KINSHIP VERIFICATION SYSTEM BASED ON MID LEVEL FEATURES

T. Theresa [1], S. Veluchamy[2]

[1]Department of Electronics and Communication Engineering, Anna University Regional Campus Madurai
Madurai, India. E-Mail: [1]theresalingam@gmail.com, [2]Pvs1834@gmail.com

ABSTRACT

Recognizing of human face from image set has recently seen its prosperity because of its effectiveness in dealing with variations in illumination, expressions, or poses. Unlike most previous kinship verification methods which apply low-level hand-crafted descriptors such as local binary pattern and Gabor features for face representation. This paper investigates about kinship verification system based on mid-level features. Here Scale invariant feature transform (SIFT) descriptors have been extracted. For better performance two features such as blob and corners are detected via Surf and Harris algorithms. Further processed features are undergone for feature learning algorithm continued with classification. Finally, accuracy rate is to be evaluated for verification of kinfaces.

Index Terms— Feature Extraction, Feature Learning, Kinship Verification, Mid-Level features, Bio metrics.

I. INTRODUCTION

Facial image analysis has been widely investigated in the computer vision and multimedia computing community, and a large number of such methods have been proposed for various practical applications, such as face recognition, facial expression recognition, gender classification ethnic classification and human age estimation. While great successes have been made in these areas, there are a few attempts on automatic face analysis for kinship verification, possibly due to lacking of such publicly available databases and great challenges of this problem. Facial images convey many important human characteristics, such as identity, gender, expression, age, ethnicity and so on [3].

In this paper, a novel kinship verification system based on mid-level features was proposed. In most preceding kinship systems, the low-level feature descriptors such as local binary patterns (LBP), DAISY descriptors and Gabor features are applied for face representation. For better characterization of face image two features are extracted. To perform this, constructed a set of face samples with unlabeled kinship relation from the labeled face in the wild (LFW) datasets as the reference set. The set of face image pairs are processed for kinship verification. First the face pairs mid-level features are displayed after conversion process and filtering technique. Experimental results on four publicly available face kinship databases are presented to determine the ability of the proposed method. Lastly, the human ability in kinship verification and our experimental results show that this method is comparable to that of human observers.

The rest of the paper is organized as follows. Section II discusses related work. Section III describes Proposed system. Section IV provides the experimenttal results, and Section V concludes the paper.

II.RELATED WORK

In this section, the review of related kinfaces and their performances are explained.

Many metric learning algorithms have been proposed over the previous decade, and some of them have been successfully applied to address the issues of face verification in the wild. Discriminative Deep Metric Learning [1] and Discriminative Multimetric Learning [2] methods were learned a set of hierarchical nonlinear transformations to deal face pairs into one feature space in a deep architecture, where the nonlinear mappings are explicitly obtained. Existing deep learning methods can be mainly categorized three classes: unsupervised, supervised and semi-supervised, and they have been successfully applied to many visual analysis applications such as object recognition, human action recognition .The demerits was about the mean verification rate and standard error rate are at lesser for some local features compared to combined datasets, the current verification rate was lesser in social media mining .The experimental performance of DDML and DML methods were around 80%.First attempt to investigate kinship verification on largest kinship data-set in metric learning algorithms. There are two short-comings among most prevailing methods: 1) Some training samples are more descriptive in learning the distance metric than others, and most existing metric learning methods consider them equally and neglect potentially different contributions of the samples to gain the distance metric; 2) most existing metric learning methods only learn a distance metric from single view data and cannot handle multi-view data directly. Exploring more discriminative features and combined facial images were explained by Neighborhood Repulsed Metric Learning method. The important application is social media analysis, such as understanding the relationships of people in a photo. For this application, there are usually many face images in a photo and we need to determine the kinship relation from two face photos in the same photo. The accuracy rate of metric learned UB Kinface data base were 63-65% (NRML), 67.3% (MNRML) [3].

The pairwise kinship verification for analyzing whether the two persons are kin or non-kin via extracting binary patterns using logistic regression for 322 pairs. More distant kinship relationships such as grandparents, grandchildren, cousins, and uncle/aunt-nephew/niece can also be implied if the family picture also contains the "intermediate" people with these kinship types. So, a multi-class linear logistic regressor is trained. The results of 322 pairs was 69.3% [4]. Prototype hyperplanes for face verification in the wild using vector machine classifier obtained the benefits of reduction in dimensions. But the down sampling operation of filtered images degrade the face verification performance. The experimental performance of with unlabelled additional data: PHL+SILD (this work) 0.8867 ± 0.0070 FPLBP (PHL+SILD) 0.759 ± 0.015 0.825 0.244. The optimal prototype hyperplanes by maximizing the FLD like objective on inadequately labeled data set with a sparsity constraint in each SVM model, that selects only a sparse set of support vectors from the generic data set [5].

When working with image set classification, the particular concern is how to extract set information and then effectively represent it for classification. An image set is represented as an affine hull related with the number of image samples and their mean. Although not obviously using a prototype model, in this representation, the affine hull model is used to implicitly to build up a prototype to account for unseen face images. Similarly, each image set is characterized by a convex geometric region spanned by its feature points, and set dissimilarity is allotted by geometric distances between convex models. Average pooling strategy is adopted to summarize all the comparison for face recognition. Recognizing human face based on vector machine classification for prototype formation is a critical skill for category learning in face recognition. The HONDA/ UCSD datasets raised an accuracy rate of 97.14% [6]. The general setting involves learning a Mahalanobis distance metric based on an objective function prescribed over labeled similar (intra-class) and dissimilar (inter-class) samples. The methods differ in the objective function defined and the optimization scheme used for minimization. However, the challenges from the distinctive nature of our problem and the large dataset employed resulted in lower performance of some current verification and metric learning schemes. Verifying unknown parent –offspring and sibling pairs over unrelated subject pairs was examined for VANDANA kinface by extracting local binary patterns (LBP), Gabor features. Due to higher dimension rate the data results in lack of transparency. The accuracy rate of this facial images 67%,80.2% [7].

Moreover, recently a more significant problem draws considerable attentions that common assumption of training and test data from the same feature space along with distribution is not always fully reasonable. This is a natural situation for any new classification problems. Manual labeling work is time-consuming and people want to restate the knowledge that has already been widely studied. In such a case, knowledge transfer or transfer learning is highly desirable. The experimental result of UB kinface database was of 68.5%. The basic problem is how to reuse the knowledge learned from other data or feature spaces. Context texture with transfer learning using KNN method the kinface analysis was done [8]. To automatically classify kinship relation from facial images, they are facing two questions. First, what characteristics are crucial for kinship verification? To tackle this problem spatial pyramid learning-based (SPLE) feature descriptor for face representation and applied support vector machine (SVM) for kinship verification. For 400 pairs of kin facial images, the classifier accuracy rate was 65.75 % (Human-A) and 72% (Human-B). The demerits were that the raw pixel representation for face images are not a good choice for face analysis task because it is usually suffered from the illumination and expression variations [9].

A list of facial features that potentially encompass geological information passed down from parents to descendants are demonstrated here. They collected 150 pairs using this method, with variations in age, gender, race, career, etc., to cover the wide distribution of facial overview of the facial image databases. The main facial features in an image using a simplified pictorial structures model were identified, then compute these important features and combine them into a feature vector. Vector Machine methods to train the classifier on these difference vectors. For Cornell kinface the accuracy rate was about70.7%. The main drawback was similarity leads to classification errors, especially when there was only a small subset of features that are useful for classification [10].

III PROPOSED SYSTEM

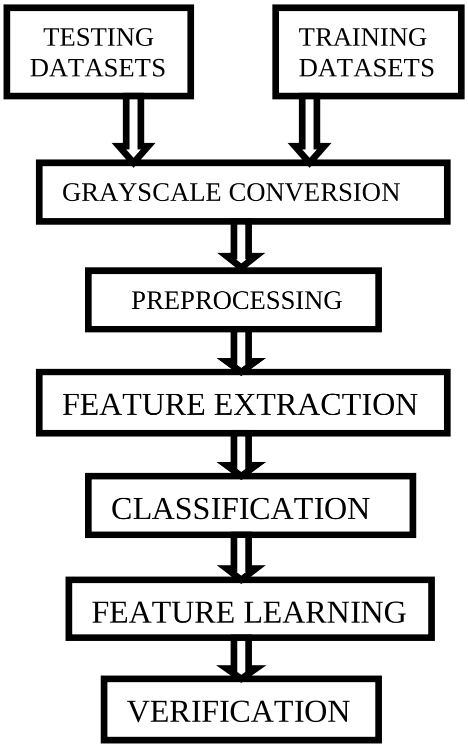


Fig1. Basic Methodology

The basic methodology of Kinship verification system is shown above blocks. Generally, the input data are divided into testing and training images. To perform pre-processing technique using certain algorithms, the datasets are converted to gray scale images from red blue green image. Then, the needed local features such as corners, blobs, uniform intensity, edges are extracted. This happens after the detection method, it is meant for detecting the eyes, nose, lips, face etc., The fifth most step in this technology is classification via KNN, SVM, Adaboost classifiers on aid to algorithms. Next is to learn the feature learning method such as, NRML, TSL, DDML DML etc., Finally the features are computed and calculated to get graphs based on pixels' rate to measure the total rates that is mean verification rate and equal error rate.

The propose system explains how the kinfaces are filtered, classified and feature learned with help of open computer vision methods. Firstly, the input face images are gray scaled by rgb to gray scale conversion. Then the pre-processing is performed with a filtering technique called Gaussian filtering technique. It's uses is to add and remove noise from the gray scaled images. Further, the local face features are extracted based on Scale Invariant Feature Transform (SIFT). The features that extracted are corners and blob based on HARRIS and SURF algorithms. Next the two features are classified with SVM classifier. After classification, the extracted features are learned via Feature Learning Algorithm. Finally, the data images are computed to get graphs based on mean verification rate and equal error rate.

A. Datasets: Two publicly available face kinship datasets, namely KinFaceW-I, KinFaceW-II were used for our evaluation. Facial images from all these datasets were collected from the internet online search. Both Kinface w1and Kinface w2 are of father-daughter (134 pairs) and father-son (156 pairs) is of mother – daughter (127 pairs) and mother-son (116 pairs). So, totally 533 pairs of face images.

B. Pre-Processing: Normally pre-processing technique is to apply for enhancing the incoming image suitable for feature extraction. The filtering, windowing techniques are used most. Here in this proposed system, the gaussian filtering method is preferred. The Gaussian blur is a type of image-blurring filter that uses a Gaussian function for calculating the transformation for implementing an each pixel in the image. The equation of a Gaussian function in two dimensions, it is the product of two equivalent Gaussians, one in each dimension:

$$G\,(x,y) = \frac{1}{2\pi\sigma^2}\, e\, \frac{x^2+y^2}{2x^2\sigma^2} \qquad (1)$$

where x is the distance from origin in the horizontal axis, y is the distance from origin in the vertical axis, and σ is the standard deviation of the Gaussian distribution.

Fig 2 Pre Processing

A Gaussian blur (also known as Gaussian smoothing) is the result of blurring an image by a Gaussian function. It is a extensively used effect in graphics software, typically to reduce image noise. Gaussian smoothing is also used as a pre-processing stage in computer vision algorithms to enhance image structures at different scales.

C. Feature Extraction:

In this work two different methods are approached. They are

1)SURF Algorithm
2)HARRIS Algorithm

The corners and blob features are extracted via these procedures.

D Local Features: Local features refer to a pattern or distinct structure found in an image, such as a point, edge, or small image patch. They are frequently associated with an image patch that alter from its immediate surroundings by texture, color, or intensity. What the feature actually represents does not matter, just that it is specific from its surroundings. Examples of local features are blobs, corners, and edge pixels.

E. Local Features Detection and Extraction: Local features and their captions, which are compressed vector representations of a local neighborhood, are the building blocks of numerous computer vision algorithms. Their applications include image registration, object classification, tracking, and motion estimation. Using local features implement these algorithms to better handle scale changes, rotation, and occlusion. The Computer Vision System provides the HARRIS method for detecting corner features, and the SURF for detecting blob features.

F. Scale Invariant Feature Transform: Scale-invariant feature transform (SIFT) is an algorithm in computer vision to detect and illustrate local features in images. The algorithm was published by David Lowe in 2000 [20].

Lowe's method for image feature generation transforms an image into a huge collection of feature vectors, each of that is invariant to image translation, scaling, and rotation, partially invariant to illumination changes along with robust to local geometric distortion. Key-point locations are characterized as maxima and minima of the result of difference of Gaussians function applied in scale space to a series of smoothed and resampled images. Low contrast candidate points and edge response points on an edge are discarded. Dominant orientations are assigned to localized key-points. These steps ensure that the key-points are more reliable for matching and recognition.

SIFT descriptors robust to local affine distortion are then received by considering pixels around a radius of the key location, blurring and resampling of local image orientation planes.

1) *Harris Algorithm:* Corner detection is an approach used within computer vision systems to extract certain kinds of features and infer the contents of an image. Corner detection is often used in motion detection, imageregistration, videotracking, imagemosaicing, panorama stitching, 3D modelling and object recognition. Corner detection extended along with the topic of interest point detection.

Harris corner detector is improved by considering the differential of the corner score with respect to direction directly, instead of using shuffled patches. (This corner score is often referred to as autocorrelation, as the term is used in the paper in which this detector is described. However, the mathematics in the paper clearly indicate that the sum of squared difference is used.

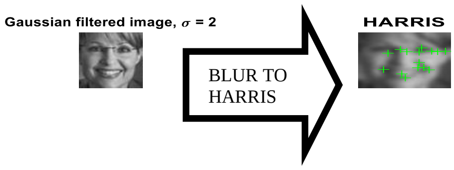


Fig 3 Corner Detection

Without loss of generality, assume a grayscale 2-dimensional image is used. Let this image be given by I . Consider taking an image patch over the area (u,v) and shifting it by (x,y). The weighted sum of squared differences (SSD) among these two patches, denoted S , is given by:

$$S(x,y)=\sum_u \sum_v w(u,v)\ I(u+x,v+y)-I(u,v)^2 \quad (2)$$

I(u+x,**v+y)** can be approximated by a Taylor expansion. Let I_x and Iy be the partial derivatives of I, such that

$$I(u+x,v+x)\approx I(u,v)+ I_x(u,v)x+Iy(u,v)y \quad (3)$$

This produces the approximation

$$S(x,y)\approx(x,y)\sum_u \sum_v w(u,v)(I(u,v)x-Iy(u,v)y)^2 (4)$$

which can be written in matrix form:

$$S(x,y)\approx(x\ \ y)\ A\begin{pmatrix}x\\y\end{pmatrix} \quad (5)$$

where *A* is the structure tensor,

$$A=\sum_u \sum_v w(u,v)\begin{pmatrix}Ix^2 & Ix\,Iy\\ Ix\,Iy & Iy^2\end{pmatrix}$$

$$=\begin{bmatrix}Ix^2 & IxIy\\ IxIy & Iy^2\end{bmatrix} \quad (6)$$

This matrix is a Harris matrix, and angle brackets stand for averaging (i.e. summation over (u,v)). If a circular window w(u,v) (or circularly weighted window, such as a Gaussian) is used, then the response will be isotropic via corner (or in general an interest point) is characterized by a large variation of 'S' in all directions of the vector (x,y). By analyzing the eigen values of A, this characterization can be expressed in the succeeding way: A should have two "large" eigen values for an interest point. Based on the magnitudes of the eigen values, the following inferences can be made established on this argument:

1. If $\lambda 1\approx 0$ and $\lambda 2\approx 0$ then this pixel (x,y)has no features of interest.
2. If $\lambda 1\approx 0$ and $\lambda 2$ has some large positive value, then an edge is found.
3. If $\lambda 1$ and $\lambda 2$ have large positive values, next a corner is found.

Harris and Stephens note that definite computation of the eigenvalues is computationally expensive, since it requires the computation of a square root, and instead suggest the following function Mc, where k is a tunable sensitivity parameter:

$$Mc=\lambda1\lambda2-k(\lambda1+\lambda2)^2$$
$$=\det(A)-k\ \mathrm{trace}^2\ (A) \quad (7)$$

Therefore, the algorithm does not have to actually calculate the eigenvalue decomposition of the matrix A and instead it is sufficient to evaluate the determinant and trace of A to find corners, or rather interest points in general. The value of k has to be determined empirically, and in the literature values in the range 0.04–0.15 have been reported as feasible.

The covariance matrix for the corner position is A^-1, i.e.

$$=\frac{1}{(Ix^2)(Iy^2)-(IxIy)^2}\begin{bmatrix}Iy^2 & -IxIy\\ -IxIy & Ix^2\end{bmatrix} \quad (8)$$

2)Surf Algorithm: In computer vision, Speeded Up Robust Features (SURF) is a local feature detector and descriptor. It that can be used for many tasks such as object recognition, image registration, classification or 3D reconstruction. It is partly inspired by the scale-invariant feature transform (SIFT) descriptor. The classic version of SURF is several times faster than SIFT and claimed by its authors to be more robust against different image transformations than SIFT.

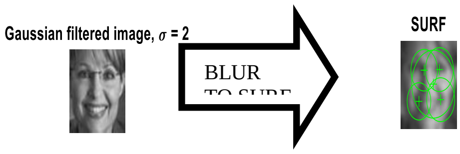


Fig 4 Blob Detection

SURF descriptors have been used to determine and recognize objects, people or faces, to reconstruct 3D scenes, to track objects and to extract points of interest. The SURF algorithm is based on the same principles and steps as SIFT; but explained in each step are different. The algorithm has three main parts: interest point detection, local neighborhood description and matching. SURF square-shaped filters as an approximation of Gaussian smoothing. (The SIFT approach is used to cascaded filters to identify scale-invariant characteristic points, where the difference of Gaussians (DoG) is computed on rescaled images progressively.) Filtering the image with a square is much faster if the integral image is used.

$$S(x,y)=\sum_{i=0}^{x}\sum_{j=0}^{y} I(i,j) \quad (9)$$

The sum of the original image within a rectangle can be estimated quickly using the integral image, involving evaluations at the rectangle's four corners. SURF uses a blob detector based on the Hessian matrix to find points of interest. The determinant of the Hessian matrix is used as a measure of limited change around the point and points are chosen where this determinant is maximal.

In contrast to the Hessian-Laplacian detector by Mikolajczyk and Schmid, SURF also uses the determinant of the Hessian for selecting the scale, as is also executed by Lindeberg. Given a point p= (x, y) in an image I, the Hessian matrix H(p, σ) at point p and scale σ, is:

$$H(p,\sigma) = \begin{bmatrix} Lxx(p,\sigma) & Lxy(p,\sigma) \\ Lyx(p,\sigma) & Lyy(p,\sigma) \end{bmatrix} \quad (10)$$

where $Lxx(p,\sigma)$ are the second-order derivatives of the grayscale image. The box filter of size 9×9 is an approximation of a Gaussian with σ=1.2 and exhibits the lowest level (highest spatial resolution) for blob-response maps.

IV EXPERIMENTAL RESULTS

A. Scale Invariant Feature Transform Features:

The results expains about how the face images of the kinfacew1 and kinface w2 are related biologically. This can be identified by simulating the face images using the matrix laboratory program. The extracted features are shown below:

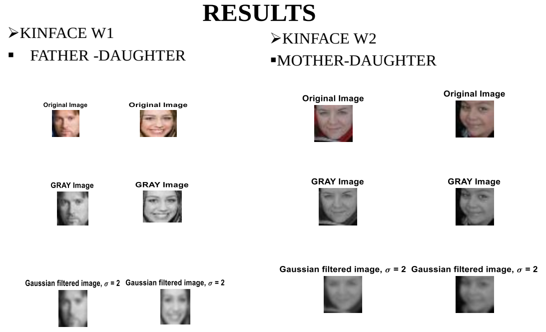
The features such as corners and blob are detected using certain harris and surf algorithms are figured.

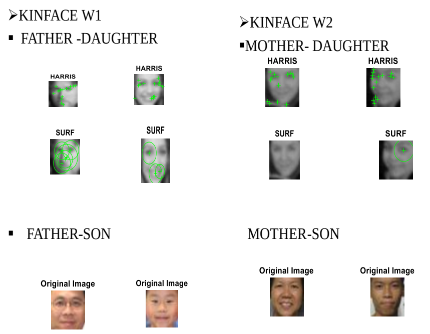
B. Kinface-W1 & Kinface-W1 Data Bases:

Kinface W1&W2 displayed are the father-daughter and father- son, mother-daughter and mother-son relations. Exhibiting of all pairs is tedious. So, among 134 pairs and 156 pairs ,127 pairs and 116 pairs, one pair of each kinface pair are viewed here.

Certain process for feature localization are, the input face images (father-daughter, father-son, mother-daughter, mother-son) are read. The original image is converted to gray scaled images via rgb to gray conversion. The obtained gray images are further pre-processed via gaussian filtered. The purpose of this filters provides a blurred image at rate of sigma value 2. Finally, the SIFT features (corners and blob) are displayed on the basis of HARRIS and SURF algorithms.

RESULTS





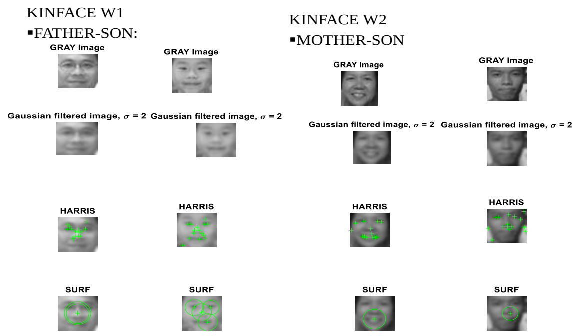


Table: Kinface Results

Sl No	Description	Measures
1	Kinfaces of Families	533 pairs
2	Feature Extraction	Corners and Blob
3	Software	Matlab
4	Rate of images read	7.3669e+05

V CONCLUSION

In this paper, we have extracted the mid-level features such as blob and corner via SIFT algorithm. The future work includes the classification of the extracted features by using SVM classifier and extracted features of the kinface system are to be learned via feature learning algorithm. The mean verification rate and equal error rate are computed based on the optimized results from the algorithm.

REFERENCES

1] J. Hu, J. Lu, and Y.-P. Tan, Discriminative deep metric learning for face verification in the wild, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog-nition, Columbus, OH, USA, Pp. 1875–1882 (2014).

2] Deng, and X. Zhou, Discriminative multimetric learning for kinship verification. IEEE Trans. Inf. Forensics Security 9(7): 1169–1178 (2014).

3] J. Lu, X. Zhou, Y.-P. Tan, Y. Shang, and J. Zhou, Neighborhood repulsed metric learning for kinship verification. IEEE Trans. Pattern Anal. Mach. Intell. 34 (2): 331–345 (2014).

4] D. H. Guo, H. Yuanhao, and L. van der Maaten, “Graph-based kinship recognition,” in Proc. Int. Conf. Pattern Recognit., Stockholm, Sweden, 2014, pp. 4287 –4292.

5] M. Kan, D. Xu, S. Shan, W. Li, and X. Chen, Learning prototype hyperplanes for face verification in the wild. IEEE Trans. Image Process 22(8): 3310–3316 (2013).

6] Mingbo Ma, Ming Shao, Xu Zhao and Yun Fu, Prototype Based Feature Learning for Face Image Set Classification. published in Automatic Face and Gesture Recognition (FG), IEEE International Conference and Workshops on year (2013).

7] G. Somanath and C. Kambhamettu, Can faces verify blood-relations? Proc. IEEE Int. Conf. Biometrics Theory Appl. Syst., Arlington, VA, USA, Pp. 105–112 (2012).

8] S. Xia, M.Shao, J. Luo, and Y. Fu, Understanding kin relationships in a photo. IEEE Trans. Multimedia 14(4): 1046–1056 (2012).
9] X. Zhou, J. Hu, J. Lu, Y. Shang and Y. Guan, Kinship verification from facial images under uncontrolled conditions, Proc. ACM Int. Conf. Multimedia, Scott-sdale AZ, USA Pp. 953–956 (2011).
10] R. Fang, K. Tang, N. Snavely, and T. Chen, Towards computational models of kinship verification, Proc. IEEE Int. Conf. Image Process., Hong Kong (2010).
11] D. Lowe, Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 60(2): 91–110 (2004).