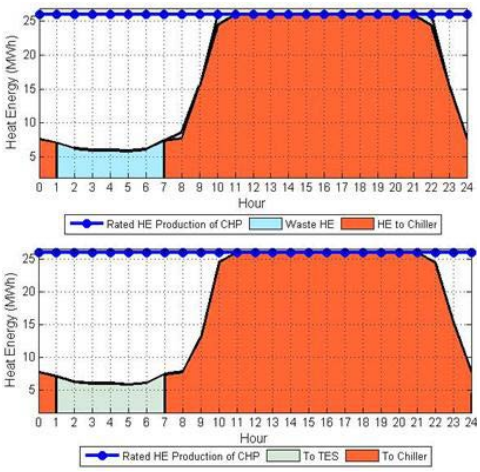


Fig.14: *Heat energy production of CHP of BEMS without TES (top) and BEMS with TES (bottom) for economic optimal operation.*

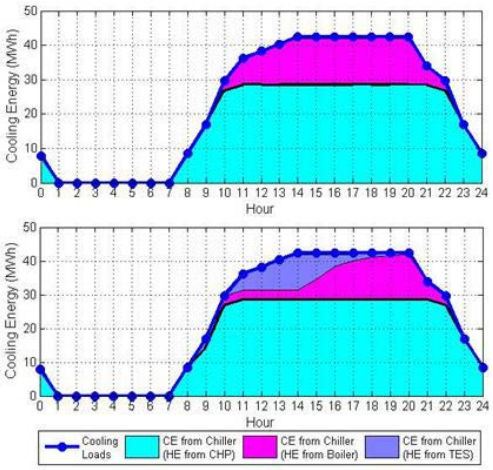


Fig.15: *Cooling energy flow to cooling load of BEMS without TES (top) and BEMS with TES (bottom) for economic optimal operation.*

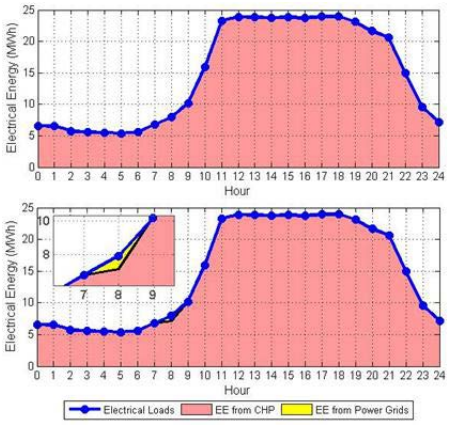


Fig.16: *Electrical energy flows to electrical load of BEMS without TES (top) and BEMS with TES (bottom) for environmental optimal operation.*

6.2 Environmental Optimal Operation

Environmental optimal operation aims to minimize TCOE. As Figure 16, CHP is the main EE supply source to electrical load. Figure 16 (bottom) shows that BEMS with TES imports EE from power grids as yellow area to supply to electrical load. Power grids take responsibility with CHP to supply EE because importing EE from grid gives CO_2 emission lower than running CHP operation during off-peak hours. In Figure 17, EE production of CHP mainly supplies to electrical load in both BEMS without TES and BEMS with TES. In Figure 17 (top), BEMS tries to earn profit of selling excessive EE of CHP production at on-peak hours. In contrary, BEMS with TES has zero exporting EE to power grid. The reason is that TES cooperates with CHP and boiler to support HE, whereas EE (proportional to HE under P2H ratio) is not produced more than electrical demand. To minimize CO_2 emission as least as possible, CHP does not produce EE to earn profit from selling EE to power grids.

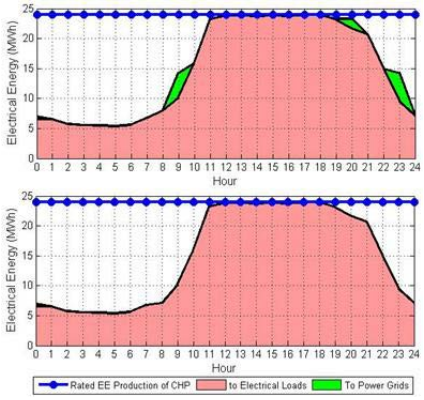


Fig.17: *Electrical energy production of CHP of BEMS without TES (top) and BEMS with TES (bottom) for environmental optimal operation.*

Heat energy production of CHP is referred in Figure 18. It is quite similar to that of economic optimal operation. CHP supplies the main HE production to absorption chiller and excessive heat during off-peak hour is released to the atmosphere as shown in Figure 18 (top). BEMS with TES utilizes excessive HE from CHP operation by charging to TES. Figure 19 shows that the main heat energy supplied to chiller comes from CHP operation in both BEMS without TES and BEMS with TES. TES cooperates with CHP and auxiliary boiler at on-peak hours to obtain minimum CO_2 emission. It is obvious that TES can reduce operation of auxiliary boiler at on-peak hours as shown in Figure 19 (bottom).

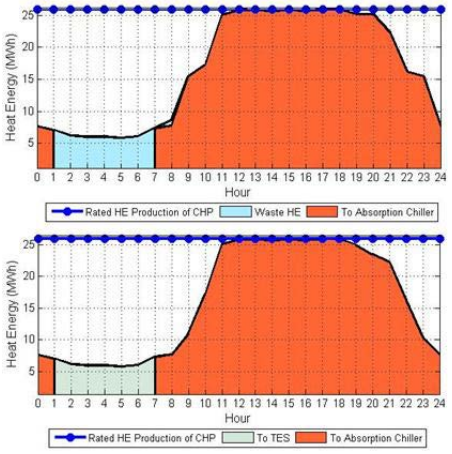


Fig.18: *Heat energy production of CHP of BEMS without TES (top) and BEMS with TES (bottom) for environmental optimal operation.*

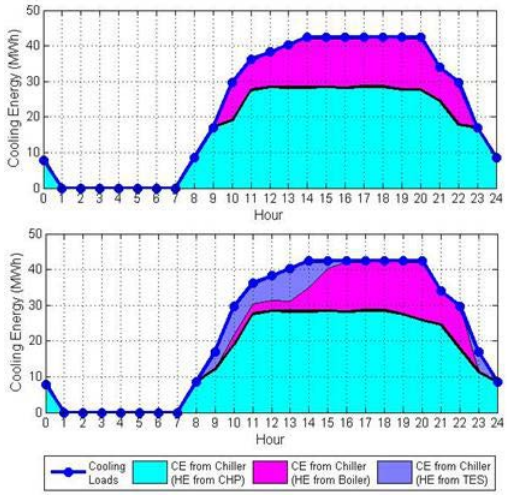


Fig.19: *Cooling energy flow to cooling load of BEMS without TES (top) and BEMS with TES (bottom) for environmental optimal operation.*

7. CONCLUSIONS

In this paper, thermal energy storage in conjunction of cogeneration utilizes excessive heat to improve energy efficiency of BEMS. TES constraints are taken into account in the energy dispatch strategy. It is observed that TES can cooperate with CHP and auxiliary boiler at on-peak hours. Moreover, BEMS with TES can significantly reduce TOC and TCOE when compared to the case of BEMS without TES. The main energy improvement comes from the reduction of heat energy production from auxiliary boiler.

References

[1] Z. Jiang and H. Rahimi-Eichi, "Design, modeling and simulation of a green building energy system," *IEEE Power & Energy Society General Meeting*, Calgary, AB, pp. 1-7, 2009.

[2] D. Estates, "Building energy management systems," *Her Majesty's Stationery Office, Ministry of Defense*, United Kingdom, Jan. 2001.

[3] S. Yokoyama, "A Guide for Biomass Production and Utilization," *The Asian Biomass Handbook*, pp. 94-97, 2008.

[4] U.S. Environmental Protection Agency and Combined Heat and Power Partnership (Dec. 2008). Catalog of CHP Technologies [Online]. Available from: www.epa.gov/chp/documents/catalog_chptech_full.pdf.

[5] Energy, Ministry. Energy Policy and Planning Office, "Thailand Power Development Plan 2015-2036 (PDP2015)," *Energy Policy and Planning Office. Ministry of Energy. Thailand. [Online]*. June 2015. Available from: www.egat.co.th/en/images/

[6] C.-C. Tseng, W.-H. Lin, S.-Y. Wu, S.-H. Chen, and S.-Y. Lin, "The transition mechanisms of type-II GaSb/GaAs quantum-dot infrared light-emitting diodes," *J. Cryst. Growth*, Vol. 323, pp. 466-469, 2011.

[7] M. T. Tsay, W. M. Lin, J. L. Lee, "Interactive best-compromise approach for operation dispatch of cogeneration systems," *IEE Proceedings: Generation, Transmission and Distribution*, vol. 148, no. 4, July 2001.

[8] R. Hashemi, "A Developed Offline Model for Optimal Operation of Combined Heating and Cooling and Power Systems," *IEEE Transactions on Energy Conversion*, vol. 24, no. 1, pp. 222-229, Mar. 2009.

[9] T. Petkajee and D. Banjerdpongchai, "Design of Cogeneration and Analysis of Economic and Environmental Optimal Operations for Building Energy Management System", *ECTI Transactions on Electrical Engineering, Electronics, and Communications (EEC)*, vol. 11, no. 2, pp. 79-94, Aug. 2013.

[10] K. Manusilp and D. Banjerdpongchai, "Optimal dispatch of cogeneration with thermal energy storage for Building Energy Management System," *2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, Chiang Mai, pp. 1-6, June 2016.

[11] "Fuel Prices in December 2015", *Department of Mineral Fuels, Ministry of Energy*, Thailand, Dec. 2015.

[12] United States Environmental Protection Agency (US EPA) and Combined Heat and Power Partnership. "Catalog of CHP Technologies. US EPA. United States of America," [Online]. December 2008. Available from: http://www.epa.gov/chp/documents/catalog_chptech_full.pdf

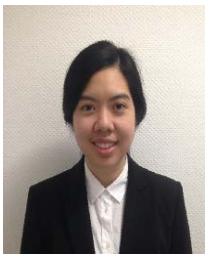
[13] L. V. Wortswinkel and W. Nijs, "Industrial Combustion Boilers," *International Energy Agency Energy Technology Systems Analysis Program (IEAETSAP) [Online]*. May 2010. Available from: http://www.ieaetsap.org/web/etechds/pdf/i01-ind_boilers-gs-ad-gct1.pdf.

[14] United States Environmental Protection Agency (US EPA). Office of Atmospheric Programs. Climate Protection Partnerships Division and Climate Change Division. "Climate Leaders Greenhouse Gas Inventory Protocol Offset Project Methodology for Project Type: Industrial Boiler Efficiency," *Climate Protection Partnerships Division and Climate Change Division. Office of Atmospheric Programs. US EPA. United States of America. [Online]*. Available from: http://www.epa.gov/climateleadership/documents/resresour/industrial_boiler_protocol.pdf

[15] Metropolitan Electricity Authority (MEA). "Electricity Tariffs for Large General Service MEA. Thailand," [Online]. Nov. 2015. Available from: http://www.mea.or.th/upload/download/file_55c49b349ab063d4bcd0d5033e699707.pdf

[16] Electricity Generating Authority of Thailand (EGAT). "Electricity Wholesale Prices for Metropolitan Electricity Authority (MEA) and Provincial Electricity Authority (PEA)". EGAT. Thailand. [Online]. www.egat.co.th

[17] Thailand Greenhouse Gas Management Organization (TGO) (Public Organization). "Summary Report: The Study of Emission Factor for an Electricity System in Thailand 2010," TGO (Public Organization). Thailand. [Online]. Available from: conference.tgo.or.th/download/tgo_or_th/ publication/GEF/2010/GEFReport_ENrevise4.pdf



Kebsiri Manusilp received the B.Eng. degree in Electrical Engineering from Khon Kaen Universiry in 2012. Currently, she is studying a master's degree in Electrical Engineering at Control System Research Laboratory (CSRL), Department of Electrical Engineering, Chulalongkorn University. Her research interest is energy management system.

David Banjerdpongchai received the B.Eng. degree from Chulalongkorn University, and the M.S. and Ph.D. degrees from Stanford University, all in Electrical Engineering. Currently, he is a professor at Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University. He is a member of ECTI Association and had served as a founding chair of Systems and Control from 2012-2015. He is a senior member of IEEE and has served as a founding chair of IEEE Control Systems Society Thailand Chapter since 2015. His research interests are energy management systems, analysis and synthesis of robust control systems, and convex optimization in robust control problems.

R Peak Detection Algorithm based on Continuous Wavelet Transform and Shannon Energy

Nantarika Thiamchoo[1] and **Pornchai Phukpattaranont**[2], Members

ABSTRACT

The R peak detection algorithm is a necessary tool for monitoring and diagnosing the cardiovascular disease. This paper presents the R peak detection algorithm based on continuous wavelet transform (CWT) and Shannon energy. We evaluate the proposed algorithm with the 48 record of ECG data from MIT-BIH arrhythmia database. Results show that the proposed algorithm gives very good DER (0.48%-0.50%) compared to those from previous publications (0.168%-0.87%). We demonstrated that the use of the CWT with a single scaling parameter is capable of removing noises. In addition, we found that Shannon energy cannot improve the DER value but it can highlight the R peak from the low QRS complex in ECG beat leading to the improvement in the robustness of the R peak detection algorithm.

Keywords: R Peak Detection Algorithm; Wavelet Transform; Shannon Energy; Electrocardiography (ECG)

1. INTRODUCTION

In order to be alive, the heart is an important organ to pump the blood containing with the oxygen and the nutrient to cells in our body. Currently, the human behaviors, such as unhealthy diet, physical inactivity, smoking and alcohol, lead to the risk of cardiovascular disease (CVD). These behaviors make the raised blood pressure, the raised blood glucose, and the raised blood lipids resulting in heart attack, stroke, or heart failure. According to the World Health Organization (WHO) report, CVDs are the first cause of death worldwide [1]. In 2012, about 17.5 million people died from CVDs. It is 31% of all global deaths. Moreover, CVDs were the second cause of death in the United Kingdom [2]. It is 27% of all deaths in 2014.

The electrocardiography (ECG) is a standard tool for monitoring and diagnosing diseases related to the CVDs for prevention and treatment. Normally, the ECG signal consists of P wave, QRS complex and T wave. The R peak is important for the heart rate calculation and is used as a preliminary step in beat segmentation algorithm. In practice, the ECG signal recording was often contaminated with a variety of noises such as baseline wandering noise, power line interference noise, muscles noise (electromyography noise), motion artefact noise, and electrode pop or contact noise. These noises cause the difficulty in detecting R peak. Then, the R peak detection algorithm that is capable of eliminating noise is very important.

For over last decades, various R peak detection algorithms were developed [3-11]. The principle of these algorithms consists of 4 main parts: signal pre-processing or noise removal, envelope detection, peak detection, and post-processing. For signal pre-processing, many techniques were applied to suppress noise. One of the well-known techniques is wavelet transform (WT). Many researches developed the R peak detection algorithm based on WT [3-4]. Several wavelet functions were studied including Harr, Gaussian, Mexican hat and Morlet wavelet functions [3]. Results show that the Mexican hat wavelet function provided the highest performance. The output from signal pre-processing part was used for calculating the envelope signal, which is used to define QRS complex duration.

Shannon energy is a successful method in detecting the envelope signal in heart sound [12]. Recently, it was applied in R peak detection algorithm [5, 6]. Results show that the good detection error rate at 0.17% [6] and 0.25% [5] tested with all ECG records from MIT-BIH arrhythmias database can be obtained.

In the peak detection processing part, the thresholding techniques are widely used [5, 7]. Normally, the thresholding techniques are divided into 2 types: the fixed thresholding and the adaptive thresholding. For fixed thresholding, the threshold value will be set to a constant value over time by experts. Although the method has a low computational complexity, it is very sensitive to noises. To improve the performance, the adaptive thresholding method was applied. The threshold value will adjust itself based on the defined mathematical model.

To increase the efficiency of the R peak detection algorithm, some researches included the post-processing part such as search back method [5-6, 9]. It is the method used for reducing the number of false

Manuscript received on September 15, 2016 ; revised on November 21, 2016.

Final manuscript received on February 21, 2017.

[1,2] The authors are with the Department of Electrical Engineering, Faculty of Engineering, Prince of Songkla University, Hat Yai, Songkhla, 90112 Thailand., E-mail: nantarika.t@gmail.com and pornchai.p@psu.ac.th

negative. When the algorithm cannot find the R peak within 150% of recent R-R interval, the threshold will be decreased to the half [5].

This paper presents the R peak detection algorithm based on continuous wavelet transform and Shannon energy. We point out the effect on using Shannon energy in our R peak detection algorithm. The rest of this paper is organized as follows. Section 2 describes the details of the proposed R peak detection algorithm and performance evaluation method. The results and discussion are presented in section 3. Finally, conclusions of this paper are given in section 4.

2. THEORY

2.1 Continuous wavelet transform

The continuous wavelet transform (CWT) was used for noise removal process. It is a method for converting a signal into another form that can be represented in time-frequency relationship. The wavelet coefficient of signal, CWT(a, b), can be expressed as

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-b}{a}\right) dt, \qquad (1)$$

where $x(t)$ is the ECG signal in this paper, $\psi(t)$ is the wavelet function, a is a scaling parameter, and b is a location parameter.

As shown in (1), the output from CWT value indicates the similarity level between the ECG signal $x(t)$ and the wavelet function. In the context of R peak detection algorithm, we need the wavelet function that has a similar shape to the pattern of QRS waveform in the ECG signal so that the QRS signal to noise ratio is maximized. Therefore, the suitable wavelet function is very important. In this work, we use the Mexican hat wavelet function, which is given by

$$\psi(t) = (1-t^2)e^{-\frac{t^2}{2}}. \qquad (2)$$

We can see from (2) that the Mexican hat wavelet function is the second derivative of a Gaussian function.

2.2 Shannon energy

Shannon energy is one of methods used to generate an envelope of the heart sound for estimating systole and diastole periods [12]. Recently, this method was applied in the field of R peak detection. The Shannon energy, $S[n]$, is formulated by

$$S[n] = -N[n]^2 \log_2 N[n]^2 \qquad (3)$$

where $N[n]$ is the normalized ECG signal after noise removal in this paper. The Shannon energy has the better capability in emphasizing the low and medium R peak amplitudes compared to the conventional square energy operation. As a result, the R-peak signal after processing with the Shannon energy does not have significant differences in terms of amplitude. This makes the use of a single thresholding technique in R peak detection algorithm possible.

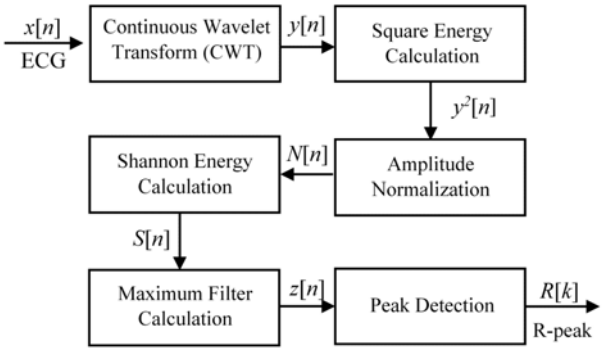


Fig.1: *The block diagram of proposed R peak detection algorithm.*

3. MATERIALS AND METHODS

3.1 Proposed algorithm

Fig. 1 shows a block diagram of the proposed algorithm. The original ECG signal, $x[n]$, was processed with CTW for noise removal. Then, the output from CWT, $y[n]$, was calculated using a square energy operation as given by

$$y^2[n] = y[n] \times y[n]. \qquad (4)$$

Subsequently, the output from the square energy operation was normalized with the maximum amplitude of $y^2[n]$, which can be expressed as

$$N[n] = \frac{y^2[n]}{\max(y^2[n])} \qquad (5)$$

Subsequently, we computed the Shannon energy of signal $N[n]$ using (3). Next step, we determine the envelope signal $z[n]$ for defining the QRS time duration used for R peak detection by the maximum filter, which is given by

$$z[n] = \max_{k\in[n-L+1,n]} S[k], \qquad (6)$$

where L is the length of maximum filter. It was varied from 50 ms to 400 ms in this paper.

In order to detect R peak, we use a single fixed thresholding method, which is given by

$$Thv = \lambda \times \max(z[n]), \qquad (7)$$

where λ is a constant. The threshold value Thv determined using λ will separate the R peak from noise. Too low threshold value results in detecting noise as

the R peak. On the other hand, too high threshold value lead to missing R peak detection. In this paper, λ was empirically varied to achieve minimum error detection rate. The QRS durations where the R peak signals locate are defined when the output signal from the maximum filter $z[n]$ is greater than the Thv value. Then, we compute the R peak location at the maximum value in $y^2[n]$.

3.2 ECG data and performance evaluation

For evaluating the performance of the proposed algorithm, the ECG signal from MIT-BIH arrhythmia databases [10] was analyzed. It consisted of 2 channels of ECG signals acquired at a sampling frequency of 360 Hz for 30 minutes. Moreover, the annotation from an expert was given. To demonstrate the performance of the proposed algorithm, we used three statistical values: the sensitivity (SEN), the positive predictive rate (PPR), and the detection error rate (DER). These values can be computed as follows:

$$\mathrm{SEN} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100\%, \tag{8}$$

$$\mathrm{PPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \times 100\%, \tag{9}$$

$$\mathrm{DER} = \frac{\mathrm{FN} + \mathrm{FP}}{\mathrm{TP} + \mathrm{FN}} \times 100\%, \tag{10}$$

where true-positive (TP) is the number of correct R peaks detected by the algorithm, false-negative (FN) is the number of missing R peaks detected, and false-positive (FP) is the number of incorrect R peaks detected by the algorithm.

4. RESULTS AND DISCUSSION

4.1 Parameter optimization

To analyze the capability of noise suppression using CWT and Shannon energy, we vary a scaling parameter a between 2 and 4 with a step size of 0.1. In addition, the maximum filter length (L) was varied between 50 and 400 ms with a step size of 25 ms. We found that the scaling parameter a 2.5 and the maximum filter length $L = 275$ ms provide the best DER value.

To investigate the effect of Shannon energy on the proposed R peak algorithm, the results from the proposed algorithm with and without Shannon energy operation are shown. Fig. 2 shows the detection error rate from proposed algorithm without Shannon energy calculation at the length $L = 275$ ms as a function of Thv between 1% and 7% of maximum signal amplitude. The minimum DER value 0.48% is obtained when the Thv is 3% of maximum signal amplitude

Fig. 3 shows the detection error rate from proposed algorithm with Shannon energy calculation at the length $L = 275$ ms as a function of Thv between

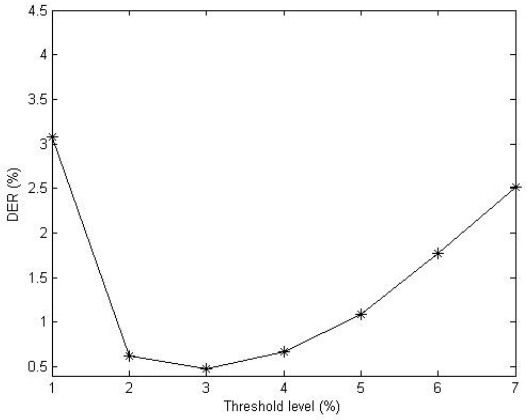


Fig.2: *The detection error rate from proposed algorithm without Shannon energy calculation at the length L = 275 ms as a function of Thv.*

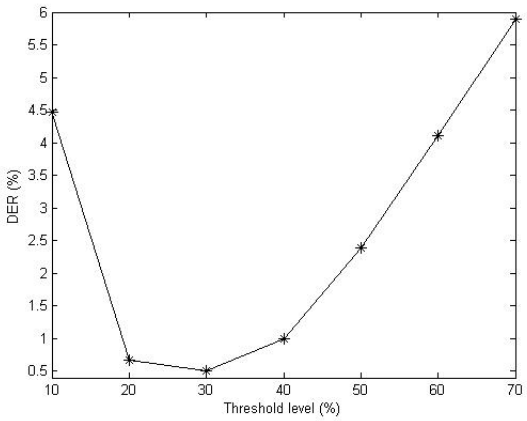


Fig.3: *The detection error rate from proposed algorithm with Shannon energy calculation at the length L = 275 ms as a function of Thv.*

10% and 70% of maximum signal amplitude. Results show that the minimum DER value 0.50% is obtained when the Thv is 30%.

Although the proposed algorithm without Shannon energy can provide lower DER value than the proposed algorithm with Shannon energy calculation, the range of its threshold value is narrower. In other words, while the range of the Thv that provides the DER value lower than 1.5% is between 2% and 5% for the proposed algorithm without Shannon energy, the wider range of those from the proposed algorithm with Shannon energy is obtained between 20% and 40%. This is very important because it allows for more flexibility when the single fixed thresholding technique in the R peak detection algorithm is used.

4.2 Signal characteristics

To compare the advantage of Shannon energy operation, example results from ECG data record num-

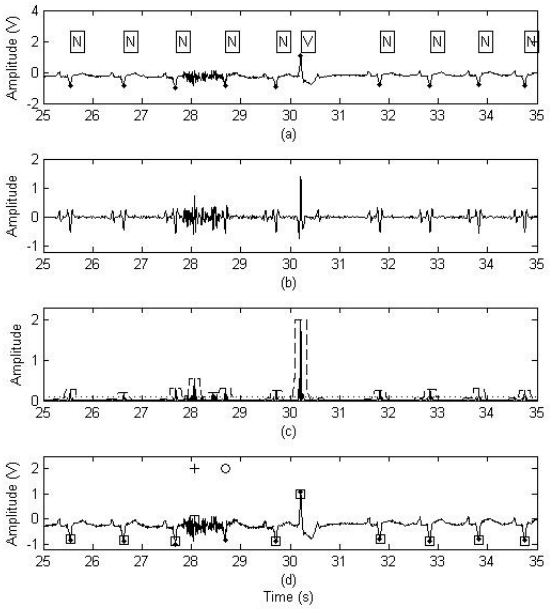


Fig.4: *The result of proposed algorithm without Shannon energy calculation applied on ECG signal record 108.*

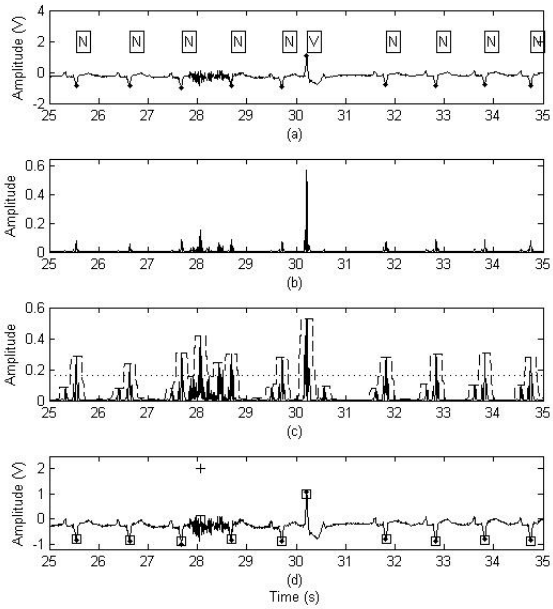


Fig.5: *The result of proposed algorithm with Shannon energy calculation applied on ECG signal record 108.*

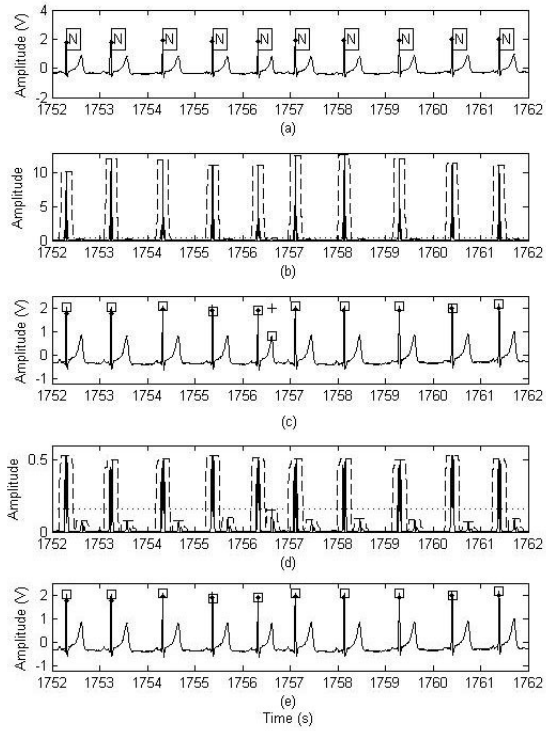


Fig.6: *The result of proposed algorithm without and with Shannon energy calculation applied on ECG signal record 113.*

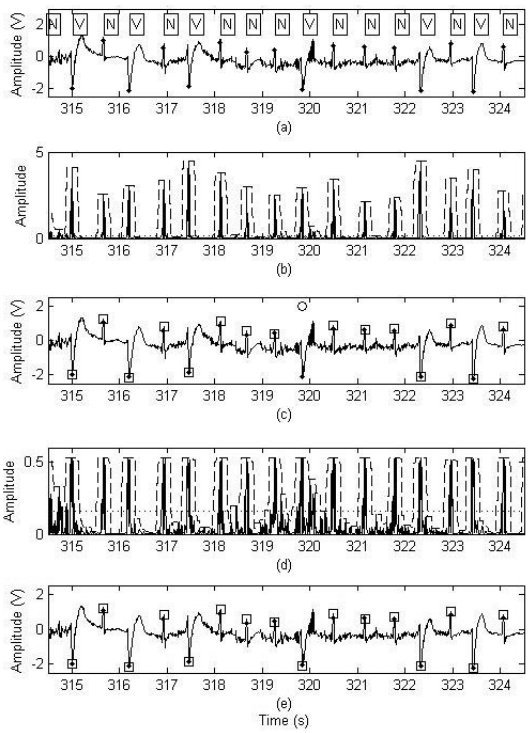


Fig.7: *The result of proposed algorithm without and with Shannon energy calculation applied on ECG signal record 200.*

ber 108, 113 and 200 are demonstrated. Fig. 4 shows the signal characteristics from the proposed algorithm without Shannon energy calculation on ECG signal record 108. Fig. 4(a) shows the original ECG signal before noise filtering overlaid by the dot markers from the expert from time 25 s to 35 s ($x[n]$). "N" stands for a normal heart beat and "V" for a premature ventricular contraction beat. The ECG signal in this record consists of noise interference, normal beats and a premature ventricular contraction (PVC) beat. Fig. 4(b) shows the ECG signal after noise removal by CWT ($y[n]$). As the result, the R peaks on signal $y[n]$ is smoother and clearer. Fig. 4(c) shows the signal from square energy calculation $y^2[n]$ in solid line, the envelope signal $z[n]$ in dashed line, and the threshold level Thv 3% in dotted line. Results show a significant difference in amplitudes of the PVC beat compared to the normal beats. Fig. 4(d) shows the original ECG signal $x[n]$ overlaid by the dot markers from the expert and square markers from the proposed algorithm. The original ECG signal $x[n]$ overlaid by the dot markers from the expert and square markers from the proposed algorithm. The positive marker ("+") and the circle marker ("o") between time 28 s to 29 s present an incorrect R peak detected and a missing R peak detected by the proposed algorithm due to the noise, respectively.

Fig. 5 shows the signal characteristics from the proposed algorithm with Shannon energy calculation on record 108. Fig. 5(a) shows the ECG signal ($x[n]$). Fig. 5(b) shows the signal from the normalized square energy calculation $N[n]$. Fig. 5(c) shows the signal from the Shannon energy calculation $S[n]$ in solid line, the envelope signal $z[n]$ in dashed line, and the threshold level 30% in dotted line. Results show that a significant difference in amplitudes of the PVC beat compared to the normal beats form the normalized square energy signal as shown in Fig. 5(b) is reduced. Fig. 5(d) shows R peaks detected by the proposed algorithm with Shannon energy. Results show that the false negative shown in Fig. 4(d) is removed.

Fig. 6 shows the signal characteristics from the proposed algorithm without and with Shannon energy calculation on ECG signal record 113. Fig. 6(a) shows the original ECG signal from time 1752 s to 1762 s ($x[n]$). The ECG signal in this record consists of high T wave amplitude in the normal heart beats resulting in the difficulty to correctly detect R peak. Fig. 6(b) shows the signal from square energy calculation $y^2[n]$ from the proposed algorithm without Shannon energy calculation in solid line, the envelope signal $z[n]$ in dashed line, and the threshold level 3% in dotted line. As the result, the R peaks on signal $y^2[n]$ is highlighted. Fig. 6(c) shows R peaks detected by the proposed algorithm without Shannon energy. Results show a false positive between time 1756 ms and 1757 ms because the threshold value is lower than the amplitude of T wave. Fig. 6(d) shows the signal from the Shannon energy calculation S[n] in solid line, the envelope signal $z[n]$ in dashed line, and the threshold level 30% in dotted line. Fig. 6(e) shows all R peaks that are correctly detected by the proposed algorithm with Shannon energy. Results show that Shannon energy calculation allows for the high threshold value resulting in the removal of the false positive due to the high T wave amplitude.

Fig. 7 shows the example result of proposed algorithm without and with Shannon energy calculation applied on ECG signal record 200. Fig. 7(a) shows the original ECG signal from time 314.5 s to 324.5 s, which is contaminated with powerline interference noise. Fig. 7(b) shows the unstable R peak amplitude that was highlighted with the square energy calculation and the threshold value 3% of maximum amplitude in dotted line. Fig. 7(c) shows R peaks detected by the proposed algorithm without Shannon energy. It consists of a false negative beat detection from a PVC beat at around 320 ms due to the use of very low threshold value resulting in the merging of two connected R peak. Fig. 7(d) shows result from Shannon energy calculation. Results show that Shannon energy operation can enhance R peak very well resulting in equal R peak amplitude. Fig. 7(e) shows all R peaks that are correctly detected by the proposed algorithm with Shannon energy. Results show that Shannon energy calculation allows for the high threshold value resulting in the removal of the false positive due to the powerline interference noise.

4.3 Performance evaluation

Table 1 shows the performance of the proposed algorithm without Shannon energy operation on 48 ECG data from MIT-BIH arrhythmia databases. Results show that the proposed algorithm can achieve detection error rate 0.48%, sensitivity 99.69% and positive predictive rate 99.83%. From the total 541 false detections, while the false negative occurs 356 times, the false positive occurs 185 times. Record 203 provide the maximum DER value of 2.89% from the 14 FP values and the 72 FN values because the ECG data from record 203 consists of various noises such as muscle artefact and baseline shift. Moreover, the ECG beat in this record contains a variety of patterns such as suddenly heart rate changing, variant R peak amplitude, and multiform of PVC beats, leading to the difficulty in detecting R peaks.

Table 2 shows the performance of the proposed algorithm with Shannon energy operation. Results show that the proposed algorithm can achieve DER value of 0.50%, sensitivity 99.66% and positive predictive rate 99.83%. From the total 562 false detections, while the false negative occurs 382 times, the false positive occurs 180 times. ECG data record 203 provides the worst DER value of 2.89%, which is the same as the proposed algorithm without Shannon energy. The DER significantly increases in the ECG

Table 1: *Performance Evaluation of the Proposed Algorithm without Shannon Energy Operation.*

Record No.	Total (beat)	TP (beat)	FN (beat)	FP (beat)	SEN (%)	PPR (%)	DER (%)
100	2273	2273	0	0	100.00	100.00	0.00
101	1865	1865	0	4	100.00	99.79	0.21
102	2187	2187	0	0	100.00	100.00	0.00
103	2084	2084	0	0	100.00	100.00	0.00
104	2228	2215	13	15	99.42	99.33	1.26
105	2572	2558	14	22	99.46	99.15	1.40
106	2027	1993	34	2	98.32	99.90	1.78
107	2136	2136	0	0	100.00	100.00	0.00
108	1763	1744	19	12	98.92	99.32	1.76
109	2532	2528	4	0	99.84	100.00	0.16
111	2124	2123	1	1	99.95	99.95	0.09
112	2539	2537	2	10	99.92	99.61	0.47
113	1795	1795	0	8	100.00	99.56	0.45
114	1879	1856	23	1	98.78	99.95	1.28
115	1953	1953	0	0	100.00	100.00	0.00
116	2412	2391	21	2	99.13	99.92	0.95
117	1535	1535	0	2	100.00	99.87	0.13
118	2278	2277	1	3	99.96	99.87	0.18
119	1987	1987	0	1	100.00	99.95	0.05
121	1863	1862	1	2	99.95	99.89	0.16
122	2476	2476	0	0	100.00	100.00	0.00
123	1518	1518	0	2	100.00	99.87	0.13
124	1619	1619	0	1	100.00	99.94	0.06
200	2601	2593	8	16	99.69	99.39	0.92
201	1963	1952	11	0	99.44	100.00	0.56
202	2136	2133	3	0	99.86	100.00	0.14
203	2980	2908	72	14	97.58	99.52	2.89
205	2656	2637	19	0	99.28	100.00	0.72
207	1860	1843	17	16	99.09	99.14	1.77
208	2955	2930	25	4	99.15	99.86	0.98
209	3005	3004	1	2	99.97	99.93	0.10
210	2650	2623	27	3	98.98	99.89	1.13
212	2748	2748	0	1	100.00	99.96	0.04
213	3251	3242	9	8	99.72	99.75	0.52
214	2262	2258	4	3	99.82	99.87	0.31
215	3363	3355	8	0	99.76	100.00	0.24
217	2208	2207	1	2	99.95	99.91	0.14
219	2154	2154	0	1	100.00	99.95	0.05
220	2047	2047	0	0	100.00	100.00	0.00
221	2427	2425	2	0	99.92	100.00	0.08
222	2483	2481	2	1	99.92	99.96	0.12
223	2605	2604	1	1	99.96	99.96	0.08
228	2053	2044	9	19	99.56	99.08	1.36
230	2256	2256	0	2	100.00	99.91	0.09
231	1571	1571	0	0	100.00	100.00	0.00
232	1780	1780	0	4	100.00	99.78	0.22
233	3079	3075	4	0	99.87	100.00	0.13
234	2753	2753	0	0	100.00	100.00	0.00
Total	**109491**	**109135**	**356**	**185**	**99.69**	**99.83**	**0.48**

data record 114 because ECG data in this record consist of many R peak amplitude variations. Although, Shannon energy operation can enhance the R peak amplitude, the threshold value 30% of maximum amplitude is too high for this record resulting in many FNs

4.4 Performance comparisons

Table 3 shows the performance comparison of the proposed algorithm with those from other papers using the ECG signal from MIT-BIH arrhythmia databases. The first six articles is sorted from the minimum DER value (0.168%, linear filtering) to the maximum DER value (0.87%, Median filter). The DER from the proposed algorithm without and with Shannon energy is 0.48% and 0.50%, respectively. We can see that the DER from the proposed algorithm without additional post processing is better than that from the algorithm given in [9], which also uses CWT for noise removal. Most errors in the proposed algorithm are caused by false negative detections. It can solved by introducing some post processing methods such as searchback method using R-R interval information and R-R interval check-up as presented in [5] and [9], respectively. We found that the false positive from the proposed algorithm is very good compared to other publications except for the algorithm given in [6] and [5].

In order to confirm the performance of our proposed algorithm, the five ECG signal records that frequently gives high DER values in other papers,

Table 2: *Performance Evaluation of the Proposed Algorithm with Shannon Energy Operation.*

Record No.	Total (beat)	TP (beat)	FN (beat)	FP (beat)	SEN (%)	PPR (%)	DER (%)
100	2273	2273	0	0	100.00	100.00	0.00
101	1865	1865	0	4	100.00	99.79	0.21
102	2187	2187	0	0	100.00	100.00	0.00
103	2084	2084	0	0	100.00	100.00	0.00
104	2228	2215	13	13	99.42	99.42	1.17
105	2572	2558	14	23	99.46	99.11	1.44
106	2027	1991	36	2	98.22	99.90	1.87
107	2136	2136	0	0	100.00	100.00	0.00
108	1763	1741	22	10	98.75	99.43	1.82
109	2532	2524	8	0	99.68	100.00	0.32
111	2124	2123	1	2	99.95	99.91	0.14
112	2539	2538	1	11	99.96	99.57	0.47
113	1795	1795	0	4	100.00	99.78	0.22
114	*1879*	*1841*	*38*	*1*	*97.98*	*99.95*	*2.08*
115	1953	1953	0	0	100.00	100.00	0.00
116	2412	2391	21	2	99.13	99.92	0.95
117	1535	1535	0	1	100.00	99.93	0.07
118	2278	2277	1	3	99.96	99.87	0.18
119	1987	1987	0	1	100.00	99.95	0.05
121	1863	1862	1	2	99.95	99.89	0.16
122	2476	2476	0	0	100.00	100.00	0.00
123	1518	1518	0	2	100.00	99.87	0.13
124	1619	1619	0	1	100.00	99.94	0.06
200	2601	2595	6	17	99.77	99.35	0.88
201	1963	1948	15	0	99.24	100.00	0.76
202	2136	2133	3	0	99.86	100.00	0.14
203	2980	2908	72	14	97.58	99.52	2.89
205	2656	2638	18	0	99.32	100.00	0.68
207	1860	1845	15	16	99.19	99.14	1.67
208	2955	2930	25	4	99.15	99.86	0.98
209	3005	3004	1	4	99.97	99.87	0.17
210	2650	2621	29	3	98.91	99.89	1.21
212	2748	2748	0	1	100.00	99.96	0.04
213	3251	3242	9	8	99.72	99.75	0.52
214	2262	2257	5	3	99.78	99.87	0.35
215	3363	3356	7	0	99.79	100.00	0.21
217	2208	2206	2	1	99.91	99.95	0.14
219	2154	2154	0	1	100.00	99.95	0.05
220	2047	2047	0	0	100.00	100.00	0.00
221	2427	2424	3	0	99.88	100.00	0.12
222	2483	2482	1	1	99.96	99.96	0.08
223	2605	2604	1	1	99.96	99.96	0.08
228	2053	2044	9	19	99.56	99.08	1.36
230	2256	2256	0	2	100.00	99.91	0.09
231	1571	1571	0	0	100.00	100.00	0.00
232	1780	1780	0	3	100.00	99.83	0.17
233	3079	3074	5	0	99.84	100.00	0.16
234	2753	2753	0	0	100.00	100.00	0.00
Total	**109491**	**109109**	**382**	**180**	**99.66**	**99.83**	**0.50**

i.e., 108, 203, 105, 223 and 207, are compared with the proposed algorithm as shown in Table 4. Results show that the proposed algorithm provides a good DER for record 108 that is one of the most difficult records to be correctly detected. The algorithm in [4] uses S-Transform (ST) for noise filtering and has the post processing operation in its R peak detection algorithm. ST is a time-frequency distribution and the output of ST is a matrix where rows represent the time and columns represent the frequency. Compared with the CWT with a single scaling parameter a used in this paper, ST requires higher computational complexity but provides worse DER value. This shows the capability of using CWT with a single scaling parameter in removing noise.

5. CONCLUSIONS

The R peak detection algorithm based on CWT and Shannon energy is presented. The proposed algorithm was evaluated with the ECG data from MIT-BIH arrhythmia database. Results show that the proposed algorithm give very good DER (0.48%-0.50%) compared to those from previous publications (0.168%-0.87%). We demonstrate that the use of the CWT with a single scaling parameter is capable of removing noises. In addition, we found that Shannon energy cannot improve the DER value but it can highlight the R peak from the low QRS complex in ECG beat after noise removal and square operation such as a PVC beat. This main advantage results in achieving the robust R peak detection algorithm. For

Table 3: *Performance Comparisons of the Proposed Algorithm with other Papers.*

Method of noise removal	Evaluation						Post processing
	TP (beat)	FN (beat)	FP (beat)	SEN (%)	PPR (%)	DER (%)	
Linear filtering [6]	109401	93	91	99.92	99.92	0.168	Yes
S-Transform [5]	108323	97	171	99.84	99.91	0.25	Yes
Whitening filter [11]	109374	109	210	99.82	99.91	0.29	No
Quadratic filter [7]	109281	202	210	99.82	99.81	0.38	No
CWT [9]	109118	376	218	99.66	99.8	0.54	Yes
Median filter [10]	108099	495	462	99.58	99.55	0.87	Yes
CWT without Shannon energy	109135	356	185	99.69	99.83	0.48	No
CWT with Shannon energy	109158	382	180	99.66	99.83	0.50	No

Table 4: *Performance Comparisons of the Proposed Algorithm with other Papers for First 5 Records that Often Give High DER.*

Method of noise removal	DER (%)				
	108	203	105	228	207
Linear filtering [6]	0.57	1.19	1.25	0.63	0.32
S-Transform [5]	2.44	0.67	1.24	0.54	1.13
Whitening filter [11]	0.51	1.93	0.23	1.03	0.65
Quadratic filter [7]	4.08	1.91	1.52	1.27	2.20
CWT [9]	4.71	2.05	2.02	3.56	1.18
Median filter [10]	2.78	4.33	4.67	1.12	0.81
CWT without Shannon energy	1.76	2.89	1.40	1.36	1.77
CWT with Shannon energy	1.82	2.89	1.44	1.36	1.67

future work, we will improve the proposed algorithm by adding the post processing operation to reduce the false negative and false positive to achieve higher performance.

ACKNOWLEDGMENT

This work is supported by Prince of Songkla University Graduate Studies Grant as well as by Department of Electrical Engineering, Faculty of Engineering, Prince of Songkla University.

References

[1] W. H. Organization. (2013, May). Cardiovascular diseases [Online]. Available: http:/ /www.who.int/cardiovascular_diseases/en/J.

[2] N. Townsend, P. Bhatnagar, E. Wilkins, K. Wickramasinghe and M. Rayner, "Mortality," in *Cardiovascular disease statistics 2015*, London, UK: British Heart Foundation, 2015, ch. 1, pp. 12-59.

[3] M. J. Rooijakkers, C. Rabotti, S. G. Oei, and M. Mischi, "Low-complexity R-peak detection for ambulatory fetal monitoring', Physiological Measurement, vol. 33, no. 7, pp. 1135-1150, Jul. 2012.

[4] B. Abibullaev and H. D. Seo, "A new QRS detection method using wavelets and artificial neural networks," *Journal of Medical Systems*, vol. 35, no. 4, pp. 683-691, Aug. 2011.

[5] Z. Zidelmal, A. Amirou, D. Ould-Abdeslam, A. Moukadem, and A. Dieterlen, "QRS detection using S-Transform and Shannon energy," *Computer Methods and Programs in Biomedicine*, vol. 116, no. 1, pp. 1-9, 2014.

[6] H. Zhu and J. Dong, "An R-peak detection method based on peaks of Shannon energy envelope," *Biomedical Signal Processing and Control*, vol. 8, no. 5, pp. 466-474, 2013.

[7] P. Phukpattaranont, "QRS detection algorithm based on the quadratic filter," *Expert Systems with Applications*, vol. 42, no. 11, pp. 4867-4877, 2015.g

[8] M. Bahoura, M. Hassani, and M. Hubin, "DSP implementation of wavelet transform for real time ECG wave forms detection and heart rate analysis," *Computer Methods and Programs in Biomedicine*, vol. 52, no. 1, pp. 35-44, 1997.

[9] S. Choi, M. Adnane, G.-J. Lee, H. Jang, Z. Jiang, and H.-K. Park, "Development of ECG beat segmentation method by combining lowpass filter and irregular R-R interval checkup strategy," *Expert Systems with Applications*, vol. 37, no. 7, pp. 5208-5218, 2010.

[10] G. Nallathambi and J. C. Principe, "Integrate

and fire pulse train automaton for QRS detection," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 2, pp. 317-326, 2014.

[11] K. Arbateni and A. Bennia, "Sigmoidal radial basis function ANN for QRS complex detection," *Neurocomputing*, vol. 145, pp. 438-450, Dec. 2014.

[12] S. Choi and Z. Jiang, "Comparison of envelope extraction algorithms for cardiac sound signal segmentation," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1056-1069, 2008.

Nantarika Thiamchoo received the B. Eng. degree with First Class Honors in Biomedical Engineering and M. Eng. degree in Electrical Engineering from Prince of Songkla University in 2014 and 2015, respectively. She is currently working toward the Ph.D. degree in Electrical Engineering in Prince of Songkla University. Her research interests are biomedical signal processing, rehabilitation system and myoelectric control.

Pornchai Phukpattaranont received the B. Eng. with First Class Honors and M. Eng. degrees in Electrical Engineering from Prince of Songkla University in 1993 and 1997, respectively, the Ph.D. degree in Electrical and Computer Engineering from the University of Minnesota, in 2004. He is currently an associate professor of Electrical Engineering at Prince of Songkla University. His research interests are signal and image analysis for medical applications and ultrasound signal processing. Examples of his ongoing research include the pattern recognition system based on Electromyographic (EMG) signal, Electrocardiographic (ECG) signal, and microscopic images of breast cancer cells. Dr. Phukpattaranont is a member of the IEEE, ECTI Association, and ThaiBME.

A Hybrid CKF-NNPID Controller for MIMO Nonlinear Control System

Adna Sento[1] and **Yuttana Kitjaidure**[2], Non-members

ABSTRACT

This paper presents a detailed study to demonstrate the online tuning dynamic neural network PID controller to improve a joint angle position output performance of 4- joint robotic arm. The proposed controller uses a new updating weight rule model of the neural network architecture using multi-loop calculation of the fusion of the gradient algorithm with the cubature Kalman filter (CKF) which can optimize the internal predicted state of the updated weights to improve the proposed controller performances, called a Hybrid CKF-NNPID controller. To evaluate the proposed controller performances, the demonstration by the Matlab simulation program is used to implement the proposed controller that connects to the 4-joint robotic arm system. In the experimental result, it shows that the proposed controller is a superior control method comparing with the other prior controllers even though the system is under the loading criteria, the proposed controller still potentially tracks the error and gives the best performances.

Keywords: Robotic Arm Control System, Neural Network PID Controller, Cubature Kalman Filters, Adaptive Learning Algorithm

1. INTRODUCTION

Nowadays, the conventional PID controller might be replaced with intelligent controller in case of the uncertain parameters of the system models and complexity of the mathematical models and especially in the nonlinear systems with disturbances. In addition, human body interface controlling the robot is recently popular which can apply in the impaired human devices such as robot arm-assisted human impaired, robot-assisted human movement, robotic therapy on the recovery in chronic stroke, etc. Consequently, these have motivated the researchers to develop the system of control, particularly in the improvement of the part planning which may use either of the inverse/forward kinematics or computational intelligence. However, these methods still require the precision of the errors of the robot angle movements. Furthermore, the best responses of the robot movement will make it reached the robot target effectively. As a result, a large number of the intelligent control techniques have been devised to improve the controller for sophisticated angle precision control problems in the fusion of dynamic neural network with the classical control such a conventional PID controller, also known as neural PID controller (NNPID controller). For example, S. Cong and Y. Liang [1] proposed the neural PID controller to improve the performance of the multi-variable control system that used a resilient back propagation neural network algorithm for updating weight rule. In 2010, Ho Phanm Huy Anh [2] provided the neural PID controller which used the back propagation algorithm to update the weight of the neural network. Wen Yu and R. Jacob [3] also proposed the Neural PID controller to the upper limb exoskeleton robot. Furthermore, Vikas Kumar also proposed the neural PID controller to control the position of the permanent magnet synchronous motor (PMSM) [4]. Recently, F. G. Rossomando and C. M. Soria proposed the neural PID controller to control the mobile robot [5]. They used the gradient algorithm to update the weight of the controller. In addition, there are other examples of the fusion of the computational intelligence with the conventional PID controller that were established such as [6], [7], and [8].

The basic idea of the conventional PID is usually known by three-term of the controller as given by the general discrete function, $u(tk)$, as follows:

$$u(t_k) = K_p e(t_k) + K_l \sum_{i=1}^{k} e(t_i)\Delta t + K_D \frac{e(t_k)-e(t_{k-1})}{\Delta t} \quad (1)$$

where K_P, K_I, and K_D are called the controller gains which are proportional gain, integral gain, and derivative gain respectively, $e(t_k)$ is a system error at the sample t_k, and Δ_t is a sampling time. In prior studies, the controller cannot effectively operate unless it must be set by specific tuning method such as conventional PID technique, LQR technique [1], trial-error technique [2], least mean square method (LMS) [3] and training technique [17]. In other words, the controllers insufficiently use for the nonlinear dynamic system unless the initial conditions are pre-set which are nearly the system stability. Furthermore, neural PID controller uses gradient algorithm to up-

Manuscript received on October 4, 2016 ; revised on November 24, 2016.

Final manuscript received on February 21, 2017.

[1,2] The authors are with the Department of Electronics Engineering, Faculty of Engineering, King Mongkuts Institute of Technology Ladkrabang, E-mail: adna@tni.ac.th and kkyuttan@kmitl.ac.th

[1] Corresponding author.

date the weights which has low speed of convergence, and is not able to tune weights for improving the error signal, the steady-state error, and the transient response of the system. As a result, we summarized that the control system which used NNPID controller can improve the performance in two objectives: 1) the initial value of the controller, and 2) the weight updating rule. Fortunately, as the growing of the computational intelligent algorithm in a last decade ago, many researchers have proposed the improvement the convergent speed using the filter model especially in the Kalman filter model which already improved the convergent speed of the neural network such as [9], [10], and [11], known as EKF algorithm. In the literature, the algorithm must form the status of the weights of the neural network in term of the dynamic model which can be given as follows [11]:

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k + \omega_k \tag{2}$$

$$\boldsymbol{y}_k = \boldsymbol{h}_k(\boldsymbol{w}_k, \boldsymbol{e}_k, \boldsymbol{b}_k) + v_k \tag{3}$$

where $\boldsymbol{w}_k$ is a state vector of weights of the neural network, $\boldsymbol{h}(\cdot)$ is a measurement model function, $\boldsymbol{e}_k$ and bk denote the input and the bias of the network, respectively, ω_k and v_k are the process noise and the measurement noise, respectively. The variable k denotes the computational iterations. The first and second equations are known as the system equation and the measurement equation, respectively. Although, EKF algorithm helps the neural network to learn the system but itself is based on linear model thus the controller may cause a problem in the linearization process. Consequently, many researchers still develop the Kalman filter model, for example, Wan et al. [12], Xiaoyu Wang and Yong Huang [13], Huizhong Yang et al. [14] to support the highly nonlinear system [11]. One of the most famous algorithms is namely cubature Kalman filter (CKF) [15]. As a result, this algorithm is widely used in the computational intelligent learning approach of the real-world applications such as the neural network time series prediction applications [18, 19, 20], the neural network fitting applications [21], and especially in the neural network control system applications [22, 23]. Moreover, this algorithm is also proved in term of the system stability [24, 25].

As aforementioned, the algorithm has several advantages such as designing as nonlinear controller, providing derivative free method, and providing an efficient solution for high-dimensional problem. Therefore, we construct the new model of the NNPID controller using the CKF algorithm, gradient algorithm, and the multi-loop tuning, called Hybrid CKF-NNPID controller to improve the output response and the initial values. The rest of this paper is organized as follows. A proposed controller design is first described in section II. Then, in section III, the experiment setup and the comparison of the control technique performances between the proposed controller and the other controllers are discussed. Finally, the conclusions are given in section IV.

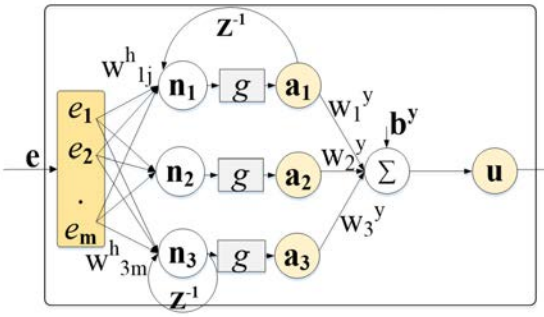


Fig.1: *Neural network controller based on PID structure.*

2. THE PROPOSED CONTROLLER DESIGN

The basic idea of the proposed controller is based on the fusion of the conventional PID controller with computational intelligent technique. The hybrid controller will make the controller increase the performances such as eliminate the restriction of the initial values setting, reduce the sensitivity of the weight update status and increase the speed of the convergence. The detail of the proposed controller will be described as follows.

2.1 The Proposed Controller Structure

The proposed controller structure comprises three layers, input layer, hidden layer, and output layer as shown in Fig. 1. In other words, the input layer receives the system error and then sends it out to the second layer which has three nodes. Each node in this layer behaves like the PID controller. Consequently, we define the first node as the integral node $\boldsymbol{a}_1$ that represents the integral gain $\mathbf{K}_\mathrm{I}$. In the second node, the $\boldsymbol{a}_2$ represents the propositional gain $\mathbf{K}_\mathrm{P}$, and the last node, the derivative node, $\boldsymbol{a}_3$ represents the derivative gain $\mathbf{K}_\mathrm{D}$ which can be expressed as follows:

$$a_{1,k} = g\left(\sum_{j=1}^{m} w_{j,k}^{h} e_{j,k} + a_{1,k-1}\Delta t\right) \tag{4}$$

$$a_{2,k} = g\left(\sum_{j=1}^{m} w_{2j,k}^{h} e_{j,k}\right) \tag{5}$$

$$a_{3,k} = g\left(\left(\sum_{j=1}^{m} w_{3j,k}^{h} e_{j,k} - \sum_{j=1}^{m} w_{3j,k-1}^{h} e_{j,k-1}\right)\frac{1}{\Delta t}\right) \tag{6}$$

where $\boldsymbol{m}$ is the number of the input nodes, $g(\cdot)$ is the nonlinear activation function that is set to tanh function. w_{ij}^{h} is the weight linked j^{th} input neuron with i^{th} hidden neuron, e_j is the j_{th} input of the neuron, and t is sampling time. The output layer is the summation of the hidden node, u_k can be expressed as

$$u_k = \sum_{i=1}^{3} w_{i,k}^{y} a_{i,k} + b^{y} \tag{7}$$

where w_i^y is the weight that links i^{th} hidden neuron with the output neuron, b^y is bias in the output layer set to zero.

2.2 The Proposed Weight Update Rule

The problem of the research is to iteratively filter the weights of the neural network PID controller which depend on the errors of the plant. Since the errors are considered as the nonlinear system signals with addition Gaussian noise, so the non-linear Kalman filter is appropriate tool for this, known as cubature Kalman filter (CKF). Therefore, in this paper, we propose the cubature Kalman filter (CKF) to update the weights of neural network PID controller, called Hybrid CKF-NNPID controller. The overall architecture of the Hybrid CKF-NNPID controller is shown in Fig. 2. It comprises two-loop operations, an external loop operation and an internal loop operation. The external loop operation is a neural network controller based on PID structure producing the control input for the plant which is subscripted by k. The other loop, internal loop, is superscripted by i that is denoted by dash line of the rectangle. This loop provides the weight update rule of the proposed controller which can improve initial values of the controller. In addition, we have redefined the system equation of the Kalman model to increase the controller performances but the measurement equation is still the same. Both of the equations are expressed as

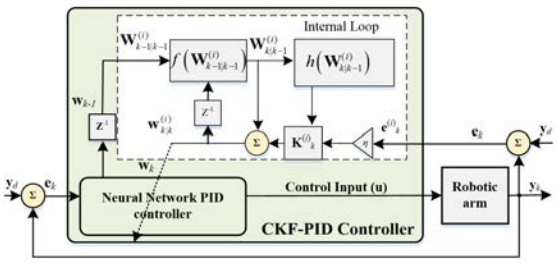


Fig.2: *The block diagram of the MIMO control system using a proposed controller.*

$$\boldsymbol{w}_{k+1} = \mathbf{f}(\boldsymbol{w}_k) + \omega_k \tag{8}$$

$$\boldsymbol{y}_k = \boldsymbol{h}_k(\boldsymbol{w}_k, \boldsymbol{e}_k, \boldsymbol{b}_k, u_k) + v_k \tag{9}$$

where $\boldsymbol{w}_k$ and $\boldsymbol{y}_k$ are a state vector and observation data, respectively, $\boldsymbol{e}_k$, $\boldsymbol{b}_k$ and u_k denote the input, the bias of the network and control input, respectively, ω_k and v_k are the process noise and the measurement noise, respectively. $\mathbf{h}(\cdot)$ is a measurement model function, and $\mathbf{f}(\cdot)$ is a process model function, known as an updated weights function, which is obtained by the gradient descent algorithm to update the internal weights. Each of the internal loop iterative operations, the updating predicted weights will be optimized by calculation the Kalman gain ($\mathbf{K}$). The algorithm will calculate until reaching the internal criteria. Then, the calculation will exit from the internal loop to the external loop and these weights are the new updated weights of the control input calculations. The external loop calculation will continually update the weights until gaining the criteria. The details of the proposed controller will be discussed as follows. The first process of the proposed controller is the formation of all of the parameters. From Fig. 1, lets $\boldsymbol{w}^y$ and $\boldsymbol{w}^h$ be a state of the weighted network in the output and the hidden layer, respectively that are expressed as

$$\boldsymbol{w}^h = \begin{bmatrix} w_{11}^h & \cdots & w_{1j}^h & b_1^h \\ w_{21}^h & \cdots & w_{2j}^h & b_2^h \\ w_{31}^h & \cdots & w_{3j}^h & b_3^h \end{bmatrix} \tag{10}$$

$$\boldsymbol{w}^y = \begin{bmatrix} w_1^y & w_2^y & w_3^y & b^y \end{bmatrix}^T \tag{11}$$

Therefore, the state vector of the Kalman model is given by the weighted network represented as

$$\boldsymbol{w}_k = \begin{bmatrix} w_{11,k}^h & \cdots & b_{3,k}^h & w_{1,k}^y & \cdots & w_{3,k}^y & b_k^y \end{bmatrix}^T \tag{12}$$

According to the basic of the dynamic neural network, the output function can be expressed as

$$u_k = \boldsymbol{w}_k^y g\left(\left[\boldsymbol{w}_k^h \boldsymbol{x}_k\right] + \mathbf{b}^h\right) + b^y \tag{13}$$

where $\mathbf{x}$ is the network input that will be discussed later. The $g(\cdot)$ is the activation function in the hidden layer which uses a tanh function. Where b^y is set to zero, and $\mathbf{b}^h$ is PID characteristic matrix that follows Eq. (4), Eq. (5), and Eq. (6), respectively, which is given by

$$\mathbf{b}^h = \left[a_{1,k-1}; 0; -\sum_{j=1}^{m} w_{3j,k-1}^h e_{j,k-1} \right] \tag{14}$$

The plant will give the observation values, also called $\mathbf{z}_k$. It will be compared with the desired value ($\mathbf{y}_d$) to calculate the error signals ($\mathbf{e}_k$) which is formed as

$$\mathbf{x}_k = \left[e_1 \ \cdots \ e_m\right]_k \tag{15}$$

where $\mathbf{x}_k$ is an inputs of the neural network. Then, weight update rule of the proposed controller shows in Fig. 2 which has the procedure as follows, First, at the beginning of the operations, all of the parameters including state vector ($\mathbf{w}$), Kalman gain ($\mathbf{K}$), error covariance ($\mathbf{P}$), and measurement noise covariance ($\mathbf{R}$) will be initiated by random to generate the control input (u) for the plant. Consequently, the external loop of the system will give the error signals ($\mathbf{e}_k$) which is the difference between current angle position and desired angle position. Next, if error is not zero, the old weights ($\mathbf{w}_{k-1|k-1}$) in the weight update rule begin to be optimized by the CKF algorithm at the internal loop as shown in Fig. 2. The procedure of CKF algorithm will be described into 2 steps as follows [15]:

Step 1: The model prediction

The predicted state ($\hat{\mathbf{w}}_{k|k-1}$) and the predicted error covariance ($\mathbf{P}_{k|k-1}$) are first given by

$$\hat{\mathbf{w}}_{k|k-1} = \frac{1}{2n}\sum_{l=1}^{2n} \boldsymbol{W}^{*}_{l,k|k-1} \tag{16}$$

$$\mathbf{P}_{k|k-1} = \tfrac{1}{2n}\sum_{l=1}^{2n}\left(\boldsymbol{W}^{*}_{l,k|k-1}\boldsymbol{W}^{*T}_{l,k|k-1}\right) - \hat{\mathbf{w}}_{k|k-1}\hat{\mathbf{w}}^{T}_{k|k-1} + \mathbf{Q}_{k-1} \tag{17}$$

where $\mathbf{Q}_k$ is the process noise error covariance matrix that assumes to be zero, and the propagated cubature points ($\mathbf{W}^{*}_{l,k|k-1}$) is given by the process model function using the back propagation neural network algorithm [16] which can be expressed as

$$\mathbf{W}^{*}_{l,k|k-1} = f\left(\mathbf{W}_{l,k-1|k-1}\right) = \mathbf{W}_{l,k-1|k-1} + \alpha\tfrac{\partial \mathbf{E}^{(i)}_{k-1}}{\partial \mathbf{W}_{l,k-1|k-1}} \tag{18}$$

where $\mathbf{E}^{(i)}_{k-1}$ is an error cost function of the internal loop at i iteration or is the instanteneous external error of time $k-1$, which will be discussed latter, α is a learning rate that is set to 0.001, $\mathbf{W}_{l,k-1|k-1}$ is a cubature points of model prediction that can be expressed as

$$\mathbf{W}_{l,k-1|k-1} = \mathbf{S}_{k-1|k-1}\zeta_l + \mathbf{w}_{l,k-1|k-1} \tag{19}$$

$$\zeta_l = \begin{cases} \sqrt{n}\mathbf{I}_l & l = 1,2,\ldots,n \\ -\sqrt{n}\mathbf{I}_{l-n} & l = n+1, n+2, \ldots, 2n \end{cases} \tag{20}$$

where n is a quatity of the state vector, and $\mathbf{I}$ is the identity matrix with l^{th} column vector, $\mathbf{w}_{l,k-1|k-1}$ is a old update weight which obtains in step 2, and the variable $\mathbf{S}_{k-1|k-1}$ is a square root of the last corresponding error covariance matrix ($\mathbf{P}_{k-1|k-1}$) expressed as

$$\mathbf{P}_{k-1|k-1} = \mathbf{S}_{k-1|k1}\mathbf{S}^{T}_{k-1|k1} \tag{21}$$

Step 2: The measurement prediction

To optimize the predicted state ($\hat{\mathbf{w}}_{k|k-1}$), we must determine the Kalman gain expressed as

$$\mathbf{K}_k = \mathbf{P}_{xy,k|k-1}\left(\mathbf{P}_{yy,k|k-1}\right)^{-1} \tag{22}$$

where innovation covariance matrix ($\mathbf{P}_{yy,k|k-1}$) and the cross-covariance matrix ($\mathbf{P}_{xy,k|k-1}$) of the state vector ($\mathbf{w}_k$) and measurement output ($\mathbf{y}_k$) are given by

$$\mathbf{P}_{yy,k|k-1} = \tfrac{1}{2n}\sum_{2n}^{1}\left[\mathbf{h}(\mathbf{w}_{l,k|k-1})\mathbf{h}(\mathbf{w}_{l,k|k-1})^{T}\right] - \hat{\mathbf{y}}_{k|k-1}\hat{\mathbf{y}}^{T}_{k|k-1} + \mathbf{R}_k \tag{23}$$

$$\mathbf{P}_{xy,k|k-1} = \tfrac{1}{2n}\sum_{2n}^{1}\left[\mathbf{w}_{l,k|k-1}\mathbf{h}(\mathbf{w}_{l,k|k-1})^{T}\right] - \hat{\mathbf{w}}_{k|k-1}\hat{\mathbf{y}}^{T}_{k|k-1} \tag{24}$$

where $\mathbf{h}(\cdot)$ is neural network PID controller function and $\mathbf{R}_k$ is the noise measurent covariance that is given by

$$\mathbf{R}_k = \mathbf{R}_{k-1} + \frac{1}{k}\left((\boldsymbol{y}_d - \boldsymbol{y}_k)(\boldsymbol{y}_d - \boldsymbol{y}_k)^{T} - \mathbf{R}_{k-1}\right) \tag{25}$$

the term of the cubature points of the measurement prediction ($\mathbf{w}_{l,k|k-1}$) and predicted output ($\hat{\mathbf{y}}_{k|k-1}$) are given by

$$\mathbf{w}_{l,k|k-1} = \mathbf{S}_{k|k-1}\zeta_l + \hat{\mathbf{w}}_{k|k-1} \tag{26}$$

$$\hat{\mathbf{y}}_{k|k-1} = \frac{1}{2n}\sum_{l=1}^{2n}\mathbf{h}\left(\mathbf{w}_{l,k|k-1}\right) \tag{27}$$

where $\mathbf{S}_{k|k-1}$ is a square root of the predicted error covariance matrix ($\mathbf{P}_{k|k-1}$) that is given by step *1*. Then, the new update of the weights and the updated corresponding error covariance in the internal loop calculation are expressed as

$$\mathbf{w}_{k|k} = \hat{\mathbf{w}}_{k|k-1} + \eta\mathbf{K}_k(\boldsymbol{y}_d - \boldsymbol{y}_{k-1}) \tag{28}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{P}_{yy,k|k}\mathbf{K}^{T}_k \tag{29}$$

where η is the Kalman scale rate that is set to 0.1. It is the optimal value in according to the section 2.3 that will be discussed later. The CKF algorithm will evaluate the level of the incremented weights, known as Kalman gain ($\mathbf{K}$). From our experiments, the Kalman gain gives a large scale so it must be multiplied by the Kalman scale rate (η) to normalize and to avoid the over fitting in the increment of the weight update. These modifications increase the controller performances such as reduce the sensitivity with noises and improve the transient response of the control system. Finally, the recursive calculation in the internal loop will continually update the weights of the network until reaching the setting criteria that is set to 20 epochs. Then, the operation

switches from the internal loop to the external loop. Namely, these weights will replace the old weights of the control input calculations. The external loop calculation will continually update the weight until the system give zero value of the system error in according to the equation (30) otherwise the operations will be back to step 2 for the new update with the new system error.

$$\mathbf{E}_k = \frac{1}{2}\sum_{j=1}^{m}(\boldsymbol{y}_{d,j} - \boldsymbol{y}_{k,j})^2 \tag{30}$$

where $\mathbf{E}_k$ is an error cost function of the external loop at time k that is a sum sqaure function of the different value between desired angles position and current angles position.

2.3 The Control System Stability Analysis

To guarantee the control system stability, we apply the Lyapunovs theory to the robotic arm controller. In fact, the Lyapunovs function is only the method for the consideration of the nonlinear control system stability of general characteristic system. In other words, the control system that corresponds to the Lyapunovs function is a defined as a stable. In order to achieve the requirement, Kalman scale rate (η) have to be defined as following steps; firstly, the Lyapunovs function $\mathbf{V}_k$ is defined as

$$\begin{aligned}\mathbf{V}_k &= \frac{1}{2}\sum_{j=1}^{m}(\boldsymbol{y}_{d,j} - \boldsymbol{y}_{j,k})^2 \\ &= \frac{1}{2}\sum_{j=1}^{m} e_{j,k}^2\end{aligned} \tag{31}$$

next, the difference of the Lyapunovs function is given by substitute the equation (31) to below equations

$$\begin{aligned}\Delta\mathbf{V}_k &= \mathbf{V}_{k+1} - \mathbf{V}_k \\ &= \frac{1}{2}\sum_{j}^{m} e_{j,k+1}^2 - \frac{1}{2}\sum_{j}^{m} e_{j,k}^2 \\ &= \frac{1}{2}\sum_{j}^{m}[(e_{j,k+1} + e_{j,k})(e_{j,k+1} - e_{j,k})] \\ &= \frac{1}{2}\sum_{j}^{m}[(e_{j,k+1} - e_{j,k} + 2e_{j,k})(\Delta e_{j,k})] \\ &= \frac{1}{2}\sum_{j}^{m}[(\Delta e_{j,k} + 2e_{j,k})(\Delta e_{j,k})] \\ &= \frac{1}{2}\sum_{j}^{m}\left[(2\Delta e_{j,k}e_{j,k}) + (\Delta e_{j,k})^2\right]\end{aligned} \tag{32}$$

In order to find the difference of the system error (Δe_k), the gradient method is utilized which can be expressed as

$$\begin{aligned}\Delta e_k &= \frac{\partial E_k}{\partial k} = \frac{\partial E_k}{\partial \mathbf{W}_k}\frac{\partial \mathbf{W}_k}{\partial k} \\ &= \frac{\partial(y_{d,k} - y_k)}{\partial \mathbf{W}_k}\Delta\mathbf{W}_k\end{aligned} \tag{33}$$

substitute the $\Delta\mathbf{W}_k$ from the equation (28) that is equal to $\eta\mathbf{K}_k\Delta e_k$ into the equation (33), then

$$\Delta e_k = \frac{\partial y_k}{\partial \mathbf{W}_k}\eta_k\mathbf{K}_k e_k \tag{34}$$

substitute the Δe_k to the equation (32). Then, the difference of the Lyapunovs function is defined as

$$\Delta V_k = \sum_{j=1}^{m}\left[\left(-\tfrac{\partial y_k}{\partial \mathbf{W}_k}\eta_k\mathbf{K}_k e_{j,k}^2\right) + \tfrac{1}{2}\left(\tfrac{\partial y_k}{\partial \mathbf{W}_k}\eta_k\mathbf{K}_k e_{j,k}\right)^2\right] \tag{35}$$

According to Lyapunovfs stability theorem, the system is stable only if $\Delta V \leq 0$ then, η is defined as the following criteria:

case 1: if $\sum_{j=1}^{m}\left[\left(\frac{\partial y_k}{\partial \mathbf{W}_k}\mathbf{K}_k e_{j,k}^2\right)\right] < 0$, then

$$\sum_{j=1}^{m}\left(\tfrac{\partial y_k}{\partial \mathbf{W}_k}\eta_k\mathbf{K}_k e_{j,k}^2\right) + \sum_{j=1}^{m}\tfrac{1}{2}\left(\tfrac{\partial y_k}{\partial \mathbf{W}_k}\mathbf{K}_k e_{j,k}\right)^2 \leq 0$$

$$\eta_k\mathbf{K}_k\sum_{j=1}^{m}\tfrac{1}{2}\left(\tfrac{\partial y_k}{\partial \mathbf{W}_k}e_{j,k}\right)^2 \leq -\sum_{j=1}^{m}\left(\tfrac{\partial y_k}{\partial \mathbf{W}_k}e_{j,k}^2\right)$$

$$\eta_k \leq -2\frac{\sum_{j=1}^{m}\left(\frac{\partial y_k}{\partial \mathbf{W}_k}e_{j,k}^2\right)}{\mathbf{K}_k\sum_{j=1}^{m}\left(\frac{\partial y_k}{\partial \mathbf{W}_k}e_{j,k}\right)^2} \tag{36}$$

case 2: if $\sum_{j=1}^{m}\left[\left(\frac{\partial y_k}{\partial \mathbf{W}_k}\mathbf{K}_k e_{j,k}^2\right)\right] \geq 0$, then

$$\sum_{j=1}^{m}\left(-\tfrac{\partial y_k}{\partial \mathbf{W}_k}\eta_k\mathbf{K}_k e_{j,k}^2\right) + \sum_{j=1}^{m}\tfrac{1}{2}\left(\tfrac{\partial y_k}{\partial \mathbf{W}_k}\mathbf{K}_k e_{j,k}\right)^2 \leq 0$$

$$\eta_k\mathbf{K}_k\sum_{j=1}^{m}\tfrac{1}{2}\left(\tfrac{\partial y_k}{\partial \mathbf{W}_k}e_{j,k}\right)^2 \leq \sum_{j=1}^{m}\left(\tfrac{\partial y_k}{\partial \mathbf{W}_k}e_{j,k}^2\right)$$

$$0 \leq \eta_k \leq 2\frac{\sum_{j=1}^{m}\left(\frac{\partial y_k}{\partial \mathbf{W}_k}e_{j,k}^2\right)}{\mathbf{K}_k\sum_{j=1}^{m}\left(\frac{\partial y_k}{\partial \mathbf{W}_k}e_{j,k}\right)^2} \tag{37}$$

where matrix $\mathbf{K}_k$ is Kalman gain that is given in the previous section.

3. EXPERIMENTS AND RESULTS

3.1 Experiment Setup

The aim of this experiment is to reduce an angle error of the arm movements with the best transient response conditions. The experiment is modeled by

the Matlab program as shown by a block diagram in Fig. 3. The demonstrated block diagram has two major components, which are the controller and the plant. In order to test the performance of the proposed controller, we create other three controllers: 1) ENNPID controller, 2) NNPID controller, controller and 3) conventional PID controller to compare with the proposed controller. All of the controllers except the conventional PID controller are created by the Matlab/Simulink using embedded Matlab function. Inside the block, it contains the listing codes based-on the C language program in according to the gradient algorithm [6] and extended Kalman filter algorithm [17], respectively. In the case of the PID controller, we use the standard libraries in the Matlab. Generally, tuning method of the PID gains the system parameters must be known, so we adds the parameters of the DC motor along with the armature inductance, stall torque, no-load speed, dc supply voltage, and the no-load current which are set with the values of 0.12 H, 0.402 Nm, 316 rad/s, 6v, and 8 mA, respectively. Consequently, we obtain K_P, K_I, and K_D, which are 25.0, 1.5, and -1.0, respectively. These values are also defined as the initial values of the weighted networks of the NNPID controller. In the case of the ENNPID controller, the initial values is set by trained values without the use of the system parameters but the controller still need the pre-training data before using it. In contrast, the initial values of the proposed controller is set by random within some interval without the use of the system parameters and pre-training data. Consequently, the proposed controller is easily applied to the control system in various conditions.

The 4-joint robotic arm structure is also created by using the Matlab/Simulinks library. The detail of the robotic arm structure is shown by the graphical view of all the joints that displays in Fig. 4. In the experiment, since system error is the error of the joint angle of the robotic arm, we put a controller for each of the joints. An angle error of that joint is the input of the controller. The sampling time of the system is equal to 0.01 second. Furthermore, the constraints of the angle movements are also defined to protect the non-sense movements of the robotic arm and we also create the multi-step input signal as the reference angles to simulate the 4-joint robotic arm movements. Then, the results of the experiments of all types of the controllers are compared the performances with each other.

3.2 Experiment Results

To verify the proposed controller, we apply the multistep input to the 4-joint robotic arm control system with various types of the controllers including CKF-NNPID controller, ENNPID controller, NNPID controller and conventional PID controller. In Fig. 5, the 4-joint robotic arm system along with (a) turntable joint, (b) bicep joint, (c) forearm joint, and

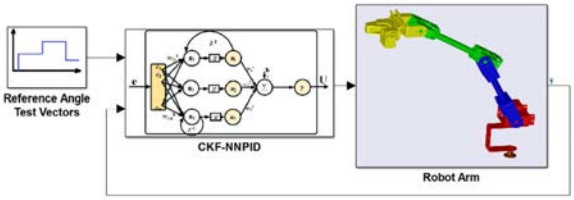


Fig.3: *The block diagram of simulation for the robotic arm control system using the proposed controller.*

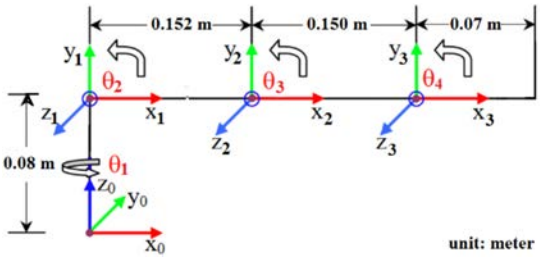


Fig.4: *The coordinate frames of the robotic arm.*

(d) wrist joint under the control of the proposed controller gives the best performance in terms of the maximum overshoot, steady state error, and rise time that are highly require in the control system evaluation because the proposed controller can regenerate the gains to create the control input to compensate the system if the faults have taken place. It clearly see that the CKF-NNPID controller is able to search the values of the gains, initial values, while the classical controller such as the conventional PID control requires the new PID gains (KI, KP, and KD) if the conditions have been changed. Moreover, simulation results show that the algorithm still gives the best result even though the robotic arm is under the loads comparing with the conventional PID control or the other neural PID controller as shown in Fig. 6.

Since we know that the controllers based on the neural network have some disadvantages such as over fitting, large data training requirements, slow convergent rate, especially in the initial value sensitivity. In order to eliminate the initial value sensitivity problem, we initiate the values for the weights of the proposed controller using the random method with 10 different values to control the robotic arm. As a result, it obviously see that the proposed controller applied with 10 different values can track the errors of each joint angle and eventually converges to zero as shown in Fig. 7, Fig. 8, Fig. 9, and Fig.10 for turntable joint, bicep joint, forearm joint, and wrist joint, respectively.

4. CONCLUSIONS

This paper investigates the online tuning neural network PID controller to a nonlinear robotic arm to

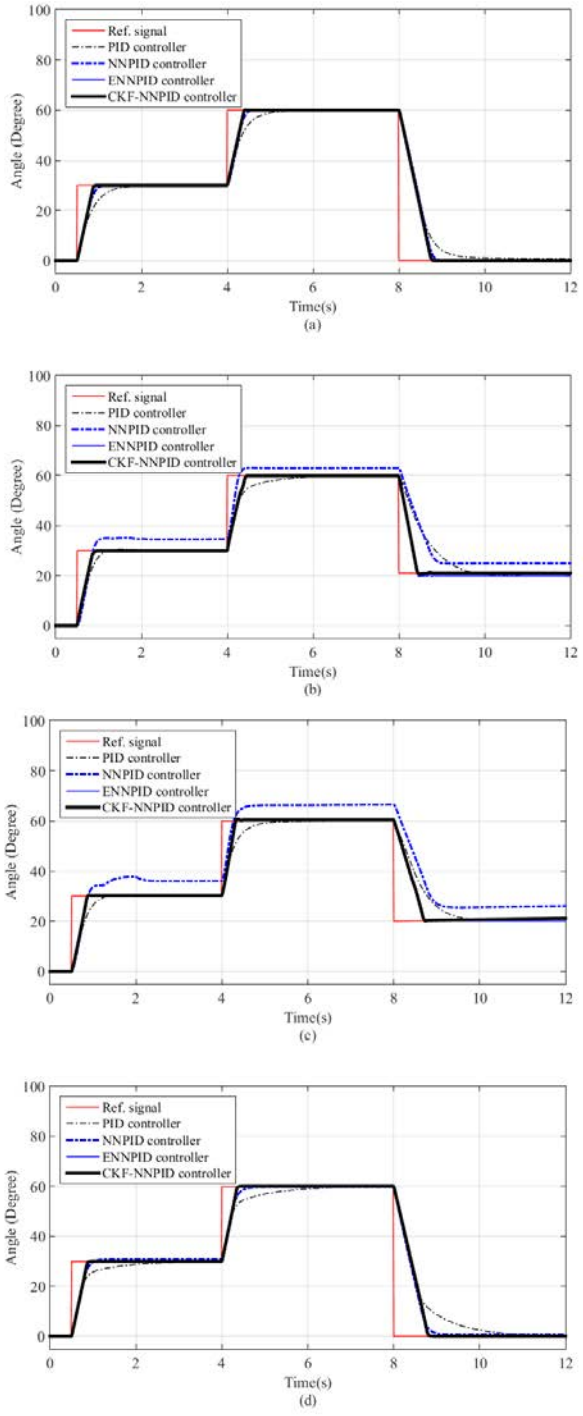


Fig.5: *The angle response comparison of the MIMO control system between the proposed controller and other controllers without load, (a) Turntable joint, (b) Bicep joint, (c) Forearm joint, and (d) Wrist joint.*

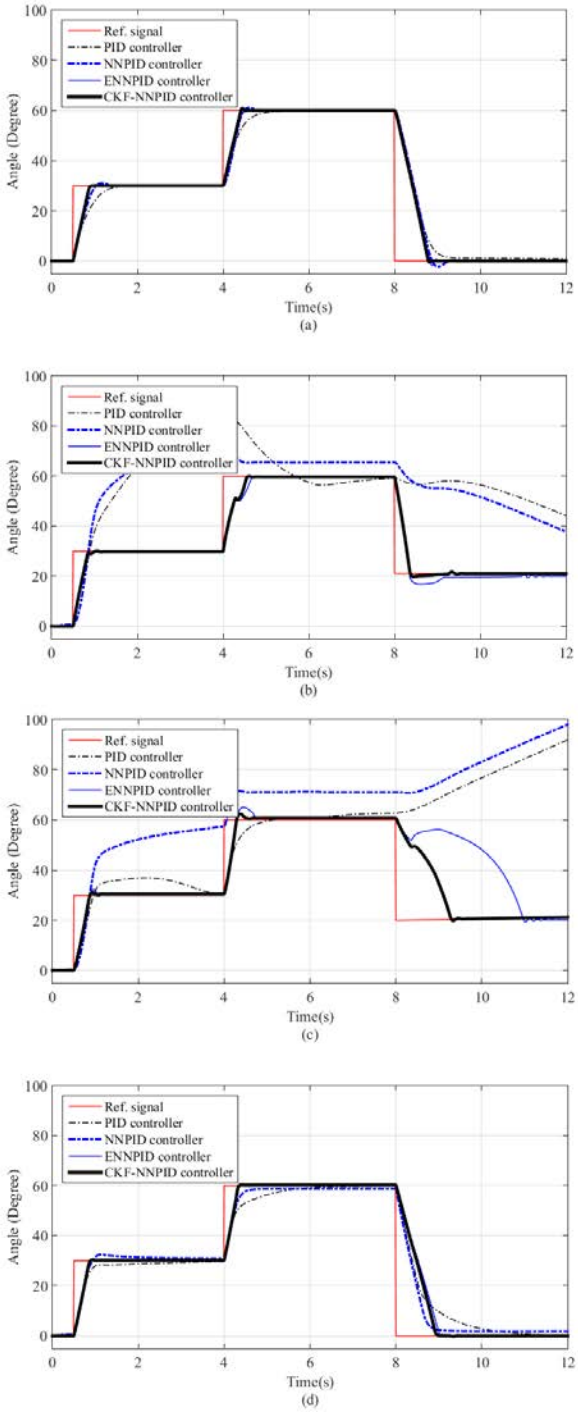


Fig.6: *The angle response comparison of the MIMO control system between the proposed controller and other controllers with load, (a) Turntable joint, (b) Bicep joint, (c) Forearm joint, and (d) Wrist joint.*

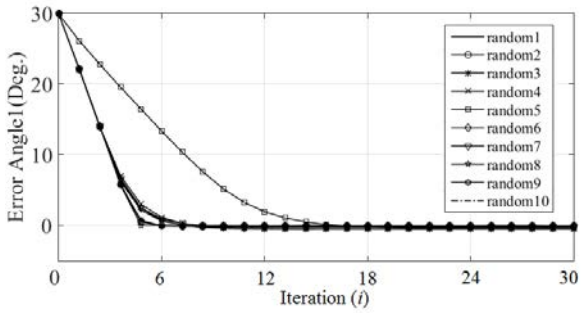


Fig.7: *The error trajectories of the turntable angle of the proposed controller with 10 of the initial values.*

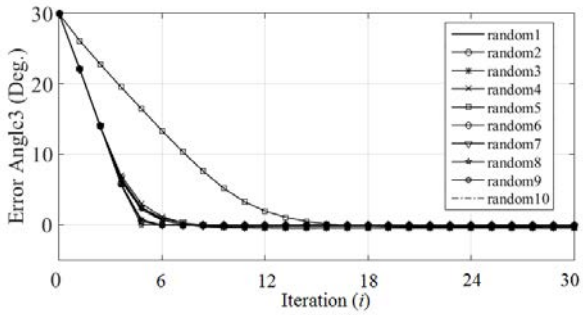


Fig.8: *The error trajectories of the bicep joint of the proposed controller with 10 of the initial values.*

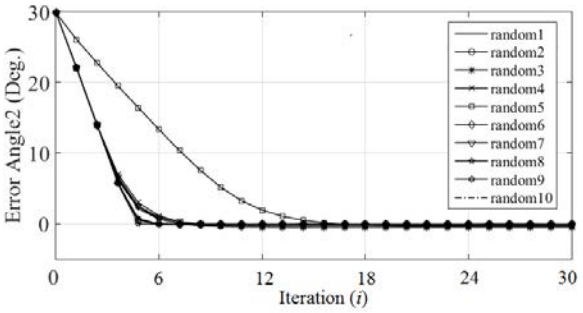


Fig.9: *The error trajectories of the forearm of the proposed controller with 10 of the initial values.*

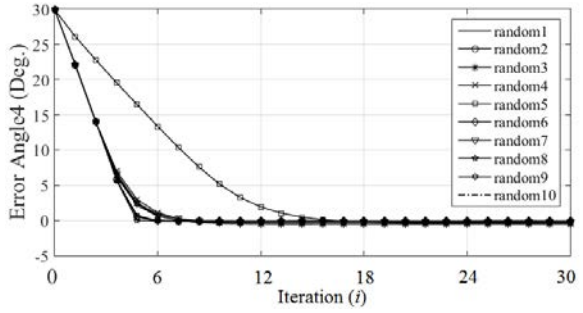


Fig.10: *The error trajectories of the wrist of the proposed controller with 10 of the initial values.*

improve its joint angle position output performance. We use multi-loop operations of the hybrid of the gradient descent algorithm with cubature Kalman filter algorithm. The controller will predict the weights in each internal iteration and then optimize them by cubature Kalman filter (CKF) algorithm until reaching the internal optimal weights, after that these weights will be used to create the compensated voltage control input of the robotic arm system, called Hybrid CKF-NNPID controller. In order to evaluate the controller performances, we apply the proposed controller to 4-joint robotic arm on the MATLAB simulation. From our experimental results, it shows that the proposed controller is a superior controller comparing with the others controller such as conventional PID controller, NNPID controller, and ENNPID controller because the proposed controller potentially tracks the error system even though the system is under the loading criteria. It also improves the convergent speed and stability sustainable. Furthermore, the proposed controller has several advantages for example, it can serve for online tuning without the need of the system parameters to tune the initial gain values while the prior controller must have the system parameters to tune the PID gains, or by trialand- error or by training sampling data.

References

[1] S. Cong and Y. Liang, "PID-Like Neural Network Nonlinear Adaptive Control for Uncertain Multivariable Motion Control Systems," in *IEEE Transactions on Industrial Electronics*, vol. 56, no.10, pp. 3872-3879, Oct. 2009.

[2] H. Huy Anh, "Online tuning gain scheduling MIMO neural PID control of the 2-axes pneumatic artificial muscle (PAM) robot arm," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6547-6560, 2010.

[3] W. Yu and J. Rosen, "Neural PID Control of Robot Manipulators With Application to an Upper Limb Exoskeleton," in *IEEE Transactions on Cybernetics*, vol. 43, no. 2, pp. 673-684, April 2013.

[4] V. Kumar, P. Gaur and A. Mittal, "ANN based self tuned PID like adaptive controller design for high performance PMSM position control," *Expert Systems with Applications*, vol. 41, no. 17, pp. 7995-8002, 2014.

[5] F. Rossomando and C. Soria, "Design and Implementation of Adaptive Neural PID for Non Linear Dynamics in Mobile Robots," in *IEEE Latin America Transactions*, vol. 13, no. 4, pp. 913-918, April 2015.

[6] J. L. Meza, R. Soto and J. Arriaga, "An Optimal Fuzzy Self-Tuning PID Controller for Robot Manipulators via Genetic Algorithm," *Artificial Intelligence, 2009. MICAI 2009. Eighth Mexican*

International Conference on, Guanajuato, 2009, pp. 21-26.

[7] J. L. Meza, V. Santibanez, R. Soto and M. A. Llama, "Fuzzy Self-Tuning PID Semiglobal Regulator for Robot Manipulators," in *IEEE Transactions on Industrial Electronics*, vol. 59, no. 6, pp. 2709-2717, June 2012.

[8] J. Armendariz, V. Parra-Vega, R. Garca-Rodrguez and S. Rosales, "Neuro-fuzzy self-tuning of PID control for semiglobal exponential tracking of robot arms," *Applied Soft Computing*, vol. 25, pp. 139-148, 2014.

[9] Y. Iiguni, H. Sakai and H. Tokumaru, "A real-time learning algorithm for a multilayered neural network based on the extended Kalman filter," in *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 959-966, Apr 1992.

[10] I. Rivals and L. Personnaz, "A recursive algorithm based on the extended Kalman filter for the training of feedforward neural models," *Neurocomputing*, vol. 20, no. 1-3, pp. 279-294, 1998.

[11] S. Haykin, *Kalman filtering and neural networks*. New York: Wiley, 2001.

[12] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, Lake Louise, Alta., 2000, pp. 153-158.

[13] X. Wang and Y. Huang, "Convergence Study in Extended Kalman Filter-Based Training of Recurrent Neural Networks," in *IEEE Transactions on Neural Networks*, vol. 22, no. 4, pp. 588-600, April 2011.

[14] H. Yang, J. Li and F. Ding, "A neural network learning algorithm of chemical process modeling based on the extended Kalman filter," *Neurocomputing*, vol. 70, no. 4-6, pp. 625-632, 2007.

[15] I. Arasaratnam and S. Haykin, "Cubature Kalman Filters," in *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254-1269, June 2009.

[16] D. Rumelhart, G. Hinton and R. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533-536, 1986.

[17] A. Sento and Y. Kitjaidure, "Neural network controller based on PID using an extended Kalman filter algorithm for multi-variable nonlinear control system," *2016 Eighth International Conference on Advanced Computational Intelligence (ICACI)*, Chiang Mai, 2016, pp. 302-309.

[18] C. Sheng, J. Zhao, Y. Liu and W. Wang, "Prediction for noisy nonlinear time series by echo state network based on dual estimation," *Neurocomputing*, vol. 82, pp. 186-195, 2012.

[19] B. Safarinejadian, M. A. Tajeddini and A. Ramezani, "Predict time series using extended, unscented, and cubature Kalman filters based on feed-forward neural network algorithm," *Control, Instrumentation, and Automation (ICCIA), 2013 3rd International Conference on*, Tehran, 2013, pp. 159-164.

[20] A. Mitchell, "Online Courses and Online Teaching Strategies in Higher Education," *Creative Education*, vol. 05, no. 23, pp. 2017-2019, 2014.

[21] I. Arasaratnam and S. Haykin, "Nonlinear Bayesian Filters for Training Recurrent Neural Networks," *MICAI 2008: Advances in Artificial Intelligence: 7th Mexican International Conference on Artificial Intelligence*, pp. 12-33, 2008.

[22] A. Chernodub, "Local control gradients criterion for selection of neuroemulators for model reference adaptive neurocontrol," *Optical Memory and Neural Networks*, vol. 21, no. 2, pp. 126-131, 2012.

[23] F. Yang, W. Sun, G. Lin and W. Zhang, "Prediction of Military Vehicles Drawbar Pull Based on an Improved Relevance Vector Machine and Real Vehicle Tests," *Sensors*, vol. 16, no. 3, p. 351, 2016.

[24] J. Zarei and E. Shokri, "Convergence analysis of nonlinear filtering based on cubature Kalman filter," in *IET Science, Measurement & Technology*, vol. 9, no. 3, pp. 294-305, 5 2015.

[25] S. Wang, J. Feng and C. Tse, "Novel cubature Kalman filtering for systems involving nonlinear states and linear measurements," *AEU - International Journal of Electronics and Communications*, vol. 69, no. 1, pp. 314-320, 2015.

Adna Sento received Bachelor of Engineering from Thammasat University, Pathumthani Thailand, and Master of Engineering from King Mongkut's Institute of Technology Ladkrabang, Bangkok Thailand. Currently, He is studying a Ph.D. degree in Electrical Engineering at King Mongkut's Institute of Technology Ladkrabang, and He is lecture at department of the Computer Engineering, Thai-Nichi Institute of Technology.

Yuttana Kitjaidure received Bachelor of Engineering and Master of Engineering from King Mongkut's Institute of Technology Ladkrabang, Bangkok Thailand, and Ph.D. in Electronic and Electrical Engineering, Imperial College, University of London, England. Currently, He is a lecture at department of the Electronics Engineering, King Mongkut's Institute of Technology Ladkrabang.

REDLE: A Platform in the Cloud for Elderly Fall Detection and Push Response Tracking†

Piyanuch Silapachote[1], **Ananta Srisuphab**[2], **Jinkawin Phongpawarit**[3], **Sirikorn Visetpalitpol**[4], and **Sirima Jirapasitchai**[5], Non-members

ABSTRACT

Caring for a rapid and ever-increasing older population, providing social support and monitoring emotional wellness, is the most immediate and most urgent challenge prompted by the global aging of baby boomers. Elderly assisted systems do not only promote independent lifestyles, enhancing their quality of life, but also reduce stress and worry of families and friends. While being physically active is beneficial and much encouraged, it does increase the risk of falls. We developed an affordable fall detection and response tracking application on the cloud platform; mobile cloud computing is a major evolution with rising impact in information technology and enterprises. Our system, named REDLE, features push notifications for fall alerts and real-time maps for tracking and providing locations and phone numbers of nearby hospitals. Implemented on Android, it captures signals from an embedded tri-axial accelerometer and a global positioning system sensor. Coupled with an efficient threshold-based fall detection algorithm for instantaneous responses, REDLE achieved a near perfect fall detection rate and accurate tracking. Users enjoyed the smoothness of our interactive interface, and complimented on its ease of use and familiarity.

Keywords: Fall Detection, Response Tracking, Eldercare, Mobile Cloud Computing, Push Notification

1. INTRODUCTION

The demand for elderly care has never been more urgent than it is today. Baby boomers, those born between the years 1943 and 1960, are between the ages of 56 and 73 years in 2016, while baby busters, those in Generation X or Y, are approximately at working ages. Proportionately, the number of elderly dependents per working individual is strikingly increasing. Aging adults naturally require support both physically and mentally. They are funtional decline and frailty. Their health condition is declining; their vision is deteriorating; their walking is becoming imbalanced and unstable. Many are more vulnerable to unintentional falls, which are the leading causes of serious injuries, disabilities, or accidental deaths worldwide. Severity of outcomes ranges from broken bones, hip fractures, wrist, arm, and ankle fractures, or upper limb injuries, to head or traumatic brain injuries. According to the World Health Organization (WHO) report [2], supported by many epidemiological studies, 40% of fatal injury in older age is a result of a fall. Among the population at the age of 65 and older, falls account for more than half of injury-related hospitalization. An estimate of 28-35% of individuals in this age group experiences a fall each year. This number increases to 32-42% for senior populations who are 70 years of age and older.

Family members and relatives are faced with difficulties and challenges caring for their vulnerable elderly parents or relatives, providing them a safe environment. Some choose a residential care or a nursing home; others opt for an in-home care service. The majority, nevertheless, embrace a traditional family-based model. Living in a multi-generation household is a rewarding experience but it could be exhausting and full with challenges. Tremendous responsibilities are upon working-age persons to care for both youngsters and elder adults. When kids are at school, their grandparents are often alone at home or out running an errand. Staying active is encouraging and proves beneficial for the elderly; it comes however at a great concern for their safety. Accidents such as falls happen and when they unfortunately do, not only is medical attention often required but also a relative or a person in charge should be promptly informed.

Addressing the need for a self-monitoring device, we have developed an ubiquitous fall detection and response tracking application on a mobile phone, that is equipped with personal confidentiality and privacy protection. REDLE automatically alerts designated and authenticated caregivers in case of a fall. Setting ours apart from the others, we utilize a push notification, processed through a Platform as a Service in the cloud, preceded over a Short Message Service (SMS or text messaging). A cloud-push posts an alert on a caregiver's device regardless of an application a user is actively using or if the device is inactive or unused. Once alerted, it quickly returns the user to our ap-

Manuscript received on October 30, 2016 ; revised on December 15, 2016.

Final manuscript received on February 13, 2017.

† The portions of this research were presented at the 13th Intl. Joint Conf. on Computer Science and Software Engineering [1]

[1,2,3,4,5] The authors are with the Faculty of ICT, Information and Communication Technology, Mahidol University, Thailand 73170, Corresponding author email: ananta.sri@mahidol.edu

plication, which readily delivers the current location where the fall has taken place. In addition to the global positioning system (GPS) coordinates of the respective elderly person, REDLE promptly features an interactive map along with contact information of nearby hospitals, clinics, or healthcare institutes.

2. BACKGROUND AND RELATED WORK

Research on fall detection and monitoring on mobile devices has gained much attention and been continuously evolved following rapid advances in networking technologies and embedded sensors. Fall detection gadgets can be broadly grouped into three categories [3]; they are wearable, ambiance, and vision-based devices. Relying only on embedded sensors, wearable approaches are the most cost effective and the easiest to setup. Ambient devices effectively combine audio and video signals, triggering events based on vibration data. Vision techniques track users using surveillance cameras, employing complex artificial intelligence algorithms. Though achieving higher accuracy, camera based only provides assistive care in a limited, often pre-allocated, and usually indoor space.

Over the past couple of decades, mobile devices, especially tablets and smartphones have become one of the basic needs for almost everyone in every age group. Carrying one or multiple phones everyday everywhere we go is a norm rather than a luxury or a burden. A cellular phone is indispensable; it is always within reach, almost never out of sight, and often in ones' pocket or around the neck. An installation of a fall detection application on a smartphone is thenceforth without a doubt a natural choice. Android operating systems is selected since it currently holds over 80% of the market share in 2014-15 with Apple iOS at less than 15% [4]. Based on Linux kernel, an Android platform allows enthusiastic developers, from amateur to professional, to customarily reconfigure both its hardware and software. Commands and controls of various sensor components are programmable; they are easily accessible and generally adaptable.

Reviewing Android-based fall detection and monitoring systems [5], an accelerometer is a primary sensor being utilized. It has come quite a very long way from being externally mounted in old, obsolete, and discontinued Nokia phones with Symbian OS to a robust tri-axial accelerometer embedded in most modern smartphones; a typical coordinate system of the three axes is illustrated in Fig.1. Raw time-series signals captured are immediately processed on an ever increasing, yet still largely limited, computing power of mobile devices. Concerned with low computational requirements on limited resources of mobile devices as well as real-time responses, we have chosen to implement a simple but effective threshold-based fall detection algorithmic scheme applied to acceleration data. Other strategies that distinguish falls from other regular everyday activities employ pattern

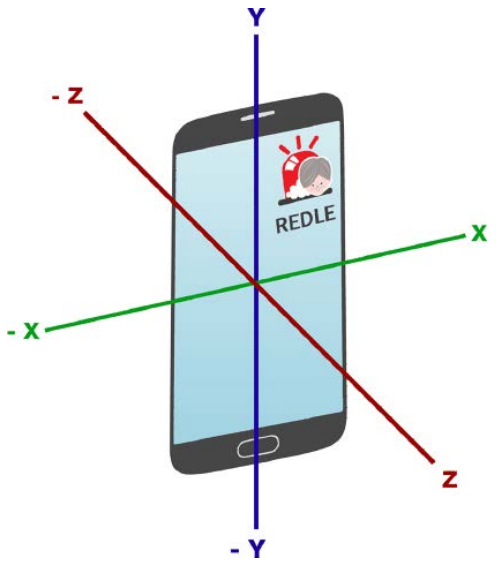


Fig.1: *The three axes associated with a tri-axial accelerometer sensor embedded in most smartphones.*

recognition methods or machine learning techniques [6]: multilayer perceptron, a feedforward neural network, a probabilistic naïve Bayes classifier, support vector machine (SVM), and decision tree (C4.5, an extension of Iterative Dichotomiser 3 or ID3). Developing of such an artificial intelligence (AI)-based detection system involves training and validating phases before conducting a test on real application data. A classifier is trained based on known, labeled samples (supervised learning) or unlabeled instances (unsupervised learning). Since training a classifier is commonly a complex time-consuming task, it hinders the feasibility of on-board smartphone processing.

Communication channels responsible for activating an alarm and sending out an alert when a fall is detected plays a crucial role in the effectiveness of the system. Physical gateways vastly limit the operative range of the applications mostly to indoor environment. Free coverage of cellular networks or wireless mobile telecommunications (3G or 4G) and Wireless Fidelity (Wi-Fi) has been extensively expanded in range; its signal quality has also been largely improved. This allows an application to operate while users are busily moving about. SMS text messaging has been dominated for alert notifications; examples include iFall [7] and Fade Fall Detector [8], among a few commercially available applications on Google Play. Putting much efforts into preventing false positive alarms, however, both iFall and Fade do not send an alert immediately after a positive indication of a fall is perceivably detected. Instead, they attempt to make contact with the user, a prompt in iFall and a verbal contact in Fade, and a fall is confirmed only after the user fails to respond in a limited time frame. Only then a text, containing the time and the GPS coordinates of the fall, is sent to selected individuals whom the user has earlier entered in the contact list.

Without remote monitoring, most fall detectors, iFall and Fade included, operate on the end-user side. The application is constantly active only on the one user (an elderly) whose falls are being monitored, and inactive on the other receiving end (a family member, a relative, or a close friend). REDLE takes a different approach, making the system two-sided and without the high-cost of setting up a remote site. Elderly persons and their caregivers are pre-paired. This avoids any delays due to a verification block when a fall happens and when every second counts (iFall attaches a password authentication with its alerts). Operating two-sided, REDLE lets a careperson query the current location of the elderly at any time, a particularly useful function especially when the elderly is traveling from one place to another for some appointments.

Distinguishing itself from existing applications and advancing a fall detection and notification system to the next level, our REDLE explores the emerging field of mobile cloud computing (MCC) [9, 10]. Clouds deliver seemingly unlimited computing resources over the internet. Access to shared storage and processing units is on demand, available from anywhere from any device connected to the internet. Application data as well as computing elements are stored and processed on either a private or a public cloud instead of a user's own device. Cloud computing solutions are generally grouped into multiple categories based on how much of the resources is disclosed to users and which level of management and functionality is provided to developers. Standing at one extreme, Software as a Service (SaaS) provides a complete package, targeted end-users. It is an application on demand with back-end services and font-end user interfaces. At the other extreme is Infrastructure as a Service (IaaS), which provides only the hardware. It offers storage, computing capabilities, and networking equipment. This is analogous to having a virtualized environment or a virtual server that developers must manage.

Closely related to SaaS is a Backend as a Service (BaaS), which is geared towards mobile applications, referred to as mobile-BaaS or mBaaS. BaaS offers plug-and-play services with the backends taken care of plus additional services such as user file management and report, email and push notifications, location services, and dashboards and social networking. Developers work with application program interfaces (APIs) generally provided in multiple languages to integrate selected elements into their own applications. Another category of cloud services is a Platform as a Service (PaaS) that provides a runtime development and deployment environment on which a customized application can be built. Users usually access PaaS hosted hardware and software via a web browser, getting a slice of cloud to develop, code, test, and deploy their applications. PaaS provides tools such as content distributed networks and load balancing, but developers must setup and built their own database.

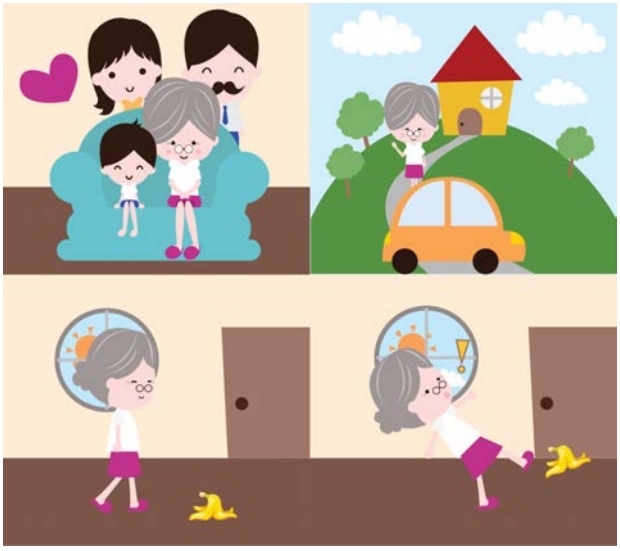


Fig.2: *A grandmother is at risk of accidental falls and it worries family and friends when she is alone.*

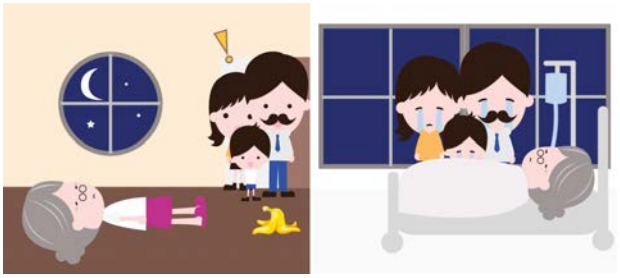


Fig.3: *Accidents happen. A grandmother slips and falls. Unabled to call for help, she lays helplessly for hours until a family arrives to find her later at night.*

These new frontiers of mobile cloud computing do offer great flexibility and excellent scalability at the minimal cost for a large number of application developers. At the core, one of the keys to our fall detection and response tracking application is the utilization of push notifications for alert messages. Flows of data, processes, implementations, and deployments of REDLE are in the cloud. Its first version effectively employed Parse [11], a notable mobile-backend-as-a-service acquired by Facebook in 2013. We have since migrated it to Heroku [12], a platform-as-a-service, following the expected retirement of Parse in 2017.

3. OVERVIEW OF REDLE

The motivation for REDLE is described in Fig.2. Advocating healthy family relationships and always keeping their bonds tight, the grandparents, the parents, and the grandchildren share a single multigeneration living household. Promoting independence of an elderly person, while everyone is busy working at an office or studying at a school, a grandmother stays home alone. She may occasionally go out to socialize with friends in the neighborhoods or to run an errand.

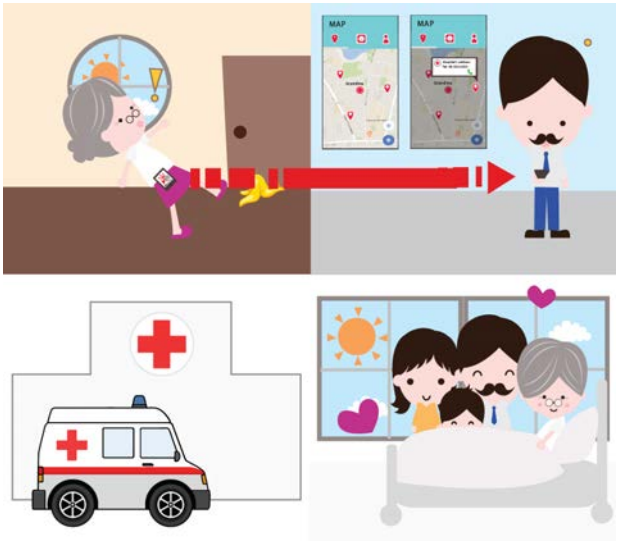


Fig.4: *REDLE alerts the family immediately after a grandmother falls. The father calls an ambulance, getting her much needed medical attentions. Informed family members promptly arrives at the hospital.*

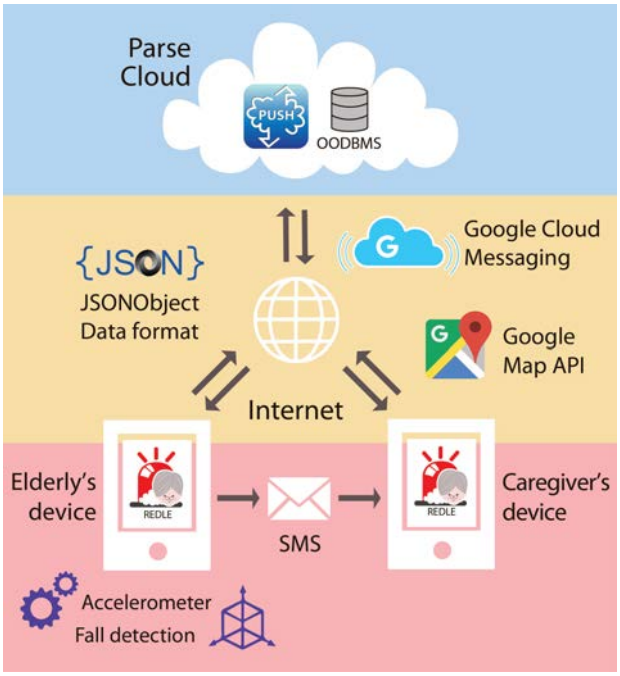


Fig.5: *The architectural platform of REDLE in three layers: mobile devices and sensors, application interfaces through the internet, and Parse Cloud.*

These activities come at a cost of accidental falls; time and again one single fall is all it takes in causing serious injuries, a forever life-changing incident. When it happens (Fig.3) a person is frequently unable to call for help. Laying on the floor, waiting helplessly until someone arrives home later in the evening or at night often leads to regrettably undesirable consequences. The injuries and associated symptoms usually progressively worsen the longer the delay of medical care.

Accidental falls, especially slipping and tripping, do worry everyone when a grandmother is relaxing at home or independently enjoying a stroll around town. Equipped with REDLE (Fig.4), nonetheless, ones can breath easier, knowing that she is being ubiquitously monitored and that they would immediately be notified if she falls. Family members together but individually register on our application. A grandmother logs in as an elderly user whose falls are to be monitored. Parents plus extended relatives and trusted friends log in and agree to join a circle of pre-designated caregivers. When our application detects a fall, it pushes a notification to relatives and friends. Along with a fall alert, it provides the location of the grandmother placed on an interactive map, addresses and contact phone numbers of nearby hospitals. A careperson can quickly call for help resulting in the grandmother receiving medical attention in no time with a comfort of having family members rush to be by her side.

4. REDLE ON PARSE IN THE CLOUD

Installable on any Android device, REDLE (Fig.5) utilizes the moblie-BaaS Parse Cloud [11]. The open source Parse Server provides data storage infrastructure, database functions, queries, and push notification services, all through a convenience of application-level hyper-text-transfer-protocol (HTTP) endpoints. Our application employs an object-oriented-database management system (OODBMS), where data is modeled as objects, supporting classes, subclasses, methods, interfaces, and inheritances. Parse objects are efficiently read and written via a lightweight JavaScript Object Notation (JSON) format. Specifically for our application, we implement five collections of stored objects. They are installation, storing the hardware model and the software version of a device, log-in sessions, user account including name, phone number, and password, the last known location of an elderly user, and a list of friend requests and responses.

Once users are registered and authenticated, their data file is stored in the cloud. Pairing of an elderly person with caregivers requires a successful exchange of push request-accept notifications. Such a request is processed automatically when an elderly person adds his or her desired caregiver. This arrangement needs to be setup only one time; still, profile preferences can be modified at any time and the contact list can be added, deleted, or edited as often as desired. A positive detection of a fall triggers flows of data and services through a Parse server, establishing necessary connections, sending and receiving alert messages and responses. Processed in parallel, the system instantly extracts and locates the geographic coordinates of the fall. It fetches the address, contact information, and phone number of local hospitals and nearby medical facilities. Integrated with Google Map API, REDLE effectively marks these locations all on one interactive map shown on a caregiver's screen.

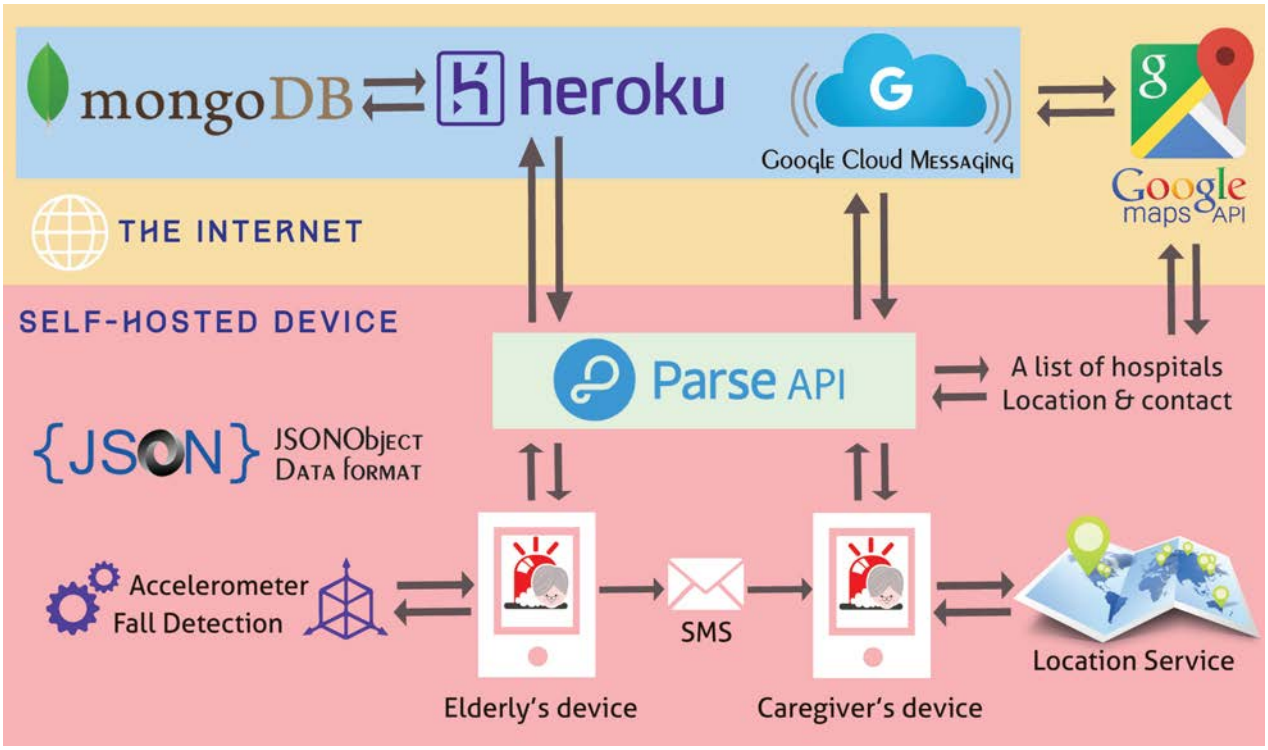


Fig.6: *The revised architecture of REDLE; it has since move from Parse to Heroku and MongoDB.*

5. REDLE ON HEROKU AND MONGODB

After a successful development on Parse, we have migrated our application to Heroku and MongoDB per Parse's recommendation as Parse hosted service is retiring in January 2017. Our revised design and architecture is diagrammed in Fig.6. Registered mobile devices, both the elderly and the caregiver side, interface with Parse API; Parse has since released its open source server, allowing developers to setup a self-hosted Parse server. Instead of a local machine, we deployed Parse server to Heroku. REDLE is then hosted by Heroku [12], a leading PaaS cloud solution.

Integrating with Heroku cloud services, our choice for database programming and management is the open-source MongoDB [13] for its speed and unprecedented scalability, adaptibility, and flexibility that is well-transparent to end users. MongoDB is not a relational database management system (RDBMS) where records are stored in tables with rows and columns. It is a cross-platform document-oriented, no-structured-query-language (NoSQL) database, where records are stored in collections. It employs a rich data model, allowing a mix of literally any type of data to be incorporated. Data is representing in JSON-like documents in a binary-encoded serialization format, called BSON. Instead of a join clause in traditional SQL for querying and retrieving data that requires combining columns of two or more related tables, MongoDB exercises a concept of embedded documents and linking. Furthermore, it features dynamic schemas. Records can be created without pre-defining a fixed schema; fields can be added, removed, or modified as needed.

6. FALL DETECTION ALGORITHM

To detect an occurrence of a fall, REDLE captures signals from a tri-axial accelerometer, embedded in most Android smartphones. An accelerometer senses acceleration forces, the rate of change of the velocity of a device. The amount of dynamic acceleration is an indication of the physical motion of a device: its vibration and how it is moving. Static or gravitational acceleration enables an analysis of the angle of a device with respect to earth: how much it is tilted. Interpreting these signals, we evaluate a velocity pattern of a fall from a measure of dynamic forces; we assess the body position of a person relative to the ground from the values and directions of static forces.

An acceleration is measured quantitatively in metre per second squared, m/s^2, reading at a sampling frequency between 50-200 Hertz [14]. The actual rate depends on the specific hardware model and the software drivers installed. A recommended resolution on a device for converting an analog voltage to a digital value (ADC) is 16-bits or at least 12-bits.

Time is vitally critical as we aim for a real-time fall detection alert and an instantaneous response tracking application. Thus, an approach that necessitates less computational complexity is preferred. This led us to a threshold-based fall detection methodology. Three parameters associated with falls are analyzed: impact, velocity, and posture. Bourke and colleagues [15, 16] demonstrated that algorithms that combine these three parameters could achieve 100% sensitivity and 100% specificity; their settings, particularly threshold values, are utilized in this work. We note

the definition of sensitivity and specificity, where TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively:

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad \text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

Sensitivity or true positive rate measures an amount of recall, that is, the proportion of positive instances that are correctly identified. Specificity or true negative rate assess such proportion of negative instances.

Fall detection parameter: impact

Impact takes into account the four phases of a fall when a person inadvertently and oftentimes suddenly comes to rest on the ground. These are pre-fall, critical, post-fall, and recovery states. A pre-fall is when one becomes unstable but may or may not be falling. A critical stage normally spans a very short period of time; it is a rapid movement of a body down towards the ground. A post-fall is a relatively inactive stage; one is coming to rest, usually in a lying position. Last but not least, during an aftermath, a person instinctively attempts to recover, sitting up or standing up, either with or without help.

Analyzing the transition between these four different phases, the dynamic Root-Sum-of-Squares (RSS) of the three signals recorded from each of the three axes of the tri-axial accelerometer are calculated [15]. The rate of change in velocity over time is captured, both its dynamic and static components. More concretely, these signals are the acceleration in the x-axis (A_x), the acceleration in the y-axis (A_y), and that in the z-axis (A_z). The mathematical formula for the total sum vector, denoted SV, follows Kangas [17]:

$$SV = \sqrt{(A_x)^2 + (A_y)^2 + (A_z)^2}$$

The contour of SV is plotted against time for evaluation. An occurrence of a fall is positively detected when RSS drops below a lower-fall-threshold (LFT) and a short time later it peeks above an upper-fall-threshold (UFT) [18]. RSS equals $+1g$ at stationary where the acceleration due to gravity g is $9.8\,m/s^2$. The value of LFT is set at $+0.65g$ and UFT $+2.8g$.

Fall detection parameter: velocity

A velocity, more explicitly, vertical velocity is estimated through a numerical integration of the Root-Sum-of-Squares over the time period of a fall.

$$v_0 = \int_{t \text{ start of a fall}}^{t \text{ impact}} SV(\tau)\,\mathrm{d}\tau$$

A fall is detected when the measured vertical velocity v_0 is lower than $-0.7\,m/s$, following an empirical analysis on activities of daily living (ADLs) data by Bourke and colleagues [15, 19].

Fall detection parameter: posture

Posture is a measure of an angle between the reference vector of gravity, $\vec{g}_{REF}$, and that of the body segment, $\vec{g}_{SEG}$, which corresponds to the z-axis of a tri-axial accelerometer. Mathematically, this angle θ at time t is the dot product of the respective vectors:

$$\theta\,(t) = \cos^{-1}\left(\frac{\vec{g}_{SEG}\,(t) \bullet \vec{g}_{REF}}{|\vec{g}_{SEG}\,(t)| \cdot |\vec{g}_{REF}|}\right)\frac{180}{\pi} \quad \text{(degrees)}$$

According to Karantonis et al. [20], a measured angle $\theta(t)$ between zero and sixty degrees implies an upright position: while zero to 20 degrees is an indicative of a person standing, 20 to 60 degrees indicates sitting. An angle between 60 to 120 degrees suggests an individual is in lying position and that beyond 120 degrees signifies an inverted position. Correspondingly, our REDLE application detects a fall when the measured angle exceeds 60 degrees.

7. PUSH NOTIFICATION FOR FALL ALERT AND RESPONSE TRACKING

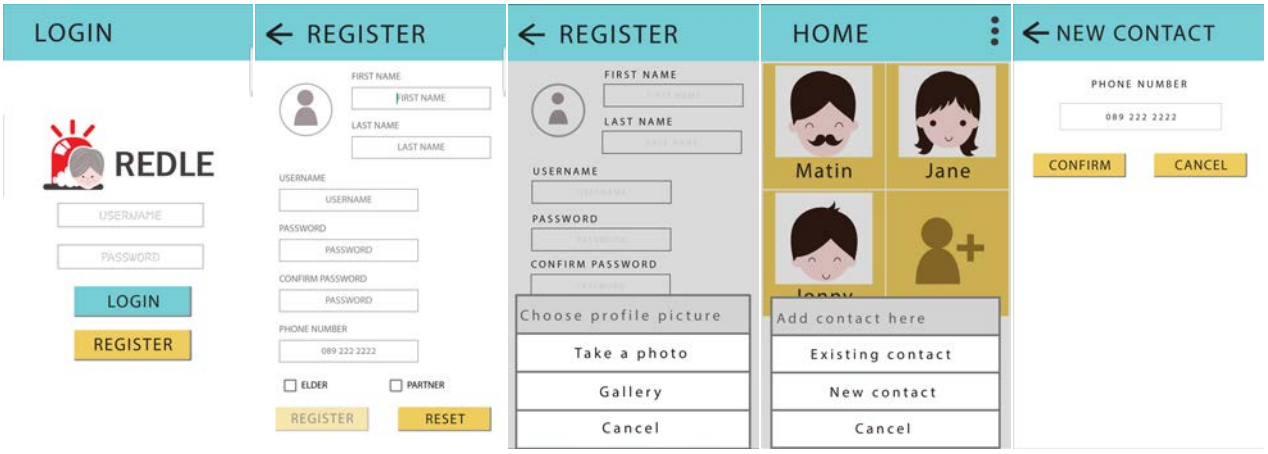
Immediately at the time a fall is detected, a push notification alert is automatically sent to every predesignated caregiver via an integrated Google Cloud Messaging (GCM). More specifically, through Parse API, using Parse Push Notification services, REDLE sets the GCM's sender ID and initiate push messages. In unfortunate circumstances, internet access may be unavailable. Instead of blocking, unabled to respond, our application is set to take an alternative route. It switches, sending a fall alert through a SMS.

The latitude and longitude geo-coordinates of the location of a fall incident are read from the embedded GPS sensor on an elderly's device. This information is then passed to a caregiver's device. Going beyond other existing applications that uses GPS signals only to provide the location of a fall, REDLE is effectively linked with Google Map APIs for maps, locations, and directions. The geo-location of a fall is used to further extract geo-information of nearby medical care centers, which is delivered directly to a careperson. Their locations are marked on a Google map, their phone numbers are shown in pops up.

Following a similar flow of data and services, REDLE includes an unparalleled real-time monitoring and tracking as another added feature. Being a two-ended application, at every instant of time, a careperson may request a map view indicating the current location of every elderly person in his or her care.

8. REDLE'S USER INTERFACE

Designing an interactive user interface for our application focuses on simplicity, user-friendliness, and ease of use. The platform of the system is built on advanced cloud computing, wireless network technologies, and embedded sensors. The front-end, notwith-



(a) Login (b) Registration (c) Profile picture (d) Home page (e) New contact

Fig.7: *Screenshots in elderly mode: login, registration, and pages for managing a contact list of caregivers.*

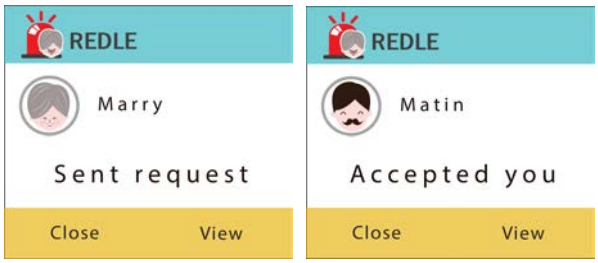


Fig.8: *REDLE's push notifications: (Left) a request to be notified of a fall is sent to a careperson's device and (Right) a response sent to an elderly's device notifying that a pairing has been successfully accepted.*

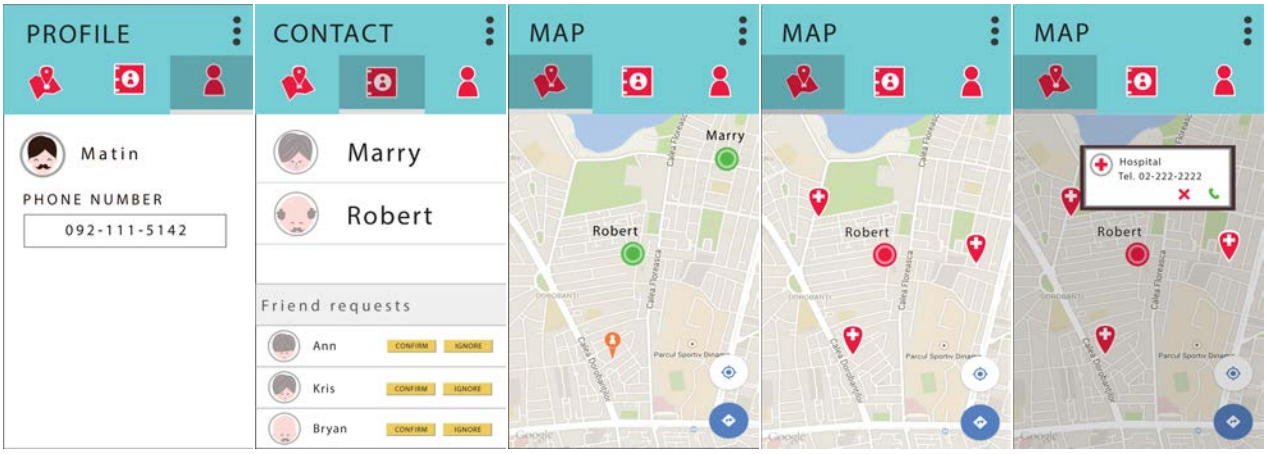
standing, must create user-experience suitable for aging baby boomers. Unlike Generation Z who are digital natives, growing up with electronics and the internet, they are digital immigrant; they are not born digitally aware. Older adults are very capable of operating mobile devices but they are not tech-savvy.

Existing applications such as iFall and Fade do not specifically target elderly users. They are intended for everyone who is at risk, individuals who are vulnerable to accidental falls. Their application user interface is therefore not designed particularly for elderly persons. Their fancy graphics and complicated functions may enhance the general look and feel, but it is not preferred by non-technical baby boomers.

In contrast, REDLE focuses on health and safely of aging individuals. The objective of our application is primarily for elderly adults to receive medical attentions as quickly as possible in all cases of unintentional falls especially in isolated environment. It is critically important to avoid much confusion during emergencies. Our interface design facilitates the urgency of medical care, ensuring that crucial elements: the geolocation of a fall, informative and interactive maps, and local hospital contacts, are quickly understandable and easily accessible, most preferably at a glance and at a fingertip. Another equally indispensable factor to the design of UI/UX (user interface and user experience) is setting up an account and keeping an application running. These must be minimal and straightforwardly natural. Communicating and interacting with REDLE should occur instinctively with a simple but elegant appearance, accessible menus, and handy commands and controls. Otherwise, many elderly people may feel fairly awkward, intimidated, or burdened, and deliberately decide to neither sign-up nor continue using the application altogether.

Examples of screenshots in elderly mode are presented in Fig.7. Keeping them simple but descriptive, these are relatively plain and uncluttered. Fig.7a is the homepage of our REDLE application. Users login with their existing account and first time users are required to register. Creating a new account (Fig.7b) is straightforward. The system asks for nothing more than necessary information: a name, a phone number, and a password. Fig.7c shows an optional profile picture, which can be added or modified at any time. Users choose a photo from a gallery or take a new one using an embedded camera. Having a picture makes it easier to identify from whom a notification comes; REDLE displays it along with every message as well as on a contact list. From the home screen (Fig.7d), adding carepersons is easy and uncomplicated. An elderly user can either select relatives and friends from his or her existing contact list or create a new one (Fig.7e) where only a phone number is needed.

When an elderly chooses or enters a caregiver, REDLE automatically sends a push request to a respective careperson's mobile device. Fig.8 depicts a notification alert a caregiver would receive. It precisely specifies an elderly relative or friend who has initiated the pairing. Once a caregiver views the request and confirms either to accept (or to reject), an elderly user is immediately informed through another accept (or reject) push message, which clearly indicates from



(a) Profile page (b) Contact list (c) GPS tracking (d) Fall location (e) Local hospitals

Fig.9: *REDLE's screenshots for caregivers, showing a profile page, a contact list, and different map features: a real-time querying and tracking, the location of a fall, and contact information of nearby medical services.*

whom the request is responded.

Fig.9 illustrates screenshots in caregiver mode. Login and registration are the same as those for elderly users. We exemplify a profile page (Fig.9a); that of an elderly mode resembles the same format. A contact page (Fig.9b) displays a list of pending requests on the bottom; each with an option either to confirm or to ignore. Confirmed requests are put onto the list at the top of the page. Once paired, by navigating to a map screen, a caregiver is able to pull the real-time location of one or more elderly persons in his or her care. Their locations are simultaneously marked on a map (Fig.9c) for easy tracking and clear visualization. This feature is particularly useful for checking if grandparents have arrived at their planned appointment or destination when they are out and about.

A series of events that REDLE spontaneously activates when a fall is detected is laid out in Fig.9d-e. A push alert mirrors those notifications shown in Fig.8. It is a message indicating that a fall has been detected along with the username and a profile picture of an elderly who has just fallen. Immediately after a caregiver responds, a map with the geo-location of a fall is displayed (Fig.9d). On the same map, a caregiver can view or browse for locations and phone numbers of hospitals, clinics, or healthcare facilities in the vicinity of the elderly person's fall location. At a fingertip, on a pop-up window detailing the contact information of a selected medical service (Fig.9e), a caregiver can conveniently tab or click on it and right away call for a much needed medical attention.

9. TESTING AND EVALUATION

A preliminary evaluation of REDLE aims towards functionality and efficacy of the application platform. We targeted reliability of MCC with push notification and interactive maps and robustness of the threshold-based fall detection algorithm. Our experiments were generally conducted in simulated environment involving only the authors of this work. The phone used is a Samsung Galaxy S6 Edge (SM-G925F) with Android 6.0.1 installed. Keeping the phone in a pocket, a common practice by most individuals, we imitated a number of fall positions frequently occurring in the elderly [21]. We tested nine positions: hard fall forward, hard fall backward, hard fall left, hard fall right, collapsing while standing, collapsing against the wall, slip, fall and holding on to a rail, and falling out of a bed.

- Hard fall forward, backward, to the left, or to the right. Walking unstably, many elderly persons trip when their feet collide. They miss a step and lose a balance. These falls are usually caused by physical hazards such as poor lighting or uneven surfaces; a person may stumble upon cluttered walkway.
- Collapse, happening unexpectedly, is typically due to existing medical conditions. Older patients may be standing freely or leaning against the wall when they suddenly become dizzy and possibly fainted.
- Slip then fall can be caused by wet, oily, or greasy floor, and inappropriate footwear (too loose or slippery shoes). Many come straight down, landed on their buttocks, resulting in painful hip injuries, sustaining a hip fracture or a dislocation hip joint.
- Involuntarily becoming shaky or wobbly, an elderly person is able to grab on to some physical structure for support such as a handrail or a bar, but cannot regain the balance and still ends-up falling.
- It is not uncommon for seniors to roll over and fall out of bed. A bad dream, anxiety or distress, struggling to get up, becoming disoriented, and misjudging the edge of a bed are prevalent reasons.

Each of the nine positions is repeated ten times for the total of ninety falls. Our application achieves a recall rate of 95.56%, detecting all but four hard falls (one with the right-side down and three backwards). Along with every fall that was detected, a real-time

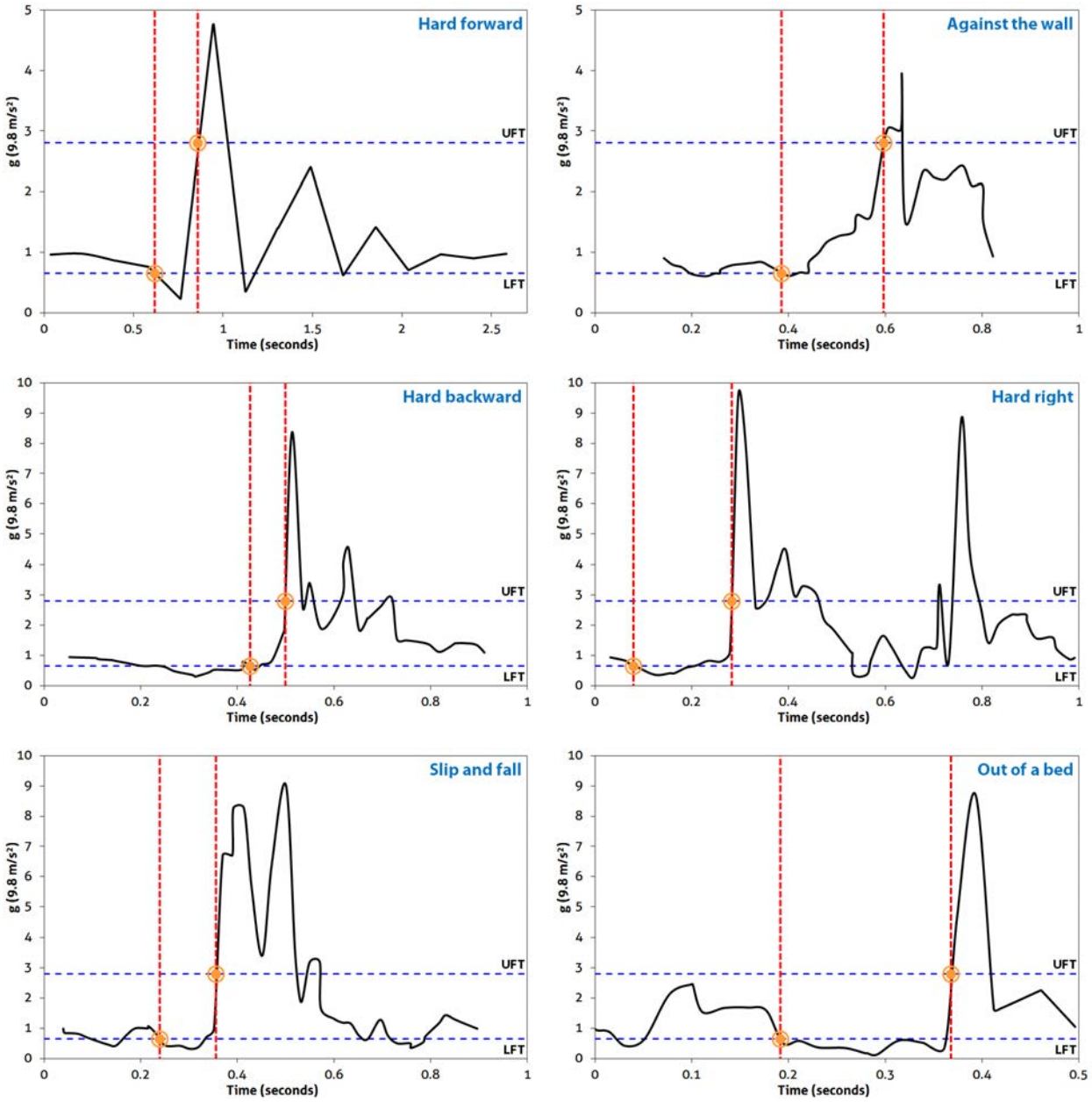


Fig.10: *Signal processing and analysis in REDLE for fall detection. Plotted are the total sum vector, SV. We showed six of the nine fall positions: hard forward, hard backward, hard right, against the wall collapse, slip, and fall out of a bed. The other three are similar: hard fall left is similar to hard fall right, standing collapse is similar to slip (both are relatively hard and sudden falls), and hold on a rail is similar to against the wall collapse (a person comes down relatively at slower speed with lesser force). Marked on every plot are the falling-edge time [15], from the last drop of the acceleration below LFT, until it peeks above UFT.*

push notification was successfully sent and received. GPS sensing and location mapping were manually assessed, thoroughly checked and verified; all of which were accurately reported and fully functioned.

Fig.10 details how REDLE analyzes the gravitational acceleration signals and detects a fall. Acceleration below $1g$ suggests an initial state of a critical phase; a person is accelerating downward, experiencing a free fall. Upon hitting the ground, one absorbs an upward force, instantaneously reversing the direction of the acceleration. During the course of an aftermath, it takes a small, but relatively longer, amount of time for an individual to come to rest. The stronger the impact, the higher the acceleration force surges and the longer the post-fall lasts. As reflected in the results, falling forward and against the wall exhibited lower impact, while falling to the side caused a more intense consequence and posed the longest post-fall period. Slip and fall was similar to falling hard to the side, but less vigorous. Last but not least, falling

out of a bed is a rather unique scenario, its mild impact followed by a brief post-fall was preceded by an extended pre-fall, which echoed a rolling motion.

Empirically, REDLE has displayed its potential to withstand variations and noises in the signals, which were partially caused by a free vibration of a mobile phone kept in a pocket, compared to previous studies [15, 19] where sensors are wrapped tightly to the body. In our experiments, though we had a few false negative, we have not experienced any false positive.

In order to collect different opinions and user recommendations on both the technical components and the usability of our application, the developer team has entered REDLE into the eighteenth Thailand National Software Contest (NSC) 2016. At the poster presentation and live demonstration, our REDLE was tested, evaluated, and criticized not only by many expert committees but also a large number of audiences, ranging from high school participants to undergraduate students and university-level faculty members. Most commented positively on the interface; they like how easy it is to use and understand. While the look and feel of our application is uniquely simple but elegantly attractive, it maintains the familiarity of commonly available mobile applications. This particular feature makes navigating the application spontaneous and effortlessly painless; users waste no time figuring out how it works or what to do. An integration of a well-known Google map makes it very straightforward to find locations and to extract needed information.

10. CONCLUSION AND FUTURE WORK

A successful development of an Android-based fall detection and response tracking mobile application in the cloud has been well-received. Distant from existing fall detectors, our platform is cloud-based; alert messages are sent via push notifications. REDLE has demonstrated a very promising potential for a deployment of the system to extended communities and an extension onto wearable devices such as instrumented wristbands. The application effectively keeps an eye on elderly adults, enabling them to stay active at the comfort of home and boosting the benefits of strong family connections across generations.

References

[1] A. Srisuphab, P. Silapachote, J. Phongpawarit, S. Visetpalitpol, and S. Jirapasitchai, "REDLE: Elderly care on clouds," in *Intl Joint Conference on Computer Science and Software Engineering*, Khon Kaen, Thailand, July 13-15, 2016.

[2] World Health Organization and Ageing and Life Course Unit, *WHO global report on falls prevention in older age*. WHO, 2008.

[3] M. Mubashir, L. Shao, and L. Seed, "A survey on fall detection: Principles and approaches," *Neurocomputing*, vol. 100, 2013.

[4] International Data Corporation Research, Inc., "IDC worldwide mobile phone tracker," Online: www.idc.com, accessed on March 18, 2016.

[5] R. Luque *et al.*, "Comparison and characterization of Android-based fall detection systems," *Sensors*, vol. 14, no. 10, October 2014.

[6] H. Kerdegari *et al.*, "Evaluation of fall detection classification approaches," in *Intl. Conf. on Intelligent and Advanced Systems*, June 2012.

[7] F. Sposaro and G. Tyson, "iFall: An android application for fall monitoring and response," in *International Conf of IEEE EMBS*, Sept 2009.

[8] Instituto Tecnológico y de Energías Renovables, S.A, "Fade: fall detector," Online: apk-dl.com, accessed on September 24, 2016.

[9] A. D. Basha, I. N. Umar, and M. Abbas, "Mobile applications as cloud computing: Implementation and challenge," *Intl. Journal of Information and Electronics Engineering*, vol. 4, Jan 2014.

[10] H. Dinh *et al.*, "A survey of mobile cloud computing: architecture, applications, and approaches," *Wireless Communications and Mobile Computing*, vol. 13, October 2011.

[11] S. Komatineni and D. MacLean, *Expert Andriod*. USA: Apress, 2013.

[12] A. Hanjura, *Heroku Cloud Application Development*. Packt Publishing Ltd., April 2014.

[13] K. Chodorow, *MongoDB: The Definitive Guide*, 2nd ed. O'Reilly Media, Inc., May 2013.

[14] Google, *Andriod Compatibility Definition*, 2015.

[15] A. K. Bourke *et al.*, "Evaluation of waist-mounted tri-axial accelerometer based fall-detection algorithms during scripted and continuous unscripted activities," *Journal of Biomechanics*, vol. 43, no. 15, pp. 3051–57, 2010.

[16] F. Bagalà *et al.*, "Evaluation of accelerometer-based fall detection algorithms on real-world falls," *PLoS ONE*, vol. 7, no. 5, May 2012.

[17] M. Kangas *et al.*, "Comparison of low-complexity fall detection algorithms for body attached accelerometers," *Gait & Posture*, vol. 28, no. 2, pp. 285–291, August 2008.

[18] A. K. Bourke, J. V. O'Brien, and G. M. Lyons, "Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm," *Gait & Posture*, vol. 26, no. 2, pp. 194–199, July 2007.

[19] A. K. Bourke *et al.*, "Fall-detection through vertical velocity thresholding using a tri-axial accelerometer characterized using an optical motion-capture system," in *IEEE EMBS*, 2008.

[20] D. M. Karantonis *et al.*, "Implementation of a real-time human movement classifier using a tri-axial accelerometer for ambulatory monitoring," *IEEE Trans Inf Technol Biomed*, vol. 10, 2006.

[21] E. R. Vieira, R. C. Palmer, and P. H. M. Chaves, "Prevention of falls in older people living in the community," *the British Medical Journal (BMJ)*, vol. 353:i1419, April 2016.

Piyanuch Silapachote received the BS Honors degree in Computer Science from Cornell University College of Engineering in 2001, the MS and PhD degrees in CS from the University of Massachusetts Amherst in 2006 and 2011. She is an instructor in the Faculty of Information and Communication Technology at Mahidol University. Her research interests are in computer vision, pattern recognition, machine learning and intelligence, and biologically-inspired computing. Besides, her work include engineering education and interdisciplinary research.

Ananta Srisuphab received the BSc, MSc, and PhD degrees in Computer Science from Mahidol University in 1990, 2002, and 2009. He had over ten years of experience in IT industry. He was an information services consultant at Unisys Corp. Thailand, and a manager at Infosoft. Presently, he is an instructor in the Faculty of Information and Communication Technology at Mahidol University. His main research interests are in computational intelligence, image and signal processing, embedded systems, and computer science and engineering education.

Jinkawin Phongpawarit received the Bachelor of Science degree in ICT from the Faculty of Information and Communication Technology, Mahidol University in 2016, where he majored in Computer Science.

Sirikorn Visetpalitpol received the Bachelor of Science degree in ICT from the Faculty of Information and Communication Technology, Mahidol University in 2016, where she majored in Multimedia Systems.

Sirima Jirapasitchai received the Bachelor of Science degree in ICT from the Faculty of Information and Communication Technology, Mahidol University in 2016, where she majored in Computer Networks.

Extended Hierarchical Extreme Learning Machine with Multilayer Perceptron

Khanittha Phumrattanaprapin [1] and **Punyaphol Horata**[2], Non-members

ABSTRACT

The Deep Learning approach provides a high performance of classification, especially when invoking image classification problems. However, a shortcomming of the traditional Deep Learning method is the large time scale of training. The hierarchical extreme learning machine (H-ELM) framework was based on the hierarchical learning architecture of multilayer perceptron to address the problem. H-ELM is composed of two parts; the first entails unsupervised multilayer encoding, and the second is the supervised feature classification. H-ELM can give a higher accuracy rate than the traditional ELM. However, there still remains room to enhance its classification performance. This paper therefore proposes a new method termed the extending hierarchical extreme learning machine (EH-ELM), which extends the number of layers in the supervised portion of the H-ELM from a single layer to multiple layers. To evaluate the performance of the EH-ELM, the various classification datasets were studied and compared with the H-ELM and the multilayer ELM, as well as various state-of-the-art such deep architecture methods. The experimental results show that the EH-ELM improved the accuracy rates over most other methods.

Keywords: Hierarchical Extreme Learning Machine, Hierarchical learning, Multilayer Perceptron

1. INTRODUCTION

The Extreme Leaning Machine (ELM) is one of the most popular algorithms for regression and classification. Guang-Bin Huang, et al., 2012 [1] proposed an ELM based on the least squares approach. The ELM is capable of learning from a simple structure called the single layer feedforward neuron networks (SLFNs) [2, 3]. Within the ELM learning process, input weights and biases are randomly generated in hidden layers which calculate and compute the output weights. The ELM offers better performance and faster training time scales than traditional algorithms; such as the back propagation (BP) [4] and the support vector machine (SVM) [5]. Many researchers and application developers have applied ELM in various applications, such as face classification [6], image segmentation [7], human action recognition [8], and power-load forecasting [9, 10].

Although ELM is a fast learning method, many researchers have developed numerous ways to improve the ELM. Kasun, et al., 2003 [11] developed an adapted version of the ELM, referred to as the Multilayer Extreme Learning Machine (ML-ELM), which increases the number of layers in each network. ML-ELM learning is achieved through a building block approach using an ELM-based autoencoder for classification problems. Each layer of the training network is represented as hierarchical learning, stacked layer by layer, suitable for learning within large sized datasets. In ML-ELM training, input weights and biases are orthogonal and randomly generated within hidden parameters. Through this method, the ML-ELM demonstrated superior performance over the traditional ELM within large datasets. The ML-ELM differs from other multilayer schemes, such as the Deep Belief Network (DBN) [12] and the Deep Boltzmann Machine (DBM) [13], because it does not need to fine-tune its parameters. In the decision making process of the ML-ELM, prior to the supervised least mean square optimiaztion, the encoding outputs are directly fed into the last layer, which calculates output weights, without random feature mapping. While the ML-ELM might not use all of the learning advantages of the ELM [1], the universal approximation capability [14] of the ELM requires the feature mapping of inputs in order to calculate the final results.

Meanwhile, the hierarchical learning schemes are designed to be usable with large datasets. Example algorithms include the deep learning (DL) approaches [15,16], in which the deep architectural [17] learning schemes are represented as multiple levels. DL is used for complicated, multi-feature extraction, in order to learn how to automatically produce features in the higher layers from the lower layers. The extraction process is based on an unsupervised learning approach, in which the output of the unsupervised parts provide the inputs of the supervised learniing parts. The DL approach based on the BP algorithm requires multiple fine-tuning of the network parameters. The complicated DL structure affects the learning speed, which is very slow in large sized datasets.

The hierarchical ELM (H-ELM) [18,19] was proposed to address this problem, which combined the ELM with the hierarchical learning approach. The

Manuscript received on October 4, 2016 ; revised on November 24, 2016.

Final manuscript received on February 8, 2017.

[1,2] The authors are with the Department of Computer Science, Faculty of Science, Khon kaen University, 123 Mitraphab Road, Nai-Meuang, Meuang, Khon kaen, 40002 Thailand, E-mail: p.khanittha@kkumail.com and punhor1@kku.ac.th

H-ELM consists of two parts: 1) the unsupervised feature extraction part: the endoding of the input dataset via the sparse autoencoder (based on the ELM autoencoder), and 2) the supervised feature classification part: the single layer used to compute the final output weights, similar to the original ELM. However, the H-ELM may modify the second part to improve classification performance.

The limitations of the supervised classification part of the single layered H-ELM have inspired the use of such deep architecture concepts, such as the two-hidden-layer extreme learning machine (T-ELM),. Qu, et al., 2016 [20], which increases the the single layer of the original ELM to two layers. The performance and testing accuracy of the T-ELM were higher than the those of the batch ELM. Therefore, in order to improve the performance of the H-ELM, we propose herein a modified version of the H-ELM, referred to as the *extended* hierarchical extreme learning machine (EH-ELM), in which the second part of the proposed method (supervised classification training) extends to two or more layers.

The rest of this paper is organized as follow: the next section introduces preliminary related works, consisting of general information and concepts of the ELM. The third section describes the proposed EH-ELM framework, and the fourth section contains the performance evaluation with which to verify the efficiency of the EH-ELM with the various classification datasets and the experimental results of other MLP learning algorithms. Eventually, we draw conclusions in the last section.

2. RELATED WORKS

In this section, we will describe the related works to prepare some essential background knowledge. There are comprised of the conventional ELM [1], H-ELM framework [18] and T-ELM [20].

2.1 Extreme Learning Machine (ELM)

ELM is a learning algorithm, which represented for SLFN with L dimension random feature space. The hidden neurons are randomly generated the parameters (input weights and biases). Let $(\boldsymbol{x_j}, \boldsymbol{t_j})$, $j = 1, \ldots, N$ be a sample of the N distinct samples where $\boldsymbol{x_j}$ is the j-th input sample, $\boldsymbol{t_j}$ is the target vector of $\boldsymbol{x_j}$. The input samples are mapped to L dimension ELM feature spaces, and the output of network is defined as follows

$$f_L(\mathbf{x}) = \sum_{j=1}^{L} g(\boldsymbol{w_j}\boldsymbol{x} + b_j) = \mathbf{h(x)}\boldsymbol{\beta} \qquad (1)$$

where $\boldsymbol{w_j}$ denotes the input weight vector between the input layer and the output layer, $b_j$is the bias of jth layer, $\boldsymbol{\beta} = [\boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_L}]^{\boldsymbol{T}}$ is the output weight matrix, g$(\cdot)$ is the activation function, $\mathbf{x}$ is the input sample and pmbh(x) $= [g_1(\mathbf{x}), \ldots, g_L(\mathbf{x})]^T$ is the vector of the output matrix $\mathbf{H}$ input $\mathbf{x}$.

ELM can resolve the learning problem is as follows

$$\mathbf{T} = \mathbf{H}\boldsymbol{\beta} \qquad (2)$$

where $\mathbf{H}$ is the hidden layer output matrix

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(\boldsymbol{x_1}) \\ \vdots \\ \mathbf{h}(\boldsymbol{x_N}) \end{bmatrix} = \begin{bmatrix} h_1(\boldsymbol{x_1}) & \cdots & h_L(\boldsymbol{x_1}) \\ \vdots & \ddots & \vdots \\ h_1(\boldsymbol{x_N}) & \cdots & h_L(\boldsymbol{x_N}) \end{bmatrix}, \qquad (3)$$

and $\mathbf{T}$ is a target matrix or the desired matrix as

$$\mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix} = \begin{bmatrix} t_{11} & \cdots & t_{1m} \\ \vdots & \ddots & \vdots \\ t_{N1} & \cdots & t_{Nm} \end{bmatrix}. \qquad (4)$$

where m is the dimension of matrix $\mathbf{T}$ for each $1, \ldots, N$ data.

ELM computes the output weights as follows

$$\boldsymbol{\beta} = \mathbf{H}^{\dagger}\mathbf{T} \qquad (5)$$

where $\mathbf{H}^{\dagger}$ is the Moore-Penrose pseudo-inverse of matrix $\mathbf{H}$.

Learning of ELM is summarized in three steps as the following:

Step 1: Randomly generated parameters consist of input weights $\mathbf{w}_{ij}$ and bias b_j for hidden layer.

Step 2: Calculate the hidden layer output matrix $\mathbf{H}$.

Step 3: Compute the output weights $\boldsymbol{\beta}$.

2.2 Hierarchical Extreme Learning Machine (H-ELM)

Hierarchical Extreme Learning Machine (H-ELM) proposed by Tang et al. [18]. H-ELM is a new multilayer perceptron (MLP) algorithm for training data. The data before learning should be transformed to the ELM random feature space, and the H-ELM algorithm has two parts as shown in Fig.1:

1) The first part is the unsupervised learning to extract features of the input samples as the hierarchical representation. This part consists of N-layers to transform the input samples to sparse high-level features, and the output of each layer is computed as

$$\mathbf{H}_i = g(\mathbf{H}_{i-1} \cdot \boldsymbol{\beta}) \qquad (6)$$

where $\mathbf{H}_i$ is the random mapping output of the i-th layer, $\mathbf{H}_{i-1}$ is a previous random mapping output of $\mathbf{H}_i$ layer, g$(\cdot)$ is an activation function, $\boldsymbol{\beta}$ is the hidden layer weights. Moreover, each layer extracts features of the input samples using the ELM sparse autoencoder, which it is defined the objective function as

$$O_\beta = \text{argmin}_\beta \left\{ \|\mathbf{H}\boldsymbol{\beta} - \mathbf{X}\|^2 + \|\boldsymbol{\beta}\|_{\ell_1} \right\} \qquad (7)$$

where $\mathbf{X}$ is the input data, $\mathbf{H}$ is the random mapping output matrix, $\boldsymbol{\beta}$ is the hidden layer weight. To ob-

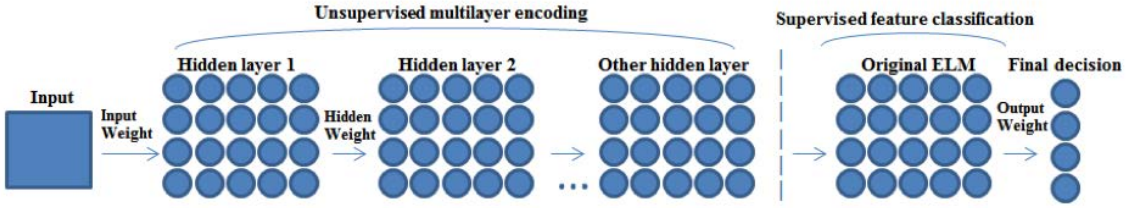


Fig.1: *The overall H-ELM two-part structure [9].*

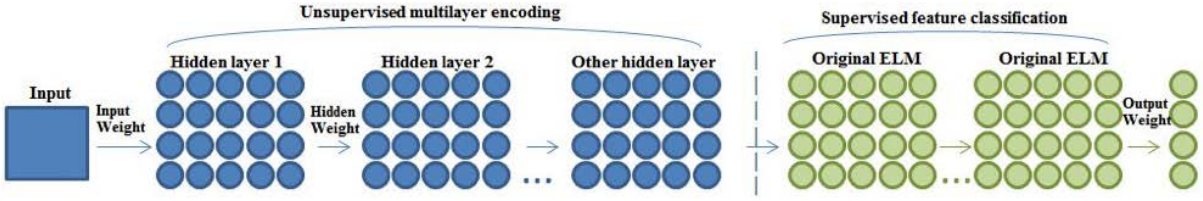


Fig.2: *The two parts of the proposed learning algorithm; the unsupervised multilayer encoding, and the supervised feature classification, respectively.*

tain $\boldsymbol{\beta}$, the ℓ_1 optimization was used in Eg. (8) [18]. Thus, the aim of the ELM sparse autoencoder is to extract features of $\mathbf{X}$ that can be reconstructed $\mathbf{X}$ from the output features without tuned parameters.

2) After the unsupervised part, this part is the supervisor learning part to compute the output weights using the traditional ELM. That is the input weights and biases are randomly generated and then they are used to compute the final output weights.

2.3 Two-hidden-layer extreme learning machine (T-ELM)

The two-hidden-layer feedforward networks (TLFNs) approach was originally proposed by Tamura and Tateishi [21] and Huang [22] to improve the TLFNs framework. T-ELM proposed the ideas of ELM into TLFNs by Qu, et al., 2016 [20], utillizing the ELM algorithm for regression and classification, but having a two-hidden-layer structure. Assume that the first and second layers have the same number of hidden nodes.

The T-ELM Learning process is summarized in the following seven steps:

Step 1: Randomly generate the input weight matrix $\mathbf{w}$ and bias $\mathbf{b}$ of each layer.

Step 2: Calculate $\mathbf{H}=g(\mathbf{w}_{IE}\mathbf{x}_E)$, where $\mathbf{w}_{IE}$ is $[\mathbf{b}\ \mathbf{w}]$ and $\mathbf{x}_E$ is $\boldsymbol{\beta}[1\ \mathbf{x}]^T$.

Step 3: Obtain the hidden weight of the first layer and the output layer, $\boldsymbol{\beta}=\mathbf{H}^{\dagger}\mathbf{T}$.

Step 4: Calculate the output of the first layer, $\mathbf{H}_1=\mathbf{T}\boldsymbol{\beta}^{\dagger}$

Step 5: Assign the hidden weight and bias of the second layer.

Step 6: Obtain the output of the second layer, $\mathbf{H}_2=g(\mathbf{w}_{HE}\mathbf{H}_E)$.

Step 7: Calculate the output weight in the output layer, $\boldsymbol{\beta}=\mathbf{H}_2^{\dagger}\mathbf{T}$.

Note, that while our proposed learning algorithm extends the layers within the supervised learning part, we have not considered the pseudo-inverse of the output weights, written as $\boldsymbol{\beta}^{\dagger}$.

3. PROPOSED LEARNING ALGORITHM

In this section, we outline our proposed framework, which extends the supervised classification part of the H-ELM, referred to as the EH-ELM. Unlike traditional DL frameworks [12, 17] which consist of two training parts, we focused on how to modify the single layer into multilayer supervised classification of the H-ELM. The EH-ELM therefore out performed and proved more efficient than the H-ELM and MLP algorithms.

In the supervised part of the proposed method, the single layer is extended into multiple layers. The output weights of each supervised layer are computed similarly to the traditional ELM, where the approximated target matrix of the previous layer is the input of the next supervised layer. The final output weights are computed from the last supervised layer of the supervised part.

According to the EH-ELM, shown in Fig 2, the input will create and feed samples into the unsupervised part. The resulting high-level features will then be extracted from the input data. The results of the next layer are further input in the multilayer supervised classification part. In the unsupervised part, the input results are computed by M-layers of the ELM based regression, and the subsequent input data is extracted and converted into hierarchical features. The output of each layer in the supervised part is

calculated as:

$$\mathbf{H}_j = g(\mathbf{H}_{j-1} \cdot \boldsymbol{\beta}) \tag{8}$$

where $\mathbf{H}_j$ denotes the output of the j-th layer of the supervised part ($j \in [i, N]$), $\mathbf{H}_{j-1}$ is the output of the (j-1)-th layer of the supervised part, g($\cdot$) represents an activation function for each hidden layer, and $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_L]^T$ is the output weight matrix of each layer. However, the parameters of each layer do not need to be fine-tuned, and each layer of the EH-ELM is an independent section.

The learning process of the EH-ELM is summarized as follows:

The unsupervised encoding part:

Input: N training sample $(\boldsymbol{x_k}, \boldsymbol{t_k})$, k=1, …, N of N distinct samples, $\boldsymbol{x_k}$ is the k-th sample, and $\boldsymbol{t_k}$ denotes the target vector of $\boldsymbol{x_k}$.

Step 1: Randomly generate the input weights $\mathbf{w}_i$ of each layer, where i is ($i \in [1, M]$).

Step 2: Insert the α for the input data, $\mathbf{H}_i$= $[x_i\ \alpha]$.

Step 3: Calculate $\mathbf{H}_i$=($\mathbf{H}_{i-1}$·$\mathbf{w}_{i-1}$).

Step 4: $\mathbf{A}_i$= $\mathbf{H}_i$ is the adjusted scale in [-1,1].

Step 5: Calculate the output weights of each layer, $\boldsymbol{\beta}_i$ = autoencoder ($\mathbf{A}_i$, $\mathbf{H}_i$) [18].

Step 6: Calculate the approximated target of each layer, $\mathbf{T}_i$=($\mathbf{H}_i\boldsymbol{\beta}_i$).

Step 7: Adjust $\mathbf{T}_i$ scale in [0,1].

Return to Step 2 until all of a layers of the unsupervised encoding part are completed.

Output: Finish training unsupervised encoding part. The result is $\mathbf{T}$.

The supervised classification part:

Input: $\mathbf{T}$ of the previous part is the input of this part, where S represents a scale factor, and C is a ℓ_2-penalty value.

Step 8: Randomly generate the input weights $\mathbf{w}_j$ of each layer, where j is ($j \in (M, P]$).

Step 9: Insert α for the input data, $\mathbf{H}_{j-1}$= $[\mathbf{T}_{j-1}\alpha]$.

Step 10: Calculate $\mathbf{T}_j$=$g(\mathbf{H}_{j-1} \cdot \mathbf{w}_j)$.

Return to Step 9 until all layers of the supervised classification part are completed.

Step 11: Calculate the expected output weights, $\boldsymbol{\beta}$=$\mathbf{H}^{\dagger}\mathbf{T}_j$.

The difference between the proposed EH-ELM and ML-ELM are as follows: 1) ML-ELM utilizes unsupervised learning, using a sample, layer by layer. In contrast, H-ELM is modified into the EH-ELM, which consists of two parts: the unsupervised encoding part and the supervised classification part. 2) The ML-ELM employs orthogonal random hidden parameters, in which to construct the weight matric of each unsupervised layer; which is not necessary in the EH-ELM, due to the various outputs of the random number of the input nodes. 3) ML-ELM uses the ℓ_2-norm ELM autoencoder, whereas the EH-ELM uses an ℓ_1 penalty autoencoder to obtain more sparsely hidden input data.

4. PERFORMANCE EVALUATION

In this section, we verify the performance of proposed EH-ELM in three ways: first, we compared the accuracy rate and the standard deviation values of the EH-ELM and H-ELM, using the various benchmark binary and multiple class classification datasets. Next, we compared the last number of hidden neurons of both the EH-ELM and H-ELM algorithms; and lastly, we compared the accuracy rates of the proposed EH-ELM with various state-of-the-art MLP algorithms. In both the second and third methods, two images classification datasets were used, shown in Tables 1 and 2.

All experiments were conducted on an Intel Core i7 3.50 GHz processor CPU, 8G ram, Windows 7, and Matlab R2014a.

4.1 Benchmark Datasets

In order to extensively verify the performance of the two comparative methods, the proposed EH-ELM and the traditional framework H-ELM, four binary classification cases and eight multiple classification case datasets were selected. Most of the datasets were obtained through the UCIdatabase (https://archive.ics.uci.edu/ml/datasets.html) [23], represented in Table 1

Table 1: *Specification of binary and multiple classes classification datasets.*

Datasets	♯train	♯test	♯ feature	♯classes
Liver	242	104	6	2
Diabetes	538	231	8	2
Leukemia	38	34	7129	2
Phoneme	3800	1622	5	2
Glass	150	65	9	6
Led7digit	350	150	7	10
Vehicle	593	254	18	4
Pendigits	537	231	256	10
Semeion	956	638	256	10
Optdigits	3823	1767	64	10
Thyroid	5040	2160	21	3
Shuttle	43500	14500	9	6

From Table 1, within the 12 classification datasets; the four binary class datasets (Liver, Diabetes, Leukemia [24], and Phoneme) and the multiple class datasets (Glass, Led7digit, Vehicle, Pendigits, Se-

meion, Optdigits, Thyroid, and Shuttle) can be classified in to six groups of data, as follows:

1) Data sets of relatively small size and low dimensions: Liver, Diabetes, Glass, Led7digit, and Vehicle.
2) Data sets of relatively small size and medium dimensions: Pengigits and Semeion.
3) Data sets of relatively small size and high dimensions: Leukemia.
4) Data sets of relatively medium size and low dimensions: Phoneme and Thyroid.
5) Data sets of relatively medium size and medium dimensions: Optdigits.
6) Data sets of relatively large size and low dimensions: Shuttle.

Table 2: *Specification classification datasets for comparison of the proposed EH-ELM and the state-of-the art MLP algorithms.*

Datasets	♯train	♯test	♯feature	♯classes
MNIST	50000	1000	784	10
NORB	24300	24300	2048	5

Table 2 demonstrates multiple classes within two datasets: MNIST and NORB, detailed as follows:

1) MNIST: the MNIST dataset [25] is the image classification dataset, which contains handwritten digits, containing 60,000 training and 10,000 testing sample composts. The digital images (28 × 28 pixels) have been sized and fixed for normalization. The original imagess are placed into a training algorithm without preprocessing. Each MNIST sample has 784 features. The number of the output layers are then assigned as the number of its like features.
2) NORB: the NORB dataset [26] represents the 3D images classification problem. It contains pictures of 50 toys, separated into five classes: four animals, human figures, airplanes, trucks, and car. We reconstructed the input pixels, as N = 2 x 1024 =2048. In this paper, sizes of NORB samples for training and testing were divided into equal groups of 24,300 units.

4.2 Comparison between the Proposed EH-ELM and H-ELM

1) Selecting the Number of Hidden Neurons: This section is the criteria for selecting the number of hidden node of each layer that can obtain from the empirical study. It described the effects of selecting hidden nodes in unsupervised autoencoder and supervised classification part of EH-ELM for each dataset problem. A case study for the Glass problem will be illustrated in this section to describe how to select the number of hidden properly. Let $L1$, $L2$ be the number of hidden nodes in the first layer and second layer of the unsupervised autoencoder part, and L3, L4 be the number of hidden nodes in the first and the second in the supervised classification part of EH-ELM. $L1$ and $L2$ were set in the range of $[10, 20, \ldots, 100]$, while $L3$ and $L4$ were set in the range of $[100, 200, \ldots, 1500]$. Each network of EH-ELM in this study, we set $L1$ is equal to $L2$, and also $L3$ is equal $L4$.

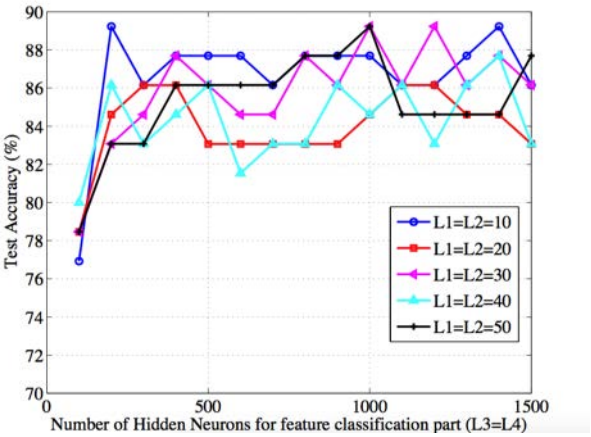


Fig.3: *The Testing Accuracy of EH-ELM in terms of a Number of Hidden Neurons L3 and L4 Feature Classification Part.*

The Fig. 3 shows the testing accuracy rate when EH-ELMs were trained by a various number of hidden nodes in each layer in range of 10 to 100. Each line is illustrated as the testing accuracy rate when $L1$ and $L2$ were fixed. $L3$ and $L4$ are various within the setting condition, $L1 = L2$, and $L3 = L4$. The graph in this figure is shown that the neural network of EH-ELM giving the highest testing accuracy rate for the Glass dataset is the network building the structure as $L1 - L2 - L3 - L4$ to be set as 10-10-200-200, respectively.

2) Performance of Learning Accuracy: We are set the number of hidden neurons in the unsupervised autoencoder and supervised classification parts of the EH-ELM equivalent to the H-ELM for fair comparison. In the EH-ELM, two-layer feature encoding is utilzed as the input for two-layer feature classification. Within the H-ELM, two-layer feature encoding is then utilzed as the input in one-layer feature classification. Performance verification of each method was made through an extensive number of datasets in both the binary class and multiple class datasets.

The training and testing experiments were repeated fifty times for validity, and the average values of the accuracy rates and standard deviations are reported in Table 3.

Table 3 illustrates the EH-ELM classification of four binary datasets (Liver, Diabetes, Leukimia, and Phoneme) and their respective accuracy rates against the H-ELM (86.22, 83.85, 84.76, and 91.92, respectively). Within the Leukemia case, the EH-ELM im-

Table 3: *Performance comparisons of the proposed EH-ELM and H-ELM with binary and multiple class benchmarks.*

Datasets	EH-ELM			H-ELM		
	Number of Hidden neurons	Accuracy (%)	Standard Deviation	Standard Deviation	Accuracy (%)	Number of Hidden neurons
Binary Classes Datasets						
Liver	20,20,200,200	**86.22**	**2.45**	3.05	85.19	20,20,200
Diabetes	50,50,400,400	**83.85**	1.73	1.61	82.96	50,50,400
Leukemia	200,200,5000,5000	**84.76**	**9.10**	10.05	82.41	200,200,5000
Phoneme	100,100,5000,5000	**91.92**	**0.36**	0.40	90.66	100,100,5000
Multiple Classes Datasets						
Glass	10,10,200,200	**84.86**	**2.56**	2.66	81.16	10,10,200
Led7digit	50,50,500,500	**75.10**	0.80	0.82	74.66	50,50,500
Vehicle	50,50,300,300	**82.54**	**1.63**	1.93	81.99	50,50,300
Pendigits	50,50,300,300	97.18	0.76	0.74	97.14	50,50,300
Semeion	300,300,500,500	**97.49**	**0.32**	0.44	97.41	300,300,500
Optdigits	50,50,490,490	97.77	0.19	0.19	97.82	50,50,490
Thyroid	200,200,5000,5000	96.00	0.17	0.17	95.93	200,200,5000
Shuttle	100,100,2000,2000	99.47	0.08	0.06	99.47	100,100,2000

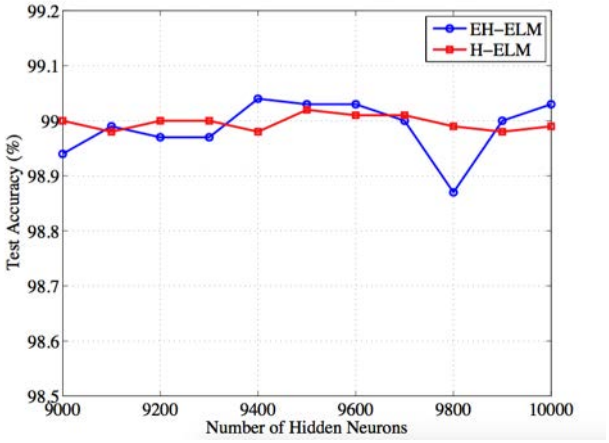
proved accuracy up to 2.3 percent of the H-ELM. The EH-ELM and H-ELM maintained comparable stanrdard derivation accuracy rates.

In the multiple class datasets (Glass, Led7digit, Vehicle, Pendigits, Semeion, Optdigits, Thyroid, and Shuttle), the EH-ELM produced accuracy rates of 84.86, 75.10, 82.54, 97.18, 97.49, 97.77, 96.00, and 99.47, respectively. Remarkably, the EH-ELM accuracy rates out performed the H-ELM in four of eight datasets (Glass, Led7digit, Vehicle, and Semeion cases. Accuracy rates and stand deviations were comparable in the Pendigets, Optdigit, Thyroid, and Shuttle datasets. Summarily, most datasets produced improved accuracy rates using the proposed H-ELM.

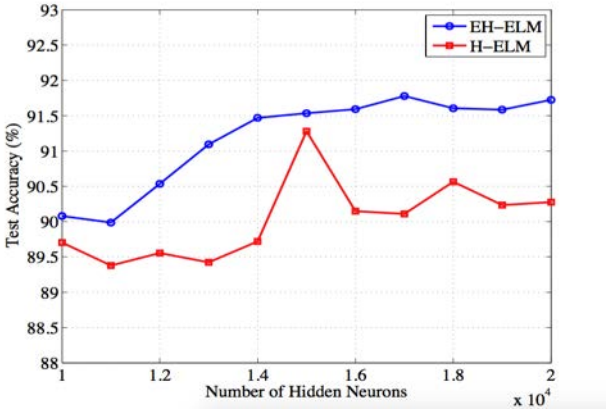
4.3 Comparison of the last number of hidden neurons between the EH-ELM and H-ELM

Comparisons were made between the proposed EH-ELM and H-ELM when the number of hidden neurons in the last supervised layer varied. The number of hidden neurons in the unsupervised autoencoder for both methods were set as equal. In the supervised feature classification of the EH-ELM, comparisons were made prior to the equal setting of the hidden neurons of the H-ELM. The results of the classification performance for testing samples are shown in Fig 4.

In the MNIST dataset, the fluctuation of the testing accuracy rates of the H-ELM were less than those of the EH-ELM, shown in Fig 4(a). The number of hidden neurons varied in range, from 9,000 to 10,000 nodes. The EH-ELM yielded the lowest accuracy rates at 9,800 hidden neuron nodes, while the highest accuracy rate was found at 9,400 hidden neuron



(*a*)



(*b*)

Fig.4: *Comparison of testing rates between EH-ELM and H-ELM: (a) MNIST (b) NORB.*

Table 4: *Learning Accuracy performance comparisons in the MNIST Dataset.*

Algorithm	Accuracy (%)	Training time (s)
SAE [27]	98.60	>17 hours
SDA [28]	98.72	>17 hours
DBN [12]	98.87	66849
MLP-BP [4]	97.39	214268.10
ML-ELM [11]	99.04	2316.19
H-ELM [18]: two unsupervised layers	98.92	281.37
H-ELM [18]: three unsupervised layers	97.81	4120.79
EH-ELM: two unsupervised layers, two supervised layers	**99.05**	1632.45
EH-ELM: two layers of unsupervised, three layer of supervised.	98.92	3552.62

Table 5: *Learning Accuracy performance comparisons in the NORB Dataset.*

Algorithm	Accuracy (%)	Training time (s)
SAE [27]	86.28	> 2 days
SDA [28]	87.62	> 2 days
DBN [12]	88.62	327960
MLP-BP [4]	84.20	34005.47
ML-ELM [11]	88.91	2344.83
H-ELM [18]: two unsupervised layers	91.28	881.43
H-ELM [18]: three unsupervised layers	90.28	8279.67
EH-ELM: two unsupervised layers, two supervised layers	**91.78**	1341.67
EH-ELM: two layers of unsupervised, three layer of supervised.	90.72	6585.57

nodes. In all cases, the accuracy rates of the EH-ELM were higher than those of the H-ELM.

In the NORB dataset, Fig 4(b), the accuracy of the EH-ELM was greater than the H-ELM in all selected number of hidden neurons, ranging from 10,000 to 20,000. Within this range, the best accuracy rate of the H-ELM at 15,000 nodes was less than the best accuracy rate of EH-ELM at 17,000 nodes. The accuracy rates of EH-ELM were higher than those of the H-ELM in every quantity of hidden nodes.

4.4 Comparision with State-of-the-Art MLP Algorithms

In this section, we apply two image classification datasets for comparison of the EH-ELM with the comparative multilayer algorithms; such as the Stacked AutoEncoder (SAE) [27], Staked Denoising Autoencoder (SDA) [28], Deep Belief Networks (DBN) [12], MLP-BP [4], ML-ELM [11], and H-ELM [18]. Within the backpropagation algorithm based MLP training algorithm (SAE, SDA, DBN, and MLP-BP), the learning rate was set at 0.1, and the decay rate at 0.95 for each learning epoch. In the case of SDA, the input corruption rate was 0.5, with a dropout rate of 0.2. The ℓ_2 penalty parameters in the ML-ELM, there were set at 10^{-1}, 10^3, and 10^8; for three layers, respectively. Within the H-ELM algorithm, the number of layers of the unsupervised encoding part was set at two and three layers. Within the proposed EH-ELM, the number of supervised layers was specified at two or three layers, in order to compare the learning performances when changing the number of supervised layers

As shown in Table 4, the testing accuracy of the proposed EH-ELM on MNIST achieves 99.05 %; similar to the ML-ELM, but with faster training times. The EH-ELM proved superior to the SAE, SDA, DBN, and MLP-BP in term of accuracy rates and the training times. Furthermore, the similar architecture of the EH-ELM's supervised part yielded accuracy rates of 99.05 and 98.90 in two and three supervised layers, which were superior to those of the H-ELM, at 98.90 and 97.81, respectively. Table 5 shows that the EH-ELM outperforms than basic H-ELM and other MLP algorithms, including the SAE, SDA, DBN, and ML-ELM. The classification performance test of the EH-ELM having two supervised layers obtained up to 91.78% accuracy, which was greater than the best accuracy rate of the H-ELM, at 91.28 [18]. Although the unsupervised layer of the H-ELM was increased to three layers, its classification rate performance of 90.20 was still less than that of the best accuracy rate of the EH-ELM.

From Tables 4 and 5, the experiment's results demonstrated that the proposed EH-ELM framework (based on H-ELM approach, having two supervised layers) significantly improved classification rates over the original H-ELM, having two or three unsupervised parts, as well as the other multi-layer algorithms in the MNIST and NORB datasets (due to the overfitting of input data). Furthermore, the train-

ing times of the EH-ELM in both the MNIST and NORB datasets (1,632.45, and 1,341.67 seconds, respectively) were slower than that of the H-ELM having two unsupervised layers (due to the greater number of supervised layers of the proposed algorithm than that of the supervised layers of the H-ELM.

5. CONCLUSIONS

In this paper, we propose a new method of high performance classification by extending the supervised classification part of the H-ELM, from one supervised layer to multiple supervised layers, referred to as the Extended H-ELM (EH-ELM). Within our experimental setting, we divided the datasets into two categories: twelve datasets consisting of four binary classes, and eight datasets consisting of multiple classes. Datasets also varied in size and dimensions. The results found that within our experimental datasets, the EH-ELM yielded greater average accuracy rates than those of the traditional H-ELM in eight of twelve datasets; with significant statistical difference ($p<0.05$) four of all eight problems consist of Phoneme, Glass, Led7digit and Thyroid datasets. Additional tests were conducted on MNIST and NORB datasets, which contained a varied number of hidden neurons in the last layer of the supervised classification parts. Within the MNIST data sets, the hidden neurons of the EH-ELM and H-ELM totaled 9400 and 9600 nodes, respectively. While both results were sufficiently high, the accuracy rates of EH-ELM were slightly superior to the H-ELM, in which the number of hidden neurons varied. Within the NORB dataset, the accuracy rates of the EH-ELM were higher than those of H-ELM, in every variable of hidden nodes, 17,000 total; of which the EH-ELM produced the best classification accuracy for this problem. The classification accuracy rates of the EH-ELM also proved superior to the state-of-the-art methods (SAE, SDA, DBN, MLP-BP, ML-ELM, and H-ELM) in both the MNIST and NORB datasets. Therefore, the EH-ELM was conclusively deemed the best classification method among all the comparative algorithms. Note that in the MNIST dataset, the accuracy of EH-ELM and H-ELM yielded 99.05% and 98.92% accuracy, with training times of 281.37 and 1632.45 seconds, respectively. In the NORB dataset, the classification accuracy of EH-ELM was 91.78% with a training time of 1341.67 seconds, while the original H-ELM yielded 90.20% with a training time of 881.43 seconds. Additionally, in comparing the effects of the various number of layers, the experimental results indicate that the EH-ELM with two-supervised layers out performed the H-ELM having three or four layers, due to the overfiting of the input data. Regarding training times, the EH-ELM proved slightly slower than the learning time scale of the H-ELM, due to its greater number of supervised layers.

ACKNOWLEDGEMENT

This research was partially supported by the Department of Computer Science, Faculty of Science, Khon Kaen University, Thailand.

References

[1] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513-529, Apr. 2012.

[2] K. Hornik, M. Stinchcombe, H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp.359-366, 1989.

[3] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Netw.*, vol. 4, no. 2, pp.251-257, 1991.

[4] C. Bishop, *Pattern Recognition and Machine Learning*, New York, NY, USA: Springer-Verlag, 2006.

[5] Suykens, J. A., and Vandewalle, J., "Least squares support vector machine classifiers," *Neural processing letters*, vol.9, no. 3,pp. 293-300, 1999.

[6] A. A. Mohammed, R. Minhas, Q. M. J. Wu, and M. A. Sid-Ahmed,"Human face recognition based on multidimensional PCA and extreme learning machine," *Pattern Recognit.*, vol. 44, nos. 10-11, pp. 2588-2597, 2011.

[7] C. Pan, D. S. Park, Y. Yang, and H. M. Yoo, "Leukocyte image segmentation by visual attention and extreme learning machine," *Neural Comput. Appl.*, vol. 21, no. 6, pp. 1217-1227, 2012.

[8] R. Minhas, A. Baradarani, S. Seifzadeh, and Q. M. J. Wu, "Human action recognition using extreme learning machine based on visual vocabularies," *Neurocomputing*, vol. 73, no. 10-12, pp. 1906-1917, 2010.

[9] S. Cheng, J. Yan, D. Zhao, et al., "Short-term load forecasting method based on ensemble improved extreme learning machine," *J. Xi'an Jiaotong University*. 2:029, 2009.

[10] L. Mao, Y. Wang, X. Liu, et al., "Short-term power load forecasting method based on improved extreme learning machine," *Power Syst. Prot. Control*, vol. 40, no. 20, pp.140-144, 2012.

[11] L. L. C. Kasun, H. Zhou, G.-B. Huang, and C. M. Vong, "Representational learning with extreme learning machine for big data," *IEEE Intell. Syst.*, vol. 28, no. 6, pp. 31-34, Nov. 2013.

[12] G. Hinton, S. Osindero, and Y. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527-1554, Jul. 2006.

[13] R. Salakhutdinov and G. Hinton, "Deep Boltzmann machines," in *Proc. 12th Int. Conf. Artif.*

Intell. Statist., Clearwater Beach, FL, USA, pp. 448-455, Jul. 2009.

[14] G.-B. Huang, L. Chen, and C.-K. Siew, "Universal approximation using incremental constructive feedforward networks with random hidden nodes," *IEEE Trans. Neural Netw.*, vol. 17, no. 4, pp. 879-892, Jul. 2006.

[15] Y. Bengio, "Learning deep architectures for AI," *Found. Trends Mach. Learn.*, vol. 2, no. 1, pp. 1-127, 2009.

[16] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798-1828, Aug. 2013.

[17] Y. Bengio, "Learning deep architectures for AI," *Found. Trends Mach. Learn.*, vol. 2, no. 1, pp. 1-127, 2009.

[18] Jiexiong Tang, Chenwei Deng, and Guang-Bin Huang, "Extreme Learning Machine for Multilayer Perceptron," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 5, no. 4, pp. 809-821, 2015.

[19] H.-G. Han, L.-D. Wang, J.-F. Qiao, "Hierarchical extreme learning machine for feedforward neural network," *Neurocomputing*, pp. 128-135. 2014.

[20] Qu, B. Y., et al. "Two-hidden-layer extreme learning machine for regression and classification," *Neurocomputing*, vol. 175, pp. 826-834, Jan. 2016.

[21] S. I. Tamura and M. Tateishi, "Capabilities of a four-layered feedforward neural network: four layers versus three," *IEEE Transactions on Neural Networks*, vol.8, no. 2, pp.251-255, 1997.

[22] G.-B. Huang, "Learning capability and storage capacity of two-hidden-layer feedforward networks," *IEEE Transactions on Neural Networks*, vol.14 no.2, pp.274-281, 2003.

[23] C. L. Blake and C. J. Merz, "UCI Repository of Machine Learning Databases," Dept. Inf. Comput. Sci., Univ. California, Irvine, CA, 1998. [Online]. Available:http://www.ics.uci.edu/ mlearn/MLRepository.html

[24] T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield, and E. S. Lander, "Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring," *science*, pp.531-537,1999.

[25] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998.

[26] Y. LeCun, F. J. Huang, and L. Bottou, "Learning methods for generic object recognition with invariance to pose and lighting," *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on.*, Vol. 2, 2004.

[27] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504-507,2006.

[28] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn.*, Helsinki, Finland, Jul. 2008, pp. 1096-1103.

Khanittha Phumrattanaprapin received the B.Sc. degree in Computer Science from Khon Kaen University, Khon Kaen, Thailand, in 2014, where she currently pursuing the M.S. degree with the Department of Computer Science. She current research interests include machine learning, hierarchical learning, deep neural networks and big data analytics.

Punyaphol Horata received the B.Sc. degree in Mathematics from Khon Kaen University, M.S. degree in Computer Science from Chulalongkorn University, and graduated Ph.D. in Computer at Khon Kaen University, Thailand. He is currently an assistant professor at computer science department, faculty of science, Khon kaen University, Thailand. His research interests include machine learning, soft computing, software engineering and pattern recognition.