Smart Computing Review

HMM-based Scheme for Smart Instructor Activity Recognition in a Lecture Room Environment

Asim Raza[1], Muhammad Haroon Yousaf[1]*, Hassan Ahmed Sial[2], and Gulistan Raja[3]

[1]Department of Computer Engineering, University of Engineering and Technology / Taxila, Pakistan / asim.raza@uettaxila.edu.pk, haroon.yousaf@uettaxila.edu.pk

[2]Department of Computer Science and Engineering, University of Engineering and Technology / Lahore, Narowal Campus, Pakistan / hassaan90@gmail.com

[3]Department of Electrical Engineering, University of Engineering and Technology / Taxila, Pakistan / gulistan.raja@uettaxila.edu.pk

*Corresponding Author: Muhammad Haroon Yousaf

Received October 17, 2015; Revised November 23, 2015; Accepted December 22, 2015; Published December 31, 2015

Abstract: Instructor activity recognition can certainly play its part as an important parameter in evaluating and improving the performance of an instructor. This paper presents a single-layered sequential approach for instructor activity recognition in the lecture room environment. A hidden Markov model (HMM) scheme is selected as a sequential approach for activity recognition. The proposed system incorporates the five major activities of the instructor in the lecture room, i.e. walking, writing, pointing towards the board, standing, and pointing towards presentations. Background/foreground modelling is carried out using a Gaussian mixture model (GMM) for instructor detection in the lecture room. Mesh features are selected to represent the instructor. After vector quantization, features are passed to the HMM for activity recognition. Time is tracked, and the occurrences of each activity are counted to elaborate on the activities the instructor performed during the lecture. The proposed scheme proved to be efficient owing to its high accuracy rate of over 90 percent in recognizing five different activities of an instructor as tested in a MATLAB simulation environment.

Keywords: Activity Recognition, Instructor Detection, Hidden Markov Model, Computer Vision

DOI: 10.6029/smartcr.2015.06.008

This research is supported and funded by Directorate of ASR & TD, University of Engineering and Technology Taxila, Pakistan. Authors are also very grateful to the peers involved in the dataset collection.

Introduction

Recognizing human activities has been a key research area for the past few years owing to its wide range of applications in many fields, including human–computer interaction, sign language interpretation, automated surveillance systems and video lecture recording systems. Instructor activity recognition can play an important role in evaluating the performance of an instructor in a lecture room, and can also be integrated as an important part in current autonomous video recording systems. We know nonverbal communication, which can involve facial expressions, physical body movements, eye contact and physical gestures, plays an important role in conveying a message to the listener. One study on student perceptions of teacher nonverbal and verbal communication analyzed the best and the worst professors from six different cultures on the basis of recalled experiences by students of those six different cultures. The study showed that the best professors were perceived to employ more nonverbal expressiveness, relaxed movement, in-class conversation, and out-of-class communication than the worst professors [1]. To keep it simple, in our research work, we chose only five physical activities to be recognized by our proposed system, which can also assess how much an instructor is physically active in a classroom. With the help of the latest research available regarding nonverbal communication and its importance in delivering lectures, we can evolve our proposed system further to help in not only evaluating but also elevating the performance of an instructor in a classroom.

Recently, a number of approaches based on these specific applications have been presented. For surveillance applications, before recognition, pedestrians must first be tracked [2]. Kalman filters have been used extensively in many domains for tracking. In visual surveillance, this method appears frequently in the literature. The problem is simple: is it possible to design a system that will automatically recognize different activities performed by people (or a person) if you have image sequences of a person or people performing those activities? But the solution is not simple, because input is a sequence of images, and people are randomly performing different activities. The system has to decide whether an activity is performed or not. Where does an activity start, and where will it end? For a machine, it is simply a sequence of images without any start or end information. Most important is recognizing the activities with maximum accuracy and minimum error.

In the past few years, several researchers have addressed this problem and tried to provide an appropriate solution for it; the most prominent among them is discussed here. Aggarwal and Cai [3] divided the action recognition system into three sub-problems: extracting the human body from the image sequence, frame tracking and recognition of action. By using 2D and 3D models, Gavrila [4] concentrated primarily on tracking hands and humans in contributing to action recognition. Moeslund et al. [5] presented a survey that addressed the problems and approaches, but largely focused on capturing human motion, human model initialization, tracking, pose estimation and activity recognition. In this paper, the words *action* and *activity* are frequently used. In this discussion, both words refer to the state of motion completed by the sequence of images collectively.

In this particular research, a hidden Markov model (HMM) is used because it was previously used by different researchers for activity recognition. Let us have a look at a few of them. Hand gesture recognition was probably the first, published by Yamato et al. [15], who used HMM in the field of activity recognition, because previously, it was used for speech recognition. In their research work, vector quantized labels were used to recognize six classes of tennis stroke. The image sequence went through many preprocessing steps before applying the HMM, including noise reduction, background subtraction, human detection, and finally, applying the HMM for stroke recognition. A real-time HMM-based system was developed to understand sentence-level American Sign Language [6], [7] and a Viterbi algorithm was used with and without strong grammar. But once the recognition process started, only the sign language could be used, because the algorithm is unable to recognize undefined hand motions, which can lead to false results. Another HMM-based model used hand localization, hand tracking and gesture spotting as preprocessing for hand gesture recognition. Graphic editor control [8] was made, in which centroids of the moving hand region are coupled to yield hand trajectory.

Instructor activity recognition in a lecture room environment was accomplished through instructor tracking, which only focused on the localization of the instructor. From 2010 to 2012, instructor activity recognition in the lecture room was done using fuzzy classification on instructor morphologic features [9]. Kameda et al. [10] presented an automatic lecture recording system for distance learning purposes. The system, named CARMUL, captures handwriting, presentation slides, audio and visual data. Habib et al. presented a gesture recognition–based framework for an autonomous lecture recording system based on the classroom scene and instructor gesture analysis [11].

In our research work, the main aim is to recognize the five types of physical activities performed by the lecturer or instructor inside the classroom. The types of activities are pointing towards the board, pointing towards a presentation, writing on the board, standing, and walking. A Gaussian mixture model was used for background modeling and instructor detection, then low-level features were extracted by dividing each frame into small boxes. Then, by using K-means clustering, unique symbols were assigned to the sequence of frames. This quantized input was used in training and testing the HMM created for each of the five activities. A maximum estimation theorem was used for activity recognition. The rest of this paper is structured as follows. In Section 2, the theoretical background is discussed. The proposed method for gesture recognition is presented in the same section. The performance of the algorithm, evaluated using different experiments, is given in Section 3. And finally, we conclude this work in Section 4.

Proposed methodology

Before going towards the methodology of the proposed research work, the theoretical background of the methodology needs to be discussed.

■ Theoretical background

Activity recognition methodologies fall into two categories [3]: single-layered approaches and hierarchical approaches. In single-layered approaches, the recognition of human activities is based on the sequence of images or features, and is very much effective. On the other hand, hierarchical approaches deal with high-level human activities by dividing the activities into sub-events. Sequential approaches are the single-layered approaches. Another sequential approach, called the state model–based approach, represents human action by constructing a model. That model is trained to generate sequences of feature vectors corresponding to the specific activity by calculating the likelihood estimation (or posterior probability) that a given sequence is generated by each activity model. The hidden Markov model is a state model–based sequential approach that is used in implementing a proposed research work [12].

■ Hidden Markov model

A hidden Markov model (HMM) consists of a number of states, each of which is assigned a probability of transition from one state to another. With time, stochastic transition of states occurs. Like Markov models, the state at a current time depends upon the previous states. According to the probability assigned to the HMM states, one symbol is generated. States of the HMM are not directly observable, but are observed using a sequence of symbols produced [13]. Figure 1 briefly illustrates the concept of the transition graph. There are three states in this example, indicated as circles. Each direction line is a transition from one state to another, where the transition probability is indicated by the character alongside the line. There are transition paths from states to themselves, which can provide the HMM with time scale invariability because they allow the HMM to stay in the same state for any duration.

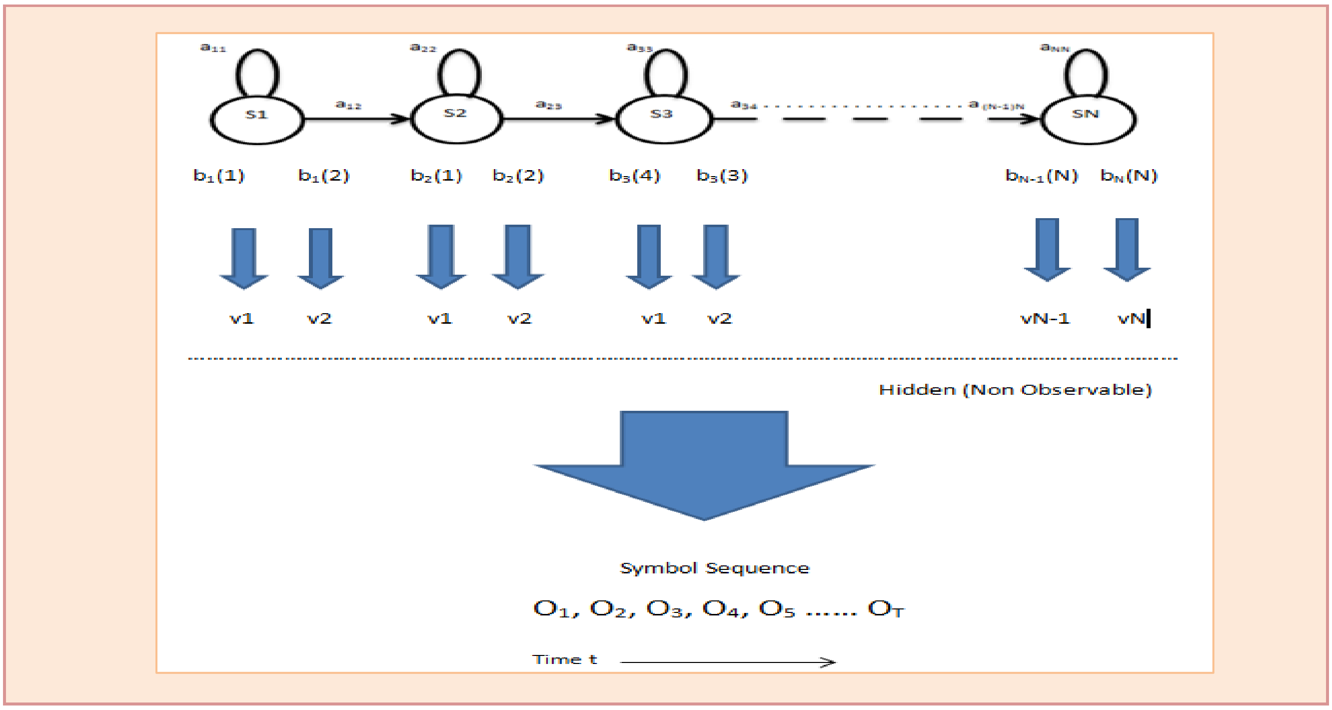


Figure 1. Hidden Markov model concept

Each HMM state outputs a symbol in any state, qj, and bj(k) is the probability of the symbol vk. If M kinds of symbols are there, bj(k) becomes an N × M matrix. The symbol sequence O = O1, O2, ..., OT for time 1 to T is observed from the HMM. Output symbol sequences of the HMM can be observed, but transition of a state sequence cannot be observed. An HMM can be characterized by three matrices: state transition probability matrix A, output symbol matrix B, and initial state probability matrix π. The parameters of these matrices are determined during the learning process, and one HMM is created for each activity to be classified, as discussed in the next sections [14].

■ General block diagram of the proposed recognition system

The proposed framework for the instructor activity recognition system in a lecture room, based on a sequential approach, is shown in Figure 2. A brief description of all the blocks is discussed in upcoming sections.

■ Background modeling and instructor detection

Different stages of background subtraction and instructor detection are shown in Figure 3. Before feature extraction, it is essential to perform some preprocessing steps. A Gaussian mixture model (GMM)-based foreground detection method is used for background subtraction, which aims to independently eliminate the instructor body from the background and then dynamically detect the body movement of the instructor. The width of the bounding box is a very useful dynamic parameter in classifying different kinds of physical activities, and it mostly depends on movement by the instructor. For removal of shadow, a series of morphologic operations is performed. And lastly, an auto white balancing algorithm is used to reduce the effect of sudden intensity changes by the camera. Finally, the desired region that exists within the bounding box is extracted.

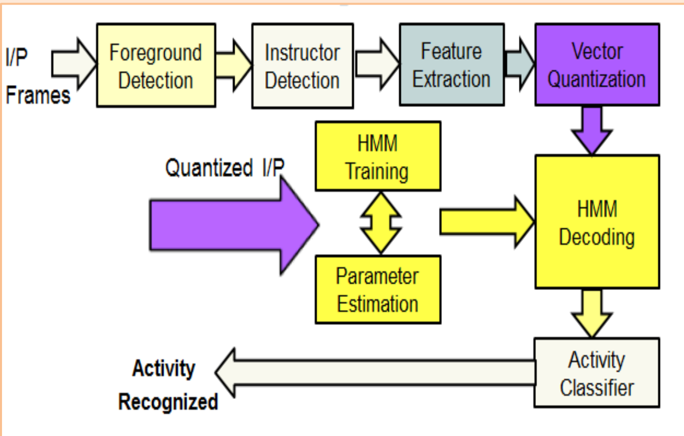


Figure 2. Block diagram of the instructor activity recognition system

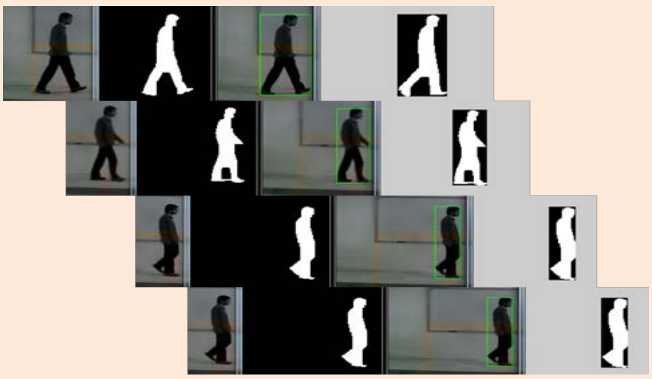


Figure 3. Different stages of the walking activity

Feature extraction

The features are extracted from variable-sized images, as extracted in the previous section, by dividing the image into small meshes. Because the image is binary, the number of 1s within each mesh is calculated and then divided by the total number of meshes formed in an image. According to Figure 4, if the total number of meshes is M × N, and the number of 1s in each mesh is n, which is 11, then an image can be transformed into a feature vector with length equal to the total number of meshes in an image. Feature vector $\mathbf{F = f_1, f_2, ... f_{MN}}$, where M × N is the number of meshes in an image [15].

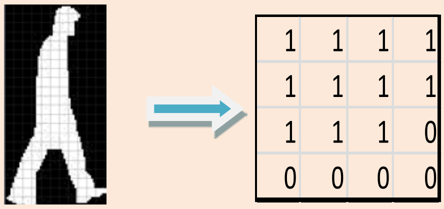


Figure 4. Mesh feature vector extraction

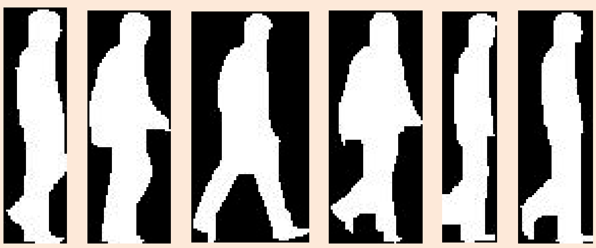
The size of an image is not always constant, but varies according to the physical activity performed by the instructor. To associate the size of a feature vector with the size of an image, a parameter is required that produces this variation in the feature vector calculation formula. For scaling, a method must be developed that will normalize these variations against one reference and that can be used to classify different activities. Maximum average width of the different activities performed is used as a reference, which scales this whole variation according to this reference. Because the maximum average width will produce the maximum possible number of meshes, by subtracting the total number of meshes, n, from the average maximum number of meshes, the difference of each frame in terms of meshes from a single reference can be calculated. By dividing the total number of 1s in a single mesh with the above calculated difference, one feature, fn, can be calculated, and this is the new feature vector, as given by equation 1.

$$\mathbf{fn} = \frac{\boldsymbol{no\ of\ 1's\ in\ a\ mesh(n)}}{\boldsymbol{Average\ Max\ no\ of\ Meshes - n}} \tag{1}$$

These low-level features against one frame are then quantized to produce the desired levels, where n is the total number of meshes in an image.

Vector quantization

K-means clustering is used to quantize a whole feature vector, f_n, as achieved in the previous section, into one distinct code word or symbol against one frame. For this, the feature vector is divided into k clusters to produce k levels (or in HMM terms, k symbols can be used). Every center of the cluster is assigned a unique symbol, and this symbol is assigned to the incoming unknown frame passing through the same procedure, which has the minimum distance from the center of the cluster. This means each feature vector f_i is transformed into symbol s_k if k = arg min $d(f_i, c_k)$, where c_k is the center of the cluster, and d(y,z) means distance between point y and z [15]. This method will produce a symbol against every frame after assigning a specific level to a frame. This is given as input to the HMM for both purposes, i.e. learning of the HMM and decoding of the correct class.



Symbol Seq.	1 1 1 2 2 3 3 4 5 5 5 6 7 8 8 9 9 10 11 12

Figure 5. Symbol assignment for the walking activity

Six different frames of the walking activity are shown in Figure 5. The symbols from 1 to 12 are assigned to each activity frame as the feature vector of each frame is divided into 12 clusters. The underlined symbols are assigned to six different frames of the walking activity.

■ HMM parameter estimation

For estimating the initial HMM parameters, i.e. transition matrix a_{ii} and emission matrix b_{ij}, the number of states and the maximum number of output symbols must be decided. One HMM classifier for activity recognition should have N states, and is 24 with output symbols at 12. Five hidden Markov models for each activity class are developed for both training and recognition purposes.

Every frame represents one of the 24 states of the HMM, and using this data, initial transition and emission matrices are calculated and further used for initial parameter estimation in the learning process.

■ HMM parameter learning

Using an estimated guess of a parameter, as described in the previous section, learned estimated parameters are produced that can be used further in re-estimating the parameters for the given set of data. This modified learned and trained transition, and emission matrices for each HMM-based activity class, is used in training each data class.

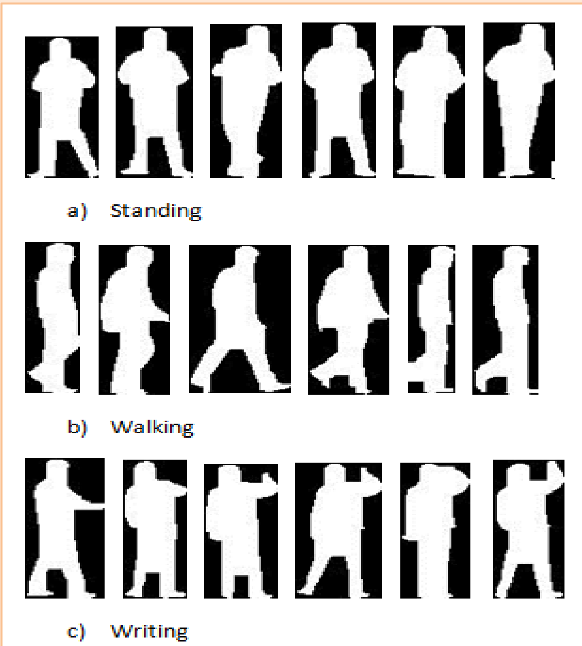


Figure 6a. Training set for standing, walking and writing

The frames of different activities are shown in figures 6a and 6b. The whole process trained the model for the sequential combination of symbols for the particular class. And using the trained matrices, i.e. transition and emission, the activity is further decoded.

■ HMM classifier and activity recognition

To decode the correct activity class, the trained transition probability and emission probability matrices are used. The sequence of symbols produced by the video is presented to all the HMM-based activity recognition classes, and by using the forward algorithm, a class with the maximum likelihood is chosen as the correct activity class. Table 1 shows the process of deciding the correct classifier on the basis of maximum likelihood.

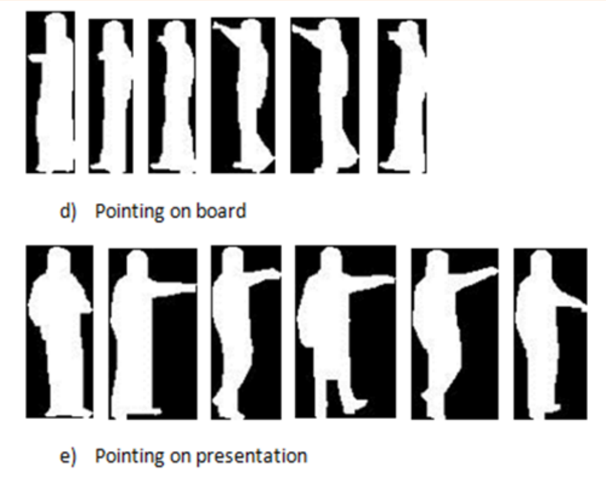


Figure 6b. Training set for pointing activities

Table 1. Activity classification based on maximum likelihood

HMM	Log-likelihood P(Seq)	Forward variable
Writing(correct class)	**-238.0008**	**0.1225**
Walking	-600.4801	0.0838
Point to board	-272.121	0.0906
Point to presentation	-292.2653	0.0859
Standing	-402.2404	0.0971

■ Proposed research work chart

Finally, the chart in Figure 7 describes the details of the research work corresponding to the goal, objectives, tasks, results and indicators of success.

Experimental results and discussions

The results drawn from different experiments are discussed in detail in this section, along with the comparison. MATLAB was used as the simulation tool to perform and evaluate these experiments.

A summary of the data set used to perform different experiments for activity recognition is given in Table 2. It includes the camera specifications as well as details of the data set used in training and testing the model.

Minimum time for an activity to be classified is 4 sec or 100 frames. Training is performed by three separate data sets, i.e. A, B and MIX. Data set A contains videos of five different activities performed by Instructor A. Data set B contains videos of five different activities performed by Instructor B. Data set MIX contains videos of five different activities performed by five different instructors.

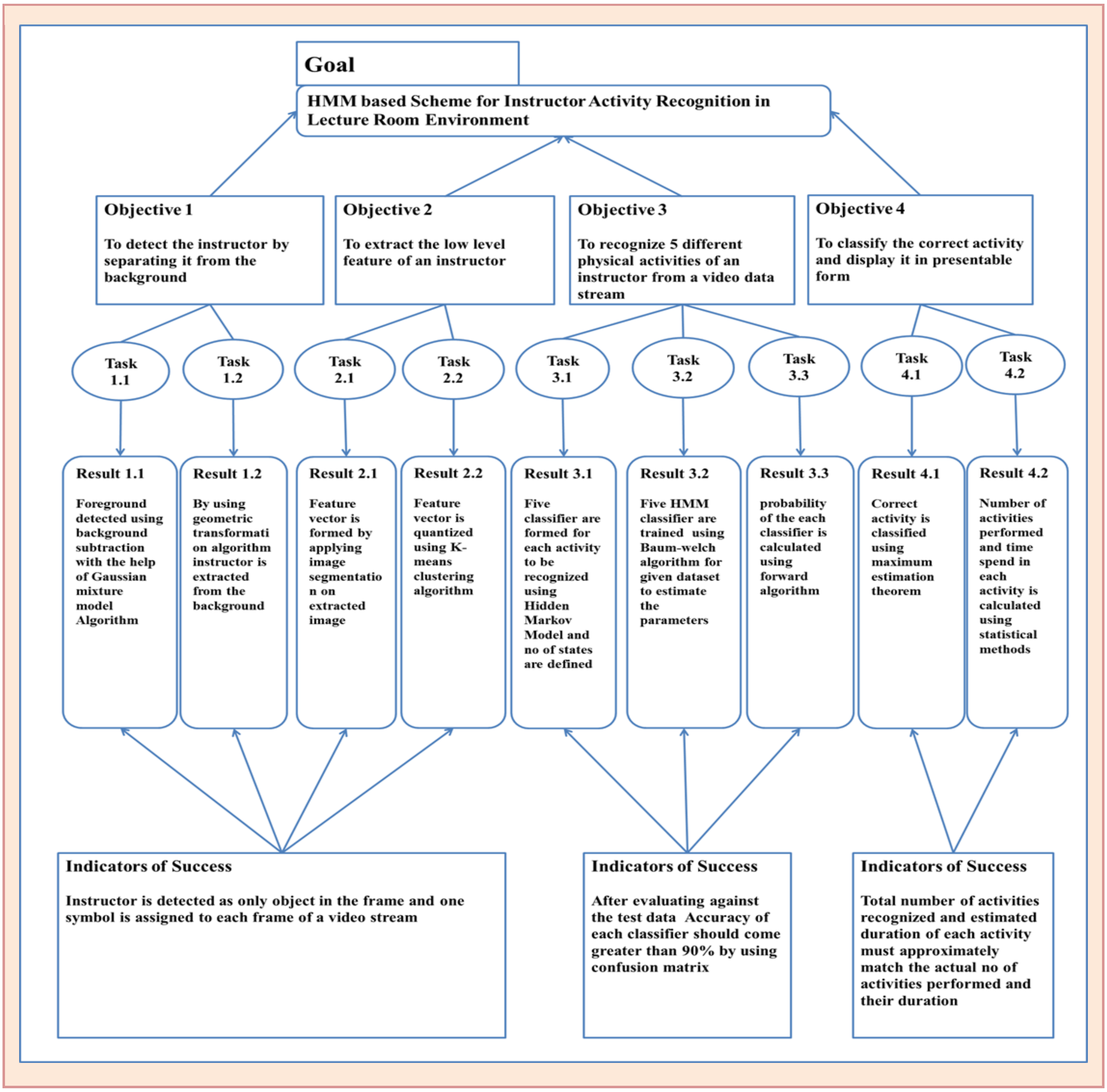


Figure 7. Proposed research work chart

Table 2. Data set details

	Frame Rate	Frame Width	Frame Length	No. of Videos Used in Training HMM Model
Subject A	25	320	240	22
Subject B	25	320	240	22
Subject MIX	25	320	240	28

Experiment 1

In experiment 1, three types of training data were used; that is, data set A and B were trained with the 22 video sequences of Instructors A and B, and data set MIX was trained with 28 video sequences of different instructors. For each activity class, only one sequence of a particular activity from all the instructors is used in training the model.

The percentage accuracy of recognition for the particular training set against the test data is tabulated in Table 3. Results show that test data set A is accurately recognized for all three training sets, whereas data set B shows accuracy in the 70s for other training data sets. For training data set A, it is normal, but for training data set MIX, it is not the best, but not bad either. The MIX data set was trained with 28 video sequences of different instructors, but only used one particular sequence of an activity from each instructor.

Table 3. Accuracy percentage of test vs. training data

Accuracy(%) Experiment 1			
	Training Data		
Test Data	**A**	**B**	**MIX**
A	100	100	100
B	71.42	100	72.85
MIX	70.25	72.85	100

Experiment 2

In experiment 2, for the MIX training data set, all three test data sets were tested. A complete video lecture sequence was presented to all classifiers, and one other class (named Reject) was introduced. If a particular sequence produced a probability that is less than the minimum possible probability of the particular class, then it was classified as a rejected sequence and placed in the Reject class. Table 4 shows the percentage accuracy results of the recognition.

Table 4. Accuracy percentage of test vs. training data

Accuracy(%) Experiment 2	
	Training Data
Test Data	**MIX**
A	99.41
B	89.51
MIX	99.02

As can be seen from the results, by introducing the Reject class, the recognition accuracy of the A and MIX data was reduced from 100 percent, but it was still above 99 percent, and recognition accuracy of class B improved to almost 90 percent. In recognition systems, the rate that a particular class is classified correctly and is not classified incorrectly, both matter. Therefore, by using true positive rates and false negative rates, the accuracy of the system is calculated. The confusion matrix for test data set A is shown in Table 5.

Experiment 3

For recognition of performance enhancement, different parameters can be used instead of a forward algorithm. For different types of training sets, the recognition performance of the same class of data produces different results. In general, for a classifier in cases of low-level feature extraction, different solutions are presented, and some of them improve feature extraction methods. By using the given parameters of the HMM for activity recognition, and deriving new parameters from them by using statistical methods, recognition performance for the generalized system can be improved. During our research work, in designing one general classifier for the given data set, an experiment is done using the derived parameters. If the training data set is small, then the proposed solution can produce very valuable results, because state model–based approaches normally require more data sets to produce the best results. The recognition results in the MIX_PS row of the derived parameters are comparatively good, compared with the forward algorithm results in the MIX row. This improvement was done for the experiment 1 results for Instructor B, and is tabulated in Table 6. The derived parameter results are an average sum of results produced by different parameters. It not only includes the correctly classified results but also the weighting of the next best possible results. If requirements of the system are more focused towards the activity performed, not on how many times it was actually performed, then this solution can be very useful.

Table 5. Confusion matrix for test data set A

	walk	write	Pt2brd	Pt2pres	stand	Reject
Walk	100	0	0	0	0	0
Write	0	97	0	0	0	3
Pt2brd	0	0	100	0	0	0
Pt2pres	0	0	0	100	0	0
stand	0	0	0	0	100	0

Table 6. Actual and derived parameter comparison

	walk	write	Pt2brd	Pt2pres	stand	Total
B_Actual	7	25	36	21	11	100
MIX	21	14	21	18	26	100
MIX_PS	18	39	36	14	18	125

The graph given below represents the enhanced performance after using derived parameters from the analysis. B_Actual is the correct result, MIX is the results of experiment 1 for Instructor B from training data set MIX, and MIX_PS is the results achieved using the weighted average sum of the derived parameters.

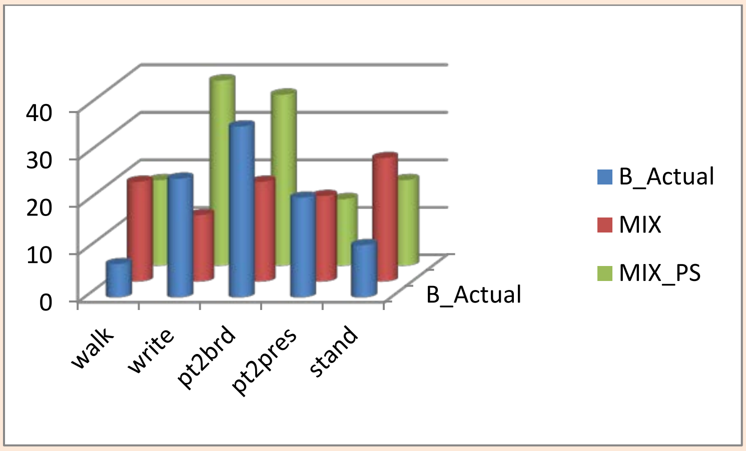


Figure 8. Actual and derived parameter comparison

Comparison

The history of research work done, and the proposed work and its comparison with the other previous works, is discussed in detail in this section. Instructor activity recognition in a lecture room environment was accomplished through instructor tracking, which only focused on the localization of the instructor. Later on, instructor activity recognition was done through template matching–based gesture recognition. The system recognized three activities of an instructor using nine different templates as training data. The system lacks accuracy due to fewer templates and the computational complexity of the algorithm. From 2010 to 2012, instructor activity recognition in the lecture room was done using fuzzy classification on instructor morphologic features [9].

The system recognized five activities of the instructor using morphologic features obtained after instructor detection. An overall efficiency comparison of the proposed research work with the above-mentioned approaches for different activities, tested on similarly trained data, is shown in Table 7 along with a graph below. The results of the HMM-based approach trained for different data sets are also comparatively good.

Table 7. Overall efficiency comparison between previous and proposed research work

	Overall efficiency percentage		
Activity performed	**Space-time based Template matching Approach Result**	**Fuzzy classification on Instructor Activity Recognition Result**	**HMM based Sequential approach for Instructor activity Recognition Result Trained for Different Data**
Walking	--	94	98.403
Writing	87	91	86.9383
Pointing to Board	83	95	91.0714
Pointing to Presentation		93	96.6173
Standing	94	98	98.2142

Constraints and limitations

Generally, every action recognition system has some constraints and limitations. In training and recognizing data, the given activity should be classified properly while it is performed for at least 4 sec. Otherwise, it gives the wrong output, as far as the number of times an activity is being performed. Multiplying it with a time duration of 4 sec will give an approximately correct time of an activity being performed in a whole sequence. But this constraint can be used as a parameter in an HMM recognition system. By giving the input as the number of frames, the assumed time for a particular system can be varied according to the requirements.

Conclusion and future work

The proposed research work was done by using a Hidden Markov model. Low-level features of the instructor activities were extracted by dividing sequential images into small meshes. This feature vector was later quantized by k-means clustering to form a set of symbols. These symbols were presented to the hidden Markov model for training and recognition. In training the HMM data sets, five different activities were used to optimize the parameters of the HMM, and later, these parameters were used in activity recognition to correctly recognize the activities performed by the instructor. The main experiment results came from recognizing different activities of an instructor in a lecture room, gathered by training the HMM mainly with three data sets. We observed that when training and testing data are of the same type, the system shows a 100 percent recognition rate, and when the data are of different types, the performance starts declining. Hence, this performance can be improved by training the model with a mixed data set. We observed in different experiment results that by increasing the number of training data sets, recognition performance increased.

To improve our current results, the feature extraction technique can be improved by introducing different geometric and morphologic features, which is really helpful in representing human actions. Furthermore, by dividing activities into subclasses, a more robust HMM can be built to overcome the problem of recognizing overlapped or physically similar, but actually different, activities. This instructor activity recognition can be used in different applications, such as patient monitoring, employee performance analysis, and surveillance systems, to recognize suspicious behavior or actions. Future work can be characterized into three major parts, where we seek improvement in pre-processing techniques, HMM-based improvements and improvements using derived parameters of HMM decoding.

References

[1] A. Georgakopoulos, Laura K. Guerrero, "Student Perceptions of Teac hers' Nonverbal and Verbal Communication: A Comparison of Best and Worst Professors across Six Cultures," *International Education Studies* Vol. 3, No. 2, pp. 3 - 16, 2010. Article (CrossRef Link)

[2] Pavan K. Turaga, Rama Chellappa, V. S. Subrahmanian, Octavian Udrea, "Machine Recognition of Human Activities: A Survey," pp. 1473-1488, 2008. Article (CrossRef Link)

[3] J. K. Aggarwal, Q. Cai, "Human motion analysis: A review," *Comput. Vis. Image Understand.*, vol. 73, no. 3, pp. 428 -440 1999. Article (CrossRef Link)

[4] D. M. Gavrila, "The visual analysis of human movement: A survey," *Comput. Vis. Image Understand.*, vol. 73, no. 1, pp. 82 -98 1999. Article (CrossRef Link)

[5] T. B. Moeslund , A. Hilton, V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Comput. Vis. Image Understand.*, vol. 104, no. 2, pp. 90 -126 2006. Article (CrossRef Link)

[6] A. Yilmaz, M. Shah, "Actions sketch: A novel action representation," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 1, pp. 984 -989 2005. Article (CrossRef Link)

[7] L. Gorelick , M. Blank , E. Shechtman , M. Irani, R. Basri, "Actions as space-time shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2247 -2253, 2007. Article (CrossRef Link)

[8] M. Belkin, P. Niyogi, "Advances in Neural Information Processing Systems," pp. 585 -591 2001 :MIT Press

[9] M. H. Yousaf, H. A. Habib, K. Azhar, "Fuzzy Classification of Instructor Morphological Features for Autonomous Lecture Recording System," *Published in Information Journal*, Vol. 16, No. 8 (B) pp. 6367-6382, August, 2013. Article (CrossRef Link)

[10] Y. Kameda, S. N. Guchi, M. Minoh "CARMUL: concurrent automatic recording for multimedia lecture," *IEEE on Multimedia and expo*, 2003, pp. 677-80, vol. 2. Article (CrossRef Link)

[11] H. A. habib, M H. Yousaf, M. mufti, "Gesture and scene recognition based autonomous video lecture acquisition framework from distance learning," *In proceeding of the 6 WSEAS International conference on distance learning and web engineering 2006*, pp. 97-100. Article (CrossRef Link)

[12] A. Ali, J. K. Aggarwal, "Segmentation and Recognition of Continuous Human Activity". In *IEEE Workshop on Detection and Recognition of Events in Video*, pp. 28-35, 2001. Article (CrossRef Link)

[13] Baum, L. E.; Petrie, T., "Statistical interference for probabilistic Functions of Finite state Markov Chains," *The Annals of Mathematical Statistics* 37(6), pp. 1554-1563. Retrieved 28, Nov. 2011. Article (CrossRef Link)

[14] S. Eickeler, A. Kosmala, G. Rigoll, "Hidden Markov Model Based Continuous Online Gesture Recognition," *In: International Conference on Patter Recognition (ICPR)*, Brisbane, vol. 2, pp. 1206-1208, 1998. Article (CrossRef Link)

[15] Yamato, J., Ohya, J., Ishii, K., "Recognizing human action in time-sequential images using hidden Markov model," *CVPR 1992*. Article (CrossRef Link)

Engr. Asim Raza received the BS (CE) degree with distinction from COMSATS Institute of Information Technology, Wah Cantt, Pakistan in 2009 and MS (CE) degree from University of Engineering and Technology, Taxila, Pakistan in 2014. He joined UET Taxila, Pakistan, as MS Full time Scholar in 2011where he is currently a Lab Engineer and Ph.D. research scholar. His main areas of research interest are computer vision, image processing and digital system design.

Dr. Muhamad Haroon Yousaf is an Assistant Professor in Computer Engineering Department of University of Engineering and Technology Taxila, Pakistan. He has completed his PhD Computer Engineering in November 2012 from the same university. His core area of research is Image Processing and Computer Vision. He has published more than 30 Publications in International and National Journals and Conferences, He is working with International Journals as Reviewer and Editor of Technical Journal UET Taxila, Pakistan. He has been the recipient of HEC BEST University Teacher Award for the year 2012-2013 by Higher Education Commission of the Pakistan. He is also member of National Curriculum Review Committee of Higher Education Commission of the Pakistan for the Computer Engineering Program.

Engr. Hassan Ahmed Sial is serving as Lecturer in computer science and engineering Department University of engineering and technology Lahore, Na rowal Campus. He has received his MS Computer Engineering (Signal and image processing) degree from university of Engineering and Technology Taxila, Pakistan in August, 2014. He received BS Computer Engineering degree with institute and campus bronze medal from Comsats Institute of information technology Islamabad in February 2012.His core areas of interest are digital image/video processing, computer vision and its application in medical. Hassan has been striving to bring innovations in the said fields through his research.

Dr. Gulistan Raja is currently working as Professor and Chairman of Electronics Engineering Department University of Engineering and Technology, Taxila. He has authored/co-authored more than seventy research publications in reputed international journals, and refereed conferences. He is a senior member of IEEE and has been member of technical program committee in many IEEE international conferences and reviewer in international journals. His research interests include digital image/video signal processing and coding, VLSI design and CAD of image/video systems, multimedia systems and communication and medical image processing.