Content Based Classification of Ethiopian Traditional Dance Videos Using Optical Flow and Histogram of Oriented Gradient

Dr. KarpagaSelvi Subramanian
Assistant Professor Ethiopian Institute of Technology-Mekelle , Computer Engineering, Mekelle University

AndargieMekonnen
PG Student , Computer Engineering, Mekelle University

Abstract - Human action recognition is a task of recognizing action performed in videos or in time ordered sequence of image frames, assigning an label to each sequence of images indicating what action has taken place. This technique is employed in content based classification of traditional Ethiopian dances by identifying gestures and movements specific to traditional dances. In this study, two optical flow estimation methods are investigated to detect motions. Histogram of Oriented Gradient (HOG) is used to detect the region of interest for detecting and locating human in sequence of time ordered images. The feature vectors extracted by two optical flow estimation methods with and without HOG are considered for classification using Support vector machine and results are compared.

Keywords: Ethiopian traditional dance, Histogram Oriented Gradient, Human action recognition, Optical flow estimation, Support Vector Machine, Video Classification Algorithm.

I. INTRODUCTION

Music has been being played, listened, and heard in every corner of the world every day. It is noted as an essential part of human life that almost all people in the world listen to music sometime in their life. For a country like Ethiopia, traditional music is a representative of country's culture serving two major roles. The traditional folk songs of various regions hold hallowed meanings and exhibit diversified customs of country and express the national solidarity. Traditional music helps young generation of the country to understand more about the country and proud of their tradition. It also motivates for further searching about cultural aspects. All traditional music in Ethiopia is associated with unique traditional dance.

Ethiopia is a widely diverse country with over eighty unique, rich ethnic, cultural, custom and linguistic groups [1]. Each ethnic group have its own traditional rhythm and dance style to transmit certain ideas, religious beliefs, historical events, ancient stories, emotions, thoughts. The younger generation of the country is more often interested in viewing video clips that contain traditional dances in entertainment websites like YouTube. It is essential to have a content based video classification method to automate indexing of traditional Ethiopian dance videos.

II. RELATED WORK

Different researches have been done on interpretation of human actions and activities in computer vision community. Especially during recent past, researchers are very much interested in human action recognition for variety of reasons. There has been a huge amount of work done to solve the problem of automatic action recognition problem for several purposes. During the study phase of this research, we concentrate on various human action recognition systems that can be useful to extract features of Ethiopian traditional dance videos.

2.1. Motion based action recognition

MD. Atiqur et al. [2] proposed an optical-flow based action recognition algorithm to recognize and cluster natural actions performed by different individuals by calculating motion magnitude and direction of movement in each video frame. The classification accuracy of MD Atiqur et al. [2] system is demonstrated by applying Weizmann database , consisting of 90 low-resolution (180x144) video sequences showing nine different people, each perform ten natural actions. The authors recommend that the problem of complex motion is not addressed in this research.

Julia Moehrmann et al. [3] investigate motion based situation recognition in group meeting to identify the actions of speaking, writing and listening. Standard face detection system is employed to identify the number of participants and skin color model for recognizing individual activities. They extract motion trajectories, perimeter, velocity and acceleration to investigate suitable features. Position data is used to extract head and hand features and individual activity is recognized using Hidden Markov Model (HMM).

2.2. Form based action recognition system

There have been a large number of researches carried out in human action recognition using shape and form features [7].Carlsson et al. viewed action recognition as a shape matching problem and demonstrated that specific actions in long video sequences can be recognized by matching the shape information extracted from individual frame to stored prototypes representing key frames of the action. The system proposed by Carlson et al. uses set of edge points and tangent lines associated with those edge points are used to represent the shape information in image. The order structure of this set is characterized by considering all combinations of the points and associated tangent lines. For each combination topological type is computed and represented by number. The effectiveness of system is demonstrated by applying it to two long video sequences of tennis players. The results reported are promising with 100% recognition without any false positive been achieved in two of the cases examined. The authors recommended the problem not addressed in their research is the automatic selection of the person specific key-frame.

Yao et al. [8] proposed human action recognition in still images by modeling human poses. They use a novel approach called 2.5D graph representation for actions in images and an exemplar based approach for action classification. During implementation they considered 15 key-points of human bodies: top head, left-middle-right shoulders and hips, left-right elbows, wrists, knees and ankles. Each key-point is characterized by view-independent 3D positions and local 2D appearance features. The key-point locations are then normalized such that the center of the torso is at (0, 0, 0), and the height of the torso (distance between middle shoulder and middle hip) is 100 pixels. Exemplar-based approaches which allow multiple exemplars to represent an action class are used for enabling more flexibility in overcoming the challenge of large within-action pose variations. To evaluate their systems Yao et al. carried out experiments on two publicly available datasets: the people playing musical instrument (PPMI) dataset and the PASCAL VOC 2011 action classification dataset. It was observed that their classification accuracy is 65.8% on first dataset and 72.4% of the second data set.

2.3. Hybrid based action recognition system

Schindler et al. [9] proposed a method for human action recognition from video, which exploits both shape and motion features. Inspired by models of the human visual system [19], the two feature sets are processed independently in separate channels. The form channel extracts a dense local shape representation from every frame, while the motion channel extracts dense optical flow from the frame and its immediate predecessor. The same processing pipeline is applied in both channels. Feature maps are pooled locally, down-sampled, and compared to a collection of learnt templates, yielding a vector of similarity scores. In a final step, the two score vectors are merged and recognition is performed with a discriminative classifier. They tested their system using two data sets; KTH and Weizmann and they reported a performance of 92.7 % for KTH and 100% for Weizmann data set.

III. DANCE VIDEO CLASSIFICATION

Fitsum [10] investigates a form based feature for Ethiopian traditional video classification by extracting the shape features in sequential video frames. Each frame is preprocessed by applying: HOG (Histogram Oriented gradient) to detect actors pose and then applied Log Gabor filtering to represent shapes. Max-pooling is performed to preserve important information and discard irrelevant details. They classification accuracy of Fitsum system is evaluated using reduced shape feature of each Ethiopian Traditional dance video with SVM (support vector machine) in four scenarios using: one , five, ten and fifteen sequential frames with which, SVM classifies with76.2%,76.8%,76.8% and 77% overall classification accuracy respectively. The author recommends that the problem of video dance classification is open because accuracy of the system is not satisfactory.

3.1. Problem definition

After an intense literature review on human action recognition, it is evident that Ethiopian traditional dance video clips are not investigated for development of automated classification and indexing which aids retrieval of video based on its visual content [8].If the traditional video clips available on online, are not properly tagged and categorized based on its visual content, the benefit would be greatly reduced. Therefore, there is a high demand for traditional video clip classification based on the traditional dance of the culture.

The objectives of this study are listed as follows. (1) To analyze dance movements specific to most predominant ten different classes of Ethiopian traditional dances listed as Afar, Benshangul, Gambella, Eskesta, Gurage, Hararghe, Oromo, Somali, Tigrinya and Wolaita. (2) To address methods to extract the features related to traditional dances movements in video clips. (3) To classify video clips using SVM and (4) to check the accuracy of classification. We are setting our goal to identify feature extraction method that is generating robust feature for better classification using Support Vector Machine (SVM).

There are number of steps to be followed to do classification on Ethiopian traditional dance videos. Following steps have been employed to achieve the above stated objectives. A diagrammatical representation of the steps to be followed is shown in Fig1.

3.2. Data set preparation

The first step to start with this research is to prepare a new dataset, since no data set for similar study was available. Several dance video clips are gathered from entertainment websites. Professional dance teachers for traditional dances of Ethiopia are consulted to analysis and identify the videos that have traditional unique dance movements for each classes of interest. Thus one hundred identified dance video clips are cut into 2 second length video clips that are used in research.

3.3. *Preprocessing*

3.3.1.Video segmentation and frame extraction

Splitting the longer videos into smaller segments is necessary to get the clips with single dancer dancing only traditional dance movements. Also for motion feature extraction by optical flow requires sequential frames of same intensity. Each selected clip is having 30 sequential frames extracted at a rate of 15 frames per second and thus a total of 3000 sequential frames is prepared and used as dataset for the research.

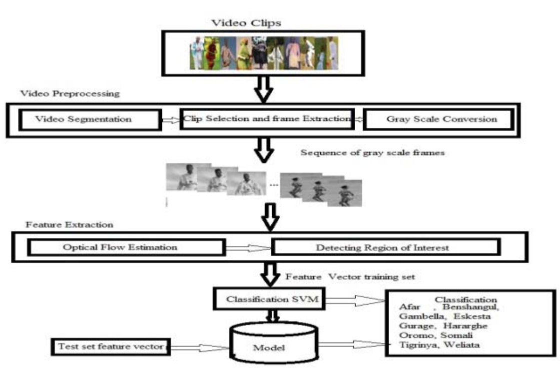

Fig 1. Block diagram of Ethiopian traditional dance video classification system.

3.3.2. Gray scale conversion

The conversion of frames from 24 bit RGB format to 8 bit gray scale format reduces computational complexity greatly while extracting features using optical flow or HOG techniques. On the other hand, gray scale frames are enough for recognizing movements in successive frames.

3.4. Feature extraction

Typical systems for action recognition proposed a higher level of interpretation based on different low-level representations of videos. This helps to design algorithms that neglect irrelevant information like the background, color of the clothes and variation in illumination and are robust to variations in scale, viewpoint and subtle variations in actions. Features refer to a location of sudden change in image. Features are important because of its high information content, it reduces the computational burden. For the purpose of action recognition in videos, feature extraction is a major key process. Motion features have excessive power to represent actions of human or movements of object from a video. Here in this study, an attempt to extract motion features using optical flow algorithms from time ordered successive sequential video frames and detecting interested point of an object by using optical flow and HOG.

3.4.1. Optical flow estimation

Two dimensional image motions is the projection of three dimensional motions of the objects relative to visual sensor and image plane. Time ordered sequences of image frames are used to estimate two dimensional image motions either as instantaneous image velocities or discrete image displacements. These are usually called the optical flow field or image velocity field which is a reliable approximation for two dimensional image motion. Although different techniques are available for calculating optical flow, gradient based techniques provide more accurate results. In gradient based motion estimation, let fi(x) denotes the frame at time i. Assume that an object is translated by distance d between time i and i+1.

$$F_{i+1}(x) = F_i(x-d) \qquad (1)$$

by applying Taylors expansion to shifted signal $F_i(x-d)$ the first order approximation to the displacement is given by (2).

$$\hat{d} = \frac{F_i(x) - F_{i+1}(x)}{F_i'(x)} \qquad (2)$$

For linear signals first order approximation is exact whereas for nonlinear signals the accuracy of the approximation depends on displacement magnitude and the higher order signal structure. The famous gradient based estimation of optical flow is Horn- Schunck and Lucas –Kanade method is outlined in the following section.

3.4.1.1. Horn - Schunck Optical flow algorithm

Horn and Schunck viewed the problem of optical flow estimation as a variational problem, where the desired vector field h is defined as the minimizer of certain energy functional J(h) (3).

$$J(h) = \int (I_x u + I_y v + I_t)^2 + \alpha^2 \left(|\nabla u|^2 + |\nabla v|^2\right) \qquad (3)$$

This functional takes twoterms i. data attachment term, provided by the optical flow constraint, and ii.a regularity term which is based on the gradient of the flow. α is a parameter to control the weight of the smoothness term compared to the optical flow constraint. The use of quadratic functional in energy model assumes the flow derivatives are expected to follow Gaussian distribution. The minimization of the equation 3 yields the following Euler- Lagrange equations (4,5).

$$I_x^2 u + I_x I_y v = \alpha^2 (\bar{u} - u) - I_x I_t \qquad (4)$$

$$I_x I_y u + I_y^2 v = \alpha^2 (\bar{v} - v) - I_x I_t \qquad (5)$$

Solving the equations for (u,v) velocity vector for moving objects yields a system of linear equations. (4)

3.4.1.2. Lucas –Kanade Optical flow Algorithm

The widely used differential method for optical flow estimation in computer vision is Lucas–Kanade method. The basic optical flow equations are solved for all the pixels in the neighborhood (6) by least square criterion. It is local method and assumes that the flow is essentially constant in a local neighborhood of the pixel under consideration.[5]. Thus it is assumed for all pixels within a window centered at pixel p satisfy the basic optical flow equation with velocity vector (u,v). In this method, the displacement of the image contents in two nearby frames in time domain are small and mostly constant with in the neighborhood of the center pixel p.

$$I_x(p_j)u + I_y(p_j)v = -I_t(p_j) \qquad (6)$$

Where j = 1,2,3, …n and $p_1, p_2, p_3, \dots p_n$, are pixels in the window centered at p, $I_x(p_j), I_y(p_j), I_t(p_j)$ are partial derivatives of the image signal I with respect to x,y,t respectively. The matrix form of the above equation 6 is

$$AV=b. \qquad (7)$$

$$A = \begin{pmatrix} I_x p_1 & I_y p_1 \\ I_x p_2 & I_y p_2 \\ \vdots & \vdots \\ I_x p_n & I_y p_n \end{pmatrix}, V = \begin{pmatrix} u \\ v \end{pmatrix} \; b = -\begin{pmatrix} I_t p_1 \\ I_t p_2 \\ \vdots \\ I_t n \end{pmatrix}$$

Least squares principle is used to determine the solution of the equations. Equation 7 can be rearranged as ATAV = ATb and velocity vector V is given by (8).

$$V = (A^T A)^{-1} A^T b \qquad (8)$$

The Lucas Kanade algorithm returns a high pixel difference score or a state of rapid change for the motion of objects in 2d scene. Thus helpful in extracting the motion features in video.

3.4.2. Region of Interest Detection

Recognition of traditional dance movements from time ordered sequential frames are possible by optical flow estimation but not robust enough to represent right region of interest due to noise and occlusion. As a result, detecting and extracting of the interesting regions from the optical flow allows to get most accurate and robust feature from it. There are many studies in finding the region of interest from frames, Dalal and Triggs propose a very successful edge and gradient based descriptor called Histogram of Oriented Gradient (HOG) for detecting and locating humans in still images [6]. In our study we attempt to extract features only in the region of interest in order to minimize computational cost needed during training of classifier and to increase the accuracy of the dance movement recognition system. HOG is the method which is based on evaluating well-normalized local histograms of image gradient orientations in a dense grid. In this study, HOG is used as flow detector.

3.5. Support Vector Machine Classifier

The final step in Ethiopian traditional dance video classification system is to classify traditional dances based on the motion features extracted using optical flow and HOG. Support Vector Machine (SVM) is chosen as classifier in our study, because training can be done in supervised mode. Once trained on features containing some particular movements, the SVM classifier can make decisions regarding the class of the traditional dance having similar movements in additional test images.

IV. EXPERIMENTAL SETUP

Since no previous dataset on similar study is available, we select 100 Ethiopian traditional dance video clips and split them into clips of 2 second lengths for the study purpose and manually labeled them with traditional dance name. The following criteria are strictly followed in selection of Traditional dance video clips (i) video clips with only single dancer is considered as a candidate for dataset. (ii) Video clips that have more descriptive sequential frames with unique movements belongs to type of traditional dances are considered. (iii) Selection of video clips are restricted with no background movement to increase the performance of the classifier. The selected videos are segmented and sequential frames are extracted at a rate of fifteen frames per second (fps). Consecutive frames extracted at a rate more than 15fps are similar and estimation of the global motion performed by the dancer becomes difficult. Optical flow extracted failed to estimate the local motions performed by dancer if segmentation of the videos are done at a rate less than 15fps.

Time ordered sequential frames thus obtained have to be preprocessed before used for extracting the motion features. The key steps in preprocessing are (i) Cropping the image on each frame so that each frame contains same person to background proportion. This step is essential because background act as reference in estimation of optical flow between consecutive frames. Figure 2 and 3 shows the difference between frames cropped with and without proportional background. (ii) Resizing the cropped frames to80 X 80 pixels frame in order to reduce the computational cost. (iii) Converting the images from color to gray scale. For motion detection gray scale images are more than sufficient and greatly reduce the computational complexity.

Fig 2. Frames cropped with proportional background

Fig 3.Frames cropped with non-proportional background

Optimal and essential feature extraction from the frames is the challenging task for classification. In the study, optical flow on consecutive video frames is used to detect the motion of the objects from background, and consists of horizontal and vertical velocity, magnitude squared and orientation. Magnitude feature is considered for classification using Support vector machine. Fig 4 shows the optical flow magnitude square calculated using Lucas – Kanade method. Experiments are conducted by calculation optical flow using two methods (i) Locus – Kanade algorithm and (ii) Horn – Schunck Algorithm. The extracted optical flow of the training set is used to train the SVM for classification and tested with testing set.

Fig 4 Optical flow squared magnitude – Lucas- Kanade Method.

Optical flow gives features related to motion and it is not considering any features related to performer or dancer. Extracting the features relevant to region of interest would improve the classification accuracy. HOG method proposed by Dalal and Triggs is used to extract features related to the region of interest from the opticalflow.The HOG feature estimate the most interesting point of optical flow of a traditional dance video by encoding the local shape feature to make the optical flow more robust for feature representation.In HOG feature extraction, selection of right cell size to get representation of descriptive shape information of optical flow is a thought-provoking task. The figure 5 indicates that a cell size of [8 8] does not encode much shape information, on the other hand cell size of [2 2] encodes more on shape information but increases the dimensionality of the HOG feature vector significantly. A good compromise of cell size[4 4] is chosen that encodes enough spatial information to visually identify an optical flow shape as well as limiting the number of dimensions in the HOG feature vector, which reduces the training time significantly. After training the SVM with HOG of optical flow, classification is done on test set and results are obtained as confusion matrix.

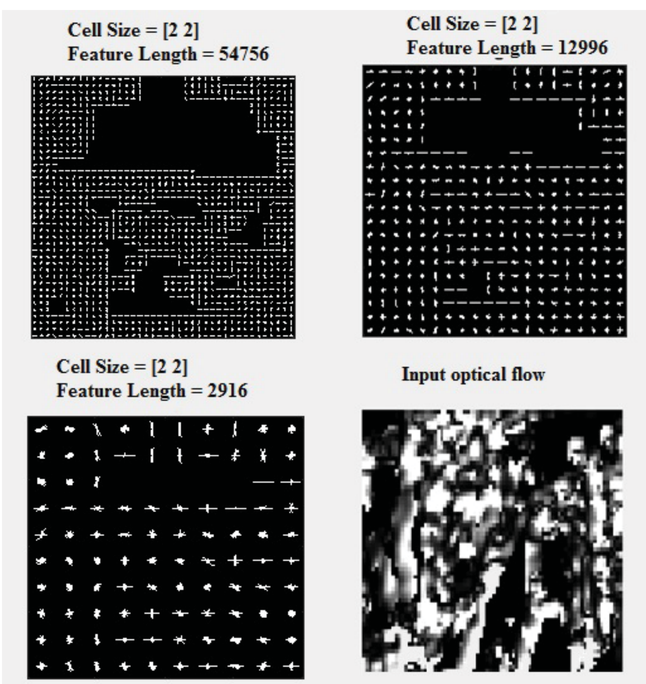

Fig 5.HOG feature extracted for different cell size for the given optical flow.

V. RESULT AND DISCUSSIONS

In this study, training was done is in four phases based on feature extracting algorithm. In all the phases 2700 frames are used during training and300 frames are used for testing. In each phase a SVM classifier model was created and trained by extracted features and later used for classification on test set. The results are represented in confusion matrix as shown in fig 6 -9 for each phase. The Diagonal of the matrix shows the percentage of correct classification highlighted with black color and misclassifications are highlighted with blue color. The overall classification accuracy of each phase is shown in table1.The study under research reveals that HOG features extracted on estimated magnitude squared of optical flow of both Lucas-Kanade and Horn-Schunck methods are improving the classification accuracy in support vector machine.

5.1. Phase 1: Classification based on Optical flow estimation by Lucas-Kanade Method

	Afar	Benshangul	Gambella	Eskesta	Gurage	Hararghe	Oromo	Somali	Tigrinya	Wolaita
Afar	100	60.0	0.0	0.0	7.0	0.0	3.0	7.0	0.0	0.0
Benshangul	0.0	27.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Gambella	0.0	0.0	100	0.0	10.0	0.0	7.0	7.0	0.0	0.0
Eskesta	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0
Gurage	0.0	0.0	0.0	0.0	17.0	0.0	10.0	0.0	0.0	0.0
Hararghe	0.0	0.0	0.0	0.0	0.0	100	3.0	0.0	0.0	0.0
Oromo	0.0	0.0	0.0	0.0	27.0	0.0	57.0	3.0	0.0	0.0
Somali	0.0	13.0	0.0	0.0	3.0	0.0	0.0	80.0	0.0	0.0
Tigrinya	0.0	0.0	0.0	0.0	37.0	0.0	10.0	0.0	100	0.0
Wolaita	0.0	0.0	0.0	0.0	0.0	0.0	10.0	3.0	0.0	100

Fig 6. Classification by SVM based on optical flow features extracted by Lucas-Kanade method

5.2 Phase 2: Classification based on Optical flow estimation by Horn- Method

	Afar	Benshangul	Gambella	Eskesta	Gurage	Hararghe	Oromo	Somali	Tigrinya	Wolaita
Afar	0.0	3.0	0.0	0.0	3.0	0.0	3.0	0.0	0.0	3.0
Benshangul	0.0	73.0	0.0	0.0	3.0	0.0	0.0	0.0	3.0	0.0
Gambella	0.0	0.0	93.0	0.0	3.0	0.0	0.0	0.0	0.0	0.0
Eskesta	0.0	0.0	0.0	70.0	0.0	0.0	0.0	0.0	0.0	0.0
Gurage	0.0	0.0	0.0	0.0	43.0	0.0	0.0	0.0	0.0	0.0
Hararghe	0.0	0.0	0.0	0.0	3.0	90.0	0.0	0.0	0.0	0.0
Oromo	0.0	0.0	0.0	0.0	0.0	0.0	97.0	0.0	0.0	0.0
Somali	100	23.0	7.0	30.0	43.0	10.0	0.0	100	13.0	97.0
Tigrinya	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	80.0	0.0
Wolaita	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0

Fig 7. Classification by SVM based on optical flow features extracted by Horn- Schunck method.

5.3. Phase 3: Classification based on HOG of Optical flow estimation by Lucas-Kanade Method

	Afar	Benshangul	Gambella	Eskesta	Gurage	Hararghe	Oromo	Somali	Tigrinya	Wolaita
Afar	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Benshangul	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Gambella	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Eskesta	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0
Gurage	0.0	0.0	0.0	0.0	97.0	0.0	0.0	0.0	0.0	0.0
Hararghe	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
Oromo	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0
Somali	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0
Tigrinya	0.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0	100	0.0
Wolaita	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100

Fig 8. Classification by SVM based on HOG of optical flow features extracted by Lucas-Kanade method

5.4. Phase 4: Classification based on HOG of Optical flow estimation by Horn-SchunckMethod

	Afar	Benshangul	Gambella	Eskesta	Gurage	Hararghe	Oromo	Somali	Tigrinya	Wolaita
Afar	100	0.0	0.0	0.0	13.0	0.0	0.0	0.0	0.0	0.0
Benshangul	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Gambella	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Eskesta	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0
Gurage	0.0	0.0	0.0	0.0	80.0	0.0	0.0	0.0	0.0	0.0
Hararghe	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
Oromo	0.0	0.0	0.0	0.0	7.0	0.0	100	0.0	0.0	0.0
Somali	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0
Tigrinya	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0
Wolaita	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100

Fig 9. Classification by SVM based on HOG of optical flow features extracted by Horn- Schunck method

Table 1. Classification Accuracy

Phase	Accuracy in Percentage
Classification based on Optical flow estimation by Lucas-Kanade Method	78.1
Classification based on Optical flow estimation by Horn-Schunck Method	64.6
Classification based on HOG of Optical flow estimation by Lucas-Kanade Method	99.7
Classification based on HOG of Optical flow estimation by Horn-Schunck Method	98

VI. CONCLUSION AND RECOMMENDATIONS

Classification was performed using a one-to-one SVM classification. Training and testing were done in four phases based on the type of feature extraction algorithm used. The performance of the system at the end is evaluated as a measure of classification accuracy using SVM classification algorithm. The overall accuracy of each classifier is, 78.1%, 64.6%, 99.7 and 98% using Lucas-Kanade, Horn-Schunck, HOG of Lucas-Kanade and HOG of Horn-Schunck methods respectively. This result implies that traditional dance classification using HOG of Lucas-Kanade method produces a high performance as compared to classification using other methods. The Lucas-Kanade method performs well than Horn-Schunck method, due most Ethiopian traditional dances focus on both on local and global motions and the capability of Lucas -Kanade method to estimate both motions types. Generally, the classification accuracy of both Lucas-Kanade and Horn-Schunck methods increase when HOG feature is implemented. The reason behind this is the ability of HOG to extract region of interest of optical flow. On the other hand, for some category, there is a high correlation between traditional dance styles. For instance, from the confusion matrix result traditional dance of Gurage highly similar with Hararghe and Tigrinya. In this study, in addition to well-known challenges of Action recognition from video, there are challenges that are unique to traditional dance video classification. Since we used the optical flow of a dancer for recognizing a particular traditional dance, factors such as traditional Clothes, traditional women's hair style and traditional objects often increased the percentage of misclassification

This work introduces our steps for classification of the traditional dance from a video sequence. This highly applicable topic is still in its infancy and much is left to be done. This work is extensible in many ways, such as

- Scale and direction features of the optical flow could also be used along with magnitude to increase the classification accuracy.
- During dataset preparation, time ordered sequential frames are used to calculate optical flow. Improvements could be achieved by calculating optical flows on video clips itself.
- Along with the generated confusion matrices, prototypes, user interface and database needs to be developed to enhance the utilization of the system.

REFERENCES

[1] TimkehetTeffera," The role of traditional music ampng east African societies: The case of selected Aerophones" in proceedings the 16th International meeting on folk musical instruments, Martin Luther university Halle-Wittenberg, Germany, 2006.

[2] Md. AtiqurRahmanAhad,"Computer vision and action recognition: A guide for image processing and computer vision community for action understanding", Atlantis of Paris, 2011

[3] Julia Moehrmann, Xin Wang and Gunther Heidemann,"Motion based situation recognition in group meetings" Proceedings SPIE 7538, Image Processing : Machine Vision Applications III, 75380N, January, 2010, DOI : 10.1117/12.838963, http://dx.doi.org/10.1117/12.838963

[4] EnricMeinhardt-Llopis, Javier Sanchez, Daniel Kondermann´ , Horn–Schunck Optical Flow with a Multi-Scale Strategy, Image Processing On Line, 3 (2013), pp. 151–172. http://dx.doi.org/10.5201/ipol.2013.20

[5] Dhara Patel,SaurabhUpadhyay,"Optical Flow Measurement using Lucas kanade Method"International Journal of Computer Applications (0975 – 8887) Volume 61– No.10, January 2013.

[6] Dalal, N.,Triggs, B. "Histograms of oriented gradients for human detection "Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference onpp886 - 893 vol. 1 June 2005, DOI 10.1109/CVPR.2005.177 ISSN 1063-6919

[7] Stefan Carlsson and Josephine Sullivan," Action recognition by shape matching to key frames",in Proceedings IEEE Computer Society Workshop on Models versus Exemplars in Computer Vision, 2001.

[8] B. Yao and L. Fei-Fei, "Action recognition with exemplar based 2.5 d graph matching," in Computer Vision–ECCV the series Lecture Notes in Computer Science, Vol., 7575 pp 173-186, 2012.

[9] K. Schindler and L. Van Gool, "Combining densely sampled form and motion for human action recognition," in Pattern Recognition,the series Lecture Notes in Computer Science Springer,Volume 5096, 2008, pp. 122- 131.

[10] FitsumTamene, "Ethiopian traditional dance video classification using visual feature", M.Sc. thesis, Gondar University,Ethiopia,2013