ISSN 2278-3083
Volume 9, No.5, September - October 2020
International Journal of Science and Applied Information Technology
Available Online at http://www.warse.org/ijsait/static/pdf/file/ijsait01952020.pdf
https://doi.org/10.30534/ijsait/2020/01952020

Database of an Indian Classical Dance Videos with Different Actions

Bhavana.R[1], Dr.Suvarna Nandyal[2]
[1]Visvesvaraya Technological University, India, sg.bhavana@gmail.com
[2]P.D.A College of Enginggring,India,suvarna_nandyal@yahoo.com

ABSTRACT

In this paper, we describe the Indian classical Dance Video Database (ICD Database), a shared database containing original dance videos with popular music genre. It consists of 100 dance videos with different actions. While much effort has been dedicated to the gathering of huge video datasets of Indian classical dance videos of Kathak and Kuchipudi containing different gestures of face, hands and legs. The video consists of only single female dancer with costume. Present day activity identification databases includes ten peculiar activity classes accumulated under good conditions. The work on these datasets is now near ceiling and thus there's a desire for the look and creation of latest computing techniques. To handle this issue we accumulated the most popular activity dance videos taken from a YouTube. We apply this database to present the dance of two machine vision systems for gesture identification and investigate the toughness of functions under different variations like camera movement, point of view, quality of video and occlusion.

Key words : Video dataset .Gesture analysis, Activity classes, Dance types.

.

1. INTRODUCTION

Now a days more than billion videos are accessible at the internet and at every 24 hours of video uploaded to youtube , it is right away needed to sturdy methods that may help to extract this large quantity of statistics. A huge amount of hard work has been done to the gathering of real dance video databases present day action reputation datasets lag a long way in the back of the maximum popular datasets, such as KTH [18], weizmann [1] or the ixmas dataset [23],includes 6-11 movements for every dance . A videos in these database includes a single staged dancer and not using a occlusion and really restricted area as they're also confined in terms of brightness and digital camera and function difference, those data are not good at illustration of actual critical movement video. Video is a rich medium with dynamic information that can be used to determine, what is happening in a scene. In this work, we consider highly dynamic video, video that requires the parameterization of motion over extended sequences to identify the activity being performed. The main challenge with highly dynamic video is that a single frame cannot provide sufficient information to understand the action being performed. Therefore, multiple frames, leading to an extended sequence of frames, need to be analyzed for scene classification.

One of the disadvantages of current gesture classification in research is both a lack of perspective that can be applied to extended/big sequences and datasets inadequate in such highly dynamic videos. Our aim is to establish which methods best represent motion as opposed to methods that use a single (properly picked) frame [6] to identify the activity, as we feel such approaches devalue the necessity for video data. In this work we introduce a 1,000 video dataset and evaluate methods that focuses on highly dynamic videos requiring motion analysis for classification. We choose the domain of dance videos as (a) there is large amount of dance videos available online and (b) the diversity of dynamics in these videos provides us with a challenging space of problems for highly dynamic video analysis. This enables us to conduct a focused study on the relevance of motion in dancing classification and the broader value of motion in improving video classification. The core challenge of this task is attaining an adequate representation of human motion across a 10-second clip. In order to highlight the trajectory of this work, we will evaluate the current approaches and demonstrate the value of isolating motion for properly evaluating these approaches and this dataset.

2. RELATED WORK

The present database is illustrated in Table 1. The Hollywood [9] and UCF50 database made a huge amount of hard work to develop more real activity identification datasets by choosing videos captured from original cinemas and YouTube. These databases are very promising because of huge differences in camera movement, dancer occurrence also differences in the location, point of view of the dancer, as well as complicated background. The UCF50 database contains the 10 activity classes from the UCF YouTube dataset. The UCF Sports dataset [14], and the Olympic Sports dataset [12], includes sports videos from youtube. Using leave-one-clip-out method, the identification accuracy is 98%. Using constant joint position is enough for identifying activities rather than joint kinematics. Based on proposed HMDB51 they done one simple experiment where they took 10 activity classes same as UCF50 and done manual labeling for the 14 joint position over 1,100 motion clips. The results of classification of features represented from the joint positions at one frames as resulted only 35% and is very low than the 54% resulted by movement features They also implemented the classification results of the 10 pose classes of the UCF50 by applying movement features and resulted at 66%. The HMDB51 is an activity dataset whose poses classes have different movements so it has a valid contribution for the examination of activity identification machine . The HMDB51 includes 51 different poses classes taken from internet.

Table 1: Existing Dataset

Dataset	Year	Actions	Clips
KTH	2004	6	100
Weizmann	2005	9	9
IXMAS	2006	11	33
Hollywood	2008	8	30-129
UCF Sports	2009	9	14-35
Hollywood2	2009	12	61-278
UCF YouTube	2009	11	100
Olympic	2010	16	50
UCF50	2010	50	min. 100
HMDB51	2011	51	min. 101
ICD Dataset	2015	100	200

3. CONTRIBUTIONS

The proposed ICD dataset contains two types of Indian classical dance videos such as kathak and kuchipudi. These videos are downloaded form youtube of ICD dataset. Each video is of 60 minutes and each class of videos contains more than 8000 frames. The proposed ICD dataset is used to examine the performance of two activity identification machines. These dance includes different gestures of face hands and legs. Depending upon the poses of the face, hands and legs we are going to identify the type of the dance.

3.1How the Dataset was Built

In this section we describe the data collection process and how videos were obtained from YouTube. Here we are creating an Indian classical dance database (ICD database) videos of Kathak and Kuchipudi. From the youtube 100 Indian classical dance videos of kathak and kuchipudi are downloaded. The videos consists of single female dancer and the dancer is wearing a constume according to the dance because the constumes are different for kathak and kuchipudi.Each video is of duration of 60 seconds.Each dance consists of different gestures. Depending upon the gesture we are going to identify the type of class.

Figure1: kathak dance images with different gestures from ICD dataset

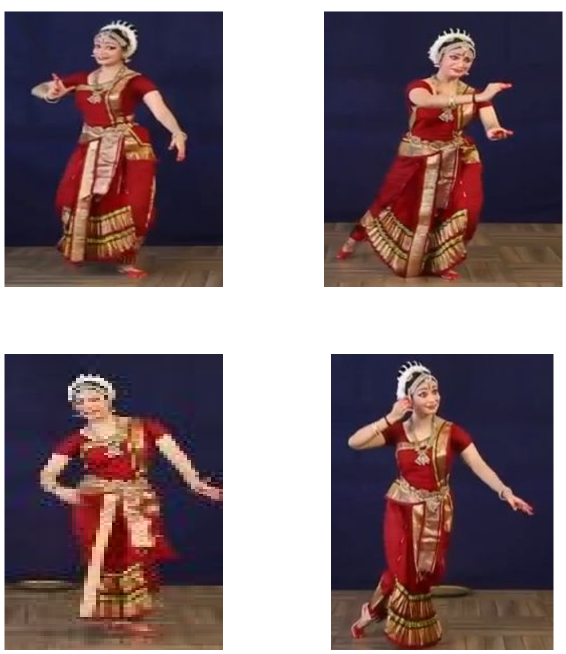

Figure 2: Kuchipudi dance images with different gestures

Table 2: Description of videos

Videos	No of frames	Bits per pixel	Frame Rate	Height	Width	Duration	Format
Kathak	14077	24	25	360	640	563.1	RGB 24
Kuchipudi	8885	24	25	360	640	355.4	RGB 24

3.2 Condition of Experiment

As a dataset for our experiments, we used 100 advanced choreography dance videos shot from the front in the ICD Database. These choreographies are different for each other. In our research we are focusing on two Indian classical dance videos. 'Kathak and Kuchipudi', these dances includes the full body movements and have complicated leg and hand movements.

4. RESULTS AND DISCUSSION

Our ICD dataset includes two dances Kathak and Kuchipudi. We have composed 4 dance videos from 2 dancers for 4 different music with different gestures. The online dance videos are also downloaded from youtube. However our database includes less dance videos , the differences in , costume, complicated background etc. makes a challenging task. We are going to make our dataset available soon. We have a plan to make the dataset bigger for each classes of dances.

REFERENCES

[1] M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri. **Actions as space-time shapes**. ICCV, 2005.

[2] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei. Imagenet:**A large-scale hierarchical image database. CVPR, 2009.**

[3] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn,and A. Zisserman. **The PASCAL Visual Object Classes Challenge 2010 (VOC2010) results**.

http://www.pascalnetwork.org/challenges/voc/voc2010/workshop/index.html.

[4] L. Fei-Fei, R. Fergus, and P. Perona. **Learning generative visual models from few training examples: an incremental bayesian approach tested on 101 object categorie**s. CVPR Workshop on Generative-Model Based Vision, 2004.

[5] H. Jhuang, E. Garrote, J. Mutch, X. Yu, V. Khilnani, T. Poggio, A. D.Steele, and T. Serre**. Automated home-cage behavioral phenotyping of mice**. Nature Communications, 1(5):1–9, 2010.

[6] H. Jhuang, T. Serre, L. Wolf, and T. Poggio. **A biologically inspired system for action recognition**. ICCV, 2007.

[7] G. Johansson, S. Bergstr¨om, and W. Epstein. **Perceiving events and objects. Lawrence Erlbaum Associates**, 1994.

[8] I. Laptev. **On space-time interest points.** Int. J. of Comput. Vision,64(2-3):107–123, 2005.

[9] I. Laptev, M. Marszałek, C. Schmid, and B. Rozenfeld. **Learning realistic human actions from movies**. CVPR, 2008.

[10] J. Liu, J. Luo, and M. Shah. **Recognizing realistic actions from videos** ”in the wild”. CVPR, 2009.

[11] M. Marszałek, I. Laptev, and C. Schmid. **Actions in context**. CVPR,2009.

[12] J. Niebles, C. Chen, and L. Fei-Fei. **Modeling temporal structure of decomposable motion segments for activity classification**. ECCV,2010.

[13] A. Oliva and A. Torralba. **Modeling the shape of the scene: A holistic representation of the spatial envelope.** Int. J. Comput. Vision,42:145–175, 2001.

[14] M. Rodriguez, J. Ahmed, and M. Shah. Action mach: **A spatio temporal maximum average correlation height filter for action recognition.**CVPR, 2008.

[15] B. Russell, A. Torralba, K. Murphy, and W. Freeman. Labelme: **a database and web-based tool for image annotation**. Int. J. Comput.Vision, 77(1):157–173, 2008.

[16] J. M. S. Maji, L. Bourdev. **Action recognition from a distributed representation of pose and appearance.** CVPR, 2011.

[17] K. Schindler and L. V. Gool. Action snippets: **How many frames does human action recognition require**. CVPR, 2008.

[18] C. Schuldt, I. Laptev, and B. Caputo. **Recognizing human actions: A local SVM approach**. ICPR, 2004.

[19] T. Serre, L.Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio. **Robust object recognition with cortex-like mechanisms**. IEEE Trans. Pattern Anal. Mach. Intell., 29(3):411–26, 2007.

[20] M. Thirkettle, C. Benton, and N. Scott-Samuel. **Contributions of form, motion and task to biological motion perception**. Journal of Vision, 9(3):1–11, 2009.

[21] A. Torralba, R. Fergus, and W. Freeman. 80 million tiny images: **A large data set for nonparametric object and scene recognition.** IEEE Trans. Pattern Anal. Mach. Intell., 11(30):1958–1970, 2008.

[22] H. Wang, M. Ullah, A. Kl¨aser, I. Laptev, and C. Schmid. **Evaluation of local spatio-temporal features for action recognition**. BMVC,2009.

[23] D. Weinland, E. Boyer, and R. Ronfard. **Action recognition from arbitrary views using 3D exemplars**. ICCV, 2007.