TNI Journal of Engineering and Technology
Vol.7 No.2 July - December 2019

Unfolding to the Past: Temporal and Dimensional Perception Enrichment for Image Impression Improvement

Kulwadee Somboonviwat[1] Thitiporn Lertrusdachakul[2*]

[1]*Faculty of Engineering at Sriracha, Kasetsart University Sriracha Campus, Chonburi, Thailand*
[2]*Faculty of Information Technology, Thai-Nichi Institute of Technology, Bangkok, Thailand*

*Corresponding Author. E-mail address: thitiporn@tni.ac.th

Received: 5 September 2019; Revised: 17 October 2019; Accepted: 29 October 2019
Published online: 25 December 2019

Abstract

An image is in essence a record of visible objects at a particular point in time. This paper proposes a method based on adaptively integrating depth cues with relative motion for enriching temporal and dimensional perception of images. The key idea of our method is to construct a natural-looking animated sequence of images from the original one by introducing differing depth cues and varying time frames to individual generated images. The differing depth cues are created by utilizing variations of visual details and color saturation, whereas the presentation time duration of each image is determined based on structural similarity of consecutive images. The resulting animated image can give a sense of increasing depth as well as a temporal dynamics from past to present. The proposed method can be applied to image impression improvement, especially in the exhibitions or presentations of historical places and artifacts.

Keywords: Image Impression, Temporal and Dimensional Perception, Depth Cues, Structural Similarity Index

I. INTRODUCTION

The evolution of digital imaging has led to an unprecedented growth of the number of digital imaging applications, ranging from industrial to daily life applications. Nowadays, digital imaging is used for diverse purposes such as medical treatment, security, law enforcement, defect detection, preventive maintenance, railway facilities and road inspection and maintenance system. Smartphones and the Internet have made everyone being able to communicate specific and meaningful moments visually. Everybody can be a photographer in their own way. Meanwhile, today's technological society has drastically changed and impacted the level of media richness. Social networking sites like Facebook and Instagram have allowed us to document and manifest our social life through images as a digital representation in the virtual world. We can easily and instantly create an extraordinary image and share it directly to our favorite social media platform.

To obtain a better and more impressive images, many software and apps have been developed to assist us in editing, enhancing and compositing digital images. Tons of smart image filters and algorithms including visual effects are currently available to provide an effective way to catch people's attention. Commercial and noncommercial software, apps and websites such as Adobe Photoshop [1], Canva [2] and BeFunky [3], not only let us adjust brightness, contrast, saturation, hue and blur but also offer photo effects and filters of inspiration, e.g., bokeh and vintage effects, mysterious and seductive filters. All of these technological advancement of digital imaging have inspired and challenged us to create an innovative approach to exploit depth and temporal perception simultaneously to render more impressive images that can naturally visualize the passage of time from past to the present.

In this paper, we utilize depth cues in relative motion to synchronize dimension of space and time (past to present) for providing naturally innovative image impression. These depth cues with reciprocal motion are automatically generated from the original color image.

II. LITERATURE REVIEW

Human brain has evolved to see the world as it is useful for survival. Therefore, our visual perception is based on learning how to see and how the brain organizes visual elements such as colors, shapes, sizes and spatial relationships [4]. The brain's response can introduce depth cues to the images to give us the illusion of three-dimensional (3D) appearance. The depth cues [5] - [8] can be constructed from shading, relative size, linear and atmospheric perspective, focus, motion, stereoscopic and physiological cues. For example, overlapping objects indicate relative distance. Larger objects and more visible detail (in focus) tend to be perceived as being closer. Close objects appear to exhibit more saturated hues and contrasting values. The technique of chiaroscuro [9] which refers to the interplay of light and shadow, can be used to achieve the illusion of three-dimensional volume on a flat surface. Qian *et al*. [10] studied the effects of two depth cues on visual working memory (VWM) performance. Their results showed that combining the coherent binocular and monocular depth cues can provide significantly better performance on the perceptually closer-in-depth items. According to [11] in stereoscopic vision, depth magnitude tends to increase with binocular disparity, while precision and realism of depth perception tend to decrease. The depth perception in disparity-defined objects has special characteristic needed to be investigated for depth estimation errors [12]. Inspired by these effects of depth on human's perception, in this paper, we put forward the idea to combine depth cues and make the realistic depth experience.

For antique and vintage images, precious retro feel can be accomplished by reducing the contrast, saturation, brightness and adjusting the temperature of color and tone [13]. The color fade will bring out some of the details and make motifs become flatter. These effects lead us to a sense of different vibe like they are from another age.

Some websites, apps and research [14] - [17] use motion as a tool for creating lively photo stylizations and effects in static and animated images. In [16], the photographs at different time points of the day are combined with geometric stencils across the scene to create layered animation of a unique impression. Different shade movement [16], [17] enhances the expression of time (day to night) and dimension in the desired moving direction. In [18], different image slices of relative time and transition from daytime to nighttime are used to capture both details and a sequence of time in a single 2D image. Unlike our proposed method which utilizes only a single original image to provide a better sense of temporal and dimensional impression, these methods require taking many photos of different time or manually adjusting the colors.

III. RESEARCH METHODOLOGY

To experience the image beyond a snapshot, we propose the method to enhance temporal perception from olden days to the present with integrated dimensional feeling through the appearance of depth cues. By introducing the depth cues of color saturation, visual details and visible time, our method can improve the perspective in time and dimension of a 2D color image. Fig. 1 describes an overview of the proposed approach. Our approach consists of three main steps: (1) image quantization, (2) structural similarity index calculation, and (3) animation creation.

First, the color image is converted into indexed images by using minimum variance quantization [19] with a series of quantized colors and dithering. The minimum variance quantization is used because it allocates more or less of the colormap entries according to the actual data and can produce more accurate colors than uniform quantization. In addition, the dithering technique of Floyd-Steinberg's error diffusion dither algorithm [20] is applied to adjust the colors of pixels in a neighborhood to make the average color in each neighborhood approximated to the original RGB color. This is to increase the apparent number of colors for a better apparent color resolution of an image. From the preliminary experiment, we obtained the most suitable numbers of quantized colors (Q) for the appropriate change of image perception as described in (1).

$$Q = 2^i \; ; i=2,3,4,\ldots,8 \qquad (1)$$

When the number of quantized colors Q equals to 4 (i=2), the image begins to have a feeling of depth. An art expert suggests that image with 8 quantized colors (i=3) evokes the most perceived sensation of former times. When the number of quantized colors Q increases, more detail and saturated hues become visible. Therefore, the images start to have more depth as Q increases and the perception of difference becomes smaller for larger Q (Q>64).

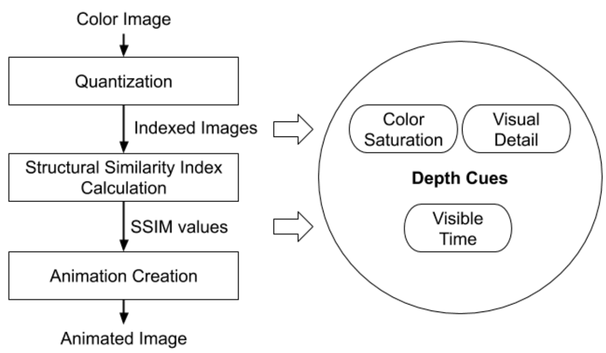


Fig. 1. The overview of the proposed approach.

Next, we integrate the motion by using visible time to animate a series of images into GIF form. The frame duration will depend on the image difference in each step of quantized colors. Specifically, it is inversely proportional to the similarity between each image. In this process, the structural similarity index (SSIM) is calculated for adjusting the visible time of each frame. SSIM [21] - [23] is selected because it can perform more accurate measurement of the perceptual difference between two similar images than the standard measure of mean square error (MSE). SSIM is based on the visual impact of three characteristics of an image, i.e., luminance, contrast and structure, and can be simplified as (2).

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2)$$

where μ_x, μ_y, σ_x, σ_y and σ_{xy} are the local means, standard deviations, and cross-covariance for images x, y. The default values of C_1 and C_2 are $(0.01L)^2$ and $(0.03L)^2$ where L is the dynamic range of the input image. The value of SSIM is standardized to a scale of -1 (very different) to 1 (very similar). By testing the duration of each frame in the preliminary experiment, we obtained the best frame duration adjusted according to the comments from an art expert to achieve the temporal and dimensional perception of the image as shown in (3).

$$tf_i = \frac{1 - SSIM(img_i, img_{i+1})}{\sum_{i=2}^{7}(1 - SSIM(img_i, img_{i+1}))} \times t_s \text{ ; } i = 2,3,4,\dots,7 \quad (3)$$

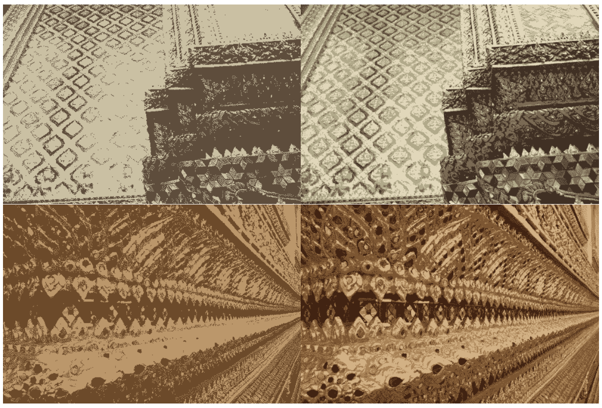
where tf_i is the frame duration of image with 2^i quantized colors (img_i),

$SSIM(img_i, img_{i+1})$ is the structural similarity index value for img_i using img_{i+1} as the reference image,

t_s is the setting time in second to control the period of animated image which refers to the total frame duration excluding the last frame. It is also set to be the duration of the last frame (the highest image quality) and the default value is 1. The duration of animated GIF and the animation speed can be changed proportionally by adjusting the value of t_s. Once the values of SSIM are determined, we then compute the frame duration of each quantized image to create an impressive animation.

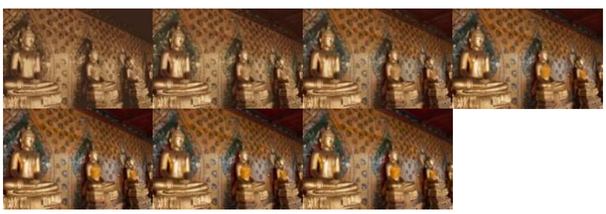
IV. RESULTS AND DISCUSSION

We applied the proposed approach to produce the images more appealing. As the temporal (past to present) and dimensional (depth) perception are the target of our method, the images of temples which carry implicit history and aesthetic, are selected to test in the experiment. These images are taken from several temples in Bangkok such as Wat Arun Ratchawararam Ratchawaramahawihan (Wat Arun), Wat Phra Chetuphon Vimolmangklararm Rajwara mahaviharn (Wat Pho) and Wat Phra Si Rattana Satsadaram (Wat Phra Kaew). The examples of images are shown in Figs. 2 and 3. Fig. 2 compares the results between 2 and 4 quantized colors. The digital images are available online at http://tiny.cc/xynhez. It can be seen that images with Q of 4 become more depth (not flat) and are used as the beginning frame of the animated images. Fig. 3 shows four examples of consecutive frames with quantized colors of 4, 8, 16, 32, 64, 128, 256 (i=2,3,4,..8) from left to right and top to bottom. It can be clearly seen that the images become more appearing and perceptible as the visual detail and color saturation are increased. The original images are available online at http://tiny.cc/mlqhez. The difference can be observed clearly when the images are viewed in large size and high quality display. When Q is small, the image difference is easily noticeable. However, it becomes almost no difference (very similar) in the last two image frames.

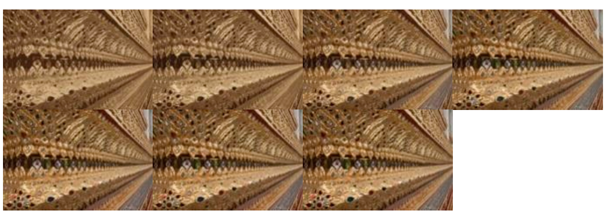


Q=2 Q=4

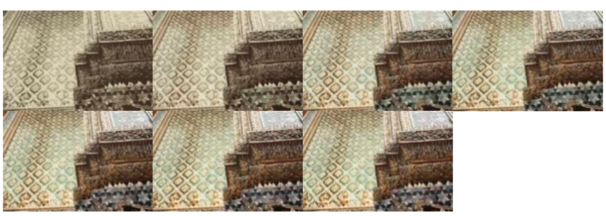
Fig. 2. The comparison of images with different quantized colors (Q).



Wat Arun 1



Wat Arun 2



Wat Phra Kaew 1

Wat Pho 1

Fig. 3. The examples of image results with 4, 8, 16, 32, 64, 128, 256 quantized colors (left to right and top to bottom).

Fig. 4 illustrates an example of local SSIM value map for Wat Arun 2. It is calculated from 2-D grayscale images of consecutive frames by using (2). The results are displayed in grayscale by scaling the minimum value to black and the maximum value to white. The higher value of local SSIM value (i.e. white color) indicates more similarity. The average value of local SSIM values will be used as a global SSIM value of the image. Fig. 5 shows the global SSIM values of the example images shown in Fig. 3. The value of 1 means very similar. The images of Wat Arun 1 have smaller SSIM values than others which are corresponding to the larger difference in visual perception between each frame. It can be obviously seen that all graphs tend to be less slope (more similarity) when Q is increased. From these variations, we can compute the frame duration (tf_i) by (3) and the results are described in Fig. 6 with the default value t_s of 1. The created GIF animation will have shorter frame duration if the frames are more similar.

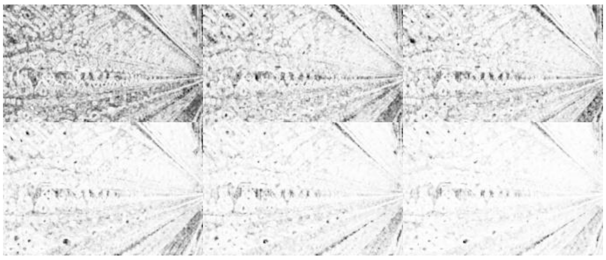


Fig. 4. Local SSIM value map of consecutive frames for i = 2, 3, 4,..., 7 in Wat Arun 2 (left to right and top to bottom).

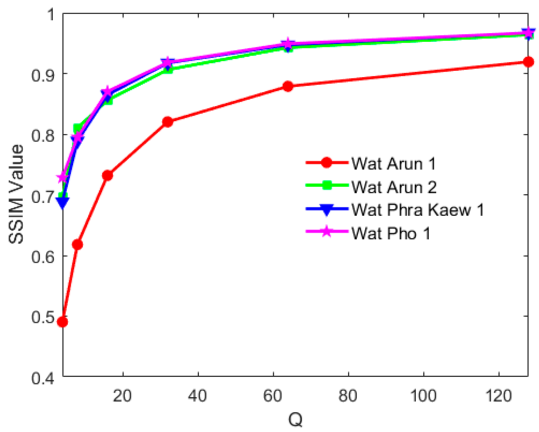


Fig. 5. The global SSIM values of the example images shown in Fig. 3.

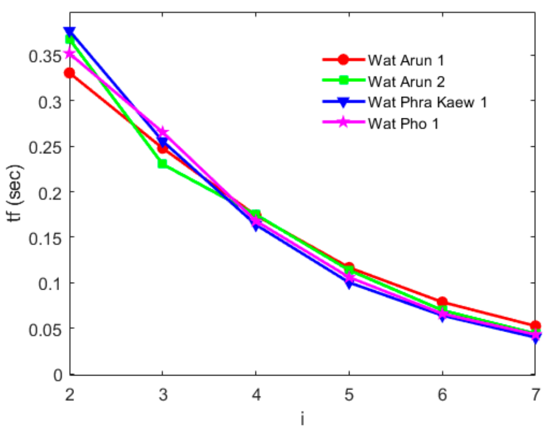


Fig. 6. The frame duration of the example images shown in Fig. 3.

Fig. 7 illustrates an example of frame duration in term of percentage of animated period. In this case by default, the total time of period is 2 seconds which 50% of them will allocate for 6 animated images and another 50% will be the duration of the last frame. The animated GIF can be viewed at http://tiny.cc/icisbz.

For the overall results, the created GIF animations can improve the perception of time (past to present) and dimension (depth) as the images are getting richer in detail and color saturation. These depth cues are enhanced by relative motion of image frame according to the image difference. An artistic specialist suggested that the proposed method would be good at image with high detail/texture/pattern, high color variation and high dimension or perspective from shadow or lightness. The difference in saturation between each frame could be used for fine tuning the frame duration. Compared to prior methods [16], [17], the proposed GIF animations generation approach utilizes a single original image to naturally provide a better sense of temporal and dimensional impression without manually adjusting the image colors. Our method can be used in various scenarios such as electronic brochures, exhibitions in the museums, visual arts, websites and social media to make the images more impression and attention.

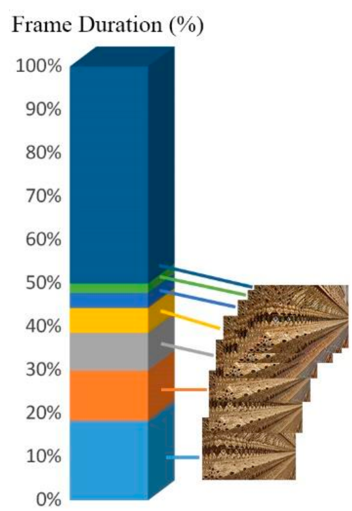


Fig. 7. The frame duration of Wat Arun 2 as percentage of animated period.

Although the indexed images of a series of quantized colors work well for the animated GIF in our approach, the frame selection and transition could be further investigated for the delicate and intricate conveying of expression.

V. CONCLUSION

In this paper, we propose an intuitive and simple method to exploit depth cues of images for successive transformation of temporal and dimensional feelings. The method is based on innate characteristics of a sequence of quantized images. The variations of visual detail and color saturation in relative motion assist the human brain to perceive more depth and time of the past to the present. With these techniques, the approach can be extended to other types of depth cues or other cues with corresponding motion to produce images that are more interesting and impressive for a variety of applications.

REFERENCES

[1] *Adobe Photoshop*. (2019). Adobe. Accessed: Sep. 3, 2019. [Online]. Available: https://www.adobe.com/sea/products/photoshopfamily.html

[2] *Canva*. (2019). [Online]. Available: https://www.canva.com/features/photo-effects/

[3] *BeFunky*. (2019). [Online]. Available: https://www.befunky.com/features/photo-effects/

[4] L. Wang. "8 mind-bending optical illusions (and what they reveal about how our brains work)." Visual Learning Center by Visme. https://visme.co/blog/best-optical-illusions/ (accessed Sep. 3, 2019).

[5] A. Hosny. "How come a 3D movie on a 2D screen ?!" https://abdelrahmanhosny.wordpress.com/2014/03/12/how-come-a-3d-movie-on-a-2d-screen/ (accessed Sep. 3, 2019).

[6] C. Jirousek. "Two dimensional illusion of three dimensional form." Art, Design, and Visual Thinking. http://char.txa.cornell.edu/language/element/form/formillu.htm (accessed Sep. 3, 2019).

[7] J.-I. Jung, J.-H. Lee, I.-Y. Shin, J.-H. Moon, and Y.-S. Ho, "Improved depth perception of single-view images," *ECTI Transactions on Electrical Eng., Electronics, and Communications*, Vol. 8, No. 2, pp. 164-172, Aug. 2010.

[8] K. Ruxpaitoon, N. Aoki, and H. Kobayashi, "Studies on perspective expression technique used in wall paintings of each Thai dynasty period," *Journal of the Society of Photographic Science and Technology of Japan*, Vol. 76, No.1, pp. 88-98, 2013.

[9] D. Scott. "Chiaroscuro." https://drawpaintacademy.com/chiaroscuro/ (accessed Sep. 3, 2019).

[10] J. Qian, J. Li, K. Wang, S. Liu, and Q. Lei, "Evidence for the effect of depth on visual working memory," *Scientific Reports*, Jul. 2017.

[11] P. B. Hibbard, A. E. Haines, and R. L. Hornsey, "Magnitude, precision, and realism of depth perception in stereoscopic vision," *Cognitive Research: Principles and Implications*, May 2017.

[12] P. Cammack and J. M. Harris, "Depth perception in disparity-defined objects: finding the balance between averaging and segregation," *Philosophical Transactions of the Royal Society B: Biological Sciences*, Jun. 2016.

[13] Lars. "4 tips for modern vintage photography." https://www.eyeem.com/blog/modern-vintage-photography (accessed Sep. 3, 2019).

[14] A. Semmo, M. Reimann, M. Klingbeil, S. Shekhar, M. Trapp, and J. Döllner, "ViVid: Depicting dynamics in stylized live photos," 2019, doi: 10.1145/3305365.3329726.

[15] *Bloggif*. (2019). [Online]. Available: https://en.bloggif.com/effect (accessed Sep. 3, 2019).

[16] F. Q. Wei. "Time in motion by Fong Qi Wei." Gestalten. https://gestalten.com/blogs/journal/time-in-motion-by-fong-qi-wei (accessed Sep. 3, 2019).

[17] fqwimages. https://fqwimages.tumblr.com/ (accessed Sep. 3, 2019).

[18] fqwimages. "Time paintings." https://fqwimages.com/ time-paintings/ (accessed Sep. 3, 2019).

[19] The MathWorks, Inc. "Reduce the number of colors in an image." https://www.mathworks.com/help/images/reduce-the-number-of-colors-in-an-image.html (accessed Sep. 3, 2019).

[20] R. W. Floyd and L. Steinberg, "An adaptive algorithm for spatial gray scale," *International Symposium Digest of Technical Papers, Society for Information Displays*, 1975, pp. 36.

[21] Z. Wang. "Image quality assessment: From error visibility to structural similarity." https://pdfs.semanticscholar.org/

66dc/7d32429e5e5906d1026706eb3a716b67bd3b.pdf (accessed Sep. 3, 2019).

[22] I. Mamun. “Image classification using SSIM.” https://towardsdatascience.com/image-classification-using-ssim-34e549ec6e12 (accessed Sep. 3, 2019).

[23] The MathWorks, Inc. “ssim.” https://www.mathworks.com/help/images/ref/ssim.html (accessed Sep. 3, 2019)