

GESTURE RECOGNITION FOR PENCAK SILAT TAPAK SUCI REAL-TIME ANIMATION

Aminudin¹, Widi Hastomo², and L.M. Rasdi Rere³

^{1,2}Department Management, Faculty of Economic, ITB Ahmad Dahlan, Jakarta, Indonesia

^{2,3}Master of Information Technology, STMIK Jakarta STI&K, Jakarta, Indonesia

E-mail: widie.has@gmail.com

Abstract

The main target in this research is a design of a virtual martial arts training system in real-time and as a tool in learning martial arts independently using genetic algorithm methods and dynamic time warping. In this paper, it is still in the initial stages, which is focused on taking data sets of martial arts warriors using 3D animation and the Kinect sensor cameras, there are 2 warriors x 8 moves x 596 cases/gesture = 9,536 cases. Gesture Recognition Studies are usually distinguished: body gesture and hand and arm gesture, head and face gesture, and, all three can be studied simultaneously in martial arts pencak silat, using martial arts stance detection with scoring methods. Silat movement data is recorded in the form of oni files using the OpenNI™ (OFW) framework and BVH (Bio Vision Hierarchical) files as well as plug-in support software on Mocap devices. Responsiveness is a measure of time responding to interruptions, and is critical because the system must be able to meet the demand.

Keywords: *Gesture Recognition, Algorithm Dynamic time warping, Algorithm Genetic, Pencak Silat*

Abstrak

Target utama dalam penelitian ini yaitu sebuah desain system virtual training pencak silat secara real-time serta sebagai alat dalam belajar seni bela diri secara mandiri menggunakan metode algoritma genetic dan dynamic time warping. Pada paper ini masih di tahapan awal yaitu terfokus pada pengambilan data set pendekar pencak silat tapak suci menggunakan animasi 3D dan kamera sensor Kinect, ada 2 pendekar x 8 jurus x 596 gerakan/gesture = 9.536 kasus. Studi Gesture Recognition biasa dibedakan: hand and arm gesture, head and face gesture, dan body gesture, ke tiganya dapat dipelajari secara bersamaan dalam seni beladiri pencak silat, menggunakan deteksi jurus pencak silat dengan metode skoring. Data pergerakan silat direkam dalam bentuk file oni menggunakan kerangka OpenNI™ (OFW) dan file BVH (Bio Vision Hierarchical) serta perangkat lunak pendukung plug-in pada perangkat Mocap. Daya tanggap merupakan tolak ukur dari waktu menanggapi interupsi, dan merupakan hal kritis karena sistem harus dapat memenuhi permintaan.

Kata Kunci: *Gesture Recognition, Algoritma Dynamic time warping, Algoritma Genetik, Pencak Silat*

1. Introduction

One of the interesting research topics to be developed lately is intelligent and interactive computing that connects computers and humans [1],[2],[3],[4],[5],[6],[7]. The study of gesture recognition is an approach in this direction, where a movement made by someone can be understood by the computer. Gesture as a symbolic movement that expresses thoughts, feelings and concerns is not new, because in daily life we often use it, such as expressing agreement by nodding the head,

expressing sadness and joy in certain facial expressions, and pointing an object for choice desires [8].

One technology that can be used for further development on this topic is the launch of the Markerless Motion Capture camera from Kinect that can feel a movement and allow one to interact without touch [9],[10],[11]. At the beginning of the launch this camera is intended only for the gaming industry, but with a variety of advantages it has, this technology has begun to be widely used for development in the field of virtual education

[12],[13], engineering [14] and health services [15],[16].

This research will benefit from studying Gesture Recognition and the application of Markerless Motion Capture technology in education for more effective and efficient human and computer interactions [17],[18],[19], however this research is focused on collecting datasets. As a case study, this research develops a Real-Time animation system that is used to study martial arts [20] in Indonesia, namely Pencak Silat Tapak Suci Putera Muhammadiyah. The selection of pencak silat as a case study research was felt to be important because silat became a tradition or part of the Indonesian of Nation, Pencak silat Tapak Suci Muhammadiyah as one of Indonesia's noble cultural heritage, generally exploiting all body movements starting from the hands, feet, body and even facial expressions in one unit, so with pencak silat the three types of motion can be applied simultaneously and as a challenge in the detection of gesture recognition that is in every movement of pencak silat moves, must be detected by the camera sensor must be precise and accurate.

Someone who will learn martial arts will be confronted with the problem of the time available to him, the costs required, and the presence of a martial arts warrior who is willing to teach it. To overcome this problem can be done by developing a system, which does not require the physical presence of a warrior. Where one can learn martial arts just by looking at the movements of a martial arts warrior who will give him immediate feedback for any errors in his movements. For this to happen, we need a media that can read its movements when crossing and then compare them to the movements of virtual warriors in the system.

Pencak silat Tapak Suci Putera Muhammadiyah as Indonesia's martial art needs to be preserved and developed at this time, because it has begun to be excluded by other martial arts as alternative options such as Karate from Japan, Kungfu from China, and Taekwando from Korea, in addition to diminishing interest in the younger generation now to learn it. One way to preserve it is by making martial arts of the Putak Suci Putera Muhammadiyah martial arts in the form of a Real-Time animation system with an animated form made interesting.

2. Literature Review

Gesture Recognition is one of the topics in computer science that aims to be able to interpret the movements of the human body through mathematical algorithms [21]. Gesture can originate from any body movement from fingers [22],[23], arms [24],[25], legs, body [26],[27], to

displaying facial expressions [28]. But in general, research on gestures can be divided into 3 basic groups namely hand [29] and arm gestures, face gestures and head and frame gestures [30],[31]. General gesture recognition requires identification of the hand and position or direction of the hand, and can pinpoint some of the features appropriate to detect movement [32].

The kinect sensor camera implementation in karate martial arts uses fast dynamic time warping algorithm with the calculation of confusion matrix resulting in an average accuracy of up to 90% in all four movements [33], whereas in Ta'i Chi gesture recognition using hidden markov models get a recognition rate of 95.2% in five movements and an average error of 8% [34], gesture recognition Karate martial arts have also been created datasets, kicks and punches using the Kinect sensor camera with different skill levels, from beginner to master level [35]. Taijiquan gesture martial art dataset collection with samples reached 2200, with 12 participants using kinect sensor cameras [36].

An animation can be made in the traditional way or using 3D software. Traditional animation uses manual techniques, where every movement must be in a picture per frame, so that the shape and location of the object being drawn changes little by little, which can give the impression the object is moving as illustrated in figure 1. For one second the movement for example, using this technique is usually requires up to dozens of images.

Animation using 3D software only requires a few keyframes, namely the position of important objects at the specified time. Blank spaces between each keyframe are calculated automatically by the computer. Illustration of an animated concept with a keyframe technique for moving balls is shown in Figure 2.

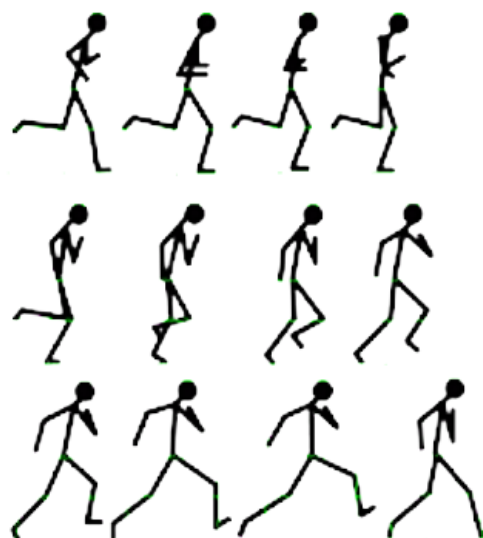


Figure 1. Illustration of Traditional Animation Concept

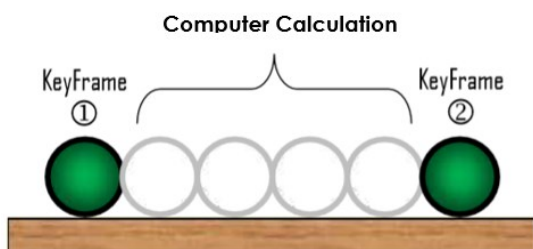


Figure 2. Illustration of The Concept of Animation With Keyframe Techniques

A real-time system is a computer system that not only requires correct computational results, but must also be in accordance with the desired deadline [34],[37]. Some characteristics that exist in real-time systems are determinism, responsiveness, reliability, safe-fail operations and user control [38]. System determinism depends on the speed of responding to interruptions and the capacity of the system to respond to requests in the time required. Responsiveness is a benchmark of time to respond to interruptions, and is critical because the system must be able to meet requests caused by external influences [39].

Advances in imaging hardware have led to the emergence of markerless motion capture technology using a camera system [40]. One of the most popular is Microsoft Kinect [41], as shown in Figure 3. which has the ability to capture depth and RGB data [42]. With its technology and relatively cheap price, its existence has caused a stir in the field of computer vision, thus triggering massive research within the box of object tracking and gesture recognition. Sensors in kinect cameras have the ability in real-time to estimate the invariance of body shape or gesture [43].

Markerless Motion Capture is a technique for taking human movements but does not use tools or special clothing used by actors [45]. This technique only uses a depth camera that will automatically determine the points of human joints [46]. Microsoft Kinect as a markerless motion capture tool has several advantages compared to other tools such as the Kinect range which is further away and

has a greater resolution [47]. As for the weakness, Kinect only has a relatively low frame rate, which ranges from 12 to 30 fps. Markerless Motion Capture is far more suitable for researchers because it can be used with simple and cheaper preparation in terms of cost [48].

Dynamic Time Warping (DTW) has long been used to measure the similarity between two signal sequences that vary in time and space [49],[50]. The DTW algorithm calculates the distance between each possible pair of points of two signals relating to the relationship of the two special values [51]. Using this distance, DTW calculates a cumulative distance matrix, to get the shortest path through this matrix [52]. The path represents the synchronization of the two signals which causes the special distance between the two points to be minimal. Usually this signal is normalized and smoothed before the distance between points is calculated [20].

Basically, DTW is a technique that permits a personal computer to search out an most advantageous fit among two given sequences with positive restrictions [53]. Non-linear sequences within the time measurement to resolve the scale in their similarity don't rely on positive non-linear permutations within the time dimension. For instance, two reference signals and a test signal, as shown in Figure 4, to compare and evaluate it with DTW are as follows: The two signals are similar in shape because they have a single peak. However, the reference signal is longer than the test signal, and the peak is one step further. In other words, the two signals are out of sync in time.

To calculate the variation among the two, imagine the gap matrix among every pattern from $x [t]$ and $y [t]$ as shown in Figure 5. The matrix calculation results are the lowest sequential numbers, close to the diagonal, indicating that the test sample $x [t]$ is close its value with the reference signal $y [t]$, shown in green. As a substitute of using a easy subtraction, it's standard to make use of a symmetrical distance measure, reminiscent of $(x [t] - y [t])^2$, so that the matrix is obtained.



Figure 3. Video Game Consoles Use Microsoft Kinect [44]

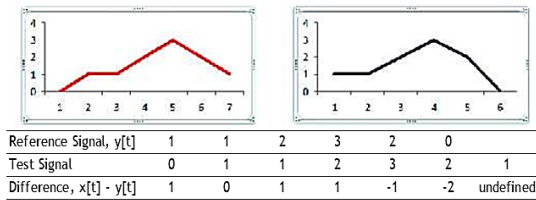


Figure 4. Comparison Between Two Patterns: (a) Reference Signal y [t] and (b) Test Signal x [t].



Figure 5. (a) Comparison of Patterns in Matrix Form, (b) Matrix Values Close to The Output. [20].

3. Method

Genetic Algorithms come from biology but are very influential in computational science in optimization. This method is very effective to get the optimal or less optimal solution to the problem because it only has a few obstacles [54]. This results in and tests a mechanism for a set of possible solutions (referred to as population in Genetic Algorithms) and brings an optimal acceptable solution. It carries out three basic operations (Reproduction, Crossover and Mutation) over and over in the population.

The first step The genetic algorithm is improved first to consider segmented image from the previous process Pixel image is an entropy of high value and quality, calculation of entropy value and pixel calculation of fitness value [55],[56].

$$X_i[x_0^{(i)}, x_1^{(i)}, \dots, x_{N_L-1}^{(i)}] \quad 0 \leq i \leq N_p - 1$$

$$0 \leq j \leq N_L - 1 \quad (1)$$

From Equation (1), $x_j^{(i)}$ represents jth chromosome gene, N_p is population pool, N_L is chromosome length. Fitness method, is the top target parameter until the value can be increased. Using equation (2), fitness capacity can be assessed.

$$F_i = \sum_{i=1}^n d_i / n \quad (2)$$

With the help of two parent chromosomes cross operation can be performed. child chromosomes choose the level of quality with the

level of hybrids [57]. Use the fitness capacity on the tyke chromosome that has just been sent after making another chromosome. In equation (3) a Chromosome is determined to calculate the hybrid rate:

$$\text{Crossover rate } (C_r) = \frac{CO_G}{C_L} \quad (3)$$

where, CO_G is the number of crossing genes, C_L .

Unified stimulated regulation is Operation of adaptive mutation, mutation operation is transmitted based on mutation rate (M_r) [58].

$$\text{Rate of mutation } (M_r) = \frac{M_p}{N_L} \quad (4)$$

where, M_p is the mutation point, N_L is the chromosome length.

In the methodology determining the selection of chromosomes will continue [58],[30], chromosomes are produced by N_p and are located on the premise, the best chromosomes are in the top position.

Dynamic Time Warping (DTW) is a well-known technique for finding optimal values between two time-dependent sequences below a certain limit. DTW has been successfully applied to overcome time deformation automatically and the different speeds associated with time-dependent data [59].

Microsoft researchers [60] argue that in detecting dance movements using the kinect sensor camera and DTW algorithm, the level of movement achieved has a fairly high accuracy of 96.9%, this means that a high degree of accuracy uses data from kinect sensor cameras using DTW algorithm can be achieved. How to calculate the distance between two points calculated through the Euclidean distance [61],

$$\text{Dist}(x, y) = |x - y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (5)$$

He path through which the grid is the best combination between these two sequences, and the distance between the two can be minimized from the total distance, this is called global distance.

To calculate the global distance must go through the grid and pass all routes, so that the overall distance can be calculated. To determine the minimum between individual elements from the total distance divided using the quantity of the weight function.

The grid will get bigger if each sequence and quantity of paths it passes. Recursive formulas can be used to obtain global distances [62].

$$GD_{xy} LD_{xy} \min (GD_{x-1}, (GD_{x-1}, GD_{x y} - 1) \tag{6}$$

Where;

GD = Global Distance (overall distance)

LD = Local Distance (Euclidean distance)

This study aims to obtain a dataset of the movements of pencak silat gestures, in the main target of making the design of the Virtual Pencak Silat Training System, so that players can develop their style and techniques in martial arts, based on direct feedback provided by virtual silat warriors in the system. Of the three Gesture Recognition characters such as; (a) hand and arm gesture, (b) face and head gesture, and (c) body gesture, can be studied simultaneously in martial arts martial arts, animation construction is made according to martial arts moves played using a combination of genetic algorithms and DTW.

As can be seen in Figure 6, the first step taken is to retrieve the movements of a martial arts warrior by using a Kinect camera. When shooting cross motion, at least 4 cameras will be placed on the front, right and left side and back. Silat movement data is recorded in the form of oni files using the OpenNI™ framework (OFW) and BVH (Bio Vision Hierarchical) files using Mocap Device Plug-In software. As an additional reference, data on pencak silat movements are also recorded with a Digital camera. The next stage is to construct a silat swordsman model construction based on the results of data retrieval using a Kinect camera, digital camera, and photo models determined as silat swordsmen. Model construction is made using Blender software that can provide 3D effects.

The expected indicator of achievement in this research is a system that can display the Silat Pendekar Silat dataset and show various movements of Tapak Suci martial arts and optimization of the kinect sensor camera capture distance. The dataset will be used as a comparison of students practicing martial arts, animation of Silat warriors is run using the Mocap Device Plug-In (MDP) software [63].

The design of the Pencak Silat Virtual Training System developed in this study, shown as shown in Figure 7, consists of several modules, namely the Kinect camera module for capturing silat movements, data processing modules, Gesture model modules as a result of silat swordsman training data, movement compatibility computation modules and real-time animation modules that can provide feedback.

The final stage to be achieved in this study is a virtual pencak silat training system using gesture recognition, is expected to be a new contribution in the field of human computer interaction (HCI) on

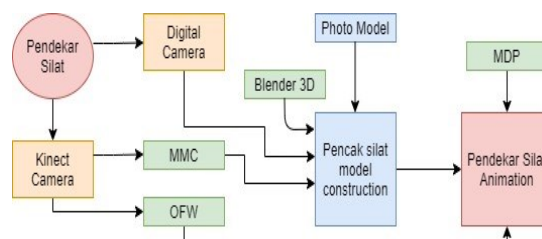


Figure 6. Stages of Gesture Model Design

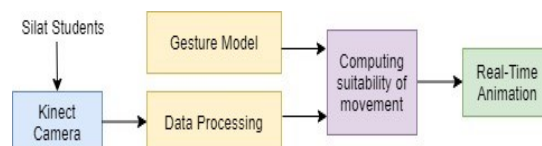


Figure 7. Virtual Training System Diagram Block

martial arts for learning, and towards the target the collection of animated swordsman models is carried out. silat using MDP, then the detection of martial arts using the scoring method, more details are as follows, the process must go through five modules. In the first module, the Kinect™ camera, which is a camera capturing motion without markers with the ability to capture the movements of someone who will learn pencak silat without intermediary media, based on tracking the framework. This camera is connected to a computer using the OpenNI™ framework and NITE middleware from PrimeSense™.

After the second Module, the data processing is generated from the previous module, based on the position of the time series body frame such as the position of the head, elbows, arms, knees and feet to a certain reference point that has been determined. The third module is derived from the gesture data model that results from the movement of a martial arts warrior, and has classified each movement according to the silat attitude it plays, where the data is stored as a database reference in the system. The fourth module is the calculation process to determine the suitability of the movement between the movement model and the input data of someone who will learn silat. Genetic Algorithm Method and DTW are used to compare two data.

The results of the movement suitability calculation will be displayed in the fifth module, which will display virtual martial arts warriors and martial arts students side by side. The system will provide an assessment of the movements performed by martial arts students with the specified values and cues. To get optimal results from this learning system, several stages must be passed. First is students learn all the basic movements by watching the silat warrior animation displayed by the system. The second stage of

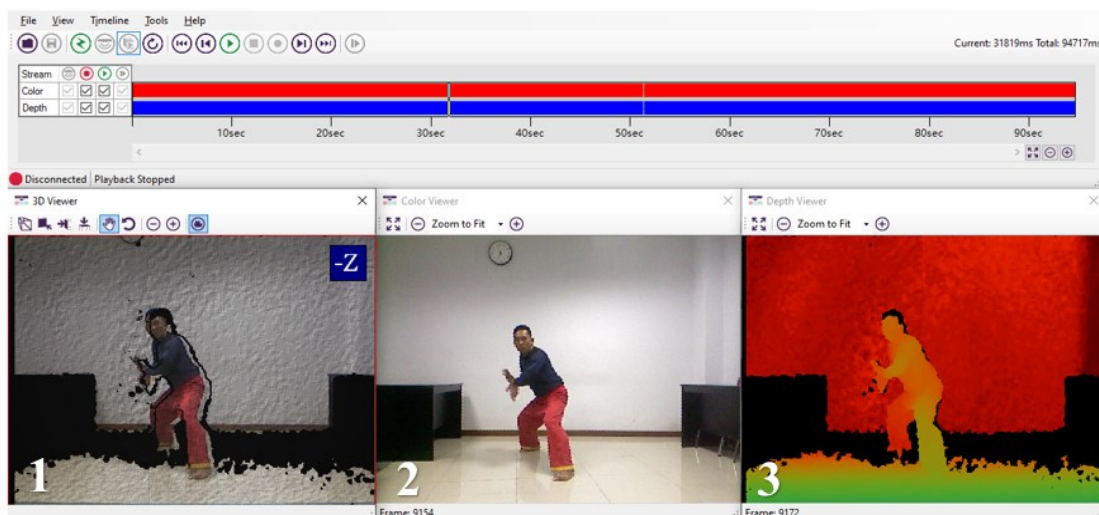


Figure 8. Kinect Studio virtual training output [1] 3D Viewer, [2] Color Viewer, [3] Depth Viewer

learning combines the basic movements that have been learned into complete silat movements in a slow tempo, and the third stage is towards the perfection of cyclic movements.

4. Result and Discussion

At this stage of the research is the retrieval of training data (data set) for silat swordsman model construction using the kinect SDK camera. Dataset that will be recorded there are eight moves not including the martial arts movements, steps, avoidance and basic moves, each stance is played by two martial tread warriors in turn. The number of datasets is 2 warriors x 8 moves x 596 cases / gesture = 9,536 cases. This data set consists of 3 visualization displays, 3D viewer, color viewer and depth viewer, as shown in Figure 8. The best distance between a kinect camera and a martial arts object is between 1.2 meters and 4 meters.

Kinect is able to produce a 32-bit RGB Color viewer stream at 30 frames per second for video resolution at 640x480 pixels, furthermore the depth viewer of this stream consists of 16 bits per pixel, and each pixel in the depth viewer uses 13 bits (high order) for data depth and 3 bits (lower order) to identify silat warriors. The infrared ray emitted on the kinect camera cannot be seen visually by the eye and is harmless, the reflection of the infrared ray is to detect the presence of silat warrior objects. For the depth range the best distance is 400mm (closest distance) to 8000mm.

The next step is to install the Kinect SDK sensor with the Mocap Device Plug-In software, the hardware requirements to install MDP namely; (a) Operating System: Windows 8 (64 bits) or above, (b) Multimedia-related API: DirectX 11 or above, (c) Internet Accessibility: Must, (d) iClone:

6.21 or above. How to install it; (1) Execute the install program on the computer where the Microsoft's Kinect (for Xbox 360 or for Windows) depth camera is installed, (2) If you encounter the message below, then please go to the Microsoft download web site to download and install the .Net Framework 2.0 (or above) on the computer, (3) Follow the instructions to finish the installation of the Mocap Device Plug-in, (4) If you are prompted with a message for installing the driver for your depth camera, then click Install unless you have installed the driver before.

To optimize the range of kinect camera catches that MDP has installed, that is, in Figure 9 the side view for the optimization point is in the range of camera height between 0,5 meters to 1 meter, the optimization point is between 2,5 meters to 5 meters. For the display from above the catch point of the optimization point in the range of 1,75 meters to 2 meters can be seen in Figure 10, the jump movements in the eagle stance are captured by the sensor better at a distance of 3 meters, so that the movement of silat warriors in this dataset records maximum at the distance 5 meters for maximum kinect sensor detection.

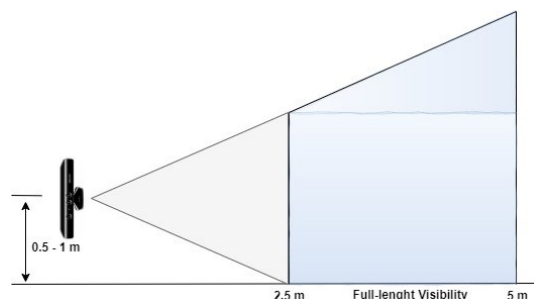


Figure 9. Optimize Kinect Camera Capture From The Side

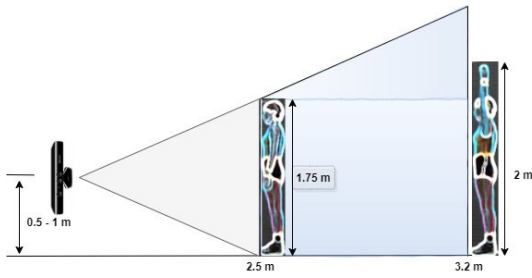


Figure 10. Optimization of Display Kinect Camera Capture from Above

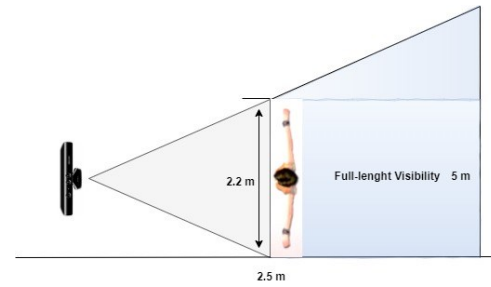


Figure 11. Optimization of Kinect Camera Capture Display from The Front

To optimize the appearance of the kinect camera capture from the front can be explained in Figure 11, the optimization point is at a distance of 2,5 meters to 5 meters, with a width of a hand span of 2,2 meters. Respect movements and greetings opening the optimization results at a distance of at least 2,5 meters, meanwhile at the movement of the tiger sleeping style the optimization point is at a distance of 3 meters. Next insert characters or avatars for virtual silat warriors, connect between the mocap device and the mocap plug-in device until the motion point pattern appears, stabilized by performing a calibration, connected to the iclone application then previewing and recording, so the character will follow the movements of martial arts warriors real time like Figure 12.

The dataset recorded at MDP and seen in figure 12 there is no evaluation of the suitability and accuracy of the movement, because in the initial stages of this study (the first grant) is still limited to dataset records, but in the second stage of research (the second grant) it might be able to use the precision, recall method , accuracy, f1-score, error etc. only limited to dataset records.

5. Conclusion

The results of this study are the sacred tread silat warrior dataset recorded using a kinect camera with MDP, the number of datasets is 2 warriors x 8 stance x 596 cases / gesture = 9.536 cases, the process of taking the dataset for 2 days. For measurement performance in future work genetic algorithms [64],[65],[66],[67],[68] and DTW [49] [50],[69],[70],[71],[72] will be used. DTW is used to compare master data (considered as a teacher, data that has been recorded in this study), with data on students who want to learn pencak silat (online). So you can find out which student movements are not in accordance with the teacher's movement. Alan and antonio [62] has argued that the use of the DTW algorithm makes steps simpler, reduces gesture detection time and increases the level of recognition. Several studies [73],[74],[75], has suggested that the use of genetic algorithms to determine the overall optimization method by building arithmetic functions to find some optimal parameters so that the level of accuracy is expected to be better.

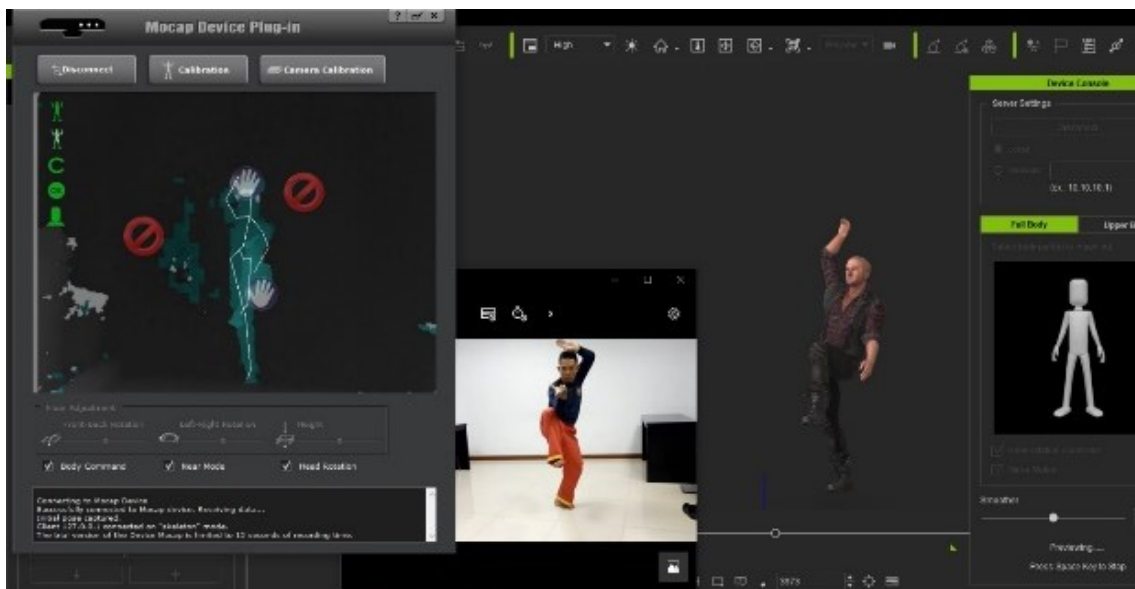


Figure 12. Pencak Silat Virtual Training

Optimization of the results of the catch distance at a distance of 2.5 meters to 5 meters, the obstacle faced is if pencak silat movement is too fast, and too close to the kinect camera, then skeleton detection does not function properly. This research creates a real-time pencak silat martial arts animation system based on the tracking skeleton captured by the Kinect camera.

The contribution of this research is very significant especially for Muhammadiyah organizations because it can introduce more of one of their "valuable heritage", namely Pencak Silat Tapak Suci, to the general public both nationally and internationally. This can be proven from the grants received to conduct this research.

Acknowledgement

The author would like to thank PP Muhammadiyah for funding this first year of research in the Batch III HibahMU program.

References

- [1] V. Ayumi and M. I. Fanany, "Distribution-Sensitive Learning on Relevance Vector Machine for Pose-Based Human Gesture Recognition," *Procedia Computer Science*, vol. 72, pp. 527–534, 2015.
- [2] G. Luo, P. Yang, M. Chen, and P. Li, "HCI on the Table: Robust Gesture Recognition Using Acoustic Sensing in Your Hand," *IEEE Access*, vol. 8, pp. 31481–31498, 2020.
- [3] Y. Li, Z. Chu, and Y. Xin, "Posture Recognition Technology Based on Kinect," *IEICE Transactions on Information and Systems*, vol. E103D, no. 3, pp. 621–630, 2020.
- [4] V. T. Hoang, "HGM-4: A new multi-cameras dataset for hand gesture recognition," *Data in Brief*, p. 105676, 2020.
- [5] S. B. and M. N. Soumi Paul, Arpan Bhattacharyya, Ayatullah Faruk Mollah, *Hand Segmentation from Complex Background for Gesture Recognition Soumi*, In: Mandal, J. K., Bhattacharya, Debika (Eds.), *Emerging Technology in Modelling and Graphics*, vol. 937, Springer Singapore, 2020.
- [6] D. D. Raymond Kirk Price, Michael Bleyer, "Multi-Spectrum Illumination and Sensor Module for Head Tracking, Gesture Recognition and Spatial Mapping." U.S. Patent, No. 10628950, 21 April 2020.
- [7] E. Escobedo-Cardenas and G. Camara-Chavez, "A Robust Gesture Recognition Using Hand Local Data And Skeleton Trajectory," *In IEEE International Conference on Image Processing*, 2015, pp. 1240–1244.
- [8] A. Butalia, A. K. Ramani, M. Ingle, and P. Kulkarni, "Classification of Sentiments through Rough Fuzzy Approach," *International Journal of Computer Science, Information Technology, & Security*, vol. 2, no. 2, 2012.
- [9] Y. Ma, D. Liu, and L. Cai, "Deep Learning-Based Upper Limb Functional Assessment Using a Single Kinect v2 Sensor," *Sensors (Switzerland)*, vol. 20, no. 7, pp. 1–26, 2020.
- [10] B. Milosevic, A. Leardini, and E. Farella, "Kinect and Wearable Inertial Sensors for Motor Rehabilitation Programs at Home: State of The Art and An Experimental Comparison," *Biomedical Engineering Online*, vol. 19, no. 1, p. 25, 2020.
- [11] S. S. and A. S. J. Potnis, "Real-Time Face Detection System for Virtual Classroom Environment", 2014.
- [12] P. K. Bun, R. Wichniarek, F. Górski, D. Grajewski, P. Zawadzki, and A. Hamrol, "Possibilities and determinants of using low-cost devices in virtual education applications," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 2, pp. 381–394, 2017.
- [13] S. S. Potnis and A. S. Jahagirdar, "Real Time Hand Gesture Recognition for Smart Classroom Environment," *International Journal of Computer Trends and Technology*, vol. 17, no. 2, pp. 78–83, 2014.
- [14] Q. Wan, Y. Li, C. Li, and R. Pal, "Gesture Recognition for Smart Home Applications Using Portable Radar Sensors," *In 36th of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 6414–6417.
- [15] M. Sathiyarayanan and S. Rajan, "MYO Armband for of the IEEE Engineering in Medicine and Biology Society," *In 8th International Conference on Communication Systems & Networks*, 2016.
- [16] S. Benatti *et al.*, "A Sub-10mW Real-Time Implementation for EMG Hand Gesture Recognition Based on a Multi-Core Biomedical SoC," *In 7th IEEE International Workshop on Advances in Sensors and Interfaces*, 2017, pp. 139–144.
- [17] Abolghasem Sadeghi-Niaraki and Soo-Mi Choi 2, "A Survey of Marker-Less Tracking and Registration Techniques for

- Health & Environmental Applications to Augmented Reality and Ubiquitous Geospatial,” *Sensors (Switzerland)*, 2020.
- [18] R. H. Abiyev and M. Arslan, “Head Mouse Control System for People With Disabilities,” *Expert Systems*, vol. 37, no. 1, pp. 1–14, 2020.
- [19] V. Kazllarof, “Automated Hand Gesture Recognition for Educational Applications,” *In Proceedings of the 20th Pan-Hellenic Conference on Informatics (PCI '16)*, 2016, Article 20, pp. 1–6, doi: 10.1145/3003733.3003746.
- [20] Keerthy, Naveen Kumar, “Virtual Kung Fu Sifu with Kinect,” *Master Thesis*, San Jose State University, 2012, doi: 10.31979/etd.d86z-hzmm
- [21] S. Kumar, N. Bansal, and S. N. Singh, “Smart Interaction Using Hand Gesture Recognition,” *In Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, 2016, pp. 1–4. doi: 10.1145/2905055.2905269
- [22] Y. Zhou, G. Jiang, and Y. Lin, “A Novel Finger And Hand Pose Estimation Technique For Real-Time Hand Gesture Recognition,” *Pattern Recognition*, vol. 49, no. August, pp. 102–114, 2016.
- [23] H. Li, W. Yang, J. Wang, Y. Xu, and L. Huang, “WiFinger: Talk to Your Smart Devices With Finger-Grained Gesture,” *In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 250–261, doi: 10.1145/2971648.2971738.
- [24] B. Milosevic, E. Farella, and S. Benatti, “Exploring Arm Posture and Temporal Variability in Myoelectric Hand Gesture Recognition,” *In 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*, 2018, pp. 1032–1037, doi: 10.1109/BIOROB.2018.8487838.
- [25] E. Mathe, A. Mitsou, E. Spyrou, and P. Mylonas, “Arm Gesture Recognition Using a Convolutional Neural Network,” *In 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, 2018, pp. 37–42, doi: 10.1109/SMAP.2018.8501886.
- [26] Y. Heryadi, M. I. Fanany, and A. M. Arymurthy, “Stochastic Regular Grammar-Based Learning for Basic Dance Motion Recognition,” *In International Conference on Advanced Computer Science and Information Systems*, 2013, pp. 419–424, doi: 10.1109/ICACSIS.2013.6761612 .
- [27] C. Corneanu, F. Noroozi, D. Kaminska, T. Sapinski, S. Escalera, and G. Anbarjafari, “Survey on Emotional Body Gesture Recognition,” *Journal of IEEE Transactions on Affective Computing*, 2018, doi: 10.1109/TAFFC.2018.2874986.
- [28] H. Joo, T. Simon, and Y. Sheikh, “Total Capture: A 3D Deformation Model for Tracking Faces, Hands, and Bodies,” *In IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8320–8329, doi: 10.1109/CVPR.2018.00868
- [29] M. E. and M. K. H. Lahiani, “Real Time Hand Gesture Recognition System For Android Devices,” *In 15th International Conference on Intelligent Systems Design and Applications*, 2015, pp. 591–596, doi: 10.1109/ISDA.2015.7489184.
- [30] S. Mitra and T. Acharya, “Gesture Recognition: A Survey”, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 3, pp. 311–324, 2007, doi: 10.1109/TSMCC.2007.893280.
- [31] V. Vishal Vadodaria , Fairfax, “Intelligent Agent/Personal Virtual Assistant With Animated 3d Persona, Facial Expressions, Human Gestures, Body Movements and Mental States,” U.S. Patent No.10178218, 08 January 2019.
- [32] T. et al. Jadooki, S., Mohamad, D., Saba, “Fused Features Mining for Depth-Based Hand Gesture Recognition To Classify Blind Human Communication,” *Neural Computing and Application* 28, 3285–3294, 2017, doi: 10.1007/s00521-016-2244-5.
- [33] B. Emad et al., “iKarate: Improving Karate Kata” *Procedia Computer Science*, vol. 170, no. 2019, pp. 466–473, 2020, doi: 10.1016/j.procs.2020.03.090.
- [34] David Alan Becker, “Sensei : A Real-Time Recognition Feedback and Training system for Ta’ici Gesture”, M.S Thesis, Program in Media Arts & Sciences, Massachusetts Institute of Technology, Massachusetts, U.S.A, 1997.
- [35] S. Bianco, F. Tisato, D. Dipartimento, and S. Comunicazione, “Karate Moves Recognition from Skeletal Motion,” *Electronic Imaging*, 2013, doi: 10.1117/12.2006229.
- [36] M. Tits, S. Laraba, E. Caulier, J. Tilmanne, and T. Dutoit, “UMONS-TAICHI: A

- Multimodal Motion Capture Dataset of Expertise In Taijiquan Gestures,” *Data in Brief*, vol. 19, pp. 1214–1221, 2018, doi: 10.1016/j.dib.2018.05.088.
- [37] R. M. Gurav and P. K. Kadbe, “Real Time Finger Tracking and Contour Detection For Gesture Recognition Using OpenCV,” *In International Conference on Industrial Instrumentation and Control*, 2015, pp. 974–977, doi: 10.1109/IIC.2015.7150886.
- [38] Fernando, “Image-based real-time gesture recognition,” 2019.
- [39] G. Paraskevopoulos, E. Spyrou, and D. Sgouropoulos, “A Real-time Approach for Gesture Recognition Using The Kinect Sensor,” *In 9th Hellenic Conference on Artificial Intelligence*, 2016, pp. 1–4, doi: 10.1145/2903220.2903241.
- [40] K. Miądlicki and M. Pajor, “Real-time gesture control of a CNC machine tool with the use Microsoft Kinect sensor,” *International Journal of Scientific & Engineering Research*, vol. 6, no. 9, pp. 538–543, 2015.
- [41] W. Zhao, “A Concise Tutorial on Human Motion Tracking and Recognition with Microsoft Kinect,” *Science China Information Sciences*, vol. 59, no. 9, pp. 1–5, 2016, doi: 10.1007/s11432-016-5604-y.
- [42] V. Ayumi and M. I. Fanany, “A Comparison of SVM and RVM for Human Action Recognition,” *Internetworking Indonesia Journal*, vol. 8, no.1, 2016.
- [43] A. Raheja, J. L., Chandra, M., & Chaudhary, “3D Gesture Based Real-Time Object Selection and Recognition”, *Pattern Recognition Letter*, vol. 115, pp. 14-19, 2018, doi: 10.1016/j.patrec.2017.09.034.
- [44] T. Berg, D. Chattopadhyay, M. Schedel, and T. Vallier, “Interactive Music: Human Motion Initiated Music Generation Using Skeletal Tracking By Kinect,” *In Society for Electro-Acoustic Music in the United States*, pp. 1–6, 2012.
- [45] E. Knippenberg, J. Verbrugghe, I. Lamers, S. Palmaers, A. Timmermans, and A. Spooren, “Markerless Motion Capture Systems as Training Device In Neurological Rehabilitation: A Systematic Review of Their Use, Application, Target Population and Efficacy”, *Journal of Neuroengineering And Rehabilitation*, vol. 14, no. 1, pp. 1–11, 2017, doi: 10.1186/s12984-017-0270-x.
- [46] A. Schmitz, M. Ye, R. Shapiro, R. Yang, and B. Noehren, “Accuracy and Repeatability of Joint Angles Measured Using a Single Camera Markerless Motion Capture System,” *Journal of Biomechanics*, vol. 47, no. 2, pp. 587–591, 2014, doi: 10.1016/j.jbiomech.2013.11.031.
- [47] A. Shingade and A. Ghotkar, “Animation of 3D Human Model Using Markerless Motion Capture Applied To Sports,” *International Journal of Computer Graphics & Animation*, vol. 4, no. 1, pp. 27–39, 2014, doi: 10.5121/ijcga.2014.4103.
- [48] S. Corazza, L. Mündermann, E. Gambaretto, G. Ferrigno, and T. P. Andriacchi, “Markerless Motion Capture Through Visual Hull, Articulated Icp And Subject Specific Model Generation,” *International Journal of Computer Vision*, vol. 87, no. 1–2, pp. 156–169, 2010, doi: 10.1007/s11263-009-0284-3.
- [49] Y. Heryadi, M. I. Fanany, and A. Murni, “A Skeleton Descriptor for Kinesthetic Element Recognition of Bali Traditional Dances,” *In Fourth International Conference on Computational Intelligence, Modelling and Simulation*, 2012, pp. 43–47, doi: 10.1109/CIMSim.2012.78.
- [50] G. Plouffe and A. M. Cretu, “Static and Dynamic Hand Gesture Recognition in Depth Data Using Dynamic Time Warping,” *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 2, pp. 305–316, 2016, doi: 10.1109/TIM.2015.2498560.
- [51] W. Ahmed, K. Chanda, and S. Mitra, “Vision Based Hand Gesture Recognition using Dynamic Time Warping for Indian Sign Language,” *In 2016 International Conference in Information Science*, 2016, pp. 120–125, doi: 10.1109/INFOSCI.2016.7845312.
- [52] A. Ribó, D. Warchol, and M. prz edu pl Oszust, “An Approach to Gesture Recognition with Skeletal Data Using Dynamic Time Warping and Nearest Neighbour Classifier,” *Journal of Intelligent Learning Systems and Applications*, vol. 8, no. 6, pp. 1–8, 2016, doi: 10.5815/ijisa.2016.06.01.
- [53] F. Jiang, S. Zhang, S. Wu, Y. Gao, and D. Zhao, “Multi-layered Gesture Recognition With Kinect,” *Journal of Machine Learning Research*, vol. 16, pp. 227–254, 2015.
- [54] S. N. Sivanandam, S. N., & Deepa, *Genetic Algorithm Implementation Using Matlab. In: Introduction to Genetic Algorithms.*

- Springer, Berlin, Heidelberg, 2008.
- [55] G. Abo Samra ; F. Khalefah, "Localization of License Plate Number Using Dynamic Image Processing Techniques and Genetic Algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 18 , issue 2, pp. 244–257, 2014, doi: 10.1109/TEVC.2013.2255611.
- [56] Y. Zhang, S. Wang, and J. Li, "Improved Particle Filtering Techniques Based on Generalized Interactive Genetic Algorithm," *Journal of Systems Engineering and Electronics*, vol. 27, no. 1, pp. 242–250, 2016.
- [57] M. Yang, M. H., Ahuja, N., & Tabb, "Extraction of 2D Motion Trajectories and Its Application to Hand Gesture Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1061–1074, 2002, doi: 10.1109/TPAMI.2002.1023803.
- [58] R. Kaluri and Pradeep Reddy, "An Overview on Human Gesture Recognition," *International Journal of Pharmacy and Technology*, vol. 8, no. 2, pp. 12037–12045, 2016.
- [59] G. García-Bautista, F. Trujillo-Romero and S. O. Caballero-Morales, "Mexican Sign Language Recognition Using Kinect And Data Time Warping Algorithm," *In International Conference on Electronics, Communications and Computers*, 2017, pp. 1-5, doi: 10.1109/CONIELECOMP.2017.7891832.
- [60] M. Raptis, D. Kirovski, and H. Hoppe, "Real-time Classification of Dance Gestures From Skeleton Animation," *In 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 2011 pp. 147–156, doi:10.1145/2019406.2019426.
- [61] S. E. Bautista M.Á. , Antonio Hern´andez-Vela, Victor Ponce, Xavier Perez-Sala, Xavier Bar´o, Oriol Pujo, Cecilio Angulo and, "Probability-Based Dynamic Time Warping for Gesture Recognition on RGB-D Data," *Lect. Notes Comput. Sci. vol 7854. Springer, Berlin, Heidelb.*, no. Advances in Depth Image Analysis and Applications, p. pp 126-135, 2013.
- [62] A. L. A. J. Alan dos Santos Soares, "Real-time 3D Gesture Recognition using Dynamic Time Warping and Simplification Methods," *Journal of WSCG*, vol. Vol.25, 20, 2017.
- [63] J. Geigel and M. Schweppe, "Motion Capture For Realtime Control Of Virtual Actors In Live, Distributed, Theatrical Performances," *Face and Gesture*, pp. 774–779, 2011, doi: 10.1109/FG.2011.5771347.
- [64] X. C. and Y. S. Y. Gong, M. Hu, "Research on Gesture Based on Genetic Algorithms - Support Vector Machine," *In International Conference on Intelligent Informatics and Biomedical, Shanghai, China*, 2019, pp. 385–390.
- [65] D. O. P. and S. C. I. B. Wijayasinghe, M. N. Saadatzi, S. Peetha, "Adaptive Interface for Robot Teleoperation Using a Genetic Algorithm," *In IEEE 14th International Conference on Automation Science and Engineering, Munich*, 2018, pp. 50–56.
- [66] F. Ullah, "American Sign Language Recognition System for Hearing Impaired People Using Cartesian Genetic Programming," *In 5th International Conference on Automation, Robotics and Application, Wellington, New Zealand, 2011*, pp. 96–99.
- [67] K. M. and S. V. S. T. R. Babu, "Optimal Skin Detection for Face Localization Using Genetic Algorithms," *In 11th International Conference On Hybrid Intelligent Systems, Melacca*, 2011, pp. 668–673.
- [68] A. El-Sawah, C. Joslin, N. D. Georganas and E. M. Petriu, "A Framework for 3D Hand Tracking and Gesture Recognition using Elements of Genetic Programming," *In Fourth Canadian Conference on Computer and Robot Vision*, 2008, pp. 495–502.
- [69] A. D. Calin, "Gesture Recognition on Kinect Time Series Data Using Dynamic Time Warping and Hidden Markov Models," *In 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, 2016, pp. 264–271.
- [70] Y. Li et al., "A Dynamic Hand Gesture Recognition Model Based on the Improved Dynamic Time Warping Algorithm," *In 25th IEEE International Conference on Automation and Computing, Lancaster, United Kingdom*, 2019, pp. 1–6.
- [71] C. Cheng, Y. Liu, J. Yang, T. Zhu and F. Ye, "Real-Time Dynamic Gesture Recognition based on Boundary-Constraint Dynamic Time Warping," *In IEEE National Aerospace & Electronics Conference, Dayton, OH, USA*, 2019, pp. 545–551.
- [72] S. Kajan, J. Goga and O. Zsíros, "Comparison of Algorithms for Dynamic Hand Gesture Recognition," *International IEEE Conference Cybernetics & Informatics, Velke Karlovice, Czech*

- Republic*, 2020, pp. 1–5.
- [73] W. Jutharee and T. Maneewarn, “Gesture Reconfiguration from Joint Failure Using Genetic Algorithm,” *In 16th International Conference on Control, Automation and Systems, Gyeongju*, 2016, pp. 1137–1142.
- [74] T. T. and K. I. T. Hashimoto, “Myogenic Potential Pattern Discernment Method Using Genetic Programming for Hand Gesture,” *In Joint 7th International Conference on Soft Computing and Intelligent Systems and 15th International Symposium on Advanced Intelligent Systems, Kitakyushu*, 2014, pp. 643–648, doi: 10.1109/SCIS-ISIS.2014.7044713.
- [75] L. Liu and L. Shao, “Synthesis of Spatio-Temporal Descriptors for Dynamic Hand Gesture Recognition Using Genetic Programming,” *In 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2013, pp. 1–7, doi: 10.1109/FG.2013. 6553765.