Green Intelligent Systems and Applications

Research Article
Volume 2(1), 2022, 42-52
https://doi.org/10.53623/gisa.v2i1.78

Sarawak Traditional Dance Motion Analysis and Comparison using Microsoft Kinect V2

Michael-Lian Gau[1], Huong-Yong Ting[1*], Jackie Tiew-Wei Ting[1], Marcella Peter[1], Khairunnisa Ibrahim[2]

[1]Drone Research and Application Centre, School of Computing and Creative Media, University of Technology Sarawak, No. 1, Jalan Universiti, 96000 Sibu, Sarawak
[2]Advanced Centre for Sustainable Socio-Economic and Technological Development, School of Computing and Creative Media, University of Technology Sarawak, No. 1, Jalan Universiti, 96000 Sibu, Sarawak

*Correspondence: alan.ting@uts.edu.my

SUBMITTED: 24 March 2022; REVISED: 10 April 2022; ACCEPTED: 13 April 2022

ABSTRACT: This research project aimed to develop a software program or an interactive dance motion analysis application that utilizes modern technology to preserve and maintain the Sarawak traditional dance culture. The software program employs the Microsoft Kinect V2 to collect the digital dance data. The proposed method analyses the collected dance data for comparison purposes only. The comparison process was executed by displaying a traditional dance on the screen where the user who wants to learn the traditional dance can follow it and obtain results on how similar the dance is compared to the recorded dance data. The comparison of the performed and recorded dance data was visualized in graph form. The comparison graph showed that the Microsoft Kinect V2 sensors were capable of comparing the dance motion but with minor glitches in detecting the joint orientation. Using better depth sensors would make the comparison more accurate and less likely to have problems with figuring out how the joints move.

KEYWORDS: Sarawak traditional; dance motion; motion analysis; motion comparison; Microsoft Kinect

1. Introduction

Sarawak is a multiracial state in Malaysia with diverse cultures and traditions. Some of the majority ethnic groups in Sarawak are the Iban (30.3%), Malays (24.4%), Chinese (24.2%), Bidayuh (8.4%), Orang Ulu (6.7%), Melanau (5.4%), Indian (0.3%) and the rest not listed (0.3%) [1]. These ethics groups have their own traditional dance culture to represent their racial identity or originality. One of the few traditional dances in Sarawak is the Iban's Ngajat and Kenyah Saga, which can be turned into video games to give space for youth to be a different person [2]. This can be an advantage in preserving the cultural heritage of Sarawak by motivating youth through modern technology such as the Microsoft Kinect V2. This involves motion that is non-rigid and singularity projection of the camera, a vast number of DOFs, and background and human clothing that cause the image noise [3]. Dancing is a form of multiple gestures expressed in a rhythmic form connected to a music piece [4]. The problem with dance recognition is the classification problem. The aim is to correctly identify the type of dance during the dancer's performance when given a frame or a sequence of structures [5]. According to [6], there is too little research on dance motion analysis due to its continuous and changeable limb movement and intensity. Furthermore, [7] shows how to plan a dance motion pattern for dancing humanoid robots using human head, hand, and leg gestures. Literature [8] uses Dynamic Time Warping (DTW) and Laban Movement analysis to evaluate dance motion based on dance poses.[9] also employs Laban Movement Analysis, which is linked to Motion Graphs and is capable of composing highly stylised dance motions. Since the introduction of real-time depth cameras, the method of obtaining body-tracking results has been greatly simplified. However, there are still limitations to the existing hardware. There was no consumer hardware that could handle a full range of body shapes and sizes interactively until the launch of the Microsoft Kinect sensor [10]. The inferred parts are projected into world space, with each part of the distribution localized to generate the 3D locations of each skeletal joint based on confidence-weighted. From a large motion capture database, artificial images of humans in various forms in terms of shape and size were used as the training data that was sampled. Thus, pose estimation is treated as object recognition using the intermediate body parts for joints of interest to be spatially localized [10]. There are two phases in motion comparison with Kinect: the recording and comparison phases [11]. One method of comparison using the k-nearest neighbor algorithm is used for classification with a non-parametric method in pattern recognition. The classification of an object is done by the majority vote of its neighbors, with the object being assigned to the class most common among its nearest neighbors. Thus, it is an instance-based type of learning in which the classification functions by approximation locally [12].

The Nave Bayes classifiers belong to a simple probabilistic classifier group that is based on the strong independence assumption between the features by the application of Bayes' theorem [13]. Nave Bayes is a conditional probability model where it assigns instance probabilities to each possible outcome. Ensembles of classifiers is an approach to combining multiple classifiers [14]. Better predictive performance can be obtained using a method that uses multiple learning algorithms rather than any part of the learning algorithm alone. Principal Component Analysis is an algorithm which sufficiently expresses the total variation in the original variable by extracting a small number of components [15]. This algorithm has been applied to common movement patterns such as locomotion. A complex and whole-body movement which involves everyday activities or expressive movement such as dancing, has not been used for this technique [16]. DTW is a notable system for finding an ideal arrangement between two given (time-dependent) sequences under specific limitations. Usually, the sequences are warped in a non-linear manner to coordinate with one another. At first, DTW was used in automatic speech recognition by comparing different speech patterns. However, the application of DTW in fields such as data mining and information retrieval has been successful as it can automatically deal with time deformation and different speeds that relate to time-dependent data [17]. Therefore, DTW was used for comparing the dance motion data in this research paper.

2. Materials and Methods

The components in the Microsoft Kinect V2 sensor, such as the RGB camera and depth sensor, are employed to collect dance data for analysis. For example, the Microsoft Kinect V2 sensor is capable of tracking the human body and detecting up to 25 joints [18]. In this research, the positions and orientations of the joints, which are the prime joints in Kenyah traditional dance, are extracted as follows: both wrists, neck, middle spine, both knees, both ankles.

The prime joints for the Kenyah traditional dance performance are the joints needed for positioning the dancers, determining the orientation of how the dancer's posture should be, and controlling the speed of the joints' motion during the dance performance, following the sound of the music. The first step of the feature extraction is assigning the joints' position and orientation into an array as a view model. The following are the features that were extracted from the Microsoft Kinect V2: Joint position (X, Y, Z coordinates of the joint) and joint orientation (Pitch, Yaw, and Roll angle orientations of the joint).

The 3-dimensional coordinates in the Microsoft Kinect V2 sensor's space are obtained using the IR sensor, which finds the joints' points. These coordinates are used to position the joints in a 3D space in this project. The center of the IR sensor on the Kinect is the origin coordinate point, which represents (x = 0, y = 0, z = 0). Joint orientations in Microsoft Kinect V2 are obtained using quaternion representation [19]. Typically, the orientation coordinates are assumed to be in the x-, y-, and z-axes, which correspond to pitch, yaw, and roll, respectively.is a 4D way of storing 3D orientation, which can later be converted into usable data in feature representation. There are two different ways of obtaining the joints' orientations. The absolute joint position and the relative joint orientation

The absolute joint position is the orientation angle placed only on the joints themselves. The algorithm for the absolute joint position is not relative to any parent or child joints. The problem with this absolute joint position is that during feature extraction, the value of the orientation received over time is not smooth due to a sudden change in a value called jittering, or in other words, "noise." This is because there is no support for the orientation angle, unlike relative joint orientation. Thus, joint smoothing filter algorithms are implemented to remove the noise during feature extraction. Relative joint orientation, also known as hierarchical joint orientation, is the joint angle relative to its parent joint. This reduces noise in the orientation value during feature extractions because there is enforcement from one joint to its parent joint. However, this type of joint orientation may cause issues such as the displayed skeletal image floating above the ground as the ankle joint is relative to the knee joint.

Feature representation involves classifying a specific sequence of gestures and actions using the extracted features of a recorded gesture or action of a joint, such as its position, orientation, and speed, which can be added into a data structure that is an array or list. In this project, the gesture or action is a sequence of dance moves. The specific basic Kenyah traditional dance moves for the Saga and the Datun Julud were recorded using the Kinect Studio application associated with the Microsoft Kinect V2 sensor. Subsequently, each of the dance moves is undergoing classification. The method to classify the sequence of specific dance moves is to add the required prime joints' features, such as position, orientation, and speed, into an array data structure. Then, the array is added into a list data structure used to add the value of the extracted features. The list function runs every frame in the recorded sequence of the dance move. For example, in one second, there are 30 frames. Therefore, if a single basic dance move uses 4 seconds, the calculation is 30 frames per second * 4 seconds. There are 120 frames of the prime joint features extracted into an array in a list data structure. The list data structure that contains the extracted features data is then exported in the Comma-Separated Value file format as illustrated in Table 1.

Table 1. Illustration of how the system collects the Kinect feature data in every frame using a List function that contains the array of the features joint data.

Frames captured	Joints Position			Joints Orientation		
	X-coordinate	Y-coordinate	Z-coordinate	Pitch	Yaw	Raw
Frame 1	*Sample x-1*	*Sample y-1*	*Sample z-1*	*Sample pitch-1*	*Sample yaw-1*	*Sample raw-1*
Frame 2	*Sample x-2*	*Sample y-2*	*Sample z-2*	*Sample pitch-2*	*Sample yaw-2*	*Sample raw-2*
Frame 3	*Sample x-3*	*Sample y-3*	*Sample z-3*	*Sample pitch-3*	*Sample yaw-3*	*Sample raw-3*
Frame n	*Sample x-n*	*Sample y-n*	*Sample z-n*	*Sample pitch-n*	*Sample yaw-n*	*Sample raw-n*

2.1. Analysis and Comparison of Saga' Kenyah Dance

As stated, Sarawak consists of many ethnicities, each having their own unique traditional dance style. In this project, the Kenyah male solo traditional dance known as Saga' Laki is used for the dance analysis and comparison. The solo dances for "Saga" are always performed spontaneously and freestyle. Thus, each performance is always different, with no absolutes. Then, the process for analyzing the traditional dances would be in place as a set of multiple dances moves into a single set, also known as the basics.

The orientation data needs to be obtained from the 13 prime joints, with each joint containing the three-orientation data: pitch, roll, and yaw. Thus, the entire set of data would be multiplied by 13 and then by 3, resulting in 39 sets of data to be extracted. Using this much data is not efficient as 39 sets are collected frame by frame. Thus, the feature extraction of the dance performance is obtained by using a particular number of joints for the dance body motion. An indexing formula is applied to the orientation data, which is as follows:

$$I = Pitch * 360 + Roll * 360^2 + Yaw$$

The index *I* value is used to compare dance body motions that require all three-orientation data, which reduces the load of the program as there is a large amount of data to compare. The solo Kenyah traditional dance has three main dance moves, as shown in Table 2. The whole solo performance contains multiple body dance motions that make up the traditional dance, as each dance motion has its own unique way of performing. As a result, the Sarawak traditional dance performance retains its uniqueness and identity.

Table 2. Segments of Solo Kenyah Saga's traditional dance motion

	Basic body dance motion	Joints Involved
Solo Kenyah Saga' Traditional Dance (Freestyle)	Arm	• Wrists
	Mid-body	• Neck • Middle Spine
	Footstep	• Ankles • Knee

Due to the three different body dance motions, the method to analyse and compare the dance motions is in three different segments. Therefore, collecting the dance motion data would require using the Kinect Studio to record dancers, focusing on the main joints involved in the dance motion required. The process of comparing traditional dances should begin with the easiest to learn and work its way up to the hardest. In the case of the Saga' Laki traditional dance, the most straightforward body dance motion to learn would be the footstep.

2.2. Saga' Laki Traditional Dance

In Figure 1A, the footstep body dance motion focuses on the lower body of the dancer, in which the dancer's joints involved would be the hip, knee, and ankle. The red lines indicate the dancer's right leg, and the white lines indicate the dancer's left leg. The image on the left is the first step of the footstep dance motion, and then the dancer lifts it backwards, bending his knee as shown in the image on the right. The same similar motion is repeated for the other legs, repeatedly following the rhythm of the music. The mid-body dance motion of the Saga's traditional dance, shown in Figure 1B, involves the performer spinning their whole body around gracefully at a slow tempo, following the rhythm of the music. Figure 1B shows the dancer making the basic body spin dance motion in which the spine, hip, and knee joints are used when turning their body around. Performing the graceful wrist spin of the Saga's traditional dance would be the most challenging dance motion compared to the footstep and body spin dance motions, as this dance motion is primarily spontaneous. However, the dancer can choose to perform it in a different direction. The initial dance motion for the arm is as shown in Figure 1C, where the performer swings their arm and flaps gracefully in a resemblance to the iconic Sarawak bird, the Hornbill. This dance motion was performed to the left and right of facing the audience.

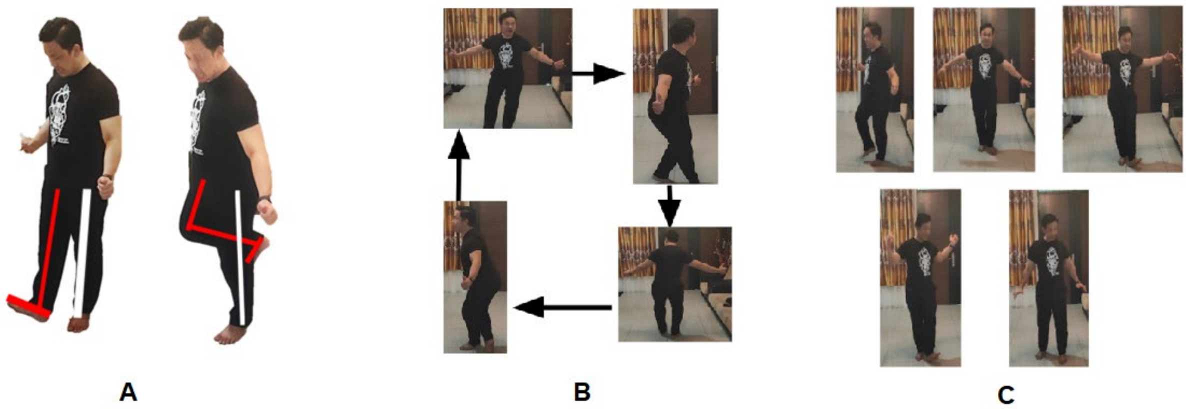

Figure 1. Images of Saga's Laki Traditional Dance: Footstep dance motion (A); Mid-body dance motion (B); Arm body dance motion (C)

3. Results and Discussion

The features are extracted using the software development kit from Microsoft to capture the orientation of the joints. Currently, the joints used for feature extraction are the orientation of both of the wrist's joints.

3.1. Saga' Kenyah Dance spontaneous action

One of the basic Kenyah Saga's dance moves is to gracefully spin with the arms posture, as shown in Figure 2A. However, from the second to the third picture of the figure, it is noticeable that the dancer's left arm starts to go haywire. However, the predictive algorithm could be used to fix this issue in future improvements. The first peak of the right wrist joint orientation in Figure 2C was due to the beginning of the dance move where the dancer moves his wrist to start performing. As observed after the peak, the orientation value changes gradually. This was because the performance does not require the dancer to orient the wrist too much. The peaks experienced in the third quarter of the graph were due to the right wrist being blocked from the view of the Microsoft Kinect V2 sensor. Thus, the right wrist orientation data starts to peak randomly as the hardware struggles to calculate and extract the data. As for the left wrist orientation, as observed in Figure 2B, the result is slightly consistent until the first quarter of the time frame in which the results start to peak randomly. The result of analyzing and comparing the Saga's Kenyah dance spontaneous action was not suitable as it combines multiple spontaneous dance motions into a single dance performance, which poses a challenge for the depth sensor to detect the joint orientation. Hence, the analysis and comparison of the traditional dance motion will be in segments, which will be explained in the next subsection.

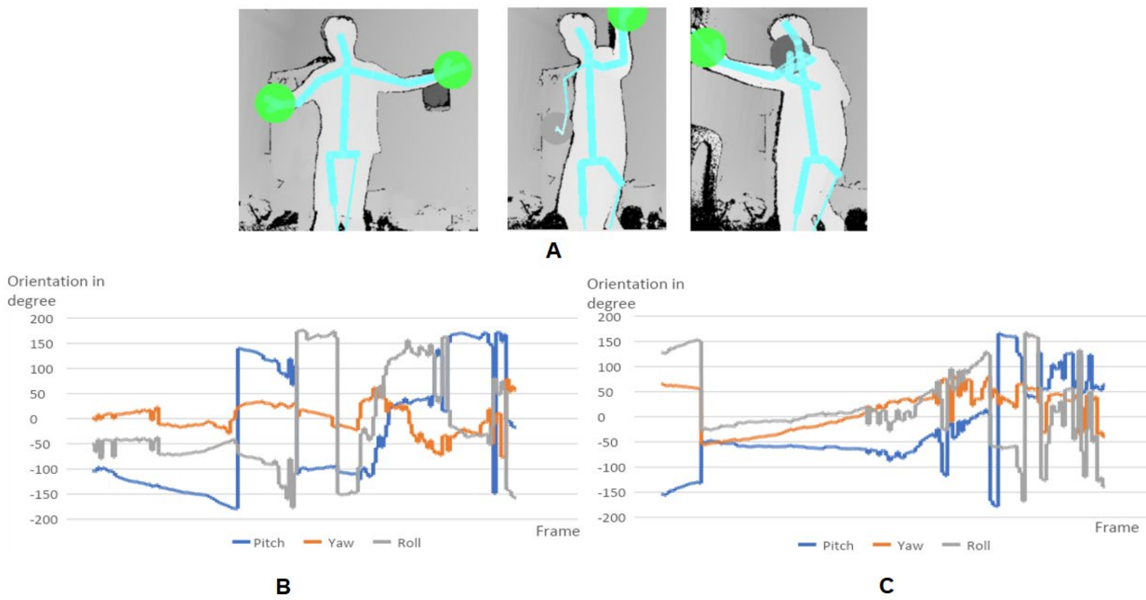

Figure 2. (A) Pre-recorded video of a Kenyah Saga' dance move, (B) left wrist orientation graph and (C) right wrist orientation graph

3.2. Analysis and Comparison of the Saga' traditional dance

The analysis and comparison for the Saga's traditional dance is separated into segments of different dance bodies: the arms, mid-body, and lower body. By having multiple segments of different dance motions, it will ease up the joint orientation detection for dance motion comparison. The program's user interface was designed to show the dance body motion selection where users can decide which dance motion, they would like to test for the dance motion comparison. The dance motion for each segment was collected by an experienced dancer, which the beginner dancer can view and perform using the program for dance motion comparison.

3.2.1. Footstep dance motion analysis and comparison

Figure 3A shows the dancer's starting position for the footstep dance motion. The dancer then performed a dance step for the left foot, then for the right foot. This sums up the basic dance moves for the footstep. The footstep dance motion focused mainly on the dancer's ankles and knee joints (refer Figures 3B and 3C). So, the ankle and knee joint orientation data are used for the analysis and comparison. Each graph has a repetitive pattern that describes the two dance steps in the Footstep dance motion. For the sake of comparison, the indexing formula was then applied to each of the joints to obtain one array of elements for each joint. After obtaining the index values of each of the joints, the comparison process is initiated. In this process, a different dancer performs the same footstep dance motion. DTW is utilized to display the comparison between the dance motion and the data obtained. Both graphs in Figure 4A and 4B of the footstep body motion show a similar pattern. The DTW has also displayed its similarity by connecting the points where the data between dances is similar.

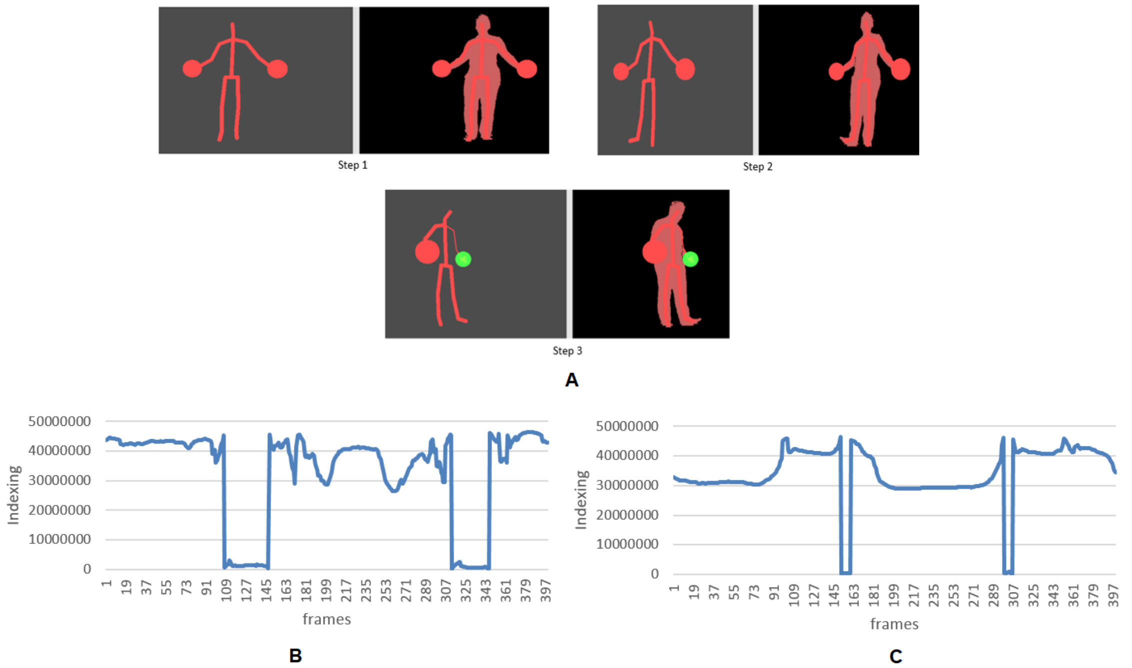

Figure 3. (A) Recorded footstep dance motion, (B) right ankle joint index graph and (C) right knee joint index

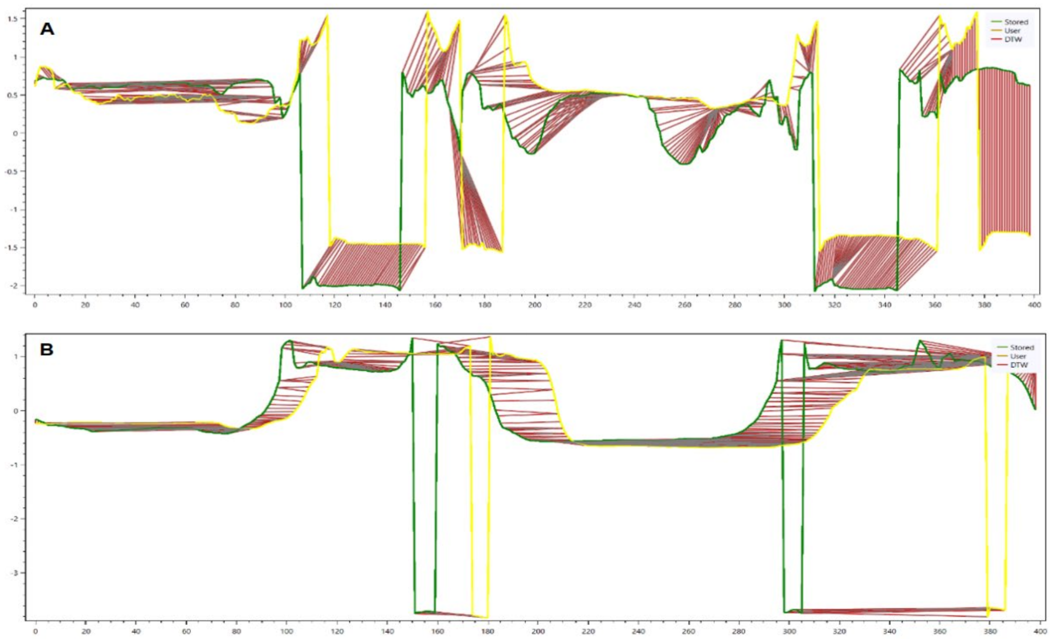

Figure 4. Comparison of right ankle (A) and knee (B) during footstep dance motion

3.2.2. Mid-body dance motion analysis and comparison

The mid-body dance motion of the Saga's traditional dance, as displayed in Figure 5A, involves the joint orientation of the spine and the neck joints of the dancer. Therefore, analysis and comparison were performed on the spine and neck joints. The data used for the comparison would be the yaw value of the joint orientations. The reason would be due to the limitations of the Microsoft Kinect V2 sensor, which is not able to properly recognize the joints when a user is facing his or her back towards the sensor. Due to this glitch, the graph shows values that spike to the top and then drop to the bottom, as shown in Figures 5B and 5C.

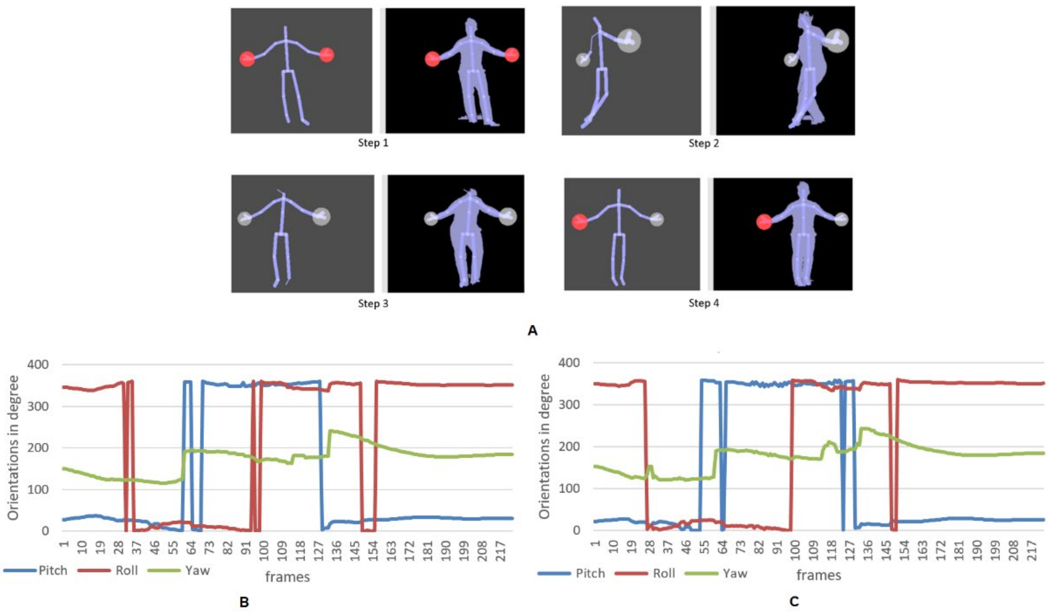

Figure 5. (A) Steps in Mid-body dance motion, (B) Neck orientation of the mid-body dance motion and (C) Spine orientation of the mid-body dance motion

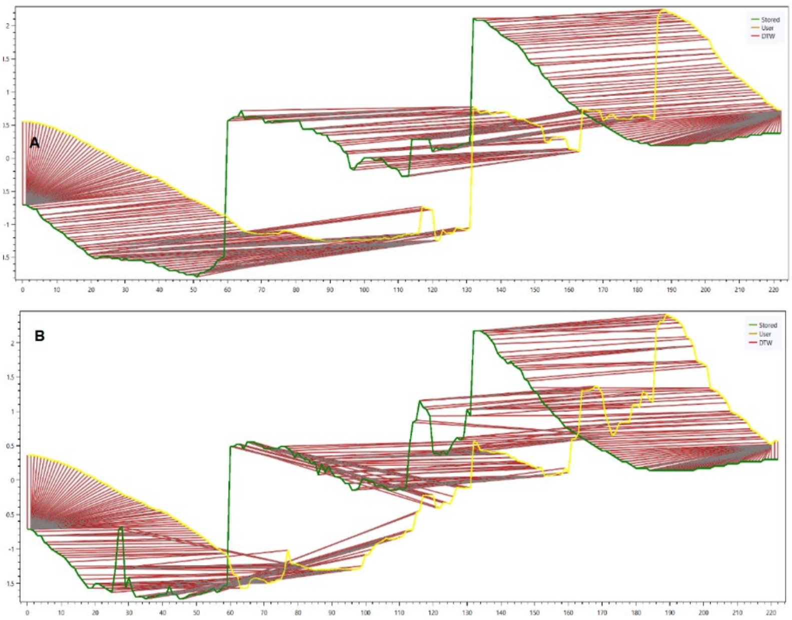

Figure 6. Comparison graph of (A) neck and (B) spine joint

Figure 6 demonstrates the comparison between the dancers in the mid-body dance motion. The array of values of the dance motion are a few frames apart, but the Dynamic Time Warping algorithm forms the line that connects both graphs that show the similarity. Thus, it shows a similar pattern between the graph and the mid-body dance motion. When the dance and the comparison are similar enough, a conclusion can be drawn that the mid-body dance motion was a good idea and that the dance was a good idea.

3.2.3. Arm body dance motion analysis and comparison

The arm-body dance motion begins with the arm of the dancer widely open, as shown in Figure 7A, and then he raises his wrist upwards to his right side as the first step of the arm-body dance. Then retract towards his body and raise his wrist upwards to his left side. Due to its major role in this dance move, the wrist joint is the only joint needed for analysis and comparison. As stated earlier, the arm-body dance was anticipated to be the most difficult to analyze and compare. Its nature is the most spontaneous dance motion compared to the footstep and mid-body dance motions. However, as observed in Figure 7B, the comparison at the end of the graph shows shorter lines of dynamic time warping, which can be used to compare the arm dance comparison.

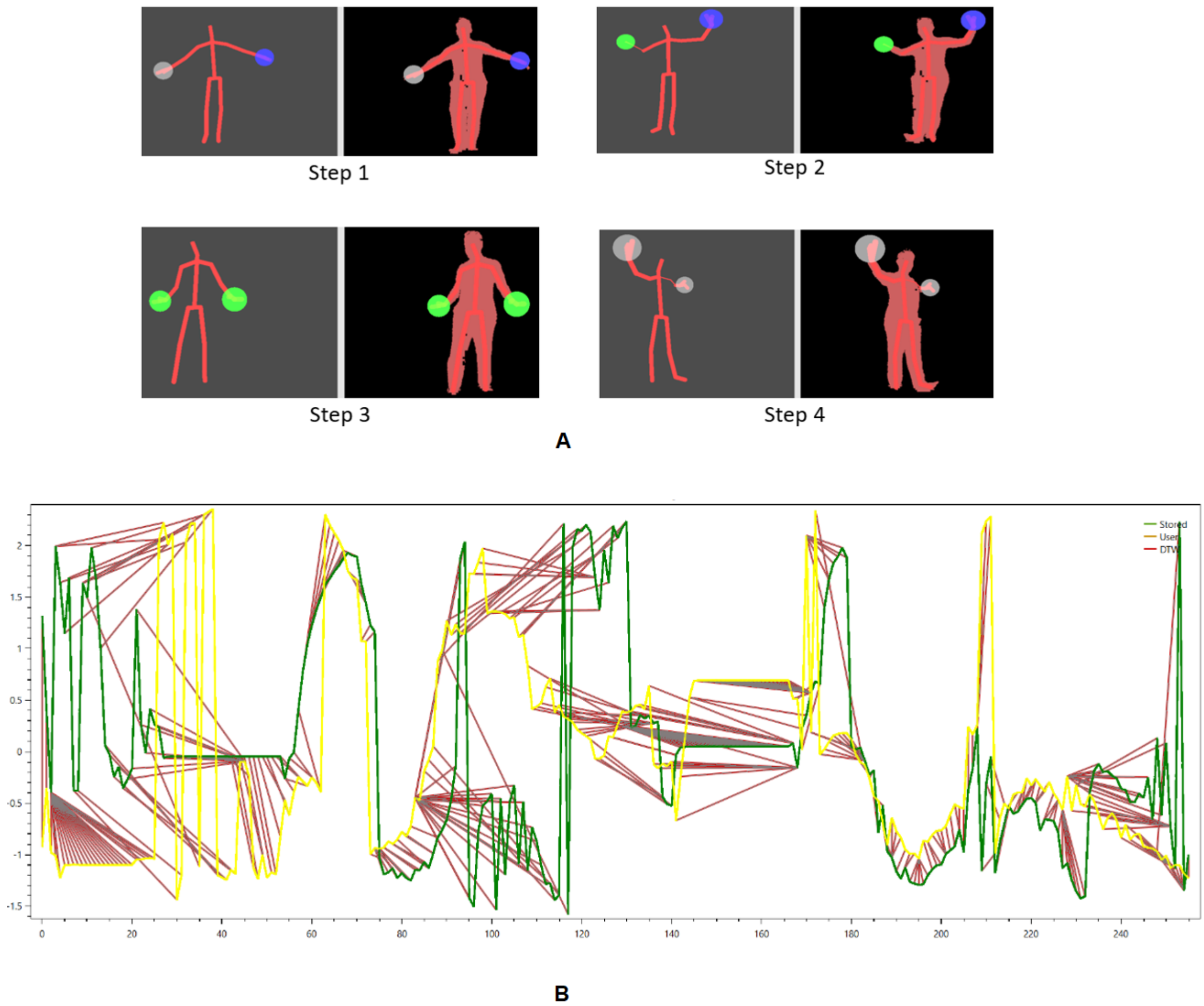

Figure 7. (A) Arm body dance motion steps and (B) left wrist comparison graph.

4. Conclusion

A Sarawak traditional dance motion analysis software prototype has been developed in this project. The proposed system is able to analyze and compare the Saga's traditional dance quantitatively. However, a few limitations were experienced during the development of this program. One of them is the limitation of the Microsoft Kinect V2 sensor itself, which cannot correctly detect the user's body when the user's back is facing towards the camera. Besides, due to extensive data, not all of the joints could be adopted for comparison. The only data reduction technique applied in this program was the indexing formula to turn the three orientation values into one index number. Due to hardware limitations, the program could not compare as much joint data as proposed due to efficiency and performance issues. Other limitations would be glitches in the collected data due to the noise. Future improvements can be made to widen the scope of the dance style to be learnt by adding other kinds of traditional dances such as the Ngajat, Bamboo Dance, and so on. Furthermore, as a new version of Kinect has recently been released, the Azure Kinect DK, with better hardware specifications, this software program can be migrated to utilize the new Kinect sensor, thus increasing efficiency in collecting dance data comparing them. In addition, data reduction techniques could be explored to reduce the load of the program when handling the comparison data in order to achieve better results.

Competing Interest

The authors declare no competing interests.

References

[1] State statistics: Malays edge past Chinese in Sarawak. (accessed on 15 March 2022) Available online: https://www.theborneopost.com/2014/02/08/state-statistics-malays-edge-past-chinese-in-sarawak/#:~:text=KUCHING%3A%20The%20Malays%20have%20overtaken,the%20Chinese%20stood%20at%20596%2C100.

[2] Liu, X. (2013). The Influence of Video Games on 21st Century Youth Identity. Master Thesis. University of Victoria, Canada.

[3] Huang, Yu; Huang, T.S. (2002). Model-based human body tracking. *2002 International Conference on Pattern Recognition*, 1, 552-555. https://doi.org/10.1109/ICPR.2002.1044791.

[4] Patra, B.G.; Das, D.; Bandyopadhyay, S. (2013). Unsupervised approach to Hindi music mood classification. Lecture Notes in Computer Science; Heidelberg, Germany; pp. 62–69. https://doi.org/10.1007/978-3-319-03844-5_7.

[5] Protopapadakis, E.; Voulodimos, A.; Doulamis, A.; Camarinopoulos, S. (2017). A Study on the Use of Kinect Sensor in Traditional Folk Dances Recognition via Posture Analysis. *Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments,* 305–310. https://doi.org/10.1145/3056540.3076200.

[6] Li, M.; Miao, Z.; Xu, W. (2021). CRNN-based attention-seq2seq model with fusion feature for automatic labanotation generation. *Neurocomputing, 454*, 430-440. https://doi.org/10.1016/j.neucom.2021.05.036.

[7] Abror, B.; Paramadihanto, D. (2017). Dance Motion Pattern Planning for K.Mei as Dancing Humanoid Robot. *International Conference on Robotics, Biomimetics, and Intelligent Computational Systems (Robionetics)*, 6-11. https://doi.org/10.1109/ROBIONETICS.2017.8203428.

[8] Jang, M.; Kim D.; Kim Y.; Kim J. (2017). Automated dance motion evaluation using Dynamic Time Warping and Laban Movement Analysis. *IEEE International Conference on Consumer Electronics (ICCE)*. https://doi.org/10.1109/ICCE.2017.7889261.

[9] Aristidou, A; Stavrakis, E.; Papaefthimiou, M; Papagiannakis, G.; Chrysanthou, Y. (2018). Style-based motion analysis for dance composition. *The Visual Computer, 34*, 12, 1725-1737. https://doi.org/10.1007/s00371-017-1452-z.

[10] Shotton, J.; Fitzgibbon, A.; Cook, M.; Sharp, T.; Finocchio, M.; Moore, R.; Blake, A. (2011). Real-time human pose recognition in parts from single depth images. *CVPR 2011*. https://doi.org/10.1109/CVPR.2011.5995316.

[11] Hemed, A. (2012). Motion Comparison using Microsoft Kinect. Project Report, Monash University, Melbourne.

[12] Nitin Bhatia, V. (2010). Survey of Nearest Neighbor Techniques. *International Journal of Computer Science and Information Security, 9*, 302-305 https://doi.org/10.48550/arXiv.1007.0085.

[13] Farid, D.M.; Zhang L.; Rahman C.M.; Hossain M.A.; Strachan, R. (2014). Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks. *Expert Systems with Applications, 41*, 1937–1946. https://doi.org/10.1016/j.eswa.2013.08.089.

[14] Rokach, L.; Schclar, A.; Itach, E. (2014). Ensemble methods for multi-label classification. *Expert Systems with Applications, 41*, 16: 7507–7523. https://doi.org/10.1016/j.eswa.2014.06.015.

[15] Daffertshofer, A; Lamoth, C.J.; Meijer, O.G.; Beek, P.J. (2004). PCA in studying coordination and variability: A tutorial. *Clinical Biomechanics, 19,* 415-428. https://doi.org/10.1016/j.clinbiomech.2004.01.005.

[16] Hollands, K.; Wing, A.; Daffertshofer, A. (2004). Principal components analysis of contemporary dance kinematics. Proceedings of the 3rd IEEE EMBSS UK & RI, University of Southampton, Southampton.

[17] Müller, M. (2007). Dynamic time warping. Information retrieval for music and motion. Springer: Berlin, Germany; pp. 69-84. https://doi.org/10.1007/978-3-540-74048-3_4.

[18] Mangal, N.K.; Tiwari, A.K. (2020). Kinect v2 tracked Body Joint Smoothing for Kinematic Analysis in Musculoskeletal Disorders. *Annual International Conference of the IEEE Engineering in Medicine & Biological Society,* 5769-5772. https://doi.org/10.1109/EMBC44109.2020.9175492.

[19] Hanson, A.J. (2006). Visualizing Quaternions (The Morgan Kaufmann Series in Interactive 3D Technology), 1st Ed.; Morgan Kaufmann: Burlington, United States.