Applying Matching Coefficients in the 2D Stick-Model to Classify Human Motion: An Experimental Study in Applied Mathematics

Shatha Abdul-Hussein Kadhum

1Department of Mathematics, college of Sciences University of Basrah, Basrah, Iraq

Abstract—Recently, different techniques have been employed toward motion estimation. Some of these approaches include image based, model based, and silhouette based estimations. Despite their promising nature and outcomes, it remains notable that the techniques rely on motion data quality before producing optimal classification with precision and accuracy. Also, most of the existing algorithms have been complex relative to motion estimation, making interpretation challenging. Therefore, this study strived to respond to these dilemmas by modeling simple human motions through which various patterns of activity behavior could be recognized and aid in classification analyses. Three body components were used to develop the framework. These components included the lower body (LB), the upper body (UB), and the backbone (BB). Indeed, it was through these parts that a simple 2D human stick figure was formed. It is also notable that upon completion, the motion estimation mathematical model was compared to the performance of real motion phases to determine its efficiency in classification. The classifiers to which the model’s performance was compared included Rules and Tress, Misc, Meta, Function, Lazy, and Bayes classifiers. From the results, it was established that the 2D stick-model matching estimation was feasible and could be used to play a crucial rule in analyzing human motion classification.

1. Introduction

In the computer vision field, increasing attention has been extended to the attribute of vision-based human motion estimation [1]. Some of the specific areas that have benefitted from this trend include motion surveillance and analysis, human computer interaction, and animation [2, 3]. However, situations involving complicated human motions imply that motion estimation is a challenging process, especially when postures vary relative to different time series [4, 5]. Despite these mixed outcomes, it remains notable that in computer vision study, human motion estimation plays a crucial role. From the previous literature, the concept of human motion estimation constitutes automated estimations and predictions of human postures relative to body segment location, joint angles, and rigid body motion [6]. Some of the previous scholarly studies avow that a central determining factor that shapes outcomes linked to motion estimation involves motion data quality [7-9]. Other studies assert that the factor that plays a moderating role in determining process outcomes involves the algorithm that is established to model and estimate the intended framework [4, 8-10]. Imperative to note is that the quality of motion affects motion estimation performance, but it is also influenced by the method used to capture the data [6], whether by marker-less or marker-based motion capture. This study strived to present a novel framework through which human motion could be estimated. The common motion resemblance that was employed involved a 2D stick model construction and simple three-body-segment components. Some of the short temporal daily activities to which the presented model was applied included jumping, running, and walking, having gained access to these activities by using a publicly available video.

2. Methodology

In the model, multiple joints connecting simple rigid objects were used to express the human framework. The selected segments that were used to present the human body included LB, UB, and
BB. The role of the BB was to link the pelvis, neck, and head. On the other hand, the UB linked the right wrist, right elbow, right shoulder, neck, left shoulder, left elbow, and left wrist. Lastly, the LB linked the right toe, the right ankle, the right knee, the pelvis, the left knee, the left ankle, and the left toe. The figure below illustrates the study’s simple 2D stick model.

Figure 1: A simple 2D stick model highlighted

Figure 2: An illustration of motion images
In the study, one of the assumptions was that for each body joint, the motion assumed a coplanar state in such a way that it would lie on the x-y plane. Another assumption was that it was at the 2D system’s origin that the human pelvis was fixed.

3. Results and Discussion

Notably, six temporal motions were used to test the 2D motion estimation model. These motions included a camel pose, a leg lock pose, a child pose, jumping, running, and walking. Notably, the datasets were gained from experimental captures and publicly available video data. Regarding the latter data, motion activities that were investigated relative to the designed framework included a single subject performing jumping, running, and walking activities. However, backgrounds varied. In relation to the dataset obtained from experimental captures, activities that were used to evaluate the performance of the designed model included a professional Yoga master’s engagement in a camel pose, leg lock pose, and child pose; with all the three translating into Yoga postures. For the first phase, 30 instances were used and they constituted five jumping instance, 16 running instances, and nine walking instances. The second phase relied on 35 camel pose instances, 53 leg lock pose instances, and 47 child pose instances.
When visual inspection was conducted, findings demonstrated that the proposed human estimation framework performed similarly as the real motion pattern. Hence, only relatively small deviations were noted regarding the activities of running and walking motions. However, the models failed relative to jumping motion estimation for the case involving public available video data. Hence, the model would estimate jumping pose sequences but there was a significant deviation compared to the results that were obtained when it was used to estimate the running and walking motions. A central trend that could explain the model’s failure was the case of jumping activity’s rapid movements that translated into occluding motion poses.

For the case investigation that relied on the experimental dataset, which was obtained from the performance of a professional Yoga trainer, there was close similarity between the performance of the actual model and the estimated stick models. Regarding the results about visual estimation, the model proved superior regarding motion estimation on Yoga and rhythmic (running and walking) motions, rather than the instance of jumping (discrete motion). As such, it was established that the estimated model’s performance was reasonable whereby it could estimate rhythmic motion accurately, proving superior to an approach such as silhouette-base estimation. However, a notable aspect is that the model was established from pure numerical interpretation. As such, it was likely to lose perceptual validity in relation to the estimation of human motion. From the previous literature, perceptual accuracy is likely to be maintained if human motion is synthesized during motion.

The estimated model was also compared to the actual model in terms of the matching rate. The matching rate was obtained using the formula:

\[
\text{% matching rate} = \left( \frac{\text{Number of “Y”}}{\text{Number of “Y”} + \text{Number of “N”}} \right) \times 100\%
\]

From the results, the matching accuracy for the walking motion was higher than the running motion. A factor that explained this variation involved the pre-processing approach that was applied to the raw data. At t4 and t11, the running motion’s LB segment exhibited anomalies. The secondary effect of these anomalies was that the actual data deviated from the estimated data’s range. On the other hand, there was 42.35% and 57.67% matching rate for the jumping motion relative to the y-coordinate and x-coordinate respectively. Combined, the model’s overall matching rate for the selected parameter stood at 50.00%. For the case of the dataset obtained from the professional Yoga trainer, the study established relatively small matching rate as there was smaller tolerance range. Hence, the actual data failed to fall within the estimated tolerance’s small range. Due to these mixed outcomes, the study proceeded to compare the classification accuracy of the estimated model with that of the actual model. From the results, the estimated and actual data classification accuracy lay between 53.33% and 100.00%.

4. Conclusion

In summary, this study developed a human motion estimation framework. The study relied on 2D movement data. The three main body segments that the 2D stick figure incorporated included: LB, UB, and BB. From the visual inspection, it was found that the framework can estimate rhythmic motion patterns with accuracy and precision. In future, there is a need to incorporate superior cyber-shooting feature video cameras to counter errors arising from rapid motions, upon which high-speed motion might be captured correctly. Overall, an implementation of the latter recommendation is poised to steer improvements in the estimated model’s accuracy.

5. References

[1]. Y. Wang, G. Baciu, Human motion estimation from monocular image sequence based on cross-entropy regularization, Pattern Recognit. Lett. 24 (2003) 315–325
[2]. B. Rosenhahn, R. Klette, G. Sommer, Silhouette based human motion estimation, in: C. Rasmussen, H. Bülthoff, B. Schölkopf, M. Giese (Eds.), Pattern Recognition, Springer, Berlin Heidelberg, 2004, pp. 294–301

[3]. B. Rosenhahn, U.G. Kersting, A.W. Smith, J.K. Gurney, T. Brox, R. Klette, A system for marker-less human motion estimation, in: W. Kropatsch, R. Sablatnig, A. Hanbury (Eds.), Pattern Recognition, Springer, Berlin Heidelberg, 2005, pp. 230–237

[4]. M. Tong, Y. Liu, T.S. Huang, 3D human model and joint parameter estimation from monocular image, Pattern Recognit. Lett. 28 (2007) 797–805

[5]. Z. Xiao, H. Nait-Charif, J. Zhang, Automatic estimation of skeletal motion from optical motion capture data, in: A. Egges, A. Kamphuis, M. Overmars (Eds.), Motion in Games, Springer, Berlin Heidelberg, 2008, pp. 144–153

[6]. M. Milanova, L. Bocchi, Video-based human motion estimation system, in: V. Duffy (Ed.), Digital Human Modeling, Springer, Berlin Heidelberg, 2009, pp. 132–139

[7]. Doost, S. N., Zhong, L., Su, B., and Morsi, Y. S. (2017). Two dimensional intraventricular flow pattern visualization using the image-based computational fluid dynamics. Comput. Methods Biomechanics Biomed. Engin. 20, 492–507

[8]. Douglas, P. S., Pontone, G., Hlatky, M. A., Patel, M. R., Norgaard, B. L., Byrne, R. A., et al. (2015). Clinical outcomes of fractional flow reserve by computed tomographic angiography-guided diagnostic strategies vs. usual care in patients with suspected coronary artery disease: the prospective longitudinal trial of FFRct: outcome and resource impacts stud. Eur. Heart J. 36, 3359–3367

[9]. Galassi, F., Alkhali, M., Lee, R., Martindale, P., Kharbanda, R. K., Channon, K. M., et al. (2018). 3D reconstruction of coronary arteries from 2D angiographic projections using nonuniform rational basis splines (NURBS) for accurate modelling of coronary stenoses. PLoS ONE 13:e0190650

[10]. Imanparast, A., Fatouraee, N., and Sharif, F. (2016). The impact of valve simplifications on left ventricular hemodynamics in a three dimensional simulation based on in vivo MRI data. J. Biomech. 49, 1482–1489