NUEJ
NARESUAN UNIVERSITY
ENGINEERING JOURNAL

การรู้จำท่ารำวงมาตรฐานโดยการวิเคราะห์ข้อมูล 3 มิติ
Recognition of Standard Thai Traditional Dance Through 3D Data Analysis

วิศรุต ขวัญคุ้ม[1], ไพศาล มุณีสว่าง[2*]

ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร
E-mail: [1]wisrutk57@nu.ac.th, [2*]paisarnmu@nu.ac.th

บทคัดย่อ

เทคนิคทางด้านการรู้จำและการแปลความหมายท่าทางการเคลื่อนไหวของมนุษย์ เป็นที่นิยมอย่างกว้างขวางในปัจจุบันและได้ถูกนำมาประยุกต์ใช้ในงานด้านการเชื่อมต่อระหว่างมนุษย์และคอมพิวเตอร์ งานวิจัยเสนอเทคนิคการรู้จำท่ารำวงมาตรฐานโดยการวิเคราะห์ข้อมูล 3 มิติ จำแนกท่ารำที่ใช้ในการรำวงมาตรฐานจำนวน 10 เพลง โดยใช้ข้อมูลจากอุปกรณ์ตรวจจับการเคลื่อนไหวที่ถูกพัฒนาขึ้นโดยบริษัท Microsoft ชื่อว่า Kinect sensor มาวิเคราะห์ร่วมกับเทคนิคโครงข่ายประสาทเทียมชนิดจัดกลุ่มเอง Self-Organizing Mapping (SOM) และสร้างรูปแบบของท่าทาง (Gesture Templates) ที่ได้จากโมเดล SOM โดยใช้เทคนิค Posture Transition Sparse Code (PTSC) พิจารณาความถี่ของการเกิดรูปแบบที่ได้จากโมเดล ผลการทดลองสรุปได้ว่าวิธีการที่นำเสนอสามารถรู้จำท่ารำวงมาตรฐานและแปลความหมายของท่ารำประกอบเพลงได้อย่างมีประสิทธิภาพ โดยผลการรู้จำท่ารำวงมาตรฐานให้ความถูกต้องโดยเฉลี่ยเท่ากับร้อยละ 82 และสามารถให้ผลการรู้จำท่ารำแบบ Real-time

คำสำคัญ: การรู้จำท่าทาง, รำวงมาตรฐาน,กล้องไคเนค, คอมพิวเตอร์วิทัศน์

Abstract

Techniques for Recognizing and describing human movements have received much attention among researchers and have been applied in a wide variety of human-computer interaction. This paper presents recognition of Thai traditional dances by analyzing 3D data. The method recognizes 10 gestures using data from Kinect sensor which is a motion sensor developed by Microsoft. Then, Self-Organizing Mapping (SOM) is applied to create gesture templates. Finally, the Posture Transition Sparse Code (PTSC) perform calculating sequence of each gesture. Experimental evaluations show that the method performed well in recognition of Thai traditional dance standards in real-time environments and provided high accuracy in classification of 10 gestures at accuracy of 82%.

Keywords: Gesture recognition, Standards Thai traditional dance, Kinect camera, Computer visions

1. บทนำ

รำวงมาตรฐานเป็นศิลปะแห่งการรำวงที่งดงาม เป็นการละเล่นพื้นบ้านอย่างหนึ่งที่บ่งบอกถึงความสนุกสนาน แสดงถึงเอกลักษณ์ของความเป็นไทย ควรค่าแก่การอนุรักษ์สืบสาน ผู้วิจัยได้ตระหนักถึงการเรียนรู้และถนอมรักษาไว้ซึ่งองค์ความรู้เหล่านี้ จึงได้เสนอแอพพลิเคชั่นในการรู้จำท่ารำวงมาตรฐาน ให้ผู้ที่สนใจสามารถเรียนรู้ฝึกฝนท่ารำวงที่ถูกต้อง และเป็นการเผยแพร่รำวงมาตรฐานให้เป็นที่รู้จักอย่างแพร่หลาย แอพพลิเคชั่นที่พัฒนาขึ้นอาศัยเทคโนโลยีคอมพิวเตอร์วิทัศน์ (Computer vision) เป็นเทคโนโลยีที่ทำให้คอมพิวเตอร์สามารถทำงานได้เหมือนกับมนุษย์สำหรับงานบางประเภท ในปัจจุบันเทคโนโลยีดังกล่าวได้ถูกนำมาประยุกต์ใช้ได้กับงานหลากหลายรูปแบบ เช่น ระบบการติดตามการเคลื่อนไหวของวัตถุที่สนใจและทำการประมวลผลข้อมูลของวัตถุนั้นจากภาพทำให้แอพพลิเคชั่นในการรู้จำกิจกรรมต่างๆ ของมนุษย์สามารถพบเห็นได้ทั่วไป เช่น การตรวจจับวัตถุในวิดีโอ การสอนเต้นรำในระบบความจริงเสมือน [1] และการสืบค้นข้อมูลจากภาพและวิดีโอ [2]

บทความนี้ได้นำเทคโนโลยีคอมพิวเตอร์วิทัศน์มาประยุกต์ใช้ในการการรู้จำท่ารำวงมาตรฐาน โดยใช้อุปกรณ์ตรวจจับความเคลื่อนไหวชนิด Kinect sensor เพื่อการบันทึกข้อมูลท่าทางการเคลื่อนไหว (ท่ารำ) แล้วนำข้อมูลไปวิเคราะห์โดยใช้เทคนิคโครงข่ายประสาทเทียมชนิดจัดกลุ่มเอง (Self-Organizing Map : SOM) เพื่อการจำแนกท่ารำของแต่ละเพลง และการสร้างรูปแบบของท่าทาง (Gesture Templates) โดยใช้วิธีทางสถิติชื่อ Posture Transition Sparse Codes (PTSC) ซึ่งจะพิจารณาความถี่ของการเกิดรูปแบบที่ได้จากโครงข่าย นอกจากนี้ผู้วิจัยได้เสนอวิธีการรู้จำแบบทันทีทันใด (Real-time recognition) ในการระบุท่ารำและชื่อเพลง

งานวิจัยนี้นำเสนอเทคนิคการรู้จำท่ารำวงมาตรฐาน เพื่อแปลความหมายและวิเคราะห์ท่ารำประกอบเพลงรำวงมาตรฐาน ระบบที่พัฒนาขึ้นจะสามารถนำไปใช้ถ่ายทอดองค์ความรู้ท่ารำนาฏศิลป์ไทยชุด “รำวงมาตรฐาน” ในการฝึกและเรียนรู้ท่ารำวงมาตรฐาน เพื่อตรวจสอบความถูกต้องของท่ารำ โดยในการฝึกเต้นท่ารำนั้นจะประกอบไปด้วย 2 ขั้นตอน คือการรู้จำท่ารำ และการวัดความถูกต้องของการแสดงท่ารำ [3] ในงานวิจัยนี้จะนำเสนอเฉพาะวิธีการรู้จำท่ารำ ซึ่งจะเป็นขั้นแรกในการสอนเต้นรำโดยใช้ระบบคอมพิวเตอร์

2. ทบทวนวรรณกรรม

การศึกษางานวิจัยที่เกี่ยวข้องในประเด็นสำคัญที่เกี่ยวกับการรู้จำท่าทางของมนุษย์แบบอัตโนมัติ มีดังนี้

Munsell และคณะ [4] ได้เสนองานวิจัยเกี่ยวกับการพิสูจน์บุคคลจากการเคลื่อนไหวทั้งเดินและวิ่งของบุคคล โดยใช้ Kinect sensor ซึ่งพบว่าเซนเซอร์ดังกล่าวสามารถแยกแยะข้อมูลทุกส่วนของร่างกายและสามารถคำนวณสัดส่วนของร่างกายได้ถูกต้องและแม่นยำ เมื่อเปรียบเทียบกับวิธีการตรวจจับการเคลื่อนไหวของร่างกายที่ถูกนำเสนอมาก่อนหน้า

Jinda-Apiraksa และคณะ [5] ได้เสนอวิธีการแยกแยะท่าทางของมือ (Hand gesture) จำนวน 10 ท่าทาง ซึ่งมีขั้นตอนหลักในการแยกแยะท่าทาง 3 ขั้นตอนคือ การแบ่งส่วนข้อมูล การคำนวณคุณลักษณะเด่น และการจัดหมวดหมู่ โดยได้ทดลองวิเคราะห์รูปภาพท่าทางของมือจำนวน 200 ภาพ พบว่าสามารถแยกแยะท่าทางของมือได้ถูกต้องถึง 182 ภาพ (หรือคิดเป็นร้อยละ 91)

Shimada และคณะ [6] ได้เสนอเทคนิคการรู้จำท่าทาง โดยใช้วิธีการรู้จำแบบโครงข่ายประสาทเทียมชนิดจัดกลุ่มเอง ชนิด SOM ร่วมกับวิธีการทำดัชนีแบบ Sparse code เพื่อการรู้จำท่าทาง ผลการวิจัยพบว่าวิธีการนี้ให้ค่าความถูกต้องในการรู้จำท่าทางสูงถึงร้อยละ 98.3 และสามารถนำไปต่อยอดพัฒนาระบบที่ทำงานแบบโต้ตอบ (Interactive system) โดยใช้การรู้จำท่าทางได้อย่างมีประสิทธิภาพ

Saha และคณะ [7] ได้เสนอวิธีการรู้จำท่าเต้นในรูปแบบนาฏศิลป์อินเดีย โดยใช้กล้อง Kinect ในการตรวจจับท่าเต้น โดยพิจารณาการรู้จำท่าเต้นจากหลายปัจจัย ได้แก่ ระยะห่างระหว่างข้อต่อ อัตราความเร็วในการเคลื่อนที่ของมือและข้อศอก มุมระหว่างศีรษะ ลำคอ และสะโพก และใช้ท่าเต้นรำต้นแบบทั้งหมด 5 ท่า ผลการวิจัยพบว่าสามารถรู้จำท่าทางการเต้นแบบนาฏศิลป์อินเดียได้ถูกต้องร้อยละ 86.8

Qian และคณะ [8] ได้เสนอวิธีการแก้ปัญหาในการรู้จำท่าทางที่ใช้การตรวจจับภาพเคลื่อนไหวจากจุด Marker โดยเสนอวิธีการ Data cleaning ที่ชื่อว่า Static-body-part (SBP) และ Rigid-body (RB) ซึ่งจะทำหน้าที่ในการซ่อมแซมหรือเพิ่มเติมข้อมูลที่ขาดหายที่ได้จากจุด Marker ก่อนที่จะนำข้อมูลไปวิเคราะห์ผลการรู้จำจากผลการวิจัยพบว่าสามารถแก้ปัญหาในการตรวจจับภาพเคลื่อนไหวจากจุด Marker ได้อย่างมีประสิทธิภาพ และสามารถนำไปประยุกต์ใช้ในการรู้จำท่าทางแบบทันทีทันใดได้

จากการทบทวนวรรณกรรมพบว่ายังไม่มีงานวิจัยใดที่นำเทคนิคในการรู้จำท่าทาง (Gesture recognition) มาใช้ร่วมกับท่ารำวงมาตรฐาน (Standard Thai Traditional Dance) โดยเฉพาะอย่างยิ่ง การใช้อุปกรณ์ตรวจจับความเคลื่อนไหว (Kinect sensor) ในการบันทึกข้อมูล และใช้เทคนิคการรู้จำแบบโครงข่ายประสาทเทียมชนิดจัดกลุ่มเอง (SOM) กับวิธี Posture Transition Sparse Codes (PTSC) ในการวิเคราะห์รูปแบบของท่ารำที่ได้จากโครงข่ายประสาทเทียม นอกจากนั้นงานวิจัยนี้ยังได้เสนอวิธีการรู้จำแบบทันทีทันใดโดยใช้ทฤษฎีความน่าจะเป็นแบบเบยส์ (Bayesian probability theory) ในการรู้จำท่ารำอีกด้วย

3. ทฤษฎีที่เกี่ยวข้อง

ในส่วนนี้จะกล่าวถึงทฤษฎีที่เกี่ยวข้องกับการรู้จำท่ารำวงมาตรฐานโดยการวิเคราะห์ข้อมูล 3 มิติ ซึ่งผู้วิจัยได้ทำการเรียบเรียงเนื้อหาให้สอดคล้องกับขั้นตอนการวิจัย ดังนี้

3.1 รำวงมาตรฐาน

รำวงมาตรฐานเป็นการแสดงที่มีวิวัฒนาการมาจาก รำโทน เป็นการรำและร้องของชาวบ้าน โดยมีโทนเป็นเครื่องดนตรีประกอบจังหวะ ลักษณะการรำ และร้องเป็นไปตามความถนัด ไม่มีแบบแผนกำหนดไว้ ซึ่งเป็นการรำและร้องอย่างง่ายๆ มุ่งเน้นที่ความสนุกสนานรื่นเริงเป็นสำคัญ โดยประกอบไปด้วยเพลงทั้งหมด 10 เพลง ซึ่งแต่งโดยกรมศิลปากร และท่านผู้หญิงละเอียด พิบูลสงคราม ซึ่งแต่ละเพลงจะมีท่ารำประกอบเพลงที่แตกต่างกันและทำการรำซ้ำท่าเดิมไปเรื่อยๆ จนกว่าจะจบเพลง ซึ่งในแต่ละเพลงจะมีท่ารำประจำเพลง ดังตารางที่ 1

ตารางที่ 1 ท่ารำประกอบในแต่ละเพลง [9]

เพลง	ท่ารำ
1) เพลงงามแสงเดือน	ท่าสอดสร้อยมาลา
2) เพลงชาวไทย	ท่าชักแป้งผัดหน้า
3) เพลงรำมาซิมารำ	ท่ารำส่าย
4) เพลงคืนเดือนหงาย	ท่าสอดสร้อยมาลาแปลง
5) เพลงดวงจันทร์วันเพ็ญ	ท่าแขกเต้าเข้ารัง และท่าผาลงเพียงไหล
6) เพลงดอกไม้ของชาติ	ท่ารำยั่ว
7) เพลงหญิงไทยใจงาม	ท่าพรหมสี่หน้า และท่านกยูงฟ้อนหาง
8) เพลงดวงจันทร์ขวัญฟ้า	ท่าช้างประสานงา และท่าจันทร์ทรงกลด
9) เพลงยอดชายใจหาญ	(หญิง) ท่าชะนีร่ายไม้ (ชาย) ท่าจ่อเพลิงกัลป์
10) เพลงบูชานักรบ	(หญิง) ท่าขัดจางนาง (ชาย) ท่าจันทร์ทรงกลด (หญิง) ท่าล่อแก้ว (ชาย) ท่าขอแก้ว

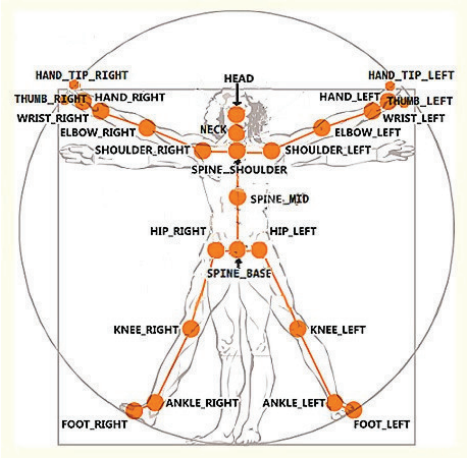
ท่าประกอบเพลงรำวงมาตรฐานมีทั้งหมด 17 ท่ารำ ซึ่งแต่ละท่าจะมีท่าทางการเคลื่อนไหวของร่างกายที่แตกต่างกันไป แสดงดังรูปที่ 1

รูปที่ 1 ท่ารำเพลงรำวงมาตรฐาน

3.2 อุปกรณ์ตรวจจับการเคลื่อนไหว (Kinect sensor)

รูปที่ 2 Microsoft Kinect sensor

กล้อง Kinect เป็นอุปกรณ์ในการบันทึกภาพเชิงลึก (Depth camera) ถูกพัฒนาโดยบริษัทไมโครซอฟต์ และมี Library functions ที่ช่วยในการแยกแยะโครงสร้างร่างกายของมนุษย์และสามารถค้นหาจุดต่างๆ บนร่างกาย (Joint) ได้มากถึง 25 จุด (ดังแสดงในรูปที่ 3) แล้วสามารถนำไปสร้างเป็นโครงร่างจำลอง (Skeleton) ในการตรวจจับการเคลื่อนไหวของร่างกาย ซึ่งสามารถช่วยลดเวลาในการหาวิธีการตรวจจับโครงสร้างของร่างกายและการตรวจจับการเคลื่อนไหว [4] ข้อดีของ Kinect คือเป็นอุปกรณ์ที่มีราคาไม่แพงเมื่อเปรียบเทียบกับอุปกรณ์ตรวจจับค่าความเคลื่อนไหวแบบอื่น พร้อมทั้งนำไปใช้งานได้อย่างสะดวกและให้ความแม่นยำในการตรวจจับความเคลื่อนไหวของร่างกายในระนาบ 3 มิติ [10] งานวิจัยนี้จึงได้นำกล้อง Kinect มาบันทึกข้อมูลท่าทางการเคลื่อนไหวของท่ารำในแต่ละท่า เพื่อนำข้อมูลไปวิเคราะห์ในขั้นตอนต่อไป



รูปที่ 3 โครงร่างจำลองที่ได้จาก Kinect sensor จำนวน 25 จุด (Microsoft SDK Skeleton) [11]

3.3 เครือข่ายประสาทเทียม (Artificial Neural Network: ANN)

เครือข่ายประสาทเทียม สามารถนำไปประยุกต์ใช้งานแก้ปัญหาในด้านต่างๆ อย่างมากมาย และประสบความสำเร็จเป็นอย่างดี โดยจุดเด่นของ ANN คือความสามารถในการเรียนรู้เป็นปัจจัยหลัก ตัวอย่างของงานที่นำ ANN ไปประยุกต์ใช้ได้แก่ การจดจำรูปแบบ (Pattern recognition) การจับกลุ่ม (Clustering) การประมาณค่าฟังก์ชัน (Function approximation) การทำนาย (Prediction) การหาค่าเหมาะที่สุด (Optimization) หน่วยความจำอ้างอิงด้วย

เนื้อหา (Content-addressable memory) ระบบควบคุม (Control system) เป็นต้น

ลักษณะการเรียนรู้ของระบบโครงข่ายประสาทเทียมสามารถแบ่งได้เป็น 2 รูปแบบดังนี้ [12]

1. การเรียนรู้แบบมีผู้สอน (Supervised Learning) เครือข่ายประสาทเทียมจะได้รับการฝึกด้วยกลุ่มฝึกที่ประกอบด้วยข้อมูลนำเข้าและผลลัพธ์ที่ต้องการ การปรับค่าน้ำหนักจะทำให้ค่าผลลัพธ์ที่ได้มีความใกล้เคียงหรือตรงกับผลลัพธ์ที่ต้องการ ตัวอย่างเครือข่ายประสาทเทียมที่ใช้วิธีการเรียนรู้แบบนี้ ได้แก่ Back Propagation และ Multilayer Perceptron เป็นต้น

2. การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) การเรียนรู้เพื่อปรับค่าน้ำหนักจะตอบสนองต่อข้อมูลนำเข้าเท่านั้น โดยไม่ได้ใช้ค่าผลลัพธ์ร่วมด้วยในการฝึก วิธีการเรียนรู้แบบนี้จะใช้การจัดแบ่งกลุ่มโดยพิจารณาตามรูปแบบของค่านำเข้า ตัวอย่างเครือข่ายประสาทเทียมที่ใช้วิธีการเรียนรู้แบบนี้ ได้แก่ Counter Propagation และ Kohonen's Self-Organizing Map (SOM) เป็นต้น

งานวิจัยนี้จะทำการรู้จำท่ารำวงมาตรฐานโดยนำข้อมูลท่ารำที่บันทึกได้จากกล้อง Kinect มาทำการจัดกลุ่มโดยเครือข่ายประสาทเทียมที่มีรูปแบบการเรียนรู้แบบไม่มีผู้สอน SOM เพื่อการจัดกลุ่มข้อมูลสำหรับทำดัชนีเพื่อจำแนกท่ารำในขั้นตอนต่อไป

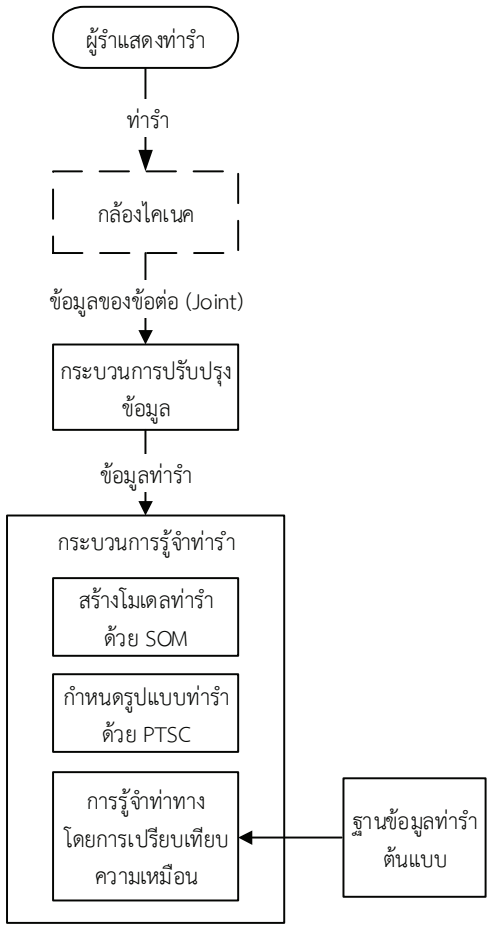
4. วิธีดำเนินงานวิจัย

ในงานวิจัยนี้ได้เลือกท่ารำที่ใช้ในการวิเคราะห์ทั้งหมด 12 ท่ารำ ดังแสดงในตารางที่ 2 โดยจะเลือกวิเคราะห์ท่ารำต้นแบบที่ไม่มีลักษณะการเคลื่อนที่แบบหมุนตัว และวิเคราะห์เฉพาะท่ารำของฝ่ายหญิงเพียงอย่างเดียว เนื่องจากกล้อง Kinect ไม่สามารถตรวจจับการเคลื่อนไหวของท่ารำที่มีลักษณะหมุนตัวกลับหรือลักษณะที่มีจุดข้อต่อทับซ้อนกันได้ และจะเลือกพิจารณาท่าทางการเคลื่อนไหวบริเวณส่วนแขนและช่วงบนของลำตัว หรือตั้งแต่จุดตำแหน่งที่ Spine base ขึ้นไป

ตารางที่ 2 ท่ารำที่ใช้ในงานวิจัยนี้

เพลง	ท่ารำ
1) เพลงงามแสงเดือน	ท่าสอดสร้อยมาลา
2) เพลงชาวไทย	ท่าชักแป้งผัดหน้า
3) เพลงรำมาซิมารำ	ท่ารำส่าย
4) เพลงคืนเดือนหงาย	ท่าสอดสร้อยมาลาแปลง
5) เพลงดวงจันทร์วันเพ็ญ	ท่าแขกเต้าเข้ารัง
6) เพลงดอกไม้ของชาติ	ท่ารำยั่ว
7) เพลงหญิงไทยใจงาม	ท่าพรหมสี่หน้า และท่านกยูงฟ้อนหาง
8) เพลงดวงจันทร์ขวัญฟ้า	ท่าช้างประสานงา และท่าจันทร์ทรงกลด
9) เพลงยอดชายใจหาญ	ท่าชะนีร่ายไม้
10) เพลงบูชานักรบ	ท่าขัดใจนาง

เมื่อได้ข้อมูลของจุดต่างๆ บนร่างกายจากกล้อง Kinect (ดังแสดงในรูปที่ 3) แล้ว สามารถนำข้อมูลดังกล่าวไปวิเคราะห์ต่อ โดยมีขั้นตอนการวิเคราะห์การรู้จำท่ารำประกอบเพลง แสดงในรูปที่ 4



รูปที่ 4 ขั้นตอนการทำงาน

4.1 กระบวนการปรับปรุงข้อมูล (Data Normalization)

ข้อมูลที่ได้มาจาก Kinect นั้นไม่สามารถนำมาใช้เปรียบเทียบได้โดยตรง เนื่องจากระยะห่างของจุดตามร่างกายแต่ละคนที่ได้มานั้นมีอัตราส่วนที่ไม่เท่ากัน ทำให้เวลานำข้อมูลไปเปรียบเทียบอาจทำให้เกิดความผิดพลาดได้ ดังนั้นจึงจำเป็นต้องมีการปรับข้อมูลให้อยู่ในอัตราส่วนเดียวกันก่อน (การนอร์มัลไลซ์ : Data Normalization) โดยกำหนดให้ตำแหน่งที่ Spine base ในรูปที่ 3 เป็นจุดเริ่มต้น และนำข้อต่อจากทุกจุดมาลบด้วยค่าของตำแหน่งของ Spine base [13] ดังสมการที่ 1

$$w_l = \frac{r_l - r_l^o}{\|r_l - r_l^o\|^2}, l = 2, \ldots, L \quad (1)$$

โดยที่ w_l คือค่าของจุดตามร่างกายที่ตำแหน่ง l

r_l^o คือค่าของตำแหน่งที่จุด Spine base

r_l คือค่าของตำแหน่งที่จุดอื่นๆ

L คือจำนวนจุดข้อต่อทั้งหมด

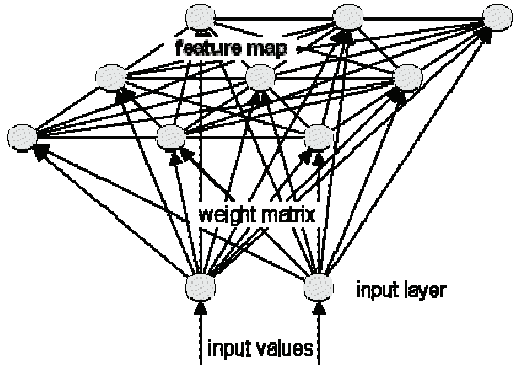
4.2 กระบวนการรู้จำ (Recognition Process)

หลักจากที่ได้ข้อมูลที่ผ่านกระบวนการนอร์มอลไลซ์มาแล้ว ก็จะเข้าสู่กระบวนการรู้จำซึ่งจะใช้ขั้นตอนดังต่อไปนี้

4.2.1 Kohonen's Self-Organizing Map (SOM)

SOM เป็นอัลกอริทึมโครงข่ายประสาทเทียมที่นิยมใช้มากที่สุด ถูกพัฒนาขึ้นในปี ค.ศ.1982 โดย Tuevo Kohonen [14] เป็นการเรียนรู้เรียนรู้แบบไม่มีผู้สอน โดยใช้โมเดลหน่วยความจำรูปแบบของการเชื่อมโยงและการเรียนรู้การปรับตัว โดยมีแนวคิดในการทำซ้ำข้อมูลเพื่อที่จะหาค่าของน้ำหนักของข้อมูลที่มีอยู่ทั้งหมดตามจำนวนกลุ่มที่ต้องการ

ในโครงข่ายประสาทเทียมแบบ SOM มีการจัดโครงสร้างของประสาทเทียมที่มีเอาต์พุตเป็น Feature map ดังรูปที่ 5 ซึ่งมีลักษณะทางกายภาพที่แต่ละโหนดหรือโปรโตไทป์จะกระจายกันอยู่ โหนดที่อยู่ใกล้กันจะทำปฏิกิริยาแตกต่างจากโหนดที่อยู่ไกลกัน [15] จุดมุ่งหมายการฝึกสอนโครงข่ายคือการทำให้เอาต์พุตที่อยู่ใกล้กันสัมพันธ์กับอินพุตที่อยู่ใกล้กัน โดยกลุ่มของข้อมูลอินพุตจะถูกกำหนดให้ผูกติดและจะถูกแทนที่ด้วยเวคเตอร์โปรโตไทป์ที่คล้ายคลึงกันมากที่สุด



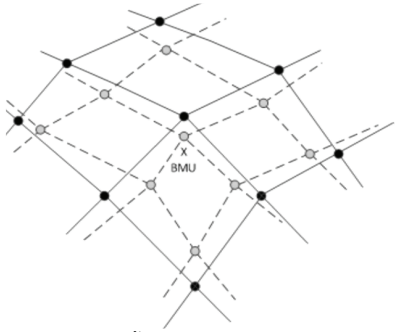
รูปที่ 5 Feature map ของ SOM

อัลกอริทึมที่ 1 แสดงการสร้างโปรโตไทป์โดย SOM สำหรับข้อมูลนำเข้าที่ใช้ในการเรียนรู้จะใช้ข้อมูลที่ผ่านกระบวนการ Data Normalization เรียบร้อยแล้วมาสร้างโมเดล และกำหนดให้มีจำนวนโปรโตไทป์ (M) เท่ากับ 100 โหนด

Algorithm 1 อัลกอริทึม SOM

1) **ค่านำเข้า:** ชุดข้อมูลลำดับของท่าทาง $X = \{\vec{x}_1, \dots, \vec{x}_N\}$
2) **ผลลัพธ์:** ชุดข้อมูลโปรโตไทป์ $Y = \{\vec{y}_1, \dots, \vec{y}_M\}$
3) **เริ่มต้น**
4) กำหนดจำนวนครั้งของการเรียนรู้ K
5) สุ่มค่าเริ่มต้นให้กับ $Y = \{\vec{y}_1, \dots, \vec{y}_M\}$
6) **จุดเริ่มต้นทำซ้ำ**
7) สุ่มเลือกโหนดอินพุต $\vec{x} \in X$
8) คำนวณระยะห่าง Euclidean ของ $\vec{x}$ กับโปรโตไทป์ทุกโหนด
9) หาโหนดผู้ชนะ $\vec{y}^*$ จากระยะห่างที่สั้นที่สุด
$d(\vec{x}, \vec{y}^*) = min\{d(\vec{x}, \vec{y}) \mid \vec{y} \in Y\}$
10) ปรับค่าน้ำหนักของโหนดผู้ชนะและโหนดใกล้เคียง $\vec{y} \in N(\vec{y}^*)$
โดยที่ $N(\vec{y}^*)$ คือเซตของโหนดใกล้เคียงโหนดผู้ชนะ
$\vec{y} = \vec{y} + \gamma(\vec{x} - \vec{y})$ โดยที่ γ คือค่าคงที่
11) **ทำซ้ำ** จนกว่าจะครบรอบที่กำหนด
12) **สิ้นสุด**

การเรียนรู้ของ SOM เริ่มต้นจากการสุ่มค่าน้ำหนักให้กับโปรโตไทป์แต่ละโหนดในเซต Y เพื่อนำไปใช้ในการเรียนรู้จากข้อมูลนำเข้า X ในแต่ละรอบของการเรียนรู้ ค่าข้อมูลนำเข้า $\vec{x}$ จะถูกเลือกแบบสุ่มค่าจากเซต X หลังจากนั้นจะคำนวณหาค่าระยะห่างระหว่างข้อมูลนำเข้ากับค่าน้ำหนักของแต่ละโหนด เมื่อโหนดใดมีค่าน้ำหนักใกล้กับค่านำเข้ามากที่สุดจะเรียกโหนดนั้นว่า Best-Matching Unit (BMU) จากนั้นจึงทำการปรับค่าน้ำหนักของโหนดที่เป็น BMU และโหนดใกล้เคียง จึงทำให้ค่าน้ำหนักที่ได้ใหม่นี้มีการเข้าใกล้ข้อมูลกันมากยิ่งขึ้น ดังรูปที่ 6



รูปที่ 6 การปรับค่าน้ำหนักเข้าใกล้จุดศูนย์กลาง คือจุด x โครงข่ายที่ด้วยเส้นทึบคือโครงข่ายก่อนการปรับค่าน้ำหนัก และโครงข่ายที่แสดงด้วยเส้นประคือโครงข่ายหลังจากที่ปรับค่าน้ำหนักแล้ว

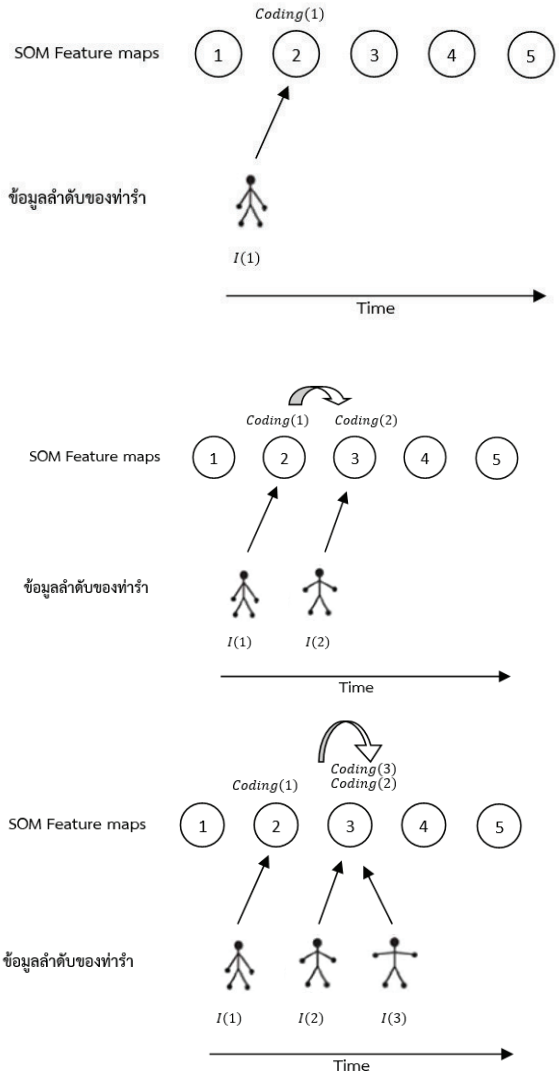
เมื่อมีการปรับน้ำหนักในแต่ละโหนดเข้าใกล้ BMU ก็จะทำกระบวนการเรียนรู้ซ้ำไปเรื่อยๆ จนกว่าจะครบจำนวนรอบที่กำหนด ซึ่งจะได้ผลลัพธ์เป็นชุดข้อมูลโปรโตไทป์ของข้อมูล

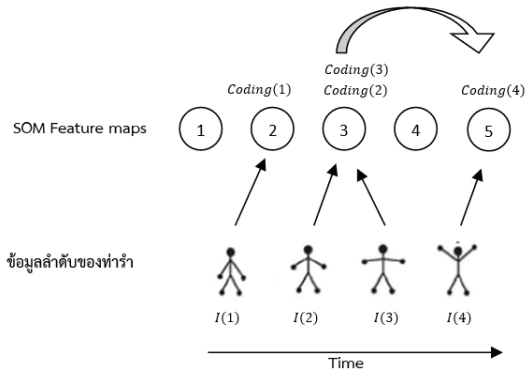
4.2.2 การทำดัชนีท่ารำด้วยวิธี Posture Transition Sparse Code (PTSC)

การทำดัชนีของท่ารำทำได้โดยนำข้อมูลท่ารำไปเข้ารหัสบนโปรโตไทป์ที่ได้มาจาก SOM โดยแต่ละท่ารำจะมีรูปแบบ (Pattern) ของโหนดที่แตกต่างกัน และการนับความถี่ของโหนดที่เกิดขึ้นจะสามารถใช้ในการทำดัชนีของท่ารำนั้นๆ ได้ ซึ่งสามารถใช้เป็นแม่แบบในการเปรียบเทียบความเหมือนท่ารำที่รับเข้ามาใหม่ วิธี PTSC จะเป็นการหาค่าความถี่ของโปรโตไทป์ที่ถูกนำมาใช้ในการเข้ารหัส โดยนับการเปลี่ยนโหนดของโปรโตไทป์ [3] แสดงการทำงานดังรูปที่ 7

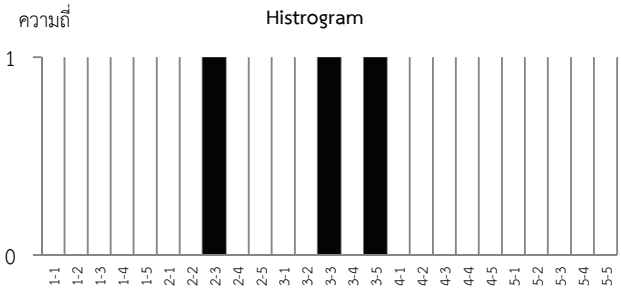
การเข้ารหัสข้อมูลจำนวน 4 เฟรม ได้แก่ $I(1), I(2), I(3), I(4)$ โดยที่ $I(1)$ จะถูกเข้ารหัสด้วยโปรโตไทป์ที่ตำแหน่งโหนดที่สอง $I(2)$ จะถูกเข้ารหัสด้วยโปรโตไทป์ที่ตำแหน่ง

โหนดที่สาม $I(3)$ จะถูกเข้ารหัสด้วยโปรโตไทป์ที่ตำแหน่งโหนดที่สาม และ $I(4)$ จะถูกเข้ารหัสด้วยโปรโตไทป์ที่ตำแหน่งโหนดที่สี่ ซึ่ง PTSC จะเป็นการคำนวณหาค่าในการเปลี่ยนจากโหนดหนึ่งไปอีกโหนดหนึ่ง กล่าวคือ จากรูปที่ 7 มีการเปลี่ยนจากตำแหน่งโหนดที่สองไปยังตำแหน่งโหนดที่สาม จำนวน 1 ครั้ง เปลี่ยนจากตำแหน่งโหนดที่สามไปยังตำแหน่งโหนดที่สาม จำนวน 1 ครั้ง และเปลี่ยนจากตำแหน่งโหนดที่สามไปยังตำแหน่งโหนดที่ห้า จำนวน 1 ครั้ง ดังนั้นค่า PTSC ที่ได้จะเป็นไปตามรูปที่ 8





รูปที่ 7 การเข้ารหัสของข้อมูลรับเข้าจำนวน 4 เฟรม ได้แก่ $I(1), I(2), I(3), I(4)$



รูปที่ 8 แผนภูมิแสดงความถี่ของการเปลี่ยนโหนดโปรโตไทป์ในแต่ละท่ารำในรูปแบบของ Histogram (h)

งานวิจัยนี้มีท่ารำที่ใช้ในการวิเคราะห์จำนวน 12 ท่ารำ (กำหนดให้จำนวนท่ารำ $C = 12$) ดังนั้นจะได้ Histogram $(h_1, h_2, \dots, h_{12})$ โดยที่ h_c คือค่า Histogram ที่เป็นต้นแบบของท่ารำลำดับที่ c จากท่ารำทั้งหมด $C = 12$

4.2.3 การรู้จำท่าทางแบบทันทีทันใด (Online Gesture Recognition)

ในงานวิจัยนี้สามารถใช้ท่ารำต่อเนื่องเป็นข้อมูลในการวิเคราะห์รู้จำท่ารำ มาประยุกต์ใช้ร่วมกับทฤษฎีความน่าจะเป็นแบบเบยส์ (Bayesian probability theory) ในการระบุเพลงประกอบท่ารำ โดยกำหนดให้ S_u เป็นท่ารำต่อเนื่องที่ได้มาจากกล้อง Kinect ค่าของ $S_u = [\vec{x}^{t0}, \dots, \vec{x}^{t}, \dots, \vec{x}^{T}]$ โดยที่ $\vec{x}^{t0}$ คือเฟรมของข้อมูลท่ารำที่เวลา t_0 ข้อมูล S_u จะถูกจัดกลุ่มให้อยู่ในกลุ่มของ c ใดๆ โดยที่ดัชนีของแต่ละท่ารำ $c = 1, \dots, C$ และ C คือจำนวนคลาสของท่ารำ ดังนั้นสามารถเขียนสมการค่าความน่าจะเป็นของแต่ละท่ารำได้ดังนี้

$$P(c|S_u) = Prior \times Likelihood \qquad (2)$$

โดยสามารถคำนวณค่า $Prior$ และ $Likelihood$ ได้จาก Algorithm 2 จะเป็นการคำนวณค่าความน่าจะเป็นของแต่ละท่ารำ ณ เวลา t ใดๆ เมื่อนำค่าความน่าจะเป็นทั้งหมดมาเปรียบเทียบกัน จะทำให้สามารถหาค่าสูงสุดของความน่าจะเป็นจากท่ารำใดท่ารำหนึ่ง และสามารถคาดเดาได้ว่าท่ารำที่เป็นข้อมูลรับเข้าเป็นท่ารำใด ดังแสดงในผลการทดลอง (รูปที่ 10)

Algorithm 2 อัลกอริทึมการรู้จำท่าทางแบบทันทีทันใด [3]

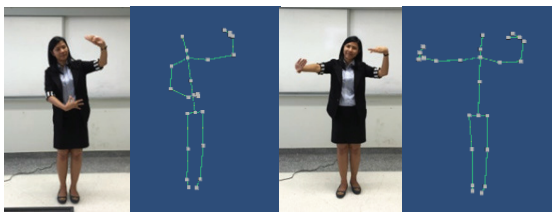
1) **ค่านำเข้า:** ลำดับของท่าทาง $S_u = [\vec{x}^{t0}, \dots, \vec{x}^{t}, \dots, \vec{x}^{T}]$
2) **ผลลัพธ์:** ความน่าจะเป็นที่เวลา t ใดๆ $P_t(c|S_u)$;
$$argmax_{c,t}\{P_t(c|S_u)\}$$
3) **กำหนดให้** $t = t_0 = 0$
4) **จุดเริ่มต้นทำซ้ำ**
5) กำหนดให้ลำดับของท่าทาง $S_u = [\vec{x}^{t0}, \dots, \vec{x}^{t}]$
6) กำหนดให้ h_s คือเวกเตอร์ PTSC ที่ได้จากการทำดัชนี S_u และ h_c คือ PTSC template ของท่ารำที่ c
7) คำนวณค่า likelihood,
$$P_t(h_s|c) = HI(h_s, h_c) = 1 - \sum_i \min[h_{s,i}, h_{c,i}]$$
8) คำนวณค่า prior,
$$P_t(c) = \begin{cases} \frac{1}{C}, & \text{ถ้า } t=t_0 \\ \frac{P_{t-1}(c|h_s)\cdot HI(h_s,h_c)}{\sum_C P_{t-1}(c|h_s)\cdot HI(h_s,h_c)}, & \text{กรณีอื่นๆ} \end{cases}$$
9) คำนวณค่า , $P_t(c|h_s)$ ของท่ารำ c
$$P_t(c|h_s) = \frac{P_t(h_s|c)P_t(c)}{P_t(h_s)} = \frac{P_t(h_s|c)P_t(c)}{\sum P_t(h_s|c)P_t(c)}$$
10) ถ้า $\max[P_t(c|h_s)] > T_r$
11) กำหนดค่า T_r = ค่าขีดแบ่งมาตรฐาน (Threshold)
12) กำหนดค่า $t = t_0$
13) กำหนดค่า prior ใหม่ $P_t(c) = \frac{1}{c}$
14) คำนวณค่า posterior $P_t(c|h_s)$
15) $t++$
16) ทำซ้ำจนกว่า $t = T$ (จนกว่าจะหมดลำดับสุดท้ายของลำดับท่าทาง)

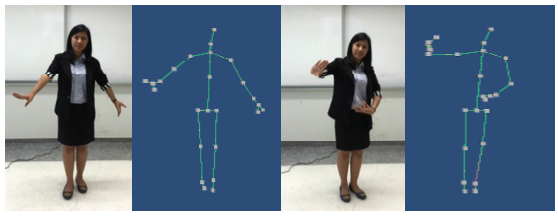
5. ผลการทดลอง

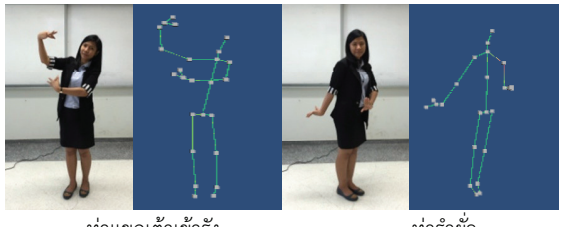
ในงานวิจัยนี้การทดสอบความถูกต้องของการจำแนกท่ารำในแต่ละเพลง จะต้องบันทึกท่ารำต้นแบบไว้ล่วงหน้าก่อน โดยให้ผู้รำต้นแบบ (ผู้รำที่มีความชำนาญ) ทำการบันทึกท่ารำซ้ำจำนวน 5 รอบต่อท่ารำ (หนึ่งรอบของท่ารำคือการเคลื่อนที่ของแขนจากด้านหนึ่งไปอีกด้านหนึ่ง) แล้วบันทึกไว้ในฐานข้อมูลในรูปแบบของข้อมูลโครงร่างจำลอง (Joint Skeleton) ดังรูปที่ 9 จากนั้นให้ผู้รำฝึกหัดทดลองทำท่ารำซ้ำต่อเนื่องไปเรื่อยๆ จนกว่าจะระบบจะให้คำตอบของท่ารำ ซึ่งผู้รำฝึกหัดจะเป็นคนละคนกับต้นแบบที่บันทึกไว้ในฐานข้อมูลซึ่งมีความแตกต่างทางด้านกายภาพ เช่น เพศ ความสูง และสัดส่วนของร่างกาย เป็นต้น เพื่อทดสอบความถูกต้องของอัลกอริทึมในการแยกแยะท่ารำของแต่ละบุคคล โดยทำการทดสอบเปรียบเทียบท่ารำจากผู้รำต้นแบบกับผู้รำฝึกหัดจำนวน 3 ตัวอย่าง ได้ผลการรู้จำท่ารำดังตารางที่ 3

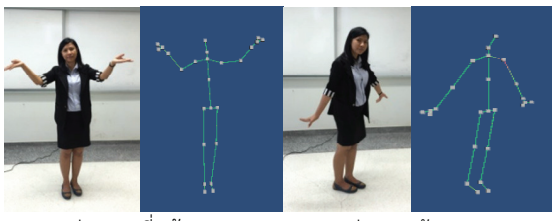
ตารางที่ 3 ผลการรู้จำที่ได้จากระบบในแต่ละเพลง

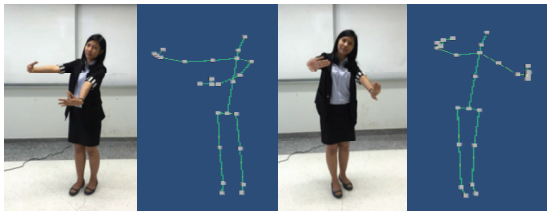
เพลง	ผู้รำฝึกหัด 1	ผู้รำฝึกหัด 2	ผู้รำฝึกหัด 3
1) เพลงงามแสงเดือน	G1	G1	G1
2) เพลงชาวไทย	G2	G2	G2,G5,G7
3) เพลงรำมาซิมารำ	G3	G3	G3
4) เพลงคืนเดือนหงาย	G4,G8	G4,G5	G5
5) เพลงดวงจันทร์วันเพ็ญ	G5	G5	G5
6) เพลงดอกไม้ของชาติ	G6	G6	G6
7) เพลงหญิงไทยใจงาม	G6,G7	G6,G7	G7
8) เพลงดวงจันทร์ขวัญฟ้า	G5,G8,G10	G6,G8	G10
9) เพลงยอดชายใจหาญ	G9	G9	G9
10) เพลงบูชานักรบ	G10	G10	G10


ท่าสอดสร้อยมาลา ท่าชักแป้งผัดหน้า

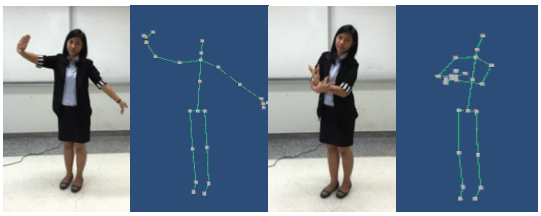

ท่ารำส่าย ท่าสอดสร้อยมาลาแปลง


ท่าแขกเต้าเข้ารัง ท่ารำยั่ว


ท่าพรหมสี่หน้า ท่านกยูงฟ้อนหาง



ท่าช้างประสานงา ท่าจันทร์ทรงกลด



ท่าชะนีร่ายไม้ ท่าขัดใจนาง

รูปที่ 9 ท่ารำต้นแบบที่บันทึกไว้ในฐานข้อมูล

ตารางที่ 4 ค่าความถูกต้องเฉลี่ยในการรู้จำแต่ละท่ารำ

เพลง	ความถูกต้องการรู้จำท่ารำ
1) เพลงงามแสงเดือน	100%
2) เพลงชาวไทย	93%
3) เพลงรำมาซิมารำ	100%
4) เพลงคืนเดือนหงาย	33%
5) เพลงดวงจันทร์วันเพ็ญ	100%
6) เพลงดอกไม้ของชาติ	100%
7) เพลงหญิงไทยใจงาม	66%
8) เพลงดวงจันทร์ขวัญฟ้า	30%
9) เพลงยอดชายใจหาญ	100%
10) เพลงบูชานักรบ	100%
ค่าเฉลี่ย	**82.20%**

จากตารางที่ 4 แสดงค่าความถูกต้องในการรู้จำท่ารำในแต่ละเพลง โดยค่าความถูกต้องคำนวณจากค่าความถี่ของคำตอบที่ได้จากการทดสอบระบบ จำนวน 10 ครั้งในแต่ละเพลง แล้วหาค่าเฉลี่ยพร้อมทั้งคำนวณเป็นค่าร้อยละของของคำตอบที่มีความถูกต้อง ดังรูปที่ 11 หากท่ารำใดมีผลความถูกต้อง ร้อยละ 100 ระบบจะให้คำตอบ Gesture number เพียงคำตอบเดียวและคำตอบเป็นท่ารำที่ตรงกับผู้รำทดสอบ แต่ถ้าระบบให้คำตอบของท่ารำไม่ตรงกับผู้รำทดสอบหรือมีความแปรปรวนของคำตอบ ก็จะทำให้ค่าความถูกต้องในการรู้จำท่ารำลดลง

จากผลการทดลองพบว่า ผลความถูกต้องในการรู้จำท่ารำประกอบเพลงที่มีความถูกต้อง ร้อยละ 100 มีจำนวน 7 เพลง เนื่องจากท่ารำในแต่ละเพลงมีการเคลื่อนไหวของแขนและลำตัวที่ชัดเจน ได้แก่ 1) เพลงงามแสงเดือน : ท่าสอดสร้อยมาลา 2) เพลงชาวไทย : ท่าชักแป้งผัดหน้า 3) เพลงรำซิมารำ : ท่ารำส่าย 4) เพลงดวงจันทร์วันเพ็ญ : ท่าแขกเต้าเข้ารัง 5) เพลงดอกไม้ของชาติ : ท่ารำยั่ว 6) เพลงยอดชายใจหาญ : ท่าชะนีร่ายไม้ 7)เพลงบูชานักรบ : ท่าขัดใจนาง

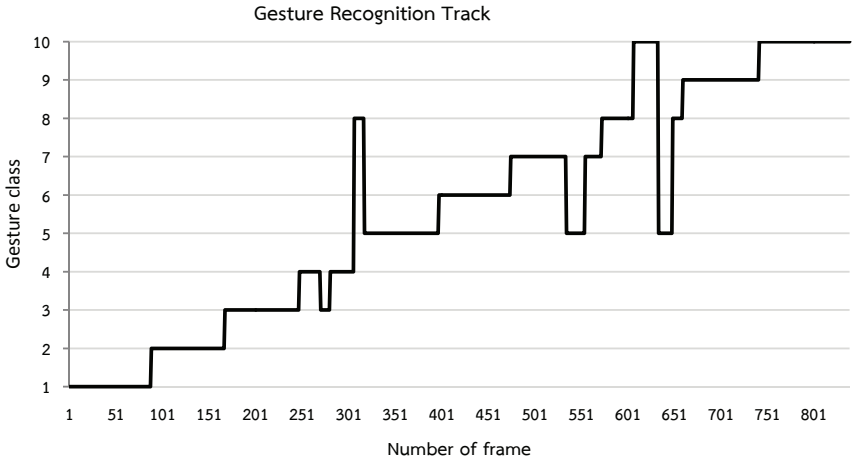
สำหรับเพลงหญิงไทยใจงาม มีผลความถูกต้องในการรู้จำท่ารำประกอบเพลง ลงลดเหลือร้อยละ 80 เนื่องจากเพลงหญิงไทยใจงาม มีท่ารำประกอบเพลงจำนวน 2 ท่า คือ ท่าพรหมสี่หน้าและท่านกยูงฟ้อนหาง ทำให้มีรูปแบบการเคลื่อนที่ของร่างกายมากกว่าเพลงอื่น โดยท่าพรหมสี่หน้ามีการเคลื่อนไหวของแขนและลำตัวที่ชัดเจน แต่ท่านกยูงฟ้อนหางมีการเคลื่อนที่ของส่วนแขนไปบริเวณด้านหลังของลำตัว ทำให้ค่าตำแหน่งของแขนที่วัดได้จากกล้อง Kinect มีความคลาดเคลื่อนเล็กน้อยจึงส่งผลค่าความถูกต้องในการรู้จำลดลง

เพลงที่มีผลความถูกต้องในการรู้จำท่ารำประกอบเพลง เพียงร้อยละ 50 คือ เพลงคืนเดือนหงาย เนื่องจากท่ารำประกอบเพลงท่าสอดสร้อยมาลาแปลง มีรูปแบบการเคลื่อนที่ของแขนและลำตัวคล้ายคลึงกับท่ารำสอดสร้อยมาลาในเพลงงามแสงเดือนนั่นเอง ส่วนเพลงที่มีผลความถูกต้องในการรู้จำท่ารำประกอบเพลงต่ำกว่าร้อยละ 50 คือ เพลงดวงจันทร์ขวัญฟ้า เนื่องจากมีท่ารำประกอบเพลงจำนวน 2 ท่า ได้แก่ ท่าช้างประสานงาและท่าจันทร์ทรงกลด ทำให้มีรูปแบบของการเคลื่อนที่ของร่างกายมากกว่าเพลงอื่น นอกจากนี้ลักษณะของท่ารำมีการหมุนตัวหันด้านข้างให้กับกล้อง Kinect ทำให้ตำแหน่งของช่วงแขนทั้งสองข้างทับซ้อนกัน และเกิดความคลาดเคลื่อนของค่าในแต่ละจุดที่วัดได้ ส่งผลให้ผลลัพธ์ของการรู้จำได้หลายคำตอบและมีไม่คงที่นั่นเอง ดังตารางที่ 3

เมื่อสรุปความถูกต้องในการรู้จำท่ารำประกอบเพลงรำวงมาตรฐานค่าเฉลี่ยรวมเท่ากับร้อยละ 82.20 ซึ่งสามารถนำไปประยุกต์ใช้ในการรู้จำท่ารำวงมาตรฐานได้

ในรูปที่ 10 จะแสดงผลการทดลองเมื่อทำท่ารำวงมาตรฐานที่แตกต่างกันแบบต่อเนื่องทั้ง 10 เพลง ระบบสามารถแสดงผลการรู้จำท่ารำของแต่ละเพลง ตามค่าคลาสของท่ารำ (Gesture class) ณ เวลาต่างๆ จำนวน 800 เฟรม

จะเห็นได้ว่าท่ารำที่มีความแปรปรวนของคำตอบ มีจำนวน 3 เพลง คือ ท่ารำในคลาสที่ 4 เพลงคืนเดือนหงาย : ท่าสอดสร้อยมาลาแปลง ท่ารำในคลาสที่ 7 เพลงหญิงไทยใจงาม : ท่าพรหมสี่หน้าและมาสนกยูงฟ้อนหาง และท่ารำในคลาสที่ 8 เพลงดวงจันทร์ขวัญฟ้า : ท่าช้างประสานงาและท่าจันทร์ทรงกลด

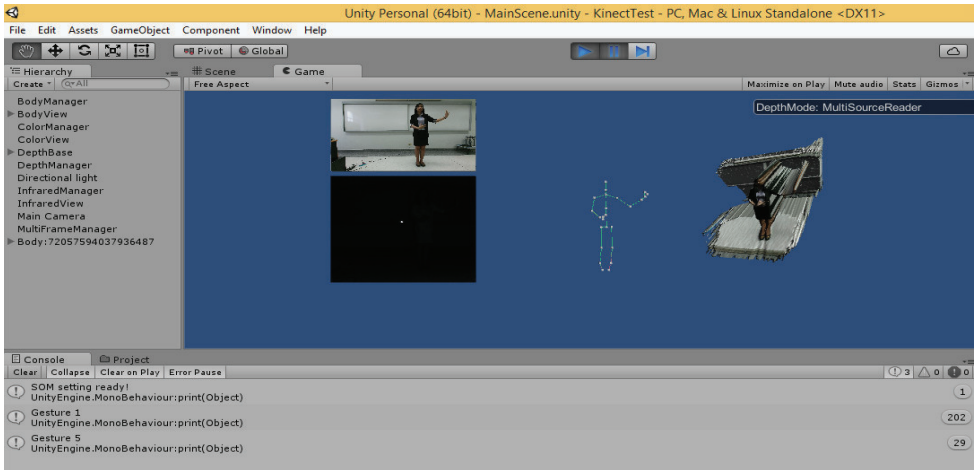


รูปที่ 10 ตัวอย่างผลการรู้จำท่ารำในแต่ละเพลง

นอกจากท่ารำประกอบเพลงที่มีรูปแบบของการเคลื่อนที่ของร่างกายที่มีความทับซ้อนกันบางจุดของข้อต่อซึ่งจะส่งผลต่อการรู้จำของระบบ ประสบการณ์หรือทักษะการรำของผู้ทดลองก็ส่งผลต่อคำตอบของการรู้จำท่ารำด้วยเช่นกัน ดังตารางที่ 3 เพลงชาวไทย ระบบสามารถให้คำตอบของการรู้จำท่ารำได้ถูกต้องสำหรับการทดลองในผู้รำฝึกหัดคนที่ 1 และ 2 แต่ผู้รำฝึกหัดคนที่ 3 ให้ผลการรู้จำที่มีความแปรปรวนของคำตอบ เนื่องจากผู้รำฝึกหัดคนที่ 3 ไม่มีความถนัดหรือมีทักษะในการรำวงมาตรฐานมาก่อนเลย จึงแสดงท่ารำที่ไม่ตรงกับท่ารำต้นแบบและเป็นสาเหตุทำให้คำตอบในการรู้จำท่ารำไม่ถูกต้องนั่นเอง

6. สรุปผลการทดลอง

งานวิจัยนี้ได้เสนอการพัฒนาระบบการรู้จำท่ารำวงมาตรฐานโดยใช้กล้อง KINECT ในการรับข้อมูลท่ารำจำนวน 12 ท่า ที่มีความแตกต่างกันในแต่ละเพลงรำวง โดยประยุกต์ใช้วิธีการจัดกลุ่มข้อมูลแบบไม่มีผู้สอนชนิด SOM และการทำดัชนีแบบ PTSC เพื่อการการรู้จำท่ารำแบบทันทีทันใด ซึ่งระบบสามารถรู้จำท่ารำประกอบเพลงได้ถูกต้องโดยไม่มีข้อผิดพลาดถึง 7 เพลง แต่อย่างไรก็ตามก็ยังมีท่ารำบางท่าที่มีความคลาดเคลื่อนในการรู้จำ เนื่องจากข้อมูลที่บันทึกได้จากโครงร่างจำลองมีความไม่ชัดเจนของข้อต่อของร่างกายที่วัดได้จากกล้อง KINECT ส่งผลให้ค่าความถูกต้องในการรู้จำลดลง ซึ่งงานวิจัยนี้สามารถวิเคราะห์ข้อมูลและรู้จำท่ารำวงมาตรฐาน โดยมีค่าเฉลี่ยความถูกต้องของการรู้จำเท่ากับ ร้อยละ 82.20 และสามารถนำไปประยุกต์ใช้การรู้จำท่ารำวงมาตรฐานได้อย่างมีประสิทธิภาพในอนาคต งานวิจัยนี้สามารถเพิ่มตัวอย่างที่ใช้ในการทดสอบการรู้จำรำให้มีหลากหลายมากยิ่งขึ้นเพื่อเพิ่มความเชื่อมั่นและความแม่นยำในการรู้จำท่ารำวงมาตรฐานต่อไป นอกจากนี้หากระบบสามารถเลือกผู้รำต้นแบบได้หลายหลายก็จะทำให้การนำไปประยุกต์ใช้กับการเรียนรู้ท่ารำวงมาตรฐานมีความน่าสนใจมากยิ่งขึ้น



รูปที่ 11 แสดงตัวอย่างหน้าจอของระบบ ซึ่งจะแสดงท่ารำ (Gesture number) ที่ระบบคำนวณได้หลังจากที่ผู้ใช้แสดงท่ารำใดๆ

7. เอกสารอ้างอิง

[1] Muneesawang Paisarn, Naimul Mefraz Khan, Matthew Kyan, R. Bruce Elder, Nan Dong, Guoyu Sun, Haiyan Li, Ling Zhong, and Ling Guan. (2015). A Machine Intelligence Approach to Virtual Ballet Training. MultiMedia, IEEE22, 2015(4), 80-92.

[2] Muneesawang Paisarn, Ning Zhang, and Ling Guan. (2014). Motion Database Retrieval with Application to Gesture Recognition in a Virtual Reality Dance Training System. Multimedia Database Retrieval. Springer International Publishing, 305-333.

[3] Kyan, Matthew, Guoyu Sun, Haiyan Li, Ling Zhong, Paisarn Muneesawang, Nan Dong, Bruce Elder, and Ling Guan. (2015). An Approach to Ballet Dance Training through MS Kinect and Visualization in a CAVE VirtualReality Environment. ACM Transactions on Intelligent Systems and Technology (TIST) 6(2), 23.

[4] Munsell, B. C., Temlyakov, A., Qu, C., & Wang, S. (2012, October). Person identification using full-body motion and anthropometric biometrics from kinect videos. In *European Conference on Computer Vision* (pp. 91-100). Springer Berlin Heidelberg.

[5] Jinda-apiraksa, A., Pongstiensak, W., & Kondo, T. (2010, May). A simple shape-based approach to hand gesture recognition. In *Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTI-CON), 2010 International Conference on* (pp. 851-855). IEEE.

[6] Shimada, A., & Taniguchi, R. I. (2008, December). Gesture recognition using sparse code of hierarchical SOM. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on* (pp. 1-4). IEEE.

[7] Saha, S., Ghosh, S., Konar, A., & Nagar, A. K. (2013, June). Gesture recognition from indian classical dance using kinect sensor. In *Computational Intelligence, Communication Systems and Networks (CICSyN), 2013 Fifth International Conference on* (pp. 3-8). IEEE.

[8] Qian, G., Guo, F., Ingalls, T., Olson, L., James, J., & Rikakis, T. (2004, June). A gesture-driven multimodal interactive dance system. In *Multimedia and Expo, 2004. ICME'04. 2004 IEEE International Conference on* (Vol. 3, pp. 1579-1582). IEEE.

[9] ศิริวัฒน์ ฉัตรเมธี. (2547). การวิเคราะห์เพลงรำวงมาตรฐาน. (วิทยานิพนธ์ปริญญามหาบัณฑิต, มหาวิทยาลัยมหิดล).

[10] Patsadu, O., Nukoolkit, C., & Watanapa, B. (2012, May). Human gesture recognition using Kinect camera. In *Computer Science and Software Engineering (JCSSE), 2012 International Joint Conference on* (pp. 28-32). IEEE.

[11] Microsoft Developer Network. Microsoft Kinect : JointType Enumeration [online]. Available: https://msdn.microsoft.com/enus/library/microsoft.kinect.jointtype.aspx

[12] Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (1996). *Neural network design* (Vol. 20). Boston: PWS publishing company.

[13] Khan, N. M., Lin, S., Guan, L., & Guo, B. (2014, December). A visual evaluation framework for in-home physical rehabilitation. In *Multimedia (ISM), 2014 IEEE International Symposium on* (pp. 237-240). IEEE.

[14] Teuvo Kohonen. (1995). Self-organizing map. Berlin : Springer-Verlag.

[15] Günter, Simon., Bunke, Horst. (2002). Self-organizing map for clustering in the graph domain. Pattern Recognition Letters 23(4), 405-417.