

Recognition of Dance Posture Using Motion Captured Images

Shubhangi D¹ & Farheen Kouser*²

¹Department of Computer Science & Engg, Visvesvaraya Technological University, India.

²Department of Computer Science & Engg, Visvesvaraya Technological University, India.

¹E-mail: drshubhangipatil1972@gmail.com ²E-mail: farhiik@gmail.com

Abstract: Country / western dance, also known as country and western dance, contains many traditional country-western music dance forms and styles that are stylistically aligned with American country and western traditions. Many of these dances were "tried and true" steps of dance that had been "aside" for many years and became popular under the names of "country-western", "cowboy" or "country". National dance in Texas is known as kicker dance. In this paper, the techniques introduced in recognizing dance types is based on motion-captured human skeleton data. In particular, the goal is to identify poses which are characteristic for each dance step performed, based on information on body joints. Here we propose to recognize the dance pose using Spatio Temporal Interest Points (STIP) and classify the same using a KNN classifier.

Keywords: Spatio Temporal Interest Points (STIP), KNN

1. Introduction

This paper provides a comprehensive survey on identification of dance poses through captured data. In last few years the automation of human activity recognition has drawn much focus in the field of video analysis technology due to the increase in the demand of many kinds of applications which includes entertainment environments, surveillance environments, and healthcare. In an entertainment environment, the recognition of activity can improve the human computer interaction (HCI), of different dancer's actions that are automatically identified during a dance. Break dancing was developed in the city of New York. In Bronx, in early 1970's, inspired from martial arts, moves were developed by street gangs. The moves were complex and athletic, eventually for self-defence that characterize modern break dancing. Break dancing is a genre of street dance combining tangled body movements, design, synchronization and aesthetics. b-boys or b-girls are those who perform this type of breakdance they are usually called as breakers. Breakdance is one of the oldest recognized form of dance in hip-hop. Musical inspirations are maestro James Brown's energetic shows. In the early days, the instrumental portion of a song frequently played by the DJ was introduced into songs. Sound is an essential ingredient in breakdancing, whereas hip-hop dance songs make an ideal soundtrack. Rap is not the only option for dancing; 70's soul, funk, and even jazz tunes all work as well. Style, fashion, concept, and technique are also crucial aspects of breakdancing.

Popular Breakdance Moves we discuss here are:

- Arm wave
- Backspin
- Floor-rock
- Up-rock
- Handglide
- Spider
- Moonwalk

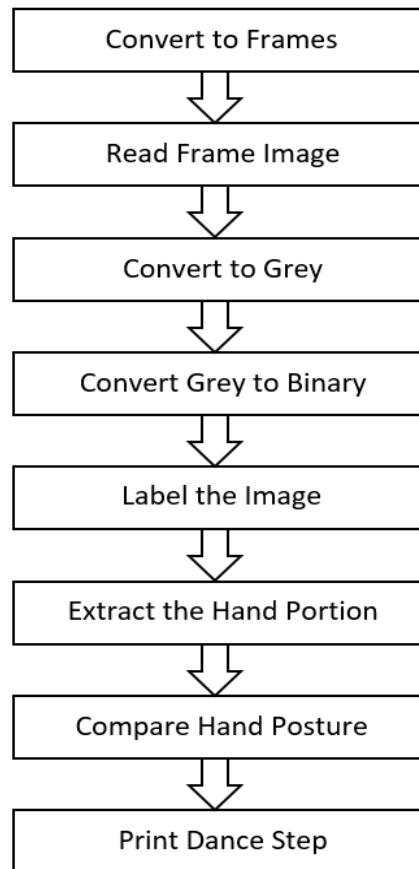


Figure 1. Proposed System

2. Related Work

We present a brief review of approaches proposed for the more general problem of human pose estimation in computer vision, one could note that many techniques are based on the detection of body parts for example, through pictorial structures explained by **Felzenszwalb, P.F.Huttenlocher [1]**. The advent of deep learning has brought forward two main groups of methods: holistic and part-based ones, which differ in the way the input images are processed by **Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep Learning on Computer Vision: A brief review on Comput. Intell. Neurosci [2]**. The holistic processing methods do not create a separate model for every part. Deep Pose is a holistic model that handles pose determination as a joint regression problem without formulating a graphical model. A drawback of holistic-based methods is that they are often inaccurate in the high-precision region due to the difficulty in learning direct regression of complicated posture vectors based on images. Part-based processing methods focus on detecting the human body parts individually, followed by a graphic model to incorporate the spatial information explained by **Toshev, A.; Szegedy, C [3]** Deep Pose: Human Pose Estimation via Deep Neural Networks proceedings on Computer Vision and Pattern Recognition. The authors, rather than training the network mistreatment the full image, use the native half patches and background patches to coach a convolutional neural network (CNN), to find out conditional possibilities of the half presence and spatial relationships. **Chen, X.; Yuille, A. [4]** Articulated create Estimation by a Graphical Model with Image Dependent Pairwise Relations. A multiresolution CNN is intended to hold out body-part-specific heat-map chance regression, that is within the

sequel succeeded by Associate in Nursing implicit graphic model for assuring joint consistency. As regards the additional specific field of dance create and move analysis, there is a comparatively restricted range of works. **Tompson, J.; Jain, A.; LeCun, Y.; Bregler, C.** [5] Joint coaching of a Convolutional Network and a Graphical Model for Human create Estimation. A gesture organization is portrayed for skeletal wireframe motion certain as shooting gestures among many dozen, in period and with high accuracy. **Raptis, M.; Kirovski, D.; Hoppe, H.** [6] amount Classification of Dance Gestures from Skeleton Animation. The authors, rather than training the network victimization the total image, use the native half patches and background patches to coach a convolutional neural network (CNN), to be told conditional possibilities of the half presence and abstraction relationships. **Tompson, J.; LeCun, Y.; Bregler, C.** [7] Joint coaching of a Convolutional Network and a Graphical Model for Human cause Estimation. A multiresolution CNN is intended to hold out body-part-specific heat-map chance regression, that is within the sequel succeeded by associate implicit graphic model for assuring joint consistency. As regards a lot of specific field of dance cause and move analysis, there is a comparatively restricted variety of works by the period of classification of Dance Gestures from Skeleton Animation. A gesture arrangement is delineated for skeletal wireframe motion certainly gestures, among many dozen, in period and with high accuracy. **Zanfir, M.; Leordeanu, M.; Sminchisescu, C.** [7] The Moving Pose: associate economical 3D mechanics Descriptor for Low-Latency Action Recognition and Detection. by **Ball, A.; Rye, D.; Ramos, F.; Velonaki, M.** [8] unsupervised agglomeration of individuals from ‘Skeleton’ knowledge. In an easy non-parametric moving cause framework is projected, for low-latency human activity and activity recognition. A technique to acknowledge individual persons from their walking gait victimization 3D skeletal knowledge from a MS Kinect device victimization the k means formula is delineated in whereas a key posture identification technique is projected. A skeletal illustration of the dancer is once more obtained by victimization knowledge from multiple depth sensors victimization this info, the dance sequence is partitioned off, first, into periods and, later, into patterns.

3. Proposed System

A three-step approach is adopted for the evaluation of dance pattern over Western dances:

- (i) Motion capturing.
- (ii) Data pre-processing and feature extraction.

Motion capturing is performed using camera/video. The still images provide as an output the position and the rotation of specific body joints at a constant frame rate. The available information is processed to form features which will be used as inputs to the dance recognition mechanism. Given a frame, or sequence of frames, during the performance of a dancer, our goal is to correctly identify which dance step is performed.

4. Methodology

All algorithms were enforced in MATLAB. In our case the knn parameterization method considers the amount of k nearest points, that was set to $k = 5$, the worth provides a decent trade-off. If $k = 3$ or less, we tend to area-unit unable to completely differentiate {to tell apart} among different dances that share a similar step. On the opposite hand, $k = 7$ or bigger leads to matching with similar steps of alternative dances.

The ensemble ways used sixteen ensemble members. the remainder of the parameters were used at the default values. The k-nearest neighbour's (KNN) algorithm is a straightforward, supervised machine learning algorithm that are often accustomed to solving each classification and regression issues.

Steps:

1. Load the information.
2. Initialize K to your chosen variety of neighbours.
3. For every example within the information.
 - 3.1 Calculate the gap between the question example and therefore, the current example from the information.
 - 3.2 Add the gap and therefore the index of the instance to an ordered assortment.
4. Type the ordered assortment of distances and indices from smallest to largest (in ascending order) by the distances.
5. Choose the primary K entries from the sorted assortment.
6. Get the labels of the chosen K entries.
7. If regression, return the mean of the K labels.
8. If classification, return the mode of the K labels.

5. Experimental Results

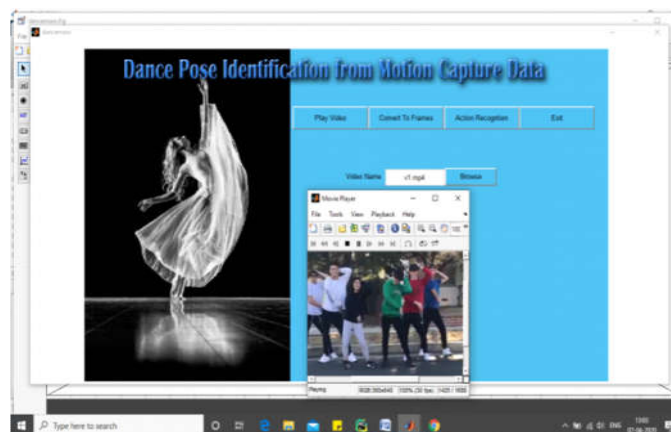


Figure 2. Read and Play Video

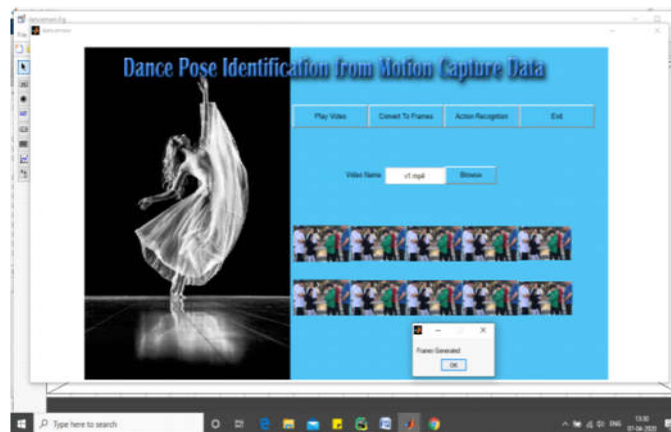


Fig.3. Convert video to frames



Figure 4. Click to get results

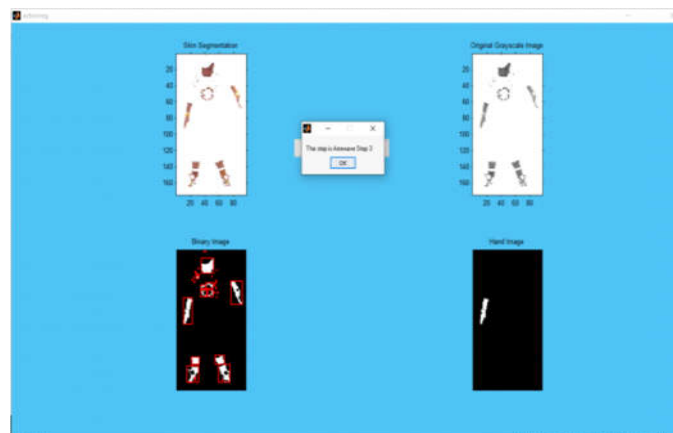


Figure 5. Displays the step (Arm wave step 3)

6. Conclusion

In this Paper, Motion capturing is performed using camera. The still images provide as an output the position and the rotation of specific body joints at a constant frame rate. The available information is processed to form features which will be used as inputs to the dance recognition mechanism. Given a frame, or sequence of frames, during the performance of a dancer, our goal is to correctly identify which dance step is performed.

References

- [1] Felzenszwalb, P.F.; Huttenlocher, D.P. Pictorial Structures for Object Recognition. *Int. J. Comput. Vis.* 2005, 61, 55–79.
- [2] Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep Learning for Computer Vision: A Brief Review. *Comput. Intell. Neurosci.* 2018, 7068349.
- [3] Toshev, A.; Szegedy, C. DeepPose: Human Pose Estimation via Deep Neural Networks. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 23–28 June 2014; pp. 1653–1660.
- [4] Chen, X.; Yuille, A. Articulated Pose Estimation by a Graphical Model with Image Dependent Pairwise Relations. In *Proceedings of the 27th International Conference on*

- Neural Information Processing Systems, Cambridge, MA, USA, 8–13 December 2014; Volume 1, pp. 1736–1744.12.
- [5] Tompson, J.; Jain, A.; LeCun, Y.; Bregler, C. Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation. arXiv, 2014.
- [6] Raptis, M.; Kirovski, D.; Hoppe, H. Real-time Classification of Dance Gestures from Skeleton Animation. In Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, New York, NY, USA, 26–28 September 2011; pp. 147–156.
- [7] Zanfir, M.; Leordeanu, M.; Sminchisescu, C. The Moving Pose: An Efficient 3D Kinematics Descriptor for Low-Latency Action Recognition and Detection. In Proceedings of the IEEE International Conference on Computer Vision, Los Angeles, CA, USA, 1–8 December 2013; pp. 2752–2759.
- [8] Ball, A.; Rye, D.; Ramos, F.; Velonaki, M. Unsupervised Clustering of People from ‘Skeleton’ Data. In Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction