International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

Feature Emulation and Cluster Analysis to Discern Gait-A Maiden towards Progression

Riant Kaur†* and Heena‡

†Department of Information Technology, Chandigarh Engineering College, Landran, Mohali, Punjab, India

Accepted 03 Jan 2016, Available online 10 Jan 2016, **Vol.6, No.1 (Feb 2016)**

Abstract

Every individual has its own style of walking. It varies from person to person. A person can be identified by its walking style which is termed as Gait. This is a better way of identification as it is associated just with the individual and not with the surpassing information from one place to another in the background. In this dissertation, we work on live home videos in MATLAB which are converted into frames and each frame is being worked upon. Hanavan's model has been introduced for feature extraction, extracting six anthropometric parameters giving clear discrimination among various individuals. Feature detection and feature matching is executed by implementing SURF detector which augments the accuracy. SVM classifier is used for separability of extracted features. It is enhanced by implementing K-means clustering analysis algorithm and MDA (multi-linear discriminant analysis framework) in recognition phase producing CCR (Correct Classification Rate), matching time, MSE (Mean Square error) and PSNR (Peak Signal to Noise ratio) comparisons. Therefore, these techniques give faster results and better accuracy.

Keywords: *Feature extraction, Gait Pal and Pal Entropy (GPPE), Hanavan's model, Speed up Robust Feature detector (SURF), Support Vector Machine (SVM), K-Means Clustering iterative algorithm, Multi-linear Discriminant Analysis (MDA).*

1. Introduction

In the present era, due to rapid development in science and technology, it becomes all the more evident to provide security against uncertain events. Today's technology has come up with various solutions to keep a check on security contraventions; one feasible solution is the biometric recognition. Biometrics is to discern identity of a discrete personality instinctively. It is to deploy human depiction for recognition. Biometric systems are more enticing for identification purposes. Various biometric systems have been implemented based on iris, DNA, voice, face, finger print, palm print identification. In this disquisition, Gait is implemented. Gait is a present-day emerging revolution in biometrics whereby it is possible to investigate the walking fashion of humans. It is a person's walking style; it is not just constrained to walking but also includes running, strolling, jogging, brisk walking etc. Here, gait is recorded using video cameras. Gait biometric has a leading edge as compared to various biometrics. It can be acquired without the target getting conspicuous about it. Also, no detailed quality of camera is required for recording it.

It gives appropriate results with any level of quality captured. An ingenuous proof of an individual from a separating distance makes it a purely engaging biometric. The potential to perceive an imaginable and possible risk in the near vicinity provides the client with a warning to proceed with further actions prior to the subject in suspicion getting involved in that possible threat. Gait is not just limited to security applications but also involves other applications like clinical and medical applications involving pathology, making diagnosis of patients suffering from diseases, cerebral paralysis and stride dissections and upgrading performance of athletes. Gait recognition includes gait cycle, gait events and gait phases.

There are two gait phases namely swing and stance phase. Swing phase is defined as the position of foot as it moves up from the ground in air and stance phase refers to the position of foot at rest on ground. There are seven gait events, four in stance and three in swing phase respectively. Gait cycle is defined as the rest period i.e., the time period from one dull occasion, when the foot is in rest on the ground to the other dull occasion when the foot achieves the contact back with the ground returning from motion in air.

Gait recognition system involves four steps starting from capturing and loading the video and subtracting the foreground area from the background moving area and producing black and white binary shadows. This is followed by extracting of features from these binary shadows and matching these features with the

*Corresponding author: **Riant Kaur**

database and recognizing various individuals using various techniques.

2. Analysis of Related Literature

Gait recognition system is strongly dependent on different walking styles. The literature (A.E. Bobick *et al*., 2001) focused on participant's two walking styles; first, when the subject lacks in movement or change giving steady body movement and second, when subject takes long steps in a specified path or direction giving tread movement. The technique was related to extracting the feature vector when the subject matter was executing a peculiar task by computing expected-confusion metric. The proposed metric operates using mutual statistics, working in a colossal population to foretell how efficiently an extracted feature works to find out the participant's identity. The test was carried out on 15 to 20 participants moving in parallel and angular views in outdoor and indoor locations in front of the camera giving improved results.

A simple methodology with a mere purpose of identifying people using gait appearance representation and classification of gender was also introduced in literature (L. Lee. *et al*., 2002). The work was based on videos capturing the statistically independent and right angled views of subject's walking styles. The proposed methodology was simple but effective enough to extract a feature vector that provides all the requisite information to discern a person's identity and classify its gender. The technique was to investigate the remembrance performance of two procedures in various identification activities to compute a collection of features over a period of time. The test was carried out on gait videos over a time period with different illumination situations. The categorization of gender was carried out by SVM classier i.e., Support Vector Machine. The experimental results gave clear gender identity discrimination among different people. The literature (Bashir Khalid *et al*., 2009) defined gait biometric as the combination of two types of data i.e., motion and mien data. The mien data could be hampered by various external and behavioral means, therefore they proposed a representation called Gait Entropy Image (Geni) which was capable of encapsulating both kinds of data and extracting out the motion data from the mien data giving better efficiency in terms of accuracy.

The proposed model was tested on three datasets namely USF human ID, SOTON and CASIA. The method of Gait using Pal and Pal Entropy (GPPE) was implemented to see its relative performance over Shannon entropy (M.Jeevan *et al.* 2013). The performance of the proposed feature was tested on CASIA Gait and Treadmill Databases and was found to fare well over Shannon entropy. The proposed feature was found to withstand under covariate conditions such as subject's walking direction, carrying a bag or wearing a thick coat and subject having different speed.

3. Proposed Gait Recognition Methodology

In other proposed Gait recognition methods, Correct Classification Rate (CCR) and computational time needs to be improved. Several parameters have not been considered for recognition and thus the system is enhanced with following objectives and methodology to overcome these issues: The main objectives of this dissertation are:

- To extract six new anthropometric parameters using hanavan's model.
- To improve Gait Recognition System using SURF (Speed Up Robust Feature) for feature selection and matching.
- To enhance Gait Recognition System using MDA (Multi-linear discriminant analysis) for recognition.
- To elevate Gait Recognition System using K-means algorithm for clustering.
- To compare results based upon four parameters CCR, Matching time, MSE and PSNR.

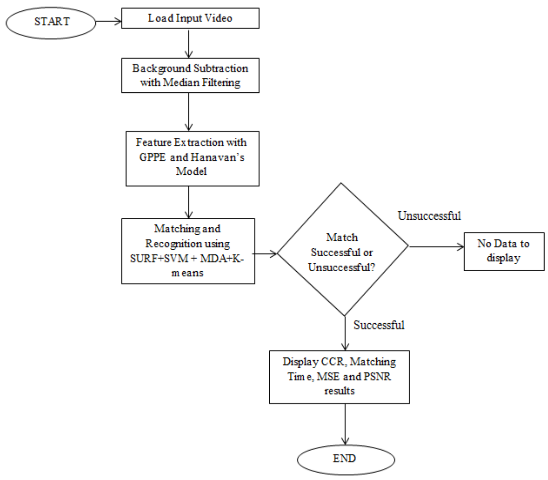


Fig.1 Flowchart of proposed Gait Recognition System

3.1 Loading of Video

In the proposed work, input query video is first loaded to be compared with the images in the database. In figure 2, the input video is being loaded. In this video, target is wearing a hat unlike target's video in the database. This is done to show that the result is not hampered by any external means.

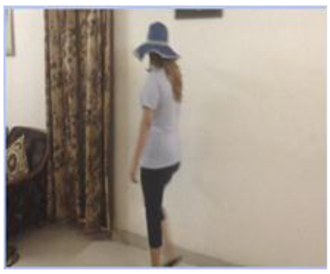


Fig.2 Loading of input query video

3.2 Background Subtraction

After the video is being loaded background subtraction is carried out. In this, the foreground moving target is subtracted from the background model for further processing. This video is converted into frames called video images and two succeeding frames get subtracted. The background model which remains constant gives zero intensity and changes into black pixels and the moving target that changes, becomes white pixels, hence producing black and white pixels called binary silhouettes as shown in figure 3. In this method, Gaussian mixture model and median filtering is used to remove noise. It is computed as in (1),

$$D_k(x,y) \begin{cases} = 1 \; if \; |F_k(x,y) - B_{k-1}(x,y)| > T \\ 0 \qquad otherwise \end{cases} \qquad (1)$$

Where $D_k(x,y)$ is the resulting difference between succeeding frames, $F_k(x,y)$ is the present frame, $B_{k-1}(x,y)$ is the background constant frame and T is the threshold that will repress the shadow relying on the value allocated.

Fig.3 Binary silhouette produced after background Subtraction

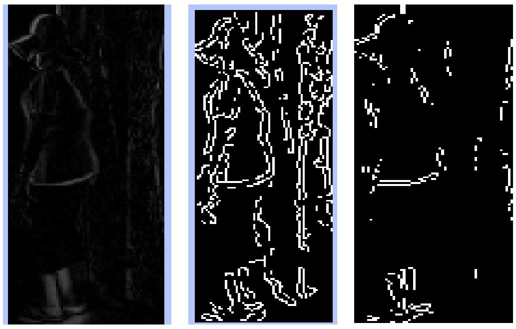
3.3 Feature Extraction

After obtaining binary silhouettes, different features are extracted from these silhouettes. Here, feature extraction is done using Gait cycle, Gait Pal and Pal Entropy (GPPE) and hanavan's model. Firstly, the gait cycle is determined followed by computing GPPE which is the unpredictability related with arbitrary variable for every pixel in the binary silhouettes. GPPE is used as a feature that is being extracted.

It is defined as in (2),

$$H_{i(m,n,\tau)} = \sum_{i=1}^{n} p_{i(m,n)} e^{(1-p_{i(m,n)})} \qquad (2)$$

Where, $p_{i(m,n)}$ is the possibility of level of gray i of position of pixel (x,y) throughout the gait cycle. Then, edge detection is performed which detects points in the image where the brightness is changing sharply and converts those points into curved line segments called edges. There are two edge detection methods used which are pre-defined in MATLAB namely prewitt and sobel presentations. Figure 4 shows GPPE, prewitt and sobel presentations.



(a) (b) (c)

Fig.4 (a) GPPE presentation, (b) Prewitt presentation, (c) Sobel presentation

Hanavan's Model: It is the human model that is designed geometrically. Initially this was brought into practice by Miller and Morrison. The model was modified in certain aspects and can be termed as modified hanavan's model. In this model, the torso of an anthropoid is segregated into a triad of trunks as follows: Upper trunk which consists of two levels

- First is omphalion, finding the centre point
- Second is xyphion, It is an oval shaped column.

Middle trunk which is an oval shaped solid and lower trunk which is also an oval shaped column. According to the model, human hand can be described as ovoidal revolution, human foot and thigh can be delineated as oval shaped solids, with a disc-shaped probable near end core and rounded crest at a distant end respectively. 41 anthropometric parameters (Clauser *et al*., 1969) are given in table 1 below,

Table 1 41 anthropometric parameters which can be measured from hanavan's model (Clauser *et al*., 1969)

No	Parameter	No	Parameter
1	Length, Hand	21	Circumference, Toe
2	Length, Wrist to Knuckle	22	Circumference, Ankle
3	Length, Forearm	23	Circumference, Shank
4	Length, Upperarm	24	Circumference, Knee
5	Length, Elbow to Acromion	25	Circumference, Upper Thigh
6	Length, Foot	26	Circumference, Head
7	Length, Shank	27	Circumference, Chest
8	Length, Thigh	28	Circumference, Xyphion Level
9	Length, Head	29	Circumference, Omphalion Level
10	Length, Upper Trunk	30	Circumference, Buttock
11	Length, Xyphion to Acromion Level	31	Width, Hand
12	Length, Middle Trunk	32	Width, Wrist
13	Length, Lower Trunk	33	Width, Foot

14	Circumference, Fist	34	Width, Toe
15	Circumference, Wrist	35	Depth, Hip
16	Circumference, Forearm	36	Width, Chest
17	Circumference, Elbow	37	Width, Xyphion Level
18	Circumference, Axillary Arm	38	Width, Omphalion Level
19	Circumference, Foot	39	Width, Coxae
20	Circumference, Ball of Foot	40	Length, Xyphion Level to Chin/Neck Intersection
41	Length, Hip to Chin/Neck Intersection = $P_{12} + P_{13} + P_{40}$		

In the proposed work, the mass of body components is computed with the help of prediction equations given in table 2, formulated by (Clauster *et al*. 1969), where M is the mass measurement of the entire body and Pi are the various anthropometric parameters stated in the above table 1.

Table 2 Prediction Equations of hanavan's model (Clauser *et al*., 1969)

Segment	Prediction Equation
Hand	$m = 0.038*P_{15} + 0.080*P_{32} - 0.660$
Forearm	$m = 0.081*M + 0.052*P_{16} - 1.650$
Upper arm	$m = 0.007*M + 0.092*P_{18} + 0.050*P_{5} - 3.101$
Foot	$m = 0.003*M + 0.048*P_{22} + 0.027*P_{6} - 0.869$
Shank	$m = 0.135*P_{23} - 1.318$
Thigh	$m = 0.074*M + 0.138*P_{25} - 4.641$
Head	$m = 0.104*P_{26} + 0.015*M - 2.189$
Trunk	$m = 0.349*M + 0.423*P_{41} + 0.229*P_{27} - 35.460$

Six anthropometric parameters have been evaluated as a result as shown in figure 5,

1. Distance between head and feet
2. Distance between both arms
3. Length of one arm
4. Length of one leg
5. Distance between right arm and left leg
6. Distance between left arm and right leg

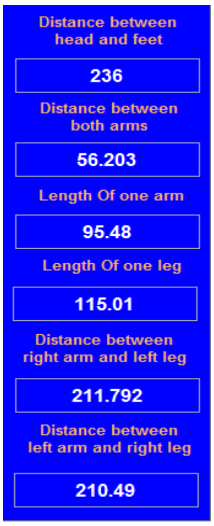


Fig.5 Six anthropometric parameters extracted using hanavan's model

3.4 Matching and Recognition

After executing feature extraction, matching and recognition phase is implemented. The input query video is matched with the videos in the database giving matched or not matched results. Various techniques are used in the entire process of this phase.

Speed Up Robust Features (SURF): SURF is a resilient feature detector which was presented by Herbert Bay. It is used in machine vision system to distinguish items from each other. In our disquisition, SURF helps to detect critical feature points that will match the images unambiguously. The focal point of SURF is a cluster of two-dimensional sequenced functions in its entirety, which are square shaped and on combining form a family of wavelets. Then, to detect the blobs in a requisite picture, it sums up the amount of approximate estimates to the Hessian blob detector. It is used to perceive persons on the premise of tread. Interest points are chosen at distinctive positions such as T-junctions, corners, centres and blobs. It consists of two parts: feature detection and feature matching. In the feature detection, SURF makes use of an intermediary object representation called integral image. This integral image is determined using the input image and boosts up the speed of evaluating the interesting feature points. Integral image IL: (x) can be calculated as in (3),

$$I_{\Sigma}(x,y) = \sum_{i=0}^{i\leq x}\sum_{j=0}^{j\leq y} I(i,j) \qquad (3)$$

Then, the Hessian Matric determinant is calculated based on the value of integral image as in (4),

$$H(x,\sigma) = \begin{matrix} L_{xx}(x,\sigma) & L_{x,y}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{matrix} \qquad (4)$$

Here $L_{xx}(x,\sigma)$ implies to the involution of the subsequent order Gaussian derivative dX2 with respect to the image which is located at $X = (x, y)$and same goes for L_{yy} andL_{xy}.

In feature matching, SURF finds out features from same positions in two images and matches two images based on these features. It is a 64- dimensional vector which will adopt nearest neighbour search using a K-Dimension Tree.

KD-tree is an objective binate tree having two set of parameters. First set of parameters is a set of points P, for which we wish to construct a KD-tree. Second set of parameters D takes into consideration the distance downwards till the root of the sub tree which stands zero at the beginning. The procedure for construction of KD-Tree [10] is explicated below:

- **Process KD-TREE**(P, D); **Input**: Set P including number of points and present depth D
- **if** P contains only one point,
- **then** return null;
- **else if** D is even,
- **then** Divide P into two subsets with a vertical line 1 through the median x-coordinate of the points in P,

Let P1 be the set of points to the left of 1 or on 1 and P2 be the set of points to the right of l;
• **else** Split P into two subsets with a horizontal line 1 through the median y-coordinate of the points in P,
Let P1 be the set of points to the below of 1 or on 1 and P2 be the set of points above l;
• **P, location:=**median;
• **P, left child**:=KDTREE(Pl , D + 1);
• **P, right child**:= KDTREE(P2, D + 1);
• return P;
• **Output**: The root of the kd-tree storing P.

PCA is employed to diminish the magnitude statistics without hampering the originality of information. Support Vector Machine (SVM) classifier is used followed by SURF detector. Once the features are detected, it acts as a separator between the matching and non- matching pixels of the two images in matching query. The matching feature points and non-matching feature points are separated by a plane giving clear distinguishing patterns between two images. Therefore, it helps to classify the patterns of the images while matching. SVM algorithm (Jindong Chen *et al*., 2010) is stated as,

Simple SVM
Step 1: Candidate SV = {closest pair from opposite classes}
Step 2: while there are violating points **do**
Find a violator
Step 3: candidate SV = candidate SV; S violator
if any αp < 0 due to addition of c to S **then**
Step 4: candidate SV = candidate SV \ p
repeat till all such points are pruned
Step 5: end if
Step 6: end while

K-Means Algorithm*:* To enhance the functionality of SVM, k-means algorithm is used. It is based upon clustering of feature points based on their colour. This algorithm works in iteration and finds the nearest neighbour. The centres of these clusters are computed and the input query is compared with these cluster centres. Based on the comparison results, these clusters are sorted in the rank of their similarity to input query. Then, the cluster which has the maximum similarity with the input query, all images in that cluster is compared with input query. The advantage of this algorithm is that, the entire database need not be searched for matching; only the most similar cluster images are matched.

Here, K is the input query and there are n image frames. These n frames are clustered into k clusters. The most similar cluster out of these k clusters is searched for matching. The frame n is assigned to that particular cluster. Centroids of these clusters are found out by the distance between the frame and the cluster centre it is closest to. The entire process keeps on iterating until the required criteria are fulfilled. Therefore, it makes the matching process faster and more accurate.

Multi-linear Discriminant Analysis (MDA): SVM classifier is enhanced by using MDA. It is a framework which uses mode optimization and higher order tensors. MDA works in lower dimensional space in each phase of feature extraction which helps to reduce the computational cost. The small sample sizes are multiplied by larger scales solving the long-time pending issue of smaller size of samples. Also, it allows any number of dimensions, adding rows or columns helping to solve the dimensionality curse. It is computed as in (5), (6), (7), (8), (9),

$$S_B = \sum_{j=1}^{\prod_{o\neq k} m_o} S_B^j \tag{5}$$

$$S_B^j = \sum_{c=1}^{N_c} n_c \left(\bar{Y}_c^{k.j} - \bar{Y}^{k.j}\right)\left(\bar{Y}_c^{k.j} - \bar{Y}^{k.j}\right)^T \tag{6}$$

$$S_W = \sum_{j=1}^{\prod_{o\neq k} m_o} S_W^j \tag{7}$$

$$S_W^j = \sum_{i=1}^{N} \left(\bar{Y}_i^{k.j} - \bar{Y}_{c_1}^{k.j}\right)\left(Y_i^{k.j} - \bar{Y}_{c_1}^{k.j}\right)^T \tag{8}$$

$$Y_i = X_i \times_1 U_1 \dots \times_{k-1} U_{k-1} \times_{k+1} U_{k+1} \dots \times_n U_n \tag{9}$$

Therefore, figure 6 shows input query is matched with database video more efficiently and accurately using the proposed work of Gait Recognition.

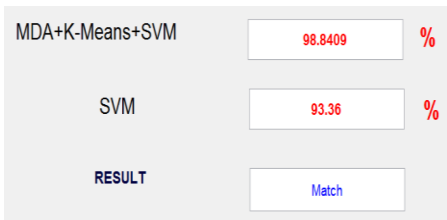


Fig.6 Accuracy percentage and matching results of proposed work

4. Experimental Results and Discussion

As this thesis work is based on walking style of individuals, we worked on live home video inputs with above proposed techniques. We have evaluated the results based upon four metrics namely:

(1). Correct Classification Rate (CCR)
(2). Mean Square Error (MSE)
(3). Matching Time
(4). Peak Signal-To-Noise Ratio (PSNR)

These four metrics are compared based on results obtained from previously solely used SVM classifier and results obtained from now used SVM with K-Means and MDA techniques. In this dissertation, we found out from the experimental results that, SVM with K-Means and MDA techniques works better and efficiently as compared to using SVM solely.

4.1 CCR

In the table 3 and figure 7, we have evaluated CCR for both methods. CCR is the measure of accuracy. It measures how accurately the query input video matches with the database video. We achieve almost 99% CCR (accuracy) with our proposed thesis work.

Table 3 Comparison of CCR between both algorithms

Parameter	SVM	SVM+K-MEANS+MDA
CCR (Accuracy in %)	93.3600	98.8409

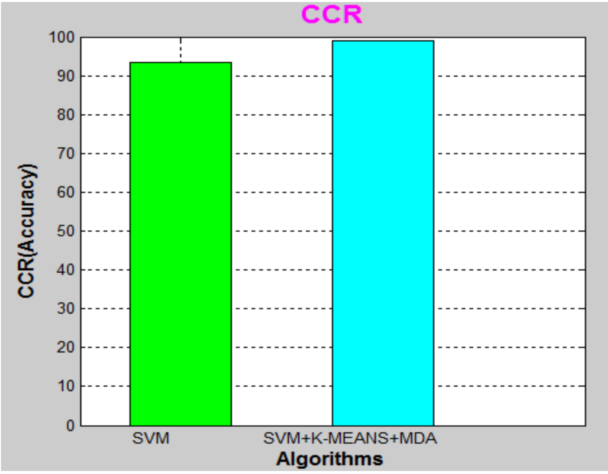


Fig.7 Bar graph showing comparison of CCR between both algorithms

4.2 Matching Time

In the table 4 and figure 8, we have evaluated matching time for both methods. It is the time taken to match the query input video with the database videos. We obtained less time with our proposed thesis work. Here,

T_{sim} is the time to evaluate similarity between input query and every video image in the database.
T_{sort} is the time evaluated to rank all video images according to the similarity.

As the images are clustered in our work using K-means in the database, the matching time is the total of three types of time; time to evaluate similarity between input query and cluster centers, time to evaluate similarity between input query and video images in nearest neighbor clusters and time to rank video images. Therefore, matching time is computed as in (10),

$$T_{cluster} = kT1_{sim} + 0(1 \log 1) \quad (10)$$

Where, k is the number of clusters; l is the images in the clusters which are nearest to the input query. Since $k \ll n$ and $1 \ll n$, then, $T_{cluster} \ll T_{total}$.

Table 4 Comparison of matching time between both algorithms

Parameter	SVM	SVM+K-MEANS+MDA
Matching Time (in ms)	1312267680769	1312267680638

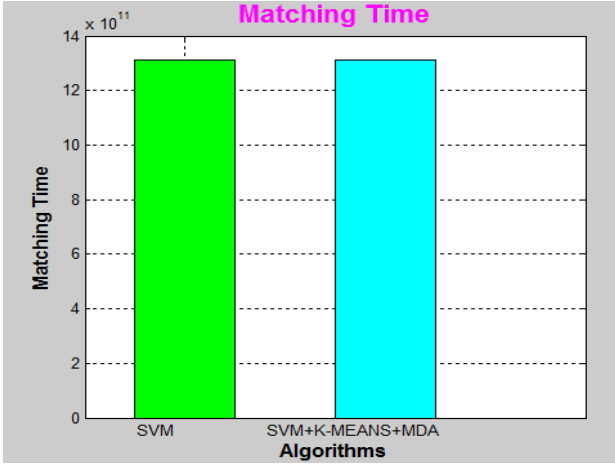


Fig.8 Bar graph showing comparison of matching time between both algorithms

4.3 MSE

In the table 5 and figure 9, mean square error is calculated for both methods. It calculates the quality of image obtained after decoding. The ideal case says it should be zero, which is practically not possible. So, it should be on minimal side as calculated in our proposed thesis work.
Therefore, MSE is computed as in (11),

$$MSE = \frac{\sum[f(i,j)-F(i,j)]^2}{N^2} \quad (11)$$

Table 5 Comparison of MSE between both algorithms

Parameter	SVM	SVM+K-MEANS+MDA
MSE	13.4767	8.2906

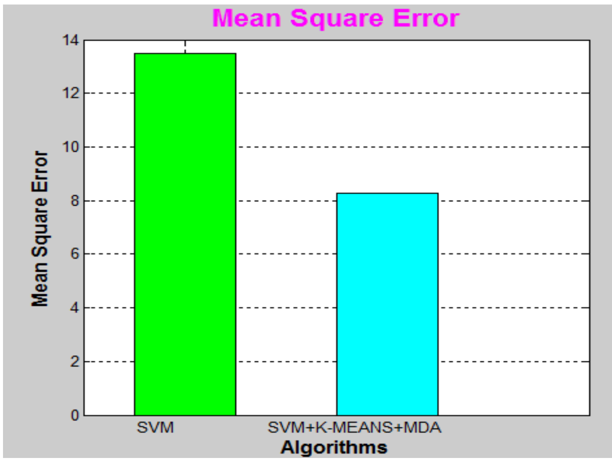


Fig.9 Bar graph showing comparison of MSE between both algorithms

4.4 PSNR

In the table 6, figure 10, we evaluated PSNR for both methods. It is the measure of quality of reconstructed image. It is inversely proportional to MSE which gives an ideal case of PSNR equals to infinity, which is practically not possible. So, it should be on maximum side as calculated in our thesis work.

Therefore, in the following equation, PSNR can be ascertained in decibels (db) as follows (12),

$$PSNR = 20 \log 10 \left(\frac{255}{RMSE}\right) \tag{12}$$

Where, the square root of MSE represents Root Mean Square Error i.e., RMSE.

Table 6 Comparison of PSNR between both algorithms

Parameter	SVM	SVM+K-MEANS+MDA
PSNR (db)	220.4727	229.8725

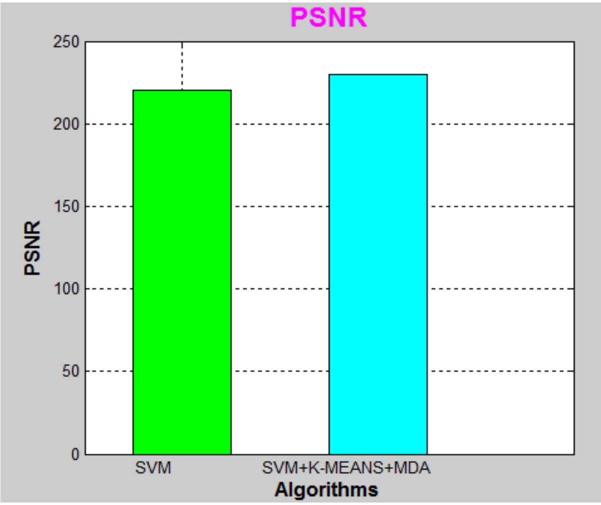


Fig.10 Bar graph showing comparison of PSNR between both algorithms

Conclusions

The present disquisition is carried out keeping in mind two objectives: First to enhance the quality of the video image obtained and secondly, to enhance the rate of matching of video image during matching and recognition phases. These two objectives were not achieved properly while solely using SVM classifier and the same were effectively achieved using new algorithms hanavan's model, K-means clustering and MDA with SURF detector along with SVM classifier. Using these algorithms, feature extraction, feature point detection, selection and separability was much more accurate and faster. There was no need to match the input query with every image in the database because of formation of clusters using K-means, hence saving the matching time and providing more scope of different dimensions using MDA. The research has given ensuring CCR, Matching Time, MSE and PSNR results. In future, Gait for very large databases and to improve occlusion may be evaluated.

References

A.E. Bobick, and A.Y. Johnson (2001), Gait Recognition Using Static, Activity Specific Parameters, *CVPR*, pp. 423-430.

Ai-Hua, Wang Ji Wei Liu (2007), Gait Recognition Method Based on positioning human body joints. International conference on wavelet analysis and pattern recognition, Beijing, China, pp. 1067-1071.

Bashir Khalid, Xiang Tao, Gong Shaogang (2009} Gait Recognition Using Gait Entropy Image, 3rd International Conference on Crime Detection and Prevention ICDP.

Clauser CE, McConville JT, T.J, Young JW (1969), Weight, Volume and Centre of Mass of Segments of the Human Body, AMRL Technical Report, Weight Patterson Air Force Base, Ohio

C. S. Lee and A. Elgammal (2006), Gait tracking and recognition using person-dependent dynamic shape model, in Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, pp. 553-559.

David G. Lowe (2004),Distinctive image features from scale-invariant key points *IJCV*, 60(2), pp.91-110.

D. Tan, K. Huang, S. Yu, and T. Tan (2007), Uni-projective features for Gait recognition, the 2nd International Conference on Biometrics

David .T.Alprt, M. S. Nixon and J. N. Carter (2003), Automatic Extraction and Description of Human Gait Models for Recognition Purposes, *Computer Vision and Image Understanding*, 90(1), pp. 1-41.

Dong Hui and Han Dian Yuan (2012), Research of image matching algorithm based on SURF features, International Conference on Computer Science and Information (CSIP), pp. 1140-1143.

Jindong Chen, Feng Pan (2010), Clustering-Based Geometric Support Vector Machines, *Life System Modeling and Intelligent Computing*, pp. 207-217.

J.E. Cutting and L.T. Kozlowski (1977),Recognition of Friends by Their Walk, *Bull. of the Psychonomic Soc.*, vol. 9, pp. 353-356.

L. Lee, and W. E. L. Grimson (2002), Gait analysis for recognition and classification, Proceedings of Automatic Face and Gesture Recognition, pp. 148-155.

Mikolajczyk K, Schmid C (2004), Scale & affine invariant interest point detectors, *IJCV*, Vol. 60,pp. 63-86.

M.Jeevan, Neha Jain, M.Hanmandlu, GirijaChetty (2013), Gait Recognition Based On Gait Pal and Pal Entropy Image, *IEEE*, pp. 4195-4199.