

หุ่นยนต์ดีบีเสมือน*

สุภัทรา ปลื้มกมล¹⁾ ณัฐพงษ์ แซ่จิ๋ว²⁾ และธวัชชัย ธาณี³⁾

¹⁾ ผู้ช่วยศาสตราจารย์ ภาควิชาวิศวกรรมเครื่องกล คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น 40002

^{2), 3)} นักศึกษา ภาควิชาวิศวกรรมเครื่องกล คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น 40002

Email : supple@kku.ac.th

บทคัดย่อ

ภาษามือเป็นภาษาที่มีความสำคัญ ที่ใช้ในการสื่อสารกับคนที่มีความพิการทางด้านการรับฟังและการพูด ซึ่งปัจจุบันบุคลากรทางด้านภาษามือยังมีอยู่น้อยมาก แต่แนวคิดทางด้านภาษามือก็ได้รับการพัฒนาให้สอดคล้องกับเทคโนโลยีที่เปลี่ยนแปลงไปอย่างหลากหลายและทันสมัย ดังนั้นงานวิจัยนี้จึงได้สร้างทำภาษามือเสมือนที่สามารถคัดแยกและจัดจำภาษามือไทยโดยใช้ระบบโครงข่ายประสาทเทียมทำงานแบบไม่มีการเชื่อมโยงกับผู้ใช้งานแบบเวลาจริง (Off-Line) ผ่านคีย์บอร์ดและเมาส์ เพื่อเป็นล่ามภาษามือให้คนพิการทางหูและคนพิการทางตาสื่อสารกันเข้าใจมากขึ้น

หุ่นยนต์ดีบีเสมือนนี้ มีขนาดครึ่งตัวประกอบด้วยมือสองข้าง ข้างละ 5 นิ้วมือและแขนสองข้าง โดยการดำเนินงานเป็นการออกแบบและสร้างมือเสมือนในโปรแกรม SolidWorks ควบคุมการเคลื่อนไหวให้เสมือนจริงที่สุดโดยใช้ SimMechanics และปรับปรุงส่วนการแสดงผลของหุ่นยนต์ให้แสดงการเคลื่อนไหวภาษามือออกมาสวยงามสมจริง โดยโปรแกรม V-Realm-Builder

ผลการทดสอบหุ่นยนต์ดีบีเสมือน ประกอบด้วย การเคลื่อนไหวของมือ 2 แบบคือการเคลื่อนไหวของมือแบบหนึ่งจังหวะและหลายจังหวะ โดยทำทางภาษามือนั้นถูกแสดงด้วยมือขวข้างเดียวและมือสองข้างตามลำดับ กรณีที่มีการเคลื่อนไหวด้วยมือขวข้างเดียวแบบหนึ่งจังหวะ จะสามารถเรียนรู้ จัดจำ และคัดแยกท่ามือของหมวดพยัญชนะ สระและตัวเลขภาษาไทย โดยแสดงท่าทางภาษามือเป็นภาพ เคลื่อนไหวพร้อมเสียงคำพูดประกอบคำที่หุ่นยนต์แสดง ตามหมวดคำที่ผู้ใช้งานต้องการเลือกชมได้อย่างรวดเร็ว และถูกต้องแม่นยำ 100% จำนวน 28 คำ โดยใช้ระบบโครงข่ายประสาทเทียมแบบป้อนไปข้าง หน้าสองชั้น ด้วยเทคนิคการเรียนรู้ของ LVQ ในกรณีที่ทำภาษามือมีการเคลื่อนไหวมือแบบหลายจังหวะและสลับซับซ้อนมากขึ้น หุ่นยนต์ดีบีเสมือนนี้สามารถแสดงการเคลื่อนไหวของมือตามหมวดคำที่ผู้ใช้งานต้องการเลือกชมได้ แต่ยังไม่สามารถเรียนรู้ จัดจำ คัดแยกท่ามือเหล่านี้ได้ เช่น หมวดคำกริยา 18 คำ คำสรรพนาม 15 คำ วันและเดือน 19 คำ และหมวดเลขจำนวน 1 เรื่องใช้เวลาประมาณ 180 วินาที แล้วเป็นภาษามือได้ประมาณ 25 ท่า ผลการแสดงผลการทำงานของหุ่นยนต์ดีบีเสมือนสามารถแสดงท่าทางภาษามือได้ค่อนข้างถูกต้อง แต่หุ่นยนต์ดีบีเสมือนนี้ยังต้องได้รับการปรับปรุงการทำงานเพื่อนำไปใช้งานที่หลากหลายมากขึ้นจนสามารถติดต่อสื่อสารกับผู้ใช้งานแบบเป็นเวลาจริงได้ต่อไป

คำสำคัญ : หุ่นยนต์ดีบีเสมือน, โครงข่ายประสาทเทียม, ภาษามือ, มือเสมือน

* รับต้นฉบับเมื่อวันที่ 19 พฤษภาคม 2548 และได้รับบทความฉบับแก้ไขเมื่อวันที่ 29 กันยายน 2548

Virtual Sign Language Robot Arm *

Supattra Plermkamon¹⁾ Nattapong Saejew²⁾ and Thawatchai Thanee³⁾

¹⁾ Assistant Professor Department Mechanical Engineering, Faculty of Engineering, Khon Kaen University 40002.

^{2), 3)} Student, Department Mechanical Engineering, Faculty of Engineering, Khon Kaen University 40002.

Email : supple@kku.ac.th

Abstract

Sign languages are the important language that have been used in communications for deaf and dumb person. Currently, there are rarely people who know how to do the sign language. However, sign languages have become an essential part of technological revolution. This research presents a new scheme for hand posture selection and recognition based on artificial neural network. Users can interact with this model by using keyboard and mouse in form of off-line operating in order to translate the virtual sign languages for deaf and dumb person to communicate friendly.

Virtual sign language robot arm model consists of an animated half human body, two arms and two hands with each five fingers. The virtual sign language robot arm was produced by SolidWorks program and using SimMechanics to control the movement of virtual hand gesture/posture. The V-Realm Builder program was used to improve and decorate the display part of this model as reality.

The testing result can be divided into 2 cases of hand movement; a single and double hand movement. The right hand side and both hand sides were used for display sign language for single and double hand movement, respectively. For a single hand movement, it can learn, recognize and classify the alphabet, vocal and number of Thai language by a two layer feed forward neural network and animate with sound track according to the required word of user. It can display 28 words of sign language very fast and 100% of accuracy. LVQ method was used for training. For double hand movement, it can display sign language of verb, pronoun, day and month, and tale with 18 words, 15 words, 19 words and 1 story, respectively. It can not learn, recognize and classify word. The time consuming for tale is 180 seconds with 25 signs. The virtual sign language robot arm can accurately display sign language, however, for more general purposes, this model has to be further developed in the filed of on-line communication with human being.

Keywords : virtual sign language, SimMechanics, V-Realm Builder, LVQ Network

* Original manuscript submitted: May 19, 2005 and Final manuscript received: September 29, 2005

บทนำ

กรอบแนวคิดของงานหุ่นยนต์ดีโอบี

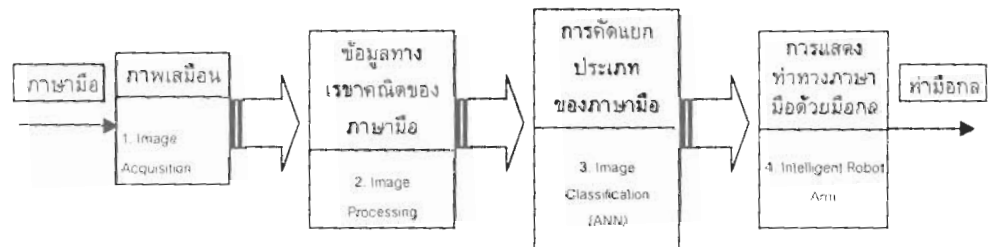
ยุคแห่งข้อมูลข่าวสารในปัจจุบันถือเป็นปัจจัยที่สำคัญยิ่งของคนทั่วโลกเพราะการสื่อสารทุกรูปแบบจะก่อให้เกิดการเรียนรู้ การพัฒนาตนเองให้มีศักยภาพเพียงพอต่อการดำรงชีวิตอยู่ได้ ความเข้าใจซึ่งกันและกัน แต่รูปแบบของการสื่อสารเหล่านั้นส่วนมากต่างก็เอื้ออำนวยต่อคนปกติ ส่วนภาษา มือ (ดีโอบี) กลับถูกใช้เป็นภาษาสัญลักษณ์สำหรับบุคคลที่มีความบกพร่องทางการได้ยินและการพูด จึงส่งผลให้กระบวนการสื่อสารระหว่างคนปกติกับผู้ที่มีความบกพร่องทางการได้ยินและการพูดยิ่งประสบความยากลำบากยิ่ง การกำหนดรูปแบบภาษามือขึ้นมาใช้แต่ละประเทศต่างมีเอกลักษณ์เฉพาะตัวโดยใช้การเลียนแบบภาษาธรรมชาติ สำหรับรูปแบบการใช้ภาษามือไทยมีสองรูปแบบ รูปแบบแรกคือการสะกดนิ้วมือ รูปแบบที่สองคือท่ามือ ท่าทาง กริยา แต่อย่างไรก็ตามองค์ความรู้ที่มีอยู่ ทางด้านนี้ของประเทศไทยยังอยู่ในลักษณะที่เป็นทักษะและประสบการณ์เฉพาะตัวของแต่ละบุคคล ยากต่อการสืบทอดให้มีมาตรฐานทางการแพทย์ เทคโนโลยีทางการศึกษา และเทคโนโลยีสารสนเทศได้อย่างสมบูรณ์ และจำนวนบุคลากรทางด้านนี้นับได้ว่ามีน้อยมาก กล่าวคืออัตราส่วนของผู้ที่ทำหน้าที่ล่ามต่อคนที่มีความบกพร่องทางการได้ยิน เฉพาะเขตกรุงเทพฯ และปริมณฑลมีค่าเท่ากับ 1:465 เท่านั้น อีกทั้งการจัดทำสื่อการเรียนรู้ภาษามือทั้งของไทย [NECTEC 2543, สยามและคณะ 2545] และต่างประเทศ โดยส่วนใหญ่ใช้คนทำหน้าที่แสดงท่าทางภาษามือที่เป็นเวลาจริง [Fillbrandt, F., et al 2003, Hamada, Y., et al 2004] แต่การใช้คนเป็นสื่อกลาง อาจจะไม่สะดวกต่อผู้ใช้งานที่จะทำการแก้ไขปรับปรุงและเพิ่มเติมท่าภาษามือให้สอดคล้องกับปัจจุบัน ที่มีความทันสมัยอยู่เสมอ ดังนั้นเพื่อลดเซกการขาดแคลนบุคลากรทางด้านนี้และผู้ใช้งานมีความสะดวกมากขึ้น แนวคิดที่จะพัฒนาเทคโนโลยีหุ่นยนต์ดีโอบีเสมือนจึงเกิดขึ้น ด้วยการนำเทคโนโลยีทางด้านปัญญาประดิษฐ์ซึ่งเป็นเทคโนโลยีขั้นสูง เช่น ระบบโครงข่ายประสาทเทียมมาผนวกผสมผสานเข้ากับศาสตร์สาขาอื่นๆ เช่น หุ่นยนต์ (Robot) การเขียนแบบขั้นสูง จักรกลวิทัศน์ (Machine Vision) จักรกลเรียนรู้แบบปรับตัวได้ (Adaptive Machine Learning) วิศวกรรม การแพทย์และหลักการสื่อสารภาษามือเพื่อก่อให้เกิดองค์ความรู้ใหม่ โดยมุ่งเน้นพัฒนาสร้างหุ่นยนต์ดีโอบีเสมือนขึ้นมาใช้เองภายในประเทศ สามารถเลียนการทำงานมือของคนติดต่อสื่อสารกับคน ในฐานะเป็นล่ามภาษามือที่มีศักยภาพ เพราะการใช้คนเป็นสื่อกลางทำหน้าที่แปลข้อมูลเช่น ข่าวสารต่างๆ ผ่านหน้าจอเครื่องรับภาพ (Television) ห้องประชุม ห้องสัมมนาเป็นภาษามือโดยตรงนั้น คนอาจเกิดความเหนื่อยล้าได้มาก ส่งผลต่อความสม่ำเสมอของผลงานให้ด้อยลงตามไปด้วย

การออกแบบและการทำงานของหุ่นยนต์ดีโอบีเสมือน

“หุ่นยนต์ดีโอบีเสมือน” เป็นหุ่นเสมือนที่ทำหน้าที่เป็นล่ามมือในระบบภาษาไทยต้นแบบ เพื่อให้ผู้บกพร่องทางการได้ยินและการพูดใช้เป็นเครื่องมือสำหรับเรียนรู้ภาษามือด้วยตนเอง และใช้ติดต่อกับ

ผู้อื่นในรูปแบบที่สนุก ไม่น่าเบื่อ สามารถเพิ่มเติมข้อมูลภาษามือได้ง่ายโดยไม่มีข้อจำกัด มีมาตรฐานการใช้งาน นำไปใช้งานที่ไหนก็ได้ ด้วยการผนวกเอาเทคโนโลยีจักรกลวิทัศน์ร่วมกับ

ระบบโครงข่ายประสาทเทียม สร้างเป็นแบบจำลองการทำงานของหุ่นยนต์ไบนีเลียนปัญหาประดิษฐ์ขึ้นดังรูปที่ 1 ทำงานในลักษณะที่ไม่มีการเชื่อมโยง (Off-Line) กับผู้ใช้งานโดยตรง แยกการทำงานเป็น 2 กรณี กรณีที่ 1 แสดงท่าภาษามือกลแบบเคลื่อนไหวจังหวะเดียว (Static) ด้วยมือขวาข้างเดียว [ทินกรและคณะ 2545] กรณีที่ 2 แสดงท่าภาษามือกลแบบเคลื่อนไหวหลายจังหวะ (Dynamic) ด้วยมือสองข้าง [ณัฐพงษ์และคณะ 2547] ซึ่งข้อกำหนดการทำงานของ computer ที่ใช้ในงานวิจัยนี้คือ Intel Celeron processor หน่วยความจำ 128 MB, ความเร็ว 463 MHz สำหรับกรณีที่ 1 ประกอบด้วย 4 ขั้นตอน เริ่มด้วยกระบวนการจัดหาข้อมูลภาพ (Image Acquisition) ของภาษามือจากการสร้างภาพ 3 มิติตามหมวดหมู่ภาษามือใน Solid Works แล้วใช้โปรแกรม SimMechanics ควบคุมตำแหน่งความเร็ว และความเร่งของการเคลื่อนที่ ขั้นตอนที่ 2 ทำการประมวลผลหลักขณะเด่นจำเพาะแต่ละท่าทางของภาษามือด้วยเทคนิคกระบวนการประมวลผลและวิเคราะห์ผลภาพ (Image Processing & Image Analysis) เพื่อใช้เป็นฐานข้อมูลนำเข้าสู่กระบวนการแปลผลภาพ ในขั้นตอนที่ 3 เป็นขั้นตอนที่ทำงานโดยระบบโครงข่ายประสาทเทียม (Artificial Neural Network) ชนิดหลายชั้นแบบไม่ย้อนกลับ (Multilayer Feed Forward) เพื่อจำแนกและจดจำข้อมูล แล้วแสดงผลท่าภาษามือเป็นภาพเคลื่อนไหวตามลักษณะของคำในหมวดต่างๆ ของภาษามือที่ต้องการเป็นขั้นตอนสุดท้าย สำหรับการทำงานกรณีที่ 2 จะแตกต่างจากกรณีที่ 1 ที่สามารถเรียนรู้ จดจำ จำแนกและแสดงตัวอักษร สระ ตัวเลขได้ ส่วนกรณีที่ 2 มีการทำงานเพียงหนึ่งขั้นตอนคือกระบวนการจัดหาข้อมูลภาพ (Image Acquisition) ท่าภาษามือที่กระบวนการสร้างภาพการเคลื่อนไหวของมือ มีหลายจังหวะและสลับซับซ้อนเพิ่มมากขึ้นจากการใช้มือขวาข้างเดียวและเคลื่อนไหวจังหวะเดียว เป็นใช้มือและท่อนแขนสองข้าง เพื่อแสดงท่าทางของหมวดคำกริยา คำสรรพนาม วัน เดือน และการเล่านิทาน



รูปที่ 1 กล้องแผนภาพแสดงลำดับการทำงานของหุ่นยนต์ไบนีเลียน

ขั้นตอนการออกแบบและสร้างหุ่นยนต์ไบนีเลียนดังรูปที่ 1 ประกอบด้วยขั้นตอนที่ 1 จัดหาข้อมูลภาพ (Image Acquisition) ด้วยการออกแบบและสร้างหุ่นยนต์ครึ่งตัวที่ประกอบด้วยชิ้นส่วนของนิ้ว 5 นิ้ว มือและท่อนแขน 2 ข้าง ลำตัว และศีรษะ แบบ 3 มิติใน SolidWorks จากนั้นเชื่อมโยงเพิ่มข้อมูลดังกล่าวกับโปรแกรม SimMechanics เพื่อควบคุมการเคลื่อนที่ของชิ้นส่วนเชิงกลเหล่านี้เช่น แรง ความเร็ว ความเฉื่อย และความเร่ง ให้มีการเคลื่อนไหวได้เสมือนจริงมากที่สุดดังรูปที่ 2 และเก็บ

บันทึกเพิ่มข้อมูลเป็น *.xml แล้วใช้โปรแกรม V-Realm-BUILDER ปรับแต่ง หุ่นยนต์ให้แสดงท่าภาษามือ ออกมาสวยงามสมจริงดังรูปที่ 3 โดยเก็บเพิ่มข้อมูลเป็น *.wrl

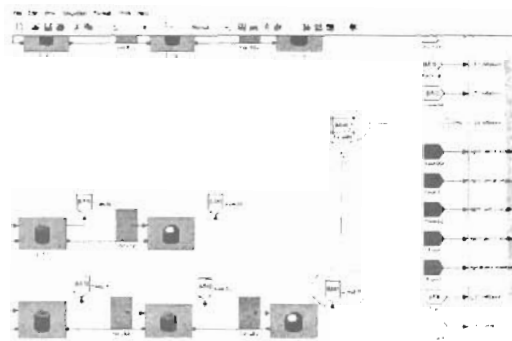


(ก)

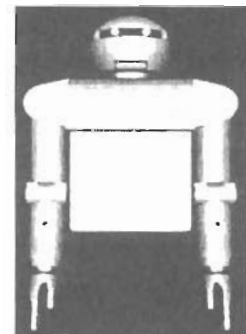


(ข)

รูปที่ 2 (ก) ภาพหุ่นยนต์ดีโบบ์เสมือนที่สร้างขึ้น (ข) ภาพหุ่นยนต์เสมือนในแฟ้มข้อมูล *. Xml



(ก)



(ข)

รูปที่ 3 (ก) ภาพหุ่นยนต์ดีโบบ์เสมือนที่ปรับปรุงจากรูปที่ 2 (ข) ภาพหุ่นยนต์เสมือนในแฟ้มข้อมูล *. wrl

ขั้นตอนที่ 2 หาลักษณะเด่นจำเพาะ **P** (Image Feature) แต่ละท่าทางของภาษามือด้วยเทคนิค กระบวนการประมวลผลและวิเคราะห์ผลภาพ จะได้ลักษณะเด่นทางเรขาคณิตของข้อมูลภาพภาษามือด้วย มือหนึ่งข้าง (กรณีนี้ 1) เช่น โมเมนต์ที่ไม่ผันแปรเชิงพีชคณิต 7 พารามิเตอร์ (Seven Invariant Moments) [Gonzalez, et al. 2002] ดังสมการที่ 1

$$\begin{aligned}
\phi_1 &= \mu_{20} + \mu_{02} \\
\phi_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \\
\phi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \\
\phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \\
\phi_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + \\
&\quad (\mu_{03} - 3\mu_{21})(\mu_{03} + \mu_{21})x[(\mu_{03} + \mu_{21})^2 - 3(\mu_{12} + \mu_{30})^2] \\
\phi_6 &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12}) \\
&\quad (\mu_{03} + \mu_{21}) \\
\phi_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\
&\quad - (3\mu_{12} - \mu_{30})(\mu_{03} + \mu_{12})x[(\mu_{03} + \mu_{21})^2 - 3(\mu_{12} + \mu_{30})^2]
\end{aligned} \tag{1}$$

โดยที่ $\phi_1, \dots, \phi_6$ เป็นโมเมนต์ที่ผันแปรเชิงพีชคณิตต่อการหมุนและเงาสะท้อนของมือกล แต่ ϕ_7 ยังคงไม่ผันแปรต่อการหมุนของมือกล แต่จะเปลี่ยนแปลงเครื่องหมายต่อเงาสะท้อนของมือกล โมเมนต์ที่ไม่ผันแปรนี้ไม่ผันแปรต่อการเปลี่ยนแปลงเชิงเรขาคณิตของมือกล แต่จะขึ้นอยู่กับ ลักษณะรูปทรงของมือกล แต่เพียงอย่างเดียวเท่านั้น พารามิเตอร์ที่แสดงลักษณะเด่นอื่นๆของมือกลที่ใช้พิจารณาประกอบเพิ่มเติมด้วยคือ พื้นที่ภายในเส้นรอบรูป (Boundary Area) เส้นรอบรูปตามลักษณะรูปทรง (Boundary Perimeter) จุดศูนย์ถ่วง (Center of Gravity) ความยาวตามแนวแกนที่สั้นที่สุด (Minor Axis Length) และความยาวตามแนวแกนที่ยาวที่สุด (Major Axis Length) กรอบของภาพ (Bounding Box) มุมจัดวาง (Orientation) เพื่อใช้พารามิเตอร์เหล่านี้ (ดังตารางที่ 1) เป็นฐานข้อมูลนำเข้าระบบฝึกการเรียนรู้ จัดจำ และแปลผลมือกล ในขั้นตอนที่ 3 ขอสั่งเกิดจากตารางที่ 1 สำหรับท่าทางมือที่มีความใกล้เคียงกันมากเช่นตัวเลข “6” กับพยัญชนะ “ว” และตัวเลข “4” กับพยัญชนะ “บ” จะมีค่าของลักษณะเด่นที่แตกต่างกันชัดเจน

ขั้นตอนที่ 3 กระบวนการจัดจำ คัดแยกและแปลผลภาพ (Image Classification) ภาษามือให้กับหุ่นยนต์ไบบีมือน โดยเน้นการแสดงผลท่าภาษามือด้วยมือขวาเพียงข้างเดียว (กรณีที่ 1) จัดหมวดหมู่ของพยัญชนะ ตัวเลขและสระ เป็น 6 กลุ่มใหญ่ (Class) และ 28 กลุ่มย่อย (Subclass) ดังตารางที่ 1 ด้วยเทคนิคของโครงข่ายประสาทเทียม (Artificial Neural Network) สถาปัตยกรรมของโครงข่ายที่เลือกใช้งานมีโครงสร้างเป็น 2 ชั้นเพราะว่าเป็นการคัดแยกสมาชิกของโครงข่ายเป็นสองกลุ่มคือ subclass เพื่อตรวจสอบจำนวนนิ้วที่ใช้แสดงท่าทาง และ class เพื่อแปลความหมายของท่าทางที่แสดงนั้น ซึ่งถือว่าเพียงพอต่อการคัดแยกข้อมูลในงานวิจัยนี้ หากมีจำนวนชั้นมากกว่านี้จะทำให้กระบวนการประมวลผลใช้เวลามากขึ้นเกินความจำเป็น โดยแบ่งโครงสร้างออกเป็น 3 ส่วน คือ 28-28-6 ดังรูปที่ 4 ส่วนที่ 1 ทำหน้าที่เป็นส่วนรับข้อมูลลักษณะเด่นจำเพาะ P ของข้อมูลภาษามือจำนวน R ตัว ส่วนที่ 2 เป็นโครงข่ายชั้นแรกเรียกว่าชั้นที่มีการแข่งขันเพื่อหาผู้ชนะเพียงหนึ่งเดียว (Competitive Layer) เขียนแทนด้วยฟังก์ชัน C มีกระบวนการเรียนรู้แบบปริมาตรผลต่างของเวกเตอร์ (Learning Vector Quantization, LVQ) โดยคำนวณหาจากระยะห่างระหว่างตัวแปรขาเข้า P และตัวแปรปรับน้ำหนัก W^i โดยตรง เขียนความสัมพันธ์ได้ดังสมการที่ 2

$$\mathbf{n}^1 = \begin{bmatrix} \|_1 \mathbf{W}^1 - \mathbf{P} \| \\ \|_2 \mathbf{W}^1 - \mathbf{P} \| \\ \vdots \\ \|_n \mathbf{W}^1 - \mathbf{P} \| \end{bmatrix} \tag{2}$$

ตัวแปรสุทธิขาออกจากโครงข่ายชั้นแรกนี้คือ

$$\mathbf{a}^1 = \text{compet}(\mathbf{n}^1) \tag{3}$$

จากสมการที่ 3 ถ้านิวรอนตัวใดมีค่าผลต่างดังสมการที่ 2 น้อยที่สุดแล้วถือว่านิวรอนตัวนี้คือผู้ชนะเพียงหนึ่งเดียว และถือว่าสมาชิกของโครงข่ายที่เป็นผู้ชนะเพียงหนึ่งเดียวนี้นี้ หมายถึงจำนวนนิ้วที่ใช้ แสดงท่าภาษามือของกลุ่มย่อยในฐานข้อมูลทั้งหมดที่ถูกสอนไว้ เขียนความสัมพันธ์ได้ดังสมการที่ 4

$$\mathbf{n}^2 = \mathbf{W}^2 \mathbf{a}^1 = \mathbf{a}^2 \tag{4}$$

โดยที่

- $\mathbf{W}^1$ = ตัวแปรปรับน้ำหนักในชั้นโครงข่ายชั้นแรก
- $\mathbf{W}^2$ = ตัวแปรปรับน้ำหนักในชั้นโครงข่ายชั้นที่สองซึ่งทราบล่วงหน้าด้วยการจัดให้สดมภ์ (column) แทนถึงกลุ่มย่อยท่าภาษามือที่สอนทั้งหมดแยกตามจำนวนนิ้วที่ใช้ ส่วนแถว (row) แทนถึงกลุ่มใหญ่ของความหมายแต่ละท่าภาษามือ

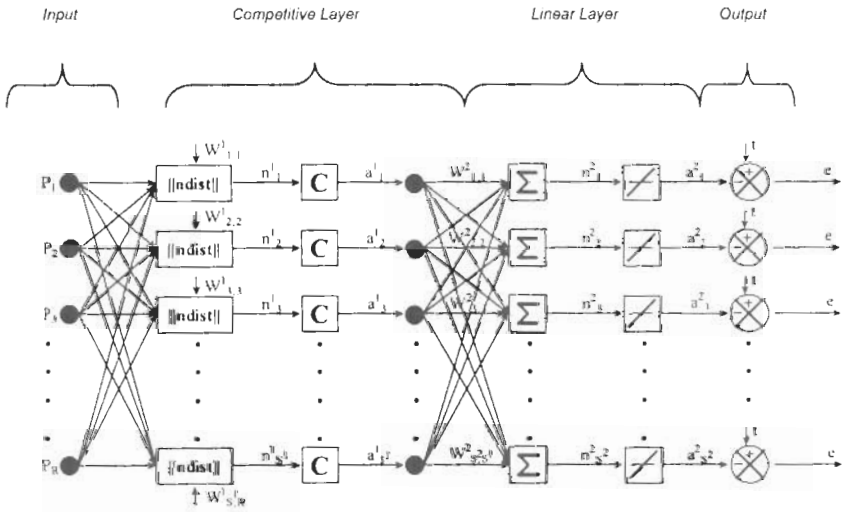
สำหรับโครงข่ายชั้นที่ 2 ส่วนที่ 3 เรียกว่าชั้นเชิงเส้น (Linear Layer) ทำหน้าที่แปลท่าภาษามือที่แสดงว่ามีความหมายอย่างไร ดังรูปที่ 5 และ 6

ขั้นตอนที่ 4 เป็นการแสดงลำดับภาพการเคลื่อนไหวของท่าภาษามือตามหมวดคำที่ต้อง การเลือกชม เช่นถ้าผู้ใช้งานต้องการเลือกชมท่าภาษามือของคำว่า “หมายเลข 2” ระบบการทำงานจะค้นหาความหมายจากระบบหน่วยความจำของหมวดคำที่สอนไว้ทั้งหมด พบว่าตรงกับหมวดคำกลุ่มใหญ่ที่ 3 กลุ่มย่อยที่ 21 พร้อมกับแสดงท่าภาษามือและออกเสียงเป็นภาษาไทยว่า “หมายเลข 2” โดยเรียงลำดับการแสดงผลเป็น 4 ลำดับ ดังรูปที่ 5 ระยะเวลาที่ใช้ในการประมวลผลจะขึ้นอยู่กับความสา มารถของ computer สำหรับงานวิจัยนี้ใช้เวลาทั้งหมดประมาณ 15 วินาทีต่อหนึ่งท่าภาษามือ ผลการทำงานของหุ่นยนต์ดีไปเสมือนนี้มีความคลาดเคลื่อนเป็นศูนย์ ดังรูปที่ 6 แสดงว่าระบบการทำงานมีความสามารถในการจดจำและคัดแยกลักษณะท่าทางของภาษามือที่ทดสอบ เพื่อแปลความ หมายท่าภาษามือได้อย่างถูกต้องแม่นยำสูงมาก รูปที่ 7 เป็นแผนภาพแสดงลำดับขั้นตอนการทำงานที่สมบูรณ์ของกรณีแรก (มือขวาข้างเดียว)

Class**	Sub Class*				Perimeter	Area	Bounding Box	Orientation
	No.	Symbol	Real	Virtual				
0	3	ม			1.797	55.708	259	69.91
1	11	สระ “ ไ ”			2.574	65.331	305	63.91
2	14	พ			2.888	70.218	501	-11.94
3	20	6			3.180	55.759	206	86.59
3	25	ว			3.946	59.331	213	87.08
4	26	4			3.604	59.092	207	81.14
4	27	ป			3.748	63.280	214	80.95
5	28	5			4.187	61.209	204	73.72

ตารางที่ 1 ลักษณะเด่น (Feature Vector) ของท่าภาษามือสำหรับมือขวาข้างเดียว

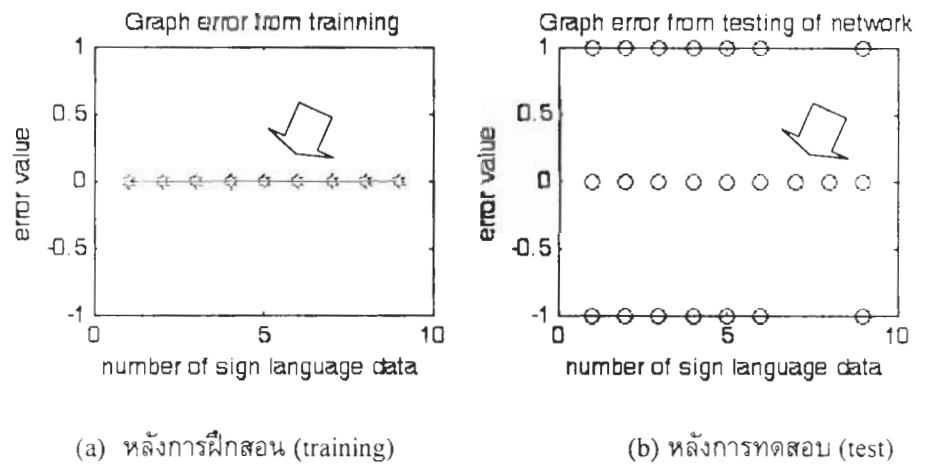
** Class = ความหมายของท่าภาษามือที่แสดงต่อจำนวนนิ้วที่ใช้ในการตีไปแต่ละท่าทาง.
* Sub Class = กลุ่มย่อยของท่าภาษามือแยกตามจำนวนนิ้วเป็น 1,2,3,4 และ 5 นิ้ว
ไม่ได้แสดงข้อมูล โมเมนต์ที่ไม่มันแปรเชิงพีชคณิต 7 พารามิเตอร์ในตารางที่1



รูปที่ 4 โครงสร้างของระบบโครงข่ายประสาทเทียมแบบมีที่ปรึกษาไม่ย้อนกลับหลายชั้น



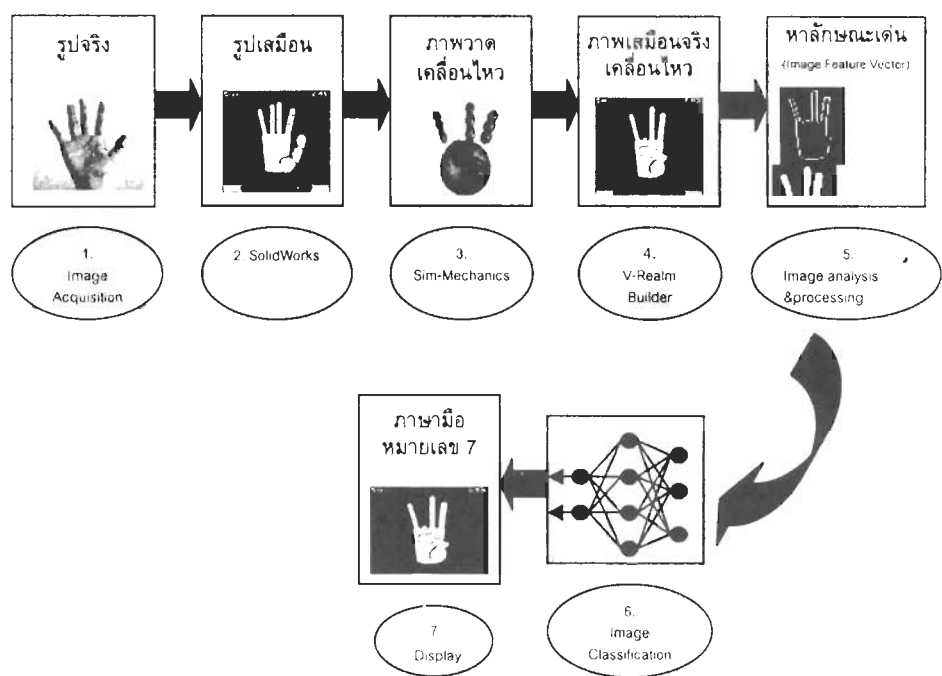
รูปที่ 5 ลำดับภาพเรียงลำดับจากหมายเลข 1 ถึง 4 แสดงแบบท่าภาษามือ "หมายเลข 2" ของหุ่นยนต์ดีโอบีเสมือน



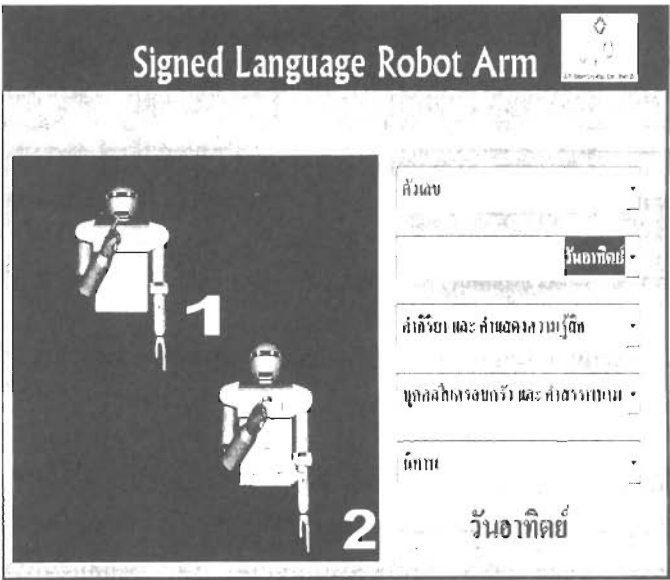
รูปที่ 6 ค่าความผิดพลาดในการจัดจำข้อมูลทั้งหมดของระบบการทำงานหุ่นยนต์ดีโอบีเสมือนในกรณีที่ 1 (มือขวาข้างเดียว)

สำหรับกรณีที่ 2 การทำงานของหุ่นยนต์ดีโอบีเสมือนจะดำเนินการคล้ายกับกรณีแรก ดังรูปที่ 7 เฉพาะขั้นตอนที่ 1 ถึงขั้นตอนที่ 4 เท่านั้นแต่การทำงานยังไม่มีระบบจดจำ คัดแยกข้อมูลท่าภาษามือ เพราะจะเน้นที่ความสามารถของการแสดงท่าภาษามือที่มือเสมือนมีการเคลื่อนไหวหลายจังหวะและมีความสลับซับซ้อนของการเคลื่อนไหวตามท่าภาษามือมากขึ้น ในหมวดแสดง วันและเดือน 19 คำ (ดังรูปที่ 8) หมวดคำกริยา 18 คำ (ดังรูปที่ 9 และรูปที่ 10) คำสรรพนาม 15 คำ และหมวดเล่าเหตุการณ์

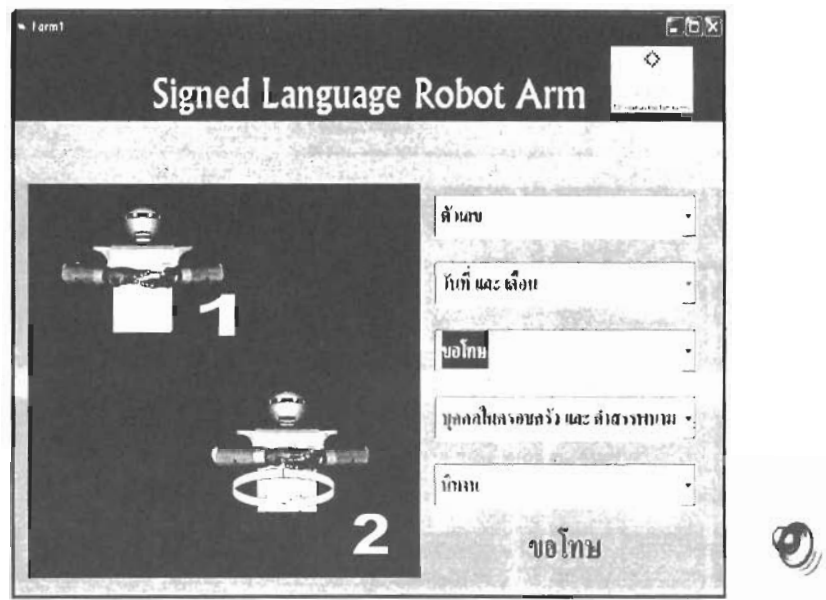
กระด่ายกับเต่า 1 เรื่อง (ดังรูปที่ 11) โดยเล่าเป็นท่าภาษามือได้ประมาณ 25 ท่าทาง ใช้เวลาประมาณ 180 วินาที



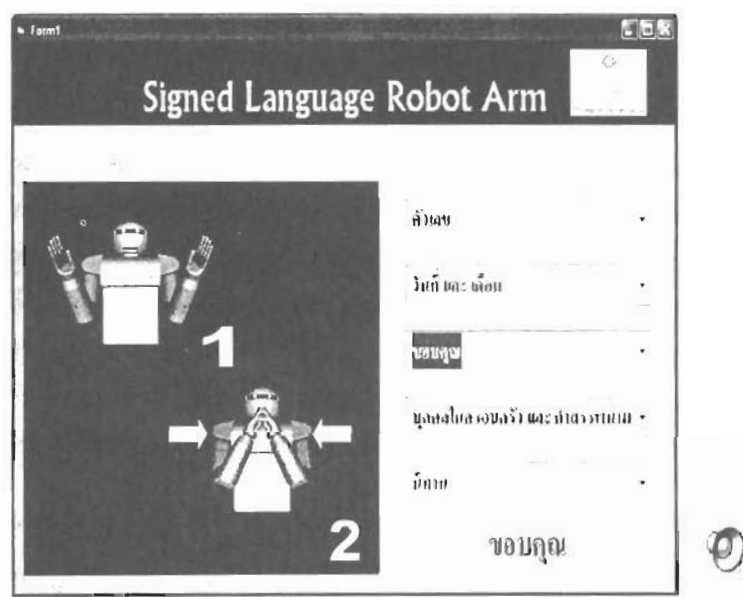
รูปที่ 7 แผนภาพกระบวนการทำงานของหุ่นยนต์ดีโอบีเสมือน



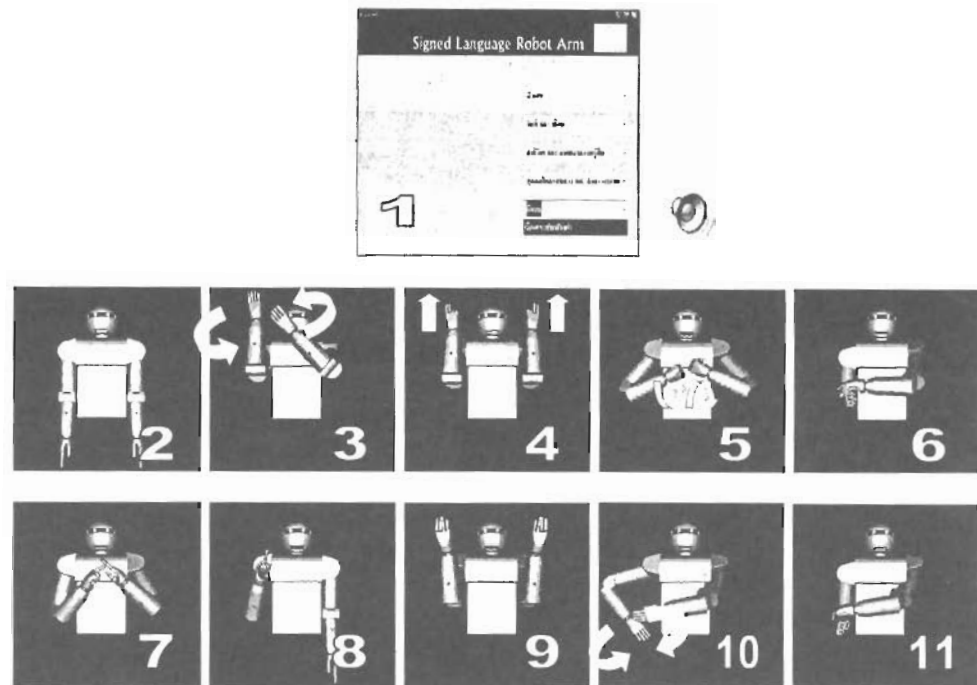
รูปที่ 8 ลำดับภาพหุ่นยนต์ดีโอบีเสมือนแบบท่ามือ หมวดวัน "วันอาทิตย์"



รูปที่ 9 ลำดับภาพหุ่นยนต์ดีดีไปเสมือนแบบทำมือหมวดคำกริยา “ขอโทษ”



รูปที่ 10 ลำดับภาพหุ่นยนต์ดีดีไปเสมือนแบบทำมือ หมวดคำกริยา “ขอบคุณ”



รูปที่ 11 ลำดับภาพหุ่นยนต์ดีป้าเสมือนแสดง หมวดนิทานเรื่อง "กระต่ายกับเต่า" (บางส่วน)

บทสรุปและข้อเสนอแนะ

ผลสำเร็จของหุ่นยนต์ดีป้าเสมือน ที่ทำงานด้วยเทคโนโลยีจักรกลวิทัศน์ปัญญาประดิษฐ์ โดยไม่มีการเชื่อมโยงกับผู้ใช้งานแบบเวลาจริง สำหรับกรณีแรกสามารถแสดงการเคลื่อนไหวมือ (Animation) ที่มีการเคลื่อนไหวมือเพียงหนึ่งจังหวะด้วยมือขวาข้างเดียวพบว่าสามารถเรียนรู้ จดจำ คัดแยกท่ามือ สำหรับแสดง พยัญชนะ สระ และตัวเลขภาษาไทย ได้ถูกต้องจำนวน 28 คำ โดยเฉพาะอย่างยิ่งท่าทาง ภาษามือที่มีความใกล้เคียงกันมากเช่นตัวเลข "6" กับพยัญชนะ "ว" และตัวเลข "4" กับพยัญชนะ "บ" ดังตารางที่ 1 ก็สามารถคัดแยกได้อย่างถูกต้อง ระยะเวลาที่ใช้ในการประมวลผลจะขึ้นอยู่กับความสามารถของ computer สำหรับงานวิจัยนี้ใช้เวลาทั้งหมดประมาณ 15 วินาทีต่อหนึ่งท่าทางภาษามือ ที่ข้อกำหนดการทำงานของ computer ที่ใช้คือ Intel Celeron processor หน่วยความจำ 128 MB, ความเร็ว 463 MHz

กรณีที่สองหุ่นยนต์ดีป้าเสมือนนี้สามารถแสดงท่าทางภาษามือที่มีการเคลื่อนไหวมือ มีหลายจังหวะ และสลับซับซ้อนมากขึ้น เช่น หมวดคำกริยา 18 คำ คำสรรพนาม 15 คำ วันและเดือน 19 คำ และหมวดเล่านิทานเรื่องกระต่ายกับเต่า 1 เรื่อง เล่าเป็นท่าทางภาษามือได้ประมาณ 25 ท่าทาง ใช้เวลาประมาณ 180 วินาที แต่ยังไม่สามารถเรียนรู้ จดจำ คัดแยกท่ามือที่สลับซับซ้อนเหล่านี้ได้ อย่างไรก็ตาม ทั้งสองกรณีสามารถทำงานโดยแสดงผลท่าทางภาษามือและเสียงออกทางจอภาพ ตามหมวดคำที่ผู้ใช้งานต้องการเลือกชมได้อย่างรวดเร็วและค่อนข้างถูกต้อง

ถึงแม้ว่าการทำงานของหุ่นยนต์ดีไปเสมือน จะสามารถทำงานตามความต้องการของผู้ใช้งานได้ในระดับหนึ่ง แต่หุ่นยนต์ดีไปก็ยังคงได้รับการพัฒนาในหลายๆส่วนเพื่อให้นำไปใช้งานได้อย่างมีประสิทธิภาพสูงขึ้นเทียบเท่าความสามารถของคนที่ทำหน้าที่เป็นล่ามภาษามือโดยตรง เช่น พัฒนาในส่วนของการแปลภาษามือในลักษณะบทสนทนากับผู้ใช้งานแบบเวลาจริง (On-line) โดยสามารถรับข้อมูลได้หลายรูปแบบเช่นเป็น ข้อความ เสียงพูดและข้อมูลภาพผ่านกล้องจับภาพเพื่อให้หุ่นยนต์ดีไปเสมือน แปลทำภาษามือแสดงผลออกมาโดยอัตโนมัติแบบเวลาจริง เป็นข้อความ เสียงและท่าทางภาษามือประกอบบนจอภาพ

ดังนั้นหากมีการนำหุ่นยนต์ดีไปเสมือน ไปประยุกต์ใช้งานในอนาคตอย่างจริงจังย่อมจะก่อให้เกิดประโยชน์ต่อผู้ที่บกพร่องทางการรับฟัง การพูดและการมองเห็น เป็นอย่างยิ่งต่อการพัฒนาคุณภาพชีวิตที่ดีขึ้นทั้งในประเด็นพื้นฐานสมรรถภาพของการเรียนรู้ การศึกษา การสื่อสาร การสาธารณสุข การทำงาน และสังคมเพื่อที่จะดำรงชีวิตได้เช่นคนปกติทั่วไป ซึ่งผลลัพธ์ที่ได้ย่อมเป็นดัชนีชี้วัดผลสำเร็จของงานหุ่นยนต์ดีไปเสมือนได้เป็นอย่างดี

กิตติกรรมประกาศ

ขอขอบพระคุณมหาวิทยาลัยขอนแก่นที่ให้การสนับสนุนงบประมาณโครงการวิจัยอุดหนุนทั่วไป ประจำปีงบประมาณ 2548 สำหรับการดำเนินงานวิจัยหุ่นยนต์ดีไปเสมือนนี้

เอกสารอ้างอิง

- ณัฐพงษ์ แซ่จิวและคณะ. 2547. หุ่นยนต์ดีไป. [ปริญญาณิพนธ์] ภาควิชาวิศวกรรมเครื่องกล คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น.
- ทินกร มินตวและคณะ. 2546. มือกลสื่อสารภาษามือด้วยระบบปัญญาประดิษฐ์. [ปริญญาณิพนธ์] ภาควิชาวิศวกรรมเครื่องกล คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น.
- นรินทร์ สันติตระกูล. 2522. ภาษามือ ฉบับประยุกต์ไทย-อังกฤษ 1. มหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร
- สยาม เจริญเสียงและวุฒิชัย วิศาลคุณา. 2545. ระบบการจดจำภาษามือไทยแบบคำเดี่ยวโดยใช้ฮาร์ดแวร์คอมพิวเตอร์. การประชุมวิชาการทางวิศวกรรมไฟฟ้าครั้งที่ 25 (EECON-25) เล่ม 2. มหาวิทยาลัยสงขลานครินทร์ สงขลา. 21-22 พฤศจิกายน 2545: 46-51.
- Fillbrandt, H., Akyol, S., and Kraiss, K.F. 2003. Extraction of 3D Hand Shape and Posture from Image Sequences for Sign Language Recognition. Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG'03)
- Gonzalez, R.C. and Woods, R.E. 2002. Digital Image Processing, 2nd ed. New Jersey : Prentice Hall, Inc.

Hamada,Y., Shimada,N. and Shirai,Y. 2004. **Hand Shape Estimation under Complex Backgrounds for Sign Language Recognition.** Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)
<http://www.nectec.or.th/courseware/signlanguage/index.html>, 13 September 2004