

# Optimization and Performance of Human Activity Using Invariant Approach of Gabor Filter In Combination with HMM

1Rajeev Shrivastava, 2Dr. Shachi Awasthi

1Research Scholar, 2 Professor

Department of Electronics and Communication Engineering

Jayoti Vidyapeeth Women's University, Jaipur

**Abstract:** We present a biometric system performing identification of automatic human activity recognition. This system is based on Gabor features extraction using Gabor filter bank construction. For feature extraction the input image is convolved with Gabor filter bank to select a set of informative and non-redundant Gabor features. The extracted features are again subjected to Discrete Radon Transform (DRT) to extract a sequence of feature vectors from an image. The HMM (Hidden Markov Model) is used for matching the input human activity image to the stored images. The hmm-based system developed in this paper matches the feature set (observation sequence) for a test image with an HMM of the claimed image, through viterbi alignment. A distance measure is obtained by calculating negative log likelihood. The purpose of this research is to develop a novel, accurate and efficient human activity verification system.

**Keywords:** Gabor, DRT, HMM.

## I. INTRODUCTION

Human detection based action recognition is a constantly expanding research area due to number of applications for surveillance (behaviour analysis), security (pedestrian detection), control (human-

Computer interfaces), content based video retrieval, etc. It is, however, a complex and difficult-to-resolve problem because of the enormous differences that exist between individuals, both in the way they move and their physical appearance, view-point and the environment where the action is carried out. Fig. 1 shows some images from the KTH's database, demonstrating the variation of the human poses w.r.t. different camera views and for different actions.

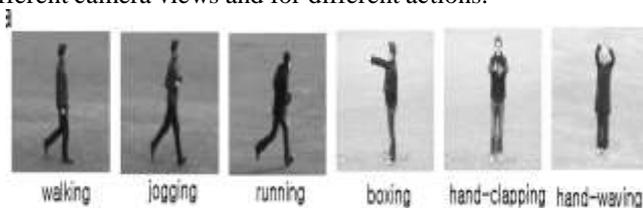


Fig 1

Toward this end, several approaches can be found in the literature. Some approaches are based on holistic body information where no attempt is made to identify individual body parts. Authors like use HMM and AdaBoost for recognition of 3-D human action considering joint position

or pose angles. However, there are actions which can be better recognized by only considering body parts, such as the dynamics of the legs for walking, running and jogging. Consequently, action recognition can be based on a prior detection of the human body parts (A. Mohan et al., 2001). Our approach introduces a technique for view-invariant human detection and subsequently learning the stochastic changes of the body part components to recognize different actions.

A fair amount of research works have been published in literature for Gabor based image recognition. Lades et al. developed a Gabor wavelet based face recognition system using dynamic link architecture (DLA) framework which recognizes faces by extracting Gabor jets at each node of a rectangular grid over the face image (M. Lades et al., 1993). Wiskott et al. Subsequently expanded on DLA and developed a Gabor wavelet based elastic bunch graph matching (EBGM) method to label and recognize facial images (L. Wiskott et al., 1997). In the EBGM algorithm, the face is represented a graph, each node of which contains a group of coefficients, known as jets. However, both DLA and EBGM require extensive amounts of computational cost. Liu and Wechsler have developed a Gabor feature based classification protocol using the Fisher linear discriminate model for dimension reduction (C. Liu et al., 2002). Shan et al. have developed an enhanced fisher model using the AdaBoost strategy for face recognition (S. Shan et al., 2005). Zhang et al proposed a face recognition method using histogram of Gabor phase pattern (B. Zhang et al., 2002).

More recently methods such as (Y. Lee et al., 2003, Gang Pan et al., 2003) have been proposed for comparing faces using pure 3D data. Pan (Gang Pan et al., 2003) uses registration error of low resolution Structured Light Scans (SLS) While Lee (Y. Lee et al., 2003) uses contour lines. However the most promising of recent research has focused on using a combination of both intensity and 3D information. In (M. Bronstein et al., 2003), Bronstein and Bronstein have patented their approach to 3D face recognition using bending-invariant canonical forms. This Technique makes use of the empirical observation that while transformations of the human face are non-rigid in nature, the set of possible transformations belong to the isometric (or

length preserving) set of transformations. They then combine the calculated canonical surface with intensity values and apply dimensionality reduction. A weighted Euclidean distance measure in this space is capable of differentiating between identical twins. In (M.G. Srinivas et al.,2003), the authors propose a system using the classical Eigen faces approach applied to 3 color channels and 1 range channel tested using the XM2VTS database, where they report a 98.5% recognition accuracy using only 15 eigenvectors. Chang *et. al.* give an excellent comparison of 2D and 3D face information in (P. Flynn et al.,2003), they also present one of the largest collections of multi-modal faces to date. A good review of the current state of the art in combined 2D- 3D face recognition can be found by the same authors in (K. Chang et al.,2003). The general trend of research in this field shows that using both 2D and 3D modalities results in a significant improvement over using either modality alone. After this Gabor Filter Bank Representation for 3D face recognition was developed by Jamie cook, Vinod Chandran, Sridharan and Clinton Fookes (Jamie cook et al.,2005). Face recognition system using Gabor filter bank with HMM model & Face recognition system using log Gabor filter bank with HMM model was developed by me(Rajeev Shrivastava et al.,2010, Ankita Nigam et al.,2011).

In this context, human body parts should ne first detected in the image. Authors like (A. Micilotta et al.,2005, D. Ramanan et al.,2007) describe human detection algorithms by probabilistic body part assembly. Authors like (L. Fengjun et al.,2006) use HMM and Adaboost for recognition of 3-D human action considering joint position or pose angles

### PROBLEM STATEMENT

Automatic human activity Recognition is a challenging problem in computer vision. One of its main goals is the understanding of the complex human visual system and the knowledge of how humans represent human activity in order to discriminate different identities with high accuracy

Two basic and conceptually independent problems have to be addressed by this kind of systems:

1. Human activity detection and
2. Recognition of the detected human activity

Our work on the recognition stage, taking the detected human activity as the input to the algorithm, this stage can be separated in two steps:

### MODELLING FLOW CHART

**1.Feature extraction.** The systems developed in this paper use similar feature extraction techniques. The bulk of the image processing and feature extraction involves the calculation of the discrete Radon transform (DRT) of each image. The DRT is obtained by calculating projections of each image at different angles. After some further image processing (normalization), each of these projections

constitutes a feature vector in an observation sequence. These features are classified as global features, since they are not extracted at stroke or sub-stroke level.

**2.Image modeling.** The systems developed in this dissertation use two very different approaches to model a specific human activity image. In the case of the HMM-based system, each Human activity image is modelled by an HMM of which the states are organized in a ring.

**3.Matching.** The distance between a test image and a model for the claimed image is obtained as follows. The HMM-based system developed in this paper matches the feature set (observation sequence) for a test image with an HMM of the claimed image, through Viterbi alignment. A distance measure is obtained by calculating a negative log likelihood.

**Verification.** When a claim is made that a test image belongs to a specific person, the extracted observation sequence is first matched with a model of the image, so that a distance measure is obtained. This distance measure is then normalized in order to compensate for the variation in the human activity image. The variation in the human activity image is estimated by matching all of the training images with the image model. In this way several distance measures are obtained. Statistics of these distance measures are then used to estimate the variation in the image training set. A global threshold, that is a threshold which is the same for all image, can therefore be used. Test images, for which the normalized distance measure is less than this threshold, are accepted the others are rejected. A schematic representation of the systems developed in this dissertation is given in figure 3.1

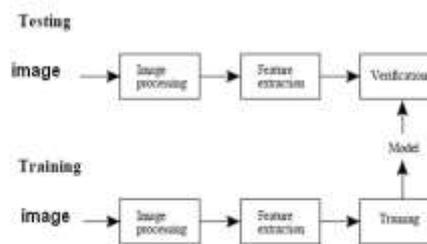


Fig 3.1 Schematic representation of the systems

Suppose we tend to area unit given a hidden mathematician model (HMM) with state area, initial chances of being in state and transition chances of transitioning from state to state. Say we tend to observe outputs  $y_1, \dots, y_T$ . The foremost seemingly state sequence that produces the observations is given by the repeat relations:

$$V_{1,k} = P(y_1 | k) \cdot \pi_k$$

$$V_{t,k} = \max_{x \in S} (P(y_t | k) \cdot a_{x,k} \cdot V_{t-1,x})$$

Here  $V_{t,k}$  is that the chance of the foremost probable state sequence  $P(x_1, \dots, x_T, y_1, \dots, y_T)$  to blame for the primary observations that have as its final state. The Viterbi path may be retrieved by saving back pointers that

bear in mind that state was utilized in the second equation. Let  $x_1, \dots, x_T$  be the perform that returns the worth of accustomed reckon if  $\text{Ptr}(k, t)$ , or if compute  $V_{t,k}$  if  $t > 1$ , or  $k$  if  $t = 1$ . Then:

$$x_T = \text{argmax}_{x \in S}(V_{T,x})$$

$$x_{t-1} = \text{Ptr}(x_t, t)$$

Here we're victimisation the quality definition of arg soap.

The complexness of this algorithmic rule is

$$O(T \times |S|^2)$$

Consider a village wherever all villagers area unit either healthy or have a fever and solely the village doctor will confirm whether or not every includes a fever. The doctor diagnoses fever by asking patients however they feel. The villagers could solely answer that they feel traditional, dizzy, or cold.

The doctor believes that the health condition of his patients operate as a distinct Markoff chain. There are a unit 2 states, "Healthy" and "Fever", however the doctor cannot observe them directly, they're hidden from him. On every day, there's a definite probability that the patient can tell the doctor he/she is "normal", "cold", or "dizzy", betting on her health condition.

The observations (normal, cold, dizzy) together with a hidden state (healthy, fever) type a hidden mathematician model (HMM), and might be depicted as follows within the Python programming language:

```

states = ('Healthy', 'Fever')
observations = ('normal', 'cold', 'dizzy')
start_probability = zero.6, 'Fever': 0.4}
transition_probability = {
    'Healthy' : zero.7, 'Fever': 0.3},
    'Fever' : zero.4, 'Fever': 0.6}
}
emission_probability = zero.5, 'cold': 0.4, 'dizzy': 0.1},
    'Fever' : zero.1, 'cold': 0.3, 'dizzy': 0.6}
}
    
```

In this piece of code, start probability represents the doctor's belief concerning that state the HMM is in once the patient 1st visits (all he is aware of is that the patient tends to be healthy). the actual chance distribution used here isn't the equilibrium one, that is (given the transition probabilities) close to zero.57, 'Fever': 0.43}. The transition probability represents the modification of the health condition within the underlying Markoff chain. during this example, there's solely a half-hour probability that tomorrow the patient can have a fever if he's healthy nowadays. The emission probability represents however seemingly the patient is to feel on every day. If he's healthy, there's a five hundredth probability that he feels normal; if he includes a fever, there's a hr probability that he feels dizzy.

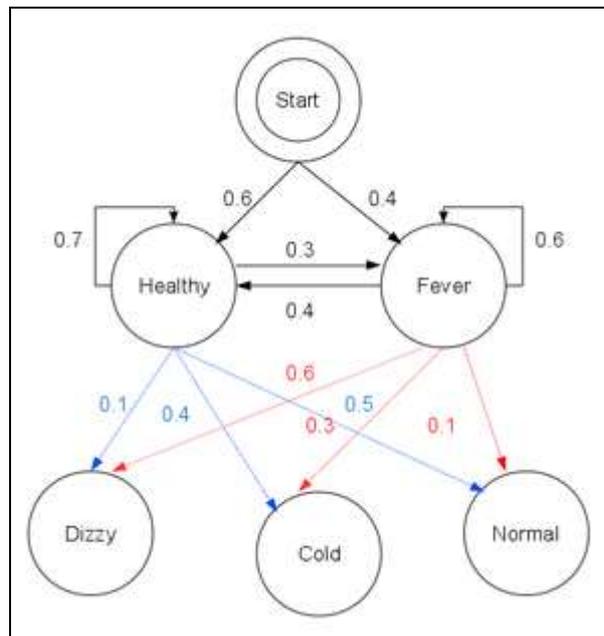


Figure 3.2: Graphical representation of the given HMM

The patient visits three days in a row and the doctor discovers that on the first day she feels normal, on the second day she feels cold, and on the third day she feels dizzy. The doctor has a question: what is the most likely sequence of health conditions of the patient that would explain these observations? This is answered by the Viterbi algorithm.

### RESULTS & DISCUSSION

In image processing, a Gabor filter, named after Dennis Gabor, is a linear filter used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. Its impulse response is defined by a sinusoidal wave (a plane wave for 2D Gabor filters) multiplied by a Gaussian function. Because of the multiplication-convolution property (Convolution theorem), the Fourier transform of a Gabor filter's impulse response is the convolution of the Fourier transform of the harmonic function and the Fourier transform of the Gaussian function. The filter has a real and an imaginary component representing orthogonal directions. The two components may be formed into a complex number or used individually.



Figure 4.1: GUI Showing the Control Panel for implementation of the proposed work



Figure 4.2: GUI Showing the Control Panel Loaded with Image and Parameters of the Gabor Filter Bank

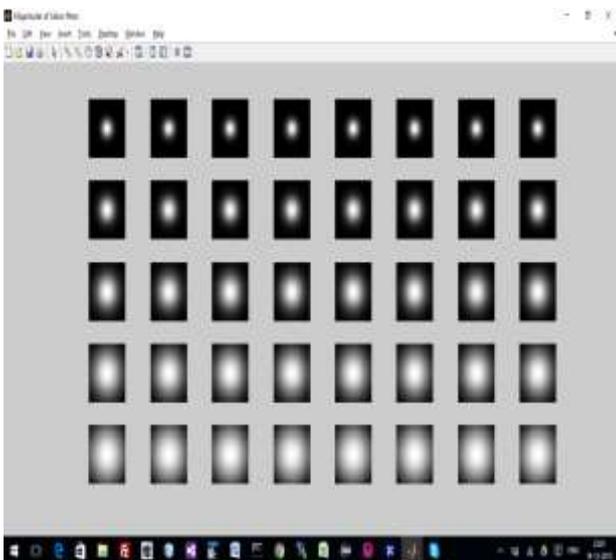


Figure 4.3: GUI Showing the Gabor Filter Bank with 5 Scales and 8 Orientations

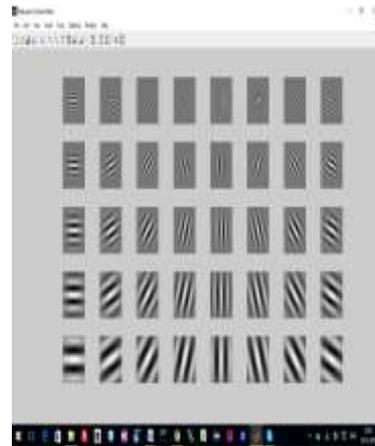


Figure 4.4: GUI Showing the Gabor Filter Bank with 5 Scales and 8 Orientations

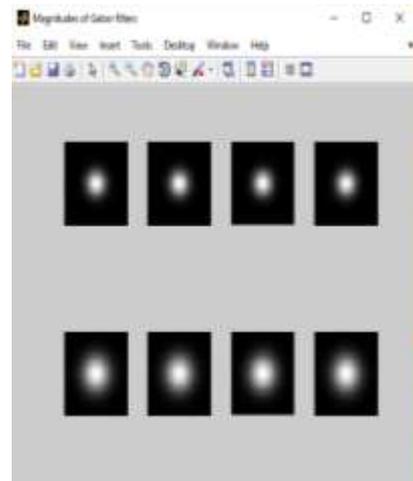


Figure 4.5: GUI Showing the Gabor Filter Bank with 2 Scales and 4 Orientations

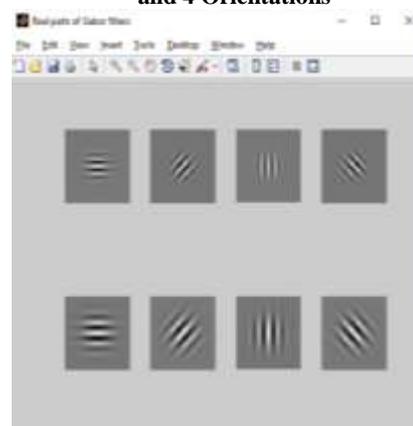
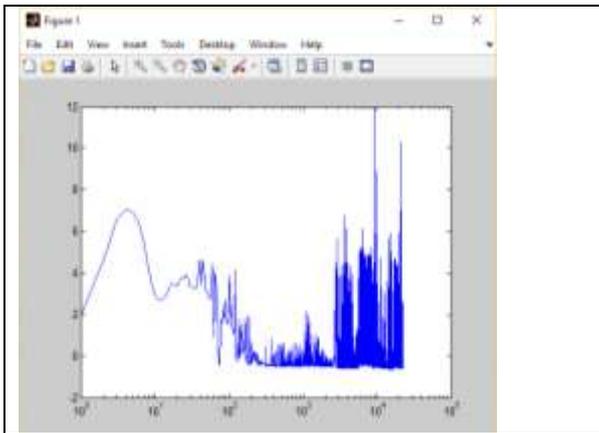
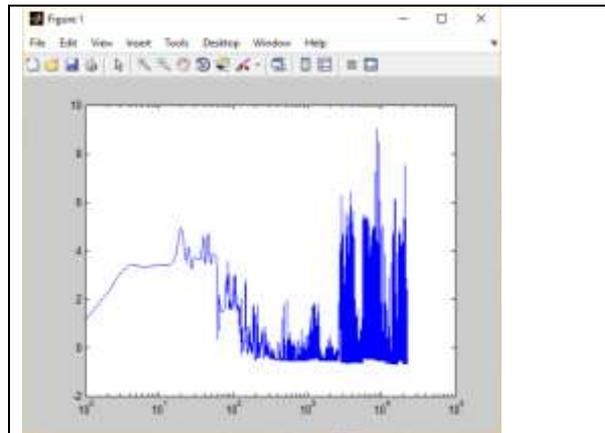


Figure 4.6: GUI Showing the Gabor Filter Bank with 2 Scales and 4 Orientations

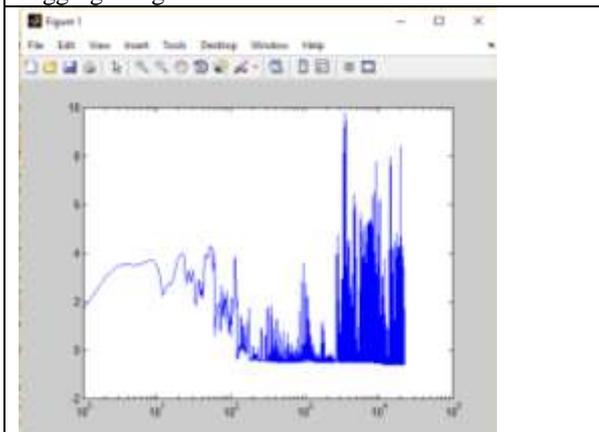
Graph Drawn from the Features obtained from Gabor Filter of different training images



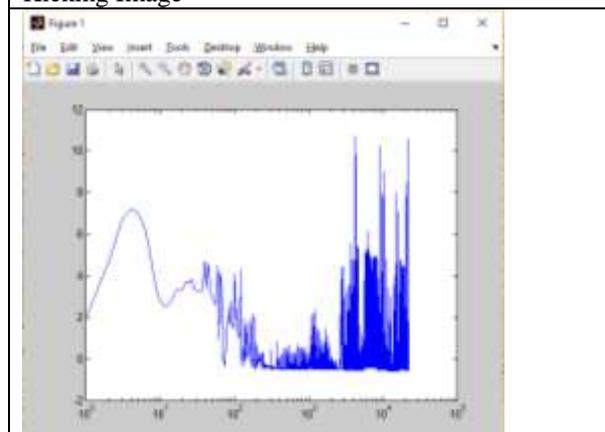
Jogging Image



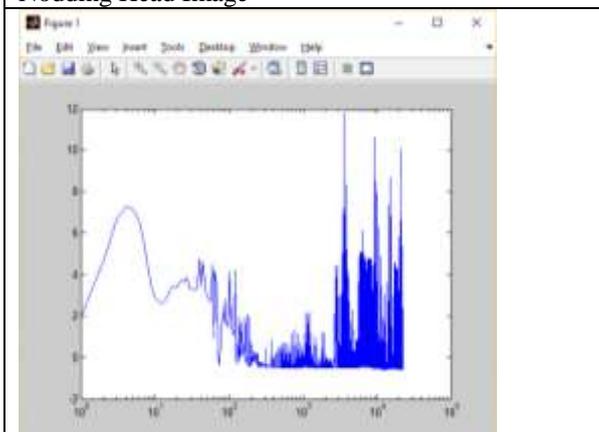
Kicking Image



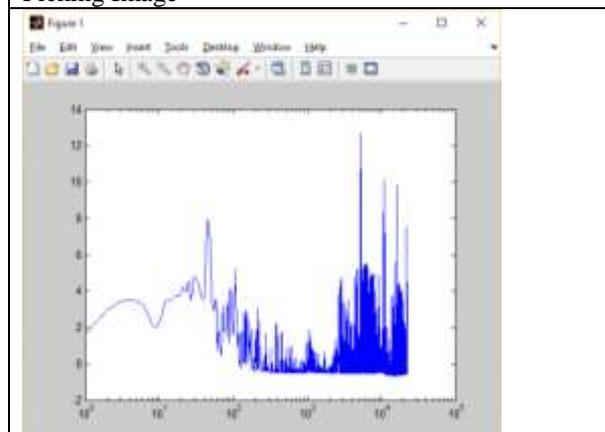
Nodding Head Image



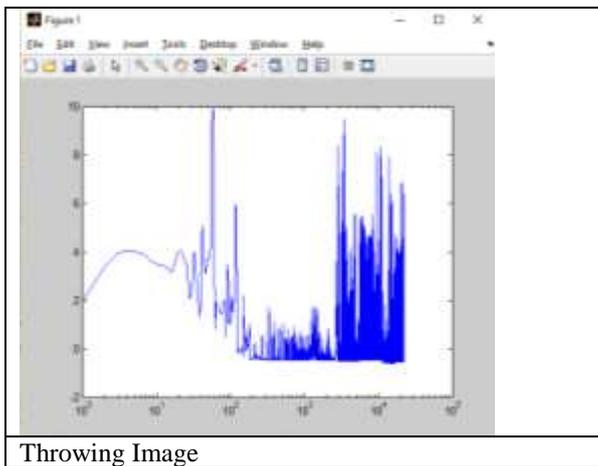
Picking Image



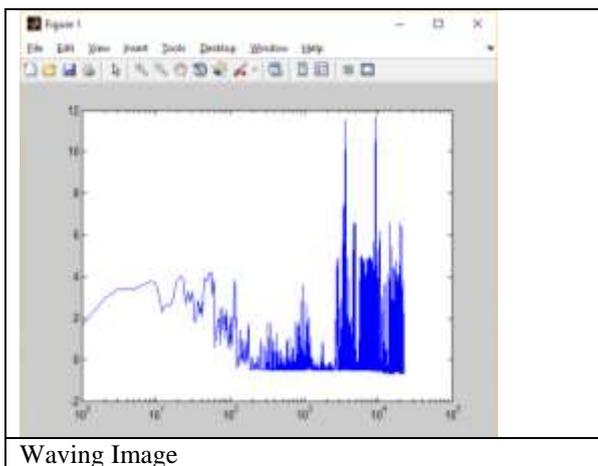
Punching Image



Standing Image



Throwing Image



Waving Image

Figure 4.7: GUI Showing the Gabor Features for different training images

Table 4.1: Reading taken from different mean values of the features used in HMM from the Gabor Features

Action	Mean Value
Jogging	3.51405765973212
Kicking	6.16537786216019
Nodding Head	0.819895681138487
Picking	3.01147006278625
Punching	3.98093690743621
Standing	0.0881943471554590
Throwing	3.58052860347304
Waving	3.58052860347304

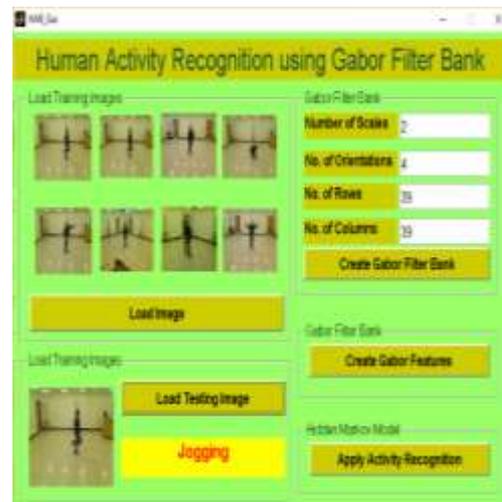


Figure 4.8: GUI Showing the Results obtained on basis of Gabor Feature Extraction, HMM and Viterbi Distance

The proposed system provides high efficiency with accuracy upto 82% i.e for different testing images the implementation done using MATLAB provides accuracy upto 82%. The algorithm has been executed for different parameter values of the Gabor Filter like number of scales, number of orientations, number of rows and number of columns. The proposed system testing snapshot shown in fig 4.8

### Conclusion and future direction

In this paper we developed human activity verification system: a HMM – based system the feature extraction method, is based on the GABOR features extraction and the calculation of the DRT. For this system Gabor filter is used for selecting Gabor feature for the human activity recognition. Small subnets of Gabor features capable of discriminating form other human activity images that are stored in the database. In this paper the system developed uses the hidden Markov model (HMM) to match a test human activity image with an appropriate reference image. this system use Gabor filter to extract the sequence of informative Gabor feature from the given human activity 1 to extracted feature are again subjected to DRT to extracted features vectors from a image the HMM-based system developed in this paper matches the feature set ( observation sequence ) for the test image with the HMM of claimed image ,through vitrerbi alignment. The same operation can be performed by using log Gabor filter with HMM and improvement of phase detection technique can be enhanced

### REFERENCES

- [1] A. Micilotta, E. Ong, and R. Bowden. Detection and tracking of humans by probabilistic body part assembly. In British Machine Vision Conference, pages 429–438, 2005.
- [2] A. Mohan, C. Papageorgiou, and T. Poggio. Example-based object detection in images by components. IEEE Transaction on Pattern Analysis and Machine Intelligence, 23(4):349–361, 2001.
- [3] Ankita Nigam and Rajeev Shrivastava “Face recognition

- using Gabor filter with Hidden Markov Model” Proceedings of the 2011 3rd IEEE International Conference on Computer Modeling and Simulation (ICCMS 2011).
- [4] B. Zhang, S. Shan, X. Chen, W. Gao, “Histogram of Gabor phase patterns: a novel object representation approach for face recognition”, IEEE Trans. Image Process., Vol. 11, No. 4, 2002, pp. 467–476.
- [5] C. Liu, and K. Wechsler, “Gabor feature based classification using the enhanced Fisher linear discriminate model for face recognition”, IEEE Trans. Image Process., Vol. 11, No. 4, 2002, pp. 467–476.
- [6] D. Ramanan, D. Forsyth, and A. Zisserman. Tracking people by learning their appearance. IEEE Transaction on PAMI, 29(1):65–81,2007.
- [7] Gang Pan, Zhaohi Wu, and Yunhe Pan, “Automatic 3D face verification from range data,” *International Conference on Acoustic, Speech & Signal Processing*, pp. 193–196, 2003.
- [8] Jamie cook, Vinod Chandran, Sridharan and Clinton Fookes “Gabor Filter Bank Representation for 3D face recognition” Proceedings of the Digital Imaging Computing: Techniques and Applications (DICTA 2005)
- [9] K. Chang K. W. Bowyer and P. J. Flynn, “A survey of approaches to three-dimensional face recognition,” Tech. Rep., University of Notre Dame Technical Report,2003.
- [10] L. Fengjun and N. Ramkant. Recognition and segmentation of 3-d human action using hmm and multi-class adaboost. In Computer Vision-ECCV, pages 359–372, 2006.
- [11] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. IEEE, 2:257–286, 1989.
- [12] L. Wiskott, J.M. Fellous, N. Kruger, and C.V. Malsburg, “Face recognition by elastic bunch graph matching”, IEEE Trans. Pattern Anal. Mach. Intel., Vol. 19, No. 7, 1997, pp. 775–779.
- [13] M. Bronstein, A. Bronstein and R. Kimmel, “Expression-invariant 3D face recognition,” in *AVBPA*, 2003.
- [14] M.G. Srintzis, F. Tsalakanidou, D. Tzovaras, “Use of depth and colour eigenfaces for face recognition”, *Pattern Recognition Letters*, vol. 24, pp. 1427–1435, 2003
- [15] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C.V. Malsburg, b C. Wurtz and W. Konen, “Distortion invariant object recognition in the dynamic link architecture”, IEEE Trans. Computers, Vol. 42, No. 3, 1993, pp. 300–311.
- [16] P. Flynn, K. Chang, K. Bowyer, “Multi-Modal 2D and 3D Biometrics for Face Recognition,” in *Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG03)*, 2003, pp. 187–194.
- [17] Rajeev Shrivastava and Ankita Nigam “Face recognition using Gabor filter with Hidden Markov Model” Proceedings of the *International Conference on TRENZ INFORMATION SCIENCES & COMPUTING* – 2010.
- [18] S. Park and J. Aggarwal. Semantic-level understanding of human actions and interactions using event hierarchy, 2004. CVPR Workshop on Articulated and Non-Rigid Motion, Washington DC, USA.
- [19] S. Shan, P. Yang, X. Chen and W. Gao, “AdaBoost Gabor Fisher classifier for face recognition”, Proc. IEEE Int. Workshop Analysis and Modeling of Faces and Gestures, 2005, pp. 278–291.
- [20] T. Moeslund, A. Hilton, and V. Krüger. A survey of advances in vision-based human motion capture and analysis. In CVIU 104, pages 90–126, 2006.
- [21] Y. Lee, K. Park, J. Shim, and T. Yi, “3D face recognition using statistical multiple features for the local depth information,” *Intl. Conf. on Multimedia and Expo, ICME '03*, vol. 3, pp. 133–136, 2003.