

Face Recognition Using Particle Filter

R.Sumathy, Dr.P.Raviraj

^{1,2}Dept. of CSE, Kalaignar Karuninidhi Institute of Technology, Coimbatore, India

Abstract

Multi view face recognition using multiple camera networks is an active research area. The main aim of this paper is to handle different pose variations in multi camera network and recognizing face from those videos. The traditional approaches handle the pose estimation explicitly, the proposed work will handle the multiple views of the poses. For a given set of multi view video sequences we use particle filter to track the 3D location of the head. The texture map is generated by back projecting the multi view video. The proposed work is developed using the Spherical Harmonic (SH) representation of the face from the texture mapped on to the sphere. A robust feature is constructed based on the properties of SH projection.

Keywords

Face recognition, still image based, video based face recognition, multi view recognition, particle filter and spherical harmonics.

I. Introduction

Face recognition is the process of identifying an individual or a person from a video or digital image and compared with the stored databases of faces. The application area of face recognition ranges from biometrics to surveillance systems. In general a face can be recognized in a still image or in video systems in a single view or in multiple views. The main drawback of using single view based face recognition is the information loss occurs during image formation. The image obtained from single view also suffers from pose variations. Another drawback is the one part of the image is self-occluded by another part of the image which is a problem in object tracking, motion capture, etc. [3].

Multi camera network provide a valuable tool for safety and security applications. Multi camera network composed of many distributed cameras with their own processors. By having so, the multiple views obtained from those cameras increases the chances of recognizing the person more efficiently. This can be achieved by estimating the pose of the person's head to an approximation. More attention has to be given on the quality of the image and camera calibration. There are many methods for multi view pose estimation [1],[2] but the pose estimation of head is a problem due to poor resolution of images, camera calibration etc.

Face recognition across multi camera network is the process of recognizing the faces from multiple views with diffuse pose and light variations. At a time we can obtain multi views of the face captured from multiple cameras that combines both frontal and non-frontal poses of the image. In this paper, a robust feature for recognizing face can be obtained For a given video sequence, the particle filter method is used to track the 3D location of the head. For each video frame, the texture map for the whole face is constructed by mapping the image intensity values onto the surface of spherical model using the back projection method. The spherical harmonic transformation of the texture map is computed and robust feature is constructed

II.Related Work

A. Face recognition based on still images

In general a face recognition process is formulated as follows: given a input still/video image is given. The detection of face can be done by segmenting the face and the features are extracted using known techniques and the face can be recognized. Fig.1. illustrates the face recognition process.

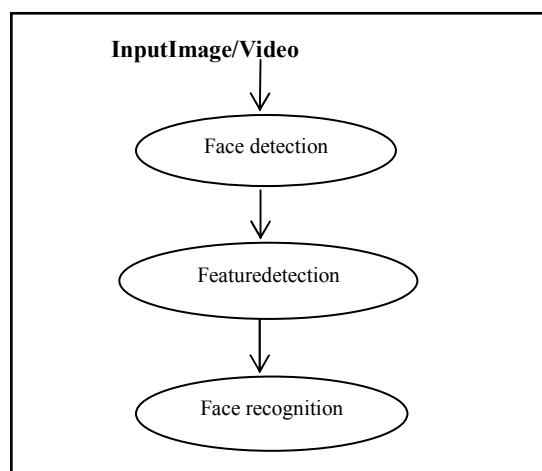


Fig. 1: Face recognition process

Many methods are proposed to resolve the problem of Face recognition. Karungaru et.al [4] proposed approach of recognizing face include template matching method that performed on genetic algorithms to automatically test several positions around the target. The information obtained is based on edge detection, YIQ ((perceived luminance), I, Q (color/luminance information)). The shape of the face is not recognized. Another approach [5] that locates the local features using Gabor mask and the feature is extracted. But the drawback is the system can be implemented for up frontal images and have low detection rate. The feature detection methods proposed by Zhang et.al [6] uses SIFT (Scale Invariant Feature Transform) that can be applied to real time systems and the method of implementing SIFT had been proposed in this paper.

Generally recognizing face from a single still image will not provide sufficient information due to unavailability of data's in those images. Also the view available in the image could not able to recognize face. Still image based multi view face recognition provides a solution. One such method [7] is to use pose estimation method and a pose estimator is constructed using support vector machine. The advantage is that data in SVM based pose estimator can be trained directly and it improves speed and accuracy. Yongbin et.al [8] proposed probabilistic appearance-based face recognition to work with multiple images uses a feature extraction algorithm to find features that describe face images of distinct people. The algorithm was robust to partial occlusions, expression and pose changes. Another approach is using view synthesis proposed by Baymer et.al [9] that creates virtual views from a given still

image using prior knowledge of 2D views of prototype faces subjected to various rotations. The combination of virtual views with the real view provides the example view where the face can be recognized. For handling the pose variation in real-life scenarios, Zhimin et.al [10] a pose-adaptive matching method that uses classifiers to deal with different pose combinations having a recognition rate of 84.45%. To obtain 3D model of face from 2D image Lu et.al [11] used feature points and warp algorithm and the texture map for the face model is constructed.

B. Face recognition based on Video

Normally video provides more information compared to still images (single or multiple views). The detection of face in video is of three processes:

- 1) Frame based detection-methods used for still images. The integration of detection and tracking i.e., detecting the face in the first frame and tracking over the sequence.
- 2) Temporal method to detect faces in multiple frames.

Most of the video sequences are collection of still images. So the algorithms applicable for still images are applied for video. One of the method is Probabilistic recognition of human faces [12]. The gallery consists of still images and the probes are video sequences. Using temporal information in a probe video, a model is proposed. As the model is non-linear to non-Gaussian noise, the motion vector and identity variable is approximated using sequential re-sampling importance algorithm. Zhao et.al [13] uses a method called Locality Preserving Projections (LPP) to recognize face sequence to discover semantic information (space and time) hidden in videos simultaneously and a comparison of V-LPP with Eigen face, Fisher faces and Laplacian faces. Zhen Lei et.al [14] used Ad boost algorithm in face detection for hausdorff distance which is a measure between face model and instance of object. The overall performance of the system is increased. Another detection approach for video based face recognition proposed by Ping et.al [15] used cascade verification technique that have three modules for eliminating non frontal portion, tilted face and frontal image is sent to the recognizer. [16] Adaptive hidden Markov models (HMM) to perform video-based face recognition uses training video sequences. The characteristics of the video sequence are analyzed and the scores provided by the HMMs are compared. The highest score provides the identity of the test video sequence. The local appearance based face recognition algorithm proposed by Stall Kamp et.al [17] classified the face. The scores are obtained and combined to estimate the entire sequence. The individual frames are weighted using distance-to-model (DTM), distance-to-second-closest (DT2ND), and their combination. Aggarwal et.al [18] possesses video-to-video face recognition by constructing a linear dynamical system ARMA model. Sub space angles are used for computing distances between probe and gallery video sequences.

C. Multi View Based Face Recognition

One of the multi view based face recognition uses [19] both gallery and probe. Frames of multi view sequence are collected together to form a gallery or probe set. The frontal or near frontal faces are picked by the pose estimator and retains and others are discarded. The recognition algorithm is frame based PCA and LCA fused by the sum rule. [20] uses AAM framework to the multi view video case. They demonstrate that when 3D constraints are imposed, the resulting 2D+3D AAM is more robust than the single view case. But recognition was not attempted. Chen et.al [21] used geometrical models to normalize pose variations. By back

projecting a face image to the surface of an elliptical head model, they obtained a texture map which was generated from different images that were compared in a probabilistic fashion.

D. Video Processing Over Distributed Networks

Tracking in video can be defined as a problem of locating a moving object (or multiple objects) over time based on observations of the object in the images. Multiple object tracking is one of the most fundamental tasks for higher level automated video content analysis. Kalman filter [22] achieves Multi target tracking by 2 steps: prediction and correction. The prediction step uses the state model to predict the new state of the variables. The correction step uses the current observations to update the object's state. A consensus algorithm is an interaction rule that specifies information exchange between a sensor and its neighbors that guarantees that all the nodes to reach a consensus. Distributed Kalman consensus algorithm defines a set of cameras communicating with its neighboring cameras and no cameras can pass information to its non-neighboring cameras. The extended Kalman Consensus allows to track targets on the ground plane using multiple measurements in the image plane taken from various cameras. Particle filter tracking approach [23] is a sequential Monte Carlo method which provides Monte Carlo approximation to the prediction distribution, which is used to search for newly observed targets. A particle filter is often used because the system and observation models are nonlinear and the posterior can temporarily become multi-model due to background clutter.

Spherical Harmonics (SH)-[24] Showed a connection between SH and PCA, where the principal components are equal to SH basis function under appropriate assumptions. Zhang et.al [25] proposed an algorithm to estimate the SH basis images for a face at a fixed pose from a single 2D image based on statistical learning. A similar strategy [26] is adopted where the distribution of SH basis images is modelled as Gaussian and its parameters are learned from a 3D face database. For this 3D face model and pose estimation to infer the face appearance is required.

III. Existing System

Many problems occur in the traditional approaches. From the single view based approach, the multiple views are generated using virtual view generation. The virtual view should have a closer appearance to the true rotated view. Using multiple view based approach, the pose estimation of head of a person is not clear due to poor resolution of images, camera calibration etc. which does not provide a robust feature that is important in the areas of surveillance. Another problem is adapting the face model parameters to customize the 3D model to the individual being tracked is being a difficult one.

IV. Proposed System

In the proposed system we can obtain multi views of the face captured from multiple cameras that combines both frontal and non-frontal poses of the image. For a given video sequence, the particle filter method is used to track the 3D location of the head. For each video frame, the texture map for the whole face is constructed by mapping the image intensity values onto the surface of spherical model using the back projection method. The spherical harmonic transformation of the texture map is computed and robust feature is constructed. Figure 2 shows the process of proposed system.

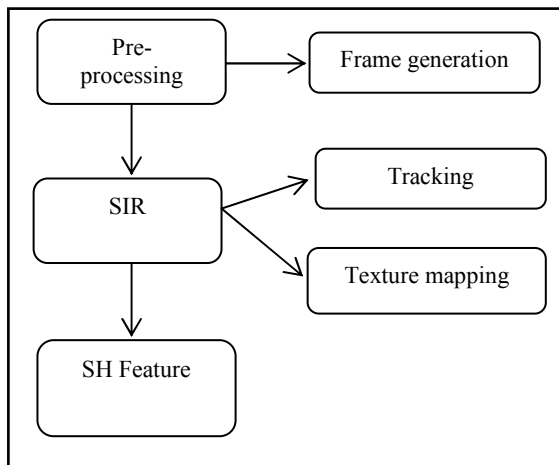


Fig. 2: Process of Proposed system

A. Pre-Processing

Pre-processing is the first phase in face recognition process which includes converting videos into frames

1. Converting videos into frames

Pre-processing is the process of converting the acquired videos into frames. The videos are acquired by any digital device like digital cameras in multiple views and the obtained video has to be optimized by performing cropping, scaling, aspect ratio adjustments, noise reduction, brightness and color corrections. It is also imperative to achieve properly sized video without motion artefacts due to interlace or film sources being transferred to video. In addition, in the production environment, we may not have control over the original creation of source material. Whether the capture is poor, or simply needs to be de-interlaced and resized, almost invariably the source video needs pre-processing in order to look its best when delivered over limited bandwidth networks



Fig.3: Shows the sample of 6 frames generated from a set of 500 frames taken from 3 cameras in multiple networks

B. Sequential Importance Resampling

1. Particle Filter

Tracking in camera based network uses [28] Particle filter (SIR) method for tracking the state of dynamic system that can be computationally traceable for large/high dimensional problems. A particle filter is often used because the system and observation models are nonlinear and the posterior can temporarily become multi-model due to background clutter. Particle filter is a Monte Carlo method for performing inference in state space models

where the state of the system evolves according to:

$$\dot{x}_k = f_k(x_{k-1}, v_{k-1}) \quad (1)$$

where x_k is a vector representing the state of the system at time k , v_{k-1} is the state of noise vector, f_k is possibly non-linear and time dependent function describing the evolution of the state vector. Information about x_k is obtained only through noisy measurement of it, z_k given by the equation

$$z_k = h_k(x_k, n_k) \quad (2)$$

where h_k is a possibly non-linear and time dependent function describing the measurement process and n_k is the measurement noise vector. The filtering problem involves the

estimation of state vector at time k , given all measurements up to and including denoted by $z_{1:k}$. In Bayesian settings the problem can be formalized by two steps:

1. Prediction step: In this step $p(x_k | z_{k-1})$ is computed from the filtering distribution $p(x_{k-1} | z_{1:k-1})$ at time $k-1$ is given by:

$$p(x_k | z_{1:k-1}) = \int p(x_k | x_{k-1}) p(x_{k-1} | z_{1:k-1}) dx_{k-1} \quad (3)$$

where $p(x_{k-1} | z_{1:k-1})$ is assumed to be known due to recursion that can be thought as prior over x_k and $p(x_k | x_{k-1})$ is given by equation (1).

2. Update Step: In this step the prior is updated with the new measurement z_k using the Bayes rule to obtain the posterior over x_k given by:

$$p(x_k | z_{1:k}) \propto p(z_k | x_k) p(x_k | z_{1:k-1}) \quad (4)$$

4.2.2 Sequential Importance Sampling (SIS)

Sequential importance sampling is based on importance sampling. In this one approximates the target p_x using the samples drawn from a proposal distribution q_x . Importance sampling is generally used when the target distribution is difficult to sample but much easier to distribute from proposal distribution. The discrepancy between the target and the proposal distribution can be overcome by assigning weight to each sample x^i by where $\pi(x)$ is a function that is proportional to $p(x)$. After applying the proposal distribution at time $k-1$ can be determined by

$$p(x_{0:k-1} | z_{1:k-1}) \approx \sum_{i=1}^N w_{k-1}^i \delta_{x_{0:k-1}} \quad (5)$$

where $\delta_{x_{0:k-1}}$ is the delta function centered at $x_{0:k-1}$

The SIS algorithm thus constructs a recursive propagation of weights and support points at each measurements. The algorithm for SIS is given below:

ALGORITHM 1: Sequential Importance Sampling

$$[\{x_k^i, w_k^i\}_{i=1}^N] = \text{SIS} [\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^N, z^k]$$

FOR $i=1:N_s$

$$\text{Draw } x_k^i \sim q(x_k | x_{k-1}^i, z^k)$$

$$\text{Assign the particle a weight, } w_k^i$$

End FOR

A common problem with SIS is *degeneracy* problem, where each particle has negligible weight and the variance of weight increases overtime. A suitable measure for degeneracy problem is the effective sample size N_{eff} defined as

$$N_{eff} = \frac{N_s}{1 + \text{var}(w_k^{*i})} \quad (6)$$

where $w_k^{*i} = p(x_k^i | z_{1:k}) / q(x_k^i | z_{1:k})$ is referred to as true weight. The estimate N_{eff} can be obtained as

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^N (w_k^i)^2} \quad (7)$$

where w_k^i is the normalized weight Small N_{eff} indicates severe degeneracy. The brute force approach is to use large N_s . This is impractical .So resampling is done to avoid degeneracy.

4.2.3 Sequential Importance Re-Sampling (SIR)

The effects of degeneracy can be reduced by resampling whenever a significant degeneracy is observed. The basic idea

is to eliminate degeneracy is to eliminate particles which have small weights and to concentrate on particles with small weights. The resampling step involves generating new set $\{x_k^{i*}\}_{i=1}^{N_s}$ by re-sampling N_s times from an appropriate discrete representation $p(x_{k-1} | z_{1:k_s})$ given by

$$p(x_{0:k-1} | z_{1:k-1}) \approx \sum_{i=1}^N k w_{k-1}^i \delta(x_k - x_k^i) \quad (8)$$

The weights are reset to $w_k^i = 1/ N_s$. This re-sampling procedure can be implemented in $O(N_s)$ times. The re-sampling algorithm is given .The algorithm indicates that the weights and state variables are considered and the sampling is made using the algorithm 1.The cumulative density function is calculated and if the value is less again the sampling is done. The sequential importance resampling algorithm is given below:

ALGORITHM 2: Sequential Importance Re-sampling

```
[{x_k^i, w_k^i} N_{i=1}^S] = SIR [{x_k^i, w_k^i} N_{i=1}^S ]
Initialize CDF: c1=0
FOR i=2:N_s
    Construct CDF: ci= ci-1+w_k^i
End FOR
Start at the bottom of CDF: i=1
Draw a starting point: u1 ~ U [0, :N_s-1]
FOR j=1: Ns
    Move along the CDF: uj= u1+N_s-1(j-1)
    WHILE uj> ci
        *i=i+1
    End WHILE
    Assign sample: x_k^{j*}=x_k^i
    Assign weight: w_k^j=:N_s-1
    Assign Parent: i^j=i
End FOR
```

The algorithm for generic particle filter is also introduced.

ALGORITHM 3: Generic Particle Filter

```
[{x_k^i, w_k^i} N_{i=1}^S] = PF [{x_{k-1}^i, w_{k-1}^i} N_{i=1}^S, z^k]
FOR i=1:N_s
    Draw x_k^i ~ q(x_k | x_{k-1}^i, z^k)
    Assign the particle a weight, w_k^i
End FOR
Calculate total weight: t=SUM [w_k^i } N_{i=1}^S]
FOR i=1:N_s
    Normalize: w_k^i = t^{-1} w_k^i
End FOR
Calculate N_{eff} using (7)
If N_{eff} < N_T
    Resample using algorithm (2)
    *[{x_k^i, w_k^i, -} N_{i=1}^S] = SIR [{x_k^i, w_k^i} N_{i=1}^S ]
End IF
```

is to eliminate degeneracy is to eliminate particles which have small weights and to concentrate on particles with small weights. The resampling step involves generating new set $\{x_k^{i*}\}_{i=1}^{N_s}$ by re-sampling N_s times from an appropriate discrete representation $p(x_{k-1} | z_{1:k_s})$ given by

$$p(x_{0:k-1} | z_{1:k-1}) \approx \sum_{i=1}^N k w_{k-1}^i \delta(x_k - x_k^i) \quad (8)$$

The weights are reset to $w_k^i = 1/ N_s$. This re-sampling procedure can be implemented in $O(N_s)$ times. The re-sampling algorithm is given .The algorithm indicates that the weights and state variables are considered and the sampling is made using the algorithm 1.The cumulative density function is calculated and if the value is less again the sampling is done. The sequential importance resampling algorithm is given below:

ALGORITHM 2: Sequential Importance Re-sampling

```
[{x_k^i, w_k^i} N_{i=1}^S] = SIR [{x_k^i, w_k^i} N_{i=1}^S ]
Initialize CDF: c1=0
FOR i=2:N_s
    Construct CDF: ci= ci-1+w_k^i
End FOR
Start at the bottom of CDF: i=1
Draw a starting point: u1 ~ U [0, :N_s-1]
FOR j=1: Ns
    Move along the CDF: uj= u1+N_s-1(j-1)
    WHILE uj> ci
        *i=i+1
    End WHILE
    Assign sample: x_k^{j*}=x_k^i
    Assign weight: w_k^j=:N_s-1
    Assign Parent: i^j=i
End FOR
```

The algorithm for generic particle filter is also introduced.

ALGORITHM 3: Generic Particle Filter

```
[{x_k^i, w_k^i} N_{i=1}^S] = PF [{x_{k-1}^i, w_{k-1}^i} N_{i=1}^S, z^k]
FOR i=1:N_s
    Draw x_k^i ~ q(x_k | x_{k-1}^i, z^k)
    Assign the particle a weight, w_k^i
End FOR
Calculate total weight: t=SUM [w_k^i } N_{i=1}^S]
FOR i=1:N_s
    Normalize: w_k^i = t^{-1} w_k^i
End FOR
Calculate N_{eff} using (7)
If N_{eff} < N_T
    Resample using algorithm (2)
    *[{x_k^i, w_k^i, -} N_{i=1}^S] = SIR [{x_k^i, w_k^i} N_{i=1}^S ]
End IF
```

4.2.4 Multi view tracking

In this stage the 3D position of the head is localized from multiple views. To describe the position and pose of an object we need 6D representation. But higher the dimensionality of the state space makes difficult in tracking problem .In SIR the number of particle grows dramatically for higher dimensional state spaces .Since we are considering only the location of the head in 3D the state space for tracking let the state space $s =$

(x, y, z) represents only the position of the head. As a result the state transition model $P(s_t | s_{t-1})$ is modelled. Similarly the observation model $P(O_t | s_t)$ of the tracker is based on multiple cues such as histogram, the gradient map and a geometric constraint.

Histogram: We assume a weak perspective camera model and calculate the image of the spherical model on the j^{th} camera image plane, which is a disk like region. A normalized histogram is built which is the first cue matching function $\Phi(O_t, s_t^i)$

Gradient map: On the circular perimeter of the model image, we select 90° arc segment on the top. This part of the boundary coincide with the arc produces a second cue matching score

$$\Phi(O_t, s_t^i) = \frac{1}{r_j^i} \sum_{m=1}^M |n_m \cdot \nabla I_m| \quad (9)$$

Geometric constraints: The detector is based on HOG. Potential human subjects are filtered using applying body size constraints and the detection result with highest confidence is picked that produces $\psi(O_t, s_t^i)$.

$$\Psi(O_t, s_t^i) = \begin{cases} 0 & \text{if } E_j^i C R_j^i \neq \emptyset \\ 1 & \text{otherwise} \end{cases} \quad (10)$$

The overall image likelihood can be calculated as

$$P(O_t | s_t^i) \propto \ln \psi(O_t, s_t^i) + \lambda_1 \ln \Phi(O_t, s_t^i) + \lambda_2 \ln \Psi(O_t, s_t^i) \quad (11)$$

where are determined by applying a logistic regression-like algorithm to independent data. The location of the head in 3D space is determined by:

$$s_t = \underset{s_t^i}{\operatorname{argmax}} P(s_t^i | O_t) \quad (12)$$

$$s_t^i = \underset{s_t^i}{\operatorname{argmax}} P(O_t | s_t^i) P(s_t^i | s_t^i) \quad (13)$$



Fig.5: Shows the result of tracking in multiple views. Our particle filter tracker tracks the location of the head in all the frames. The figure shows 6 frames out of 500

$$3. x_n = \sqrt{R^2 - z_n^2} \cos \alpha_n, y_n = \sqrt{R^2 - z_n^2} \sin \alpha_n$$

Then perform a coordinate transformation for these sample points. The original points are $\{(x^n, y^n, z^n) \mid n=0, 1, \dots, N\}$. After the transformation we obtain $\{(x'_{n,j}, y'_{n,j}, z'_{n,j})\}$. These are the coordinates in the j^{th} camera plane. The visibility of j^{th} camera is determined and an un-occluded point that (x^n, y^n, z^n) satisfies $z'_{n,j} \leq z'_{0,j}$ contribute image on the j^{th} camera image plane $z'_{0,j}$ is the head center to the j^{th} camera plane. The texture map T_j is obtained by back projection approach by creating a back projecting link between x^n, y^n, z^n and the image coordinates $\{(x''_{n,j}, y''_{n,j}, z''_{n,j})\}$. This procedure is repeated over all cameras. The associated texture map for the coordinates (x^n, y^n, z^n) is determined by:

$$T(x^n, y^n, z^n) = T^j(x^n, y^n, z^n) \quad (14)$$

where $j = \underset{j}{\operatorname{argmax}} w_{n,j} \mid j=1, 2, \dots, K$

V. SH Feature

A robust recognition system must be able to identify an object across variable lighting conditions. The robust feature is based on the theory of spherical harmonics which is set of orthogonal basis function defined over unit sphere.

The SH basis function for degree l and order m has the following form:

$$Y_{lm}(\theta, \phi) = K_{lm} P_l^m(\cos \theta) e^{im\phi} \quad (15)$$

where K_{lm} is normalization constant such that

$$\int_{\theta=0}^{\pi} Y_{lm} \int_{\phi=0}^{2\pi} Y_{lm} Y_{lm}^* d\theta d\phi = 1 \quad (16)$$

The real SH are orthonormal and they share most of the properties of the general spherical harmonics. The spherical model is different from 3D face model. A low resolution images can be dealt using SH model and hence it is suitable for camera network. Even when the face undergoes extreme pose variations the SH model extracts the feature from texture maps leading to pose robust face recognition. Another advantage of SH spectrum is ease of use. The only one parameter to be determined in SH feature is number of degrees. If the number of degrees is more means more approximation.

VI. Experimental Results

We compare the performance of our video recognition algorithm with 2 other ones:

- 1) Image-based recognition with SH spectrum feature and majority voting for video-level fusion. We use SVM with RBF kernel for every multi-view frame recognition
- 2) The Manifold-Manifold Distance (MMD)

(MMD) algorithm uses code and parameter settings. When comparing two multi-view videos, we first calculate the MMD between the sequence pairs of the same view, and then use the minimum MMD across views as the distance measure. Moreover, they both have certain ability to handle pose variations, especially the two algorithms based on manifold. However, because they are designed to work with a single camera, they are single-view in nature. Repeating these algorithms for each view does not fully utilize the multi view information. For example, we found in our experiments that mismatches made by the MMD algorithm often happens when the minimum MMD is produced between the back-off head clusters, which have similar appearance representations even for different subjects. In contrast, the proposed method based on a robust feature performs noticeably better in this experiment.

An additional advantage of the algorithm is that it requires no pose estimation or model registration. Comparison between the ensemble matching algorithm and the majority voting method which both uses the proposed feature demonstrates the superiority of a systematic fusion strategy to an ad-hoc one. Figure 6 shows the performances of various face recognition algorithms.

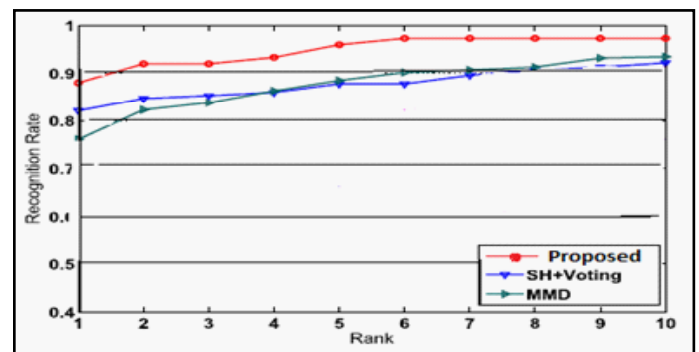


Fig.6 .Performances of various Face recognition Algorithms

VII. Conclusion

In this paper we presented a multi view based approach to recognize face. The main advantage of this paper is it does not require any pose estimation or model registration step. Under normal diffuse lighting conditions we presented a robust feature by SH theory. Also the feature acquisition is automated by multi view tracking algorithm using particle filter.

VIII. Future Work

The pose insensitivity property of SH representation relies on the assumption that the spherical function remains unchanged other than rotation. This could be affected by real world lighting conditions. Under normal lighting conditions this assumption is reasonable. In extreme lighting conditions the assumption becomes invalid because if there are any anisotropic illumination variation or strong directional lighting is casted onto the face from the side will result in large fluctuations of features and makes the performance to degrade. So the future work will focus on overcoming these drawbacks.

References

- [1] S.Ba ,J.M.Odobez "Probabilistic head pose tracking evaluation in single and multiple camera set ups", *Multimodal Techno . Perception Humans*, vol 4625, pp.276-286, June 2008
- [2] Q. Cai, A. C. Sankaranarayanan, Q. Zhang, Z. Zhang, z. Liu "Real time head pose tracking from multiple camera with generic model" in *Proc. CVPR Workshops*, pp.25-32 June 2010
- [3] Jang I. Y., Lee K. H. Depth "Video based human model reconstruction resolving self occlusion" *IEEE Transaction Consumer Electronics*, 2010, 56(3): 1933-1941
- [4] Karungaru, S., Fukumi, M.; Akamatsu, N "Face recognition using genetic algorithm based template matching" *IEEE Communications and Information Technology*, 2004. ISIT 2004. *IEEE International Symposium on* 1252 - 1257 vol.2
- [5] Er. Monika Verma Er. Pooja Rani Er. Harish Kundras "A Hybrid Approach to Human Face Detection 2010 *International Journal of Computer Application*" Vol 1 – No. 13
- [6] Jing Zhang , Hongshi Sang ; Xubang She "Overview of approaches for accelerating scale invariant feature detection algorithm" *Electric Information and Control Engineering (ICEICE)*, 2011 *International Conference on IEEE*
- [7] Wen Yi zho, Rama Chellappa "Image based face recognition issues and methods", *IEEE transactions* 1997
- [8] Yongbin Zhang, Aleix M. Martínez "A weighted probabilistic approach to face recognition from multiple images and video sequences" *Elsevier* August 2005
- [9] David Beymer Tomaso Poggio "Face recognition from one example view" *Massachusetts Institutue of Technology*, Sep 1995
- [10] Zhimin Cao, Qi Yin, Xiaou Tang, Jian Sun "Face Recognition with Learning-based Descriptor" *Computer Vision and Pattern Recognition (CVPR)*, *IEEE Conference* June 2010
- [11] Chen Lu, Yang Jie "Automatic 3D Face Model Reconstruction Using One Image" *Advances in Machine Vision, Image Processing, and Pattern Analysis Lecture Notes in Computer Science Volume 4153*, Springer 2006, pp 235-243
- [12] S. Zhou, V. Krueger, and R. Chellappa. "Probabilistic recognition of human faces from video" *Computer Vision and Image Understanding*, 91:214–245, July-August 2003
- [13] Ke Lu , Zhengming Ding , Jidong Zhao , Yue Wu "Video-based face recognition" *IEEE Image and Signal Processing (CISP)*, 2010 *3rd International Congress on (Volume:1)* Oct. 2010
- [14] Zhen Lei ,Chao Wang , Qinghai Wang , Yanyan Huang, "Real-Time Face Detection and Recognition for Video Surveillance Applications" *Computer Science and Information Engineering*, 2009 *WRI World Congress on (Volume:5)*, April 2 2009
- [15] Ping Zhang , Lorman, "A video-based face detection and recognition system using cascade face verification modules" *IEEE MS Applied Imagery Pattern Recognition Workshop, AIPR*, Oct. 2008
- [16] Xiaoming Liu ; Tsuhan Chen "Video-based face recognition using adaptive hidden Markov models", *Computer Vision and Pattern Recognition*, 2003. *Proceedings. 2003 IEEE Computer Society Conference on (Volume:1)* June 2003
- [17] Stallkamp, Ekenel, H.K.; Stiefelhagen, J. "Video-based Face Recognition on Real-World Data". *Computer Vision*, 2007. *ICCV 2007. IEEE 11th International Conference on* Oct. 2007
- [18] Aggarwal, G.; Chowdhury, A.K.R.; Chellappa, R. "A system identification approach for video-based face recognition" *Pattern Recognition*, 2004. *ICPR 2004. Proceedings of the 17th International Conference (vol 4)* Aug 2004
- [19] A. Pnevmatikakis, L. Polymenakos, "video-to-video face recognition", *Far field Intech Chennai* 2007, pp.468-486
- [20] 20. K. Ramnath, S. Koterba, J. Xiao, C. Hu, I. Mathews, s. Baker "Multi view AAM fitting and construction", *International Journal of Computer Vision* Vol 76, PP.183-204, Feb 2008
- [21] X. Liu, T. Chen "Pose-robust face recognition using geometry assisted probabilistic modeling" *IEEE Conference Computer Vision Pattern Recognition*, vol 1 pp.502-509, June 2005
- [22] R. E. Kalman "A new approach to linear filtering and prediction problems" *Transactions of the ASME – Journal of Basic Engineering*, No. 82 (Series D). (1960), pp. 35-45
- [23] Olfati-Saber, R., Fax, J. A., Murray, R. M. (2007). "Consensus and cooperation in networked multi-agent systems." *Proceedings of the IEEE*, pp.215–233
- [24] R. Ramamoorthi "Analytic PCA construction for theoretical analysis of lighting variability in images of a convex lambertian objects" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, pp.1322–1333 Oct 2002
- [25] L. Zhang, D. Samaras, "Face Recognition from a single training image under arbitrary unknown lighting using spherical Harmonics" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.28 pp.351-363, Mar. 2006
- [26] Z. Yue, W. Zhao, R. Chellappa "Pose encoded spherical Harmonics for face recognition and synthesis using a single image", *EURASIP J. Adv. Signal Process.*, pp.1-18, Jan 2008
- [27] R. Olfati-Saber, J. Fax, and R. Murray. *Consensus and Cooperation in Networked Multi-Agent Systems. Proceedings of the IEEE*, 95(1):215–233, Jan. 2007
- [28] M. Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking" *IEEE Transactions on signal processing*, vol. 50, , February

2002

[29] *Amit.K.RoyChowdry,Bi.song “CameraNetworks-The
Acquisition and Analysis of Videos over Wide Areas” San
Manteo,CA,USA,Morgan 2012*