International Journal of Pure and Applied Mathematics
Volume 119 No. 14 2018, 1243-1253
ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

SURVEY ON THE GOOD, THE BAD AND THE UGLY FACE RECOGNITION TECHNIQUES

J.Jayachitra[1], Dr.S.Matilda[2], Dr.K.Suganya Devi[3], Dr.P.Srinivasan[4]
[1]Department of IT,IFET College of Engg, Villupuram, India,
[2]Dept of CSE, IFET College of Engg, Villupuram, India
[3]Department of CSE, University College of Engg,Panruti, India
[4]Department of Physics, University College of Engg,Panruti, India

Abstract - A large number of face recognition algorithms have been developed in last decades. Over the past four decades, performance of Face Recognition on frontal faces in controlled environment has improved significantly; but frontal faces with uncontrolled environment and expression remains achallenge. The Good Bad Ugly(GBU) based Face Recognitiontechniques focuses on the attention of the fundamental problem of comparing single frontal face images across changes in appearance. This paper introduces many of the recent success and future challenges that are faced by the researchers in GBU based Face recognition.

Keywords -:Local Region Principle Component Analysis, Cohort Linear Discriminant Analysis, Partial Least Square Regression and Sparse Representation based Classification, Face Recognition.

1. INTRODUCTION

Face Recognition (FR) is a biometric method of identifying an individual by comparing live capture or digital image data with the stored record for that person. Facial recognition systems are commonly used for security purposes but are increasingly being used in a variety of other applications. The following four-stage process illustrates the way biometric systems operate:

1. **Capture -** a physical or behavioral sample is captured by the system during enrollment
2. **Extraction -** unique data is extracted from the sample and a template is created
3. **Comparison -** the template is then compared with a new sample
4. **Matching -** the system then decides if the features extracted from the new sample are matching or not

The four big Problems in FR are Aging,Pose,Illumination and Expression(APIE)[1].There have been several recent efforts to perform large scale independent evaluations of FR systems in the form of Grand challenge and

Vendor tests(Ex FRGC, FRVT).The recently introduced GBU builds on the success of these evaluations to encourage development of algorithms that are robust to variations in frontal face images. Partition is the most prominent feature of the GBU based face recognition.Partitioning the challenge by levels of difficulty is most prominent feature of the GBU problem design and other one is controlling for the “recognizability“ of people by selecting images of the people[1].

Several studies have considered the design of a quality metrics that may describe what makes an image easy or hard to recognize given a single subject .Till date, the primary insights from these works are that quality is inherent to image pairs rather than a single image. Now days, all the images taken from mobile studio environments and ambient illumination environments have been used in face recognition.

We can make three important observations in this field after surveying the research,viz., work that define quality measures for individual images ,works that measures the quality based upon properties of the compared image pair and works that investigate effects of various covariates on the matching performance.

Based on the survey, they were considered the various features that affect the face recognition. Some of them were classified the features as face specific and image specific characters and some other were treated as a covariates. The face-specific characters are illumination, expressions and poses. The image specific characters are space color, Sharpness, hue and saturation

Normally the covariates are classified into three major categories, viz., subject specific covariates, environment specific covariates and image specific covariates. The subject specific covariates are Person ID, Image ID, Gender, race, initial subject age, facial expression and glasses. The environment covariates arelighting, setting and location. The image specific covariates are eye location, resolution, tilt, edge density and focus [2]

The remaining section of this paper isorganizedasfollows.The related works are given in Section 2. The generation of GBU partitions is described in Section 3. The fundamental of algorithm is discussed in Section 4. The GBU based face recognition techniques details are presented in Section 5. Finally, the conclusion is provided in Section 6.

2. RELATED WORKS

P.Jonathon et al[1], introduces the GBU challenge problem. They constructed GBU partitions from the Notre Dame multibiometric data set used in the FRVT 2006. Among the four APIE factors, they considered only the illumination and expressions. In this work, they constructed a good partition which contains the most number of same lighting and same expression match pairs, followed by the bad and then the ugly partitions. Similarly, the Ugly partition contained the most number of different lighting, different expression match pairs, followed by the Bad and then the Good partitions. They developed ocular based LRPCA which provides best results when compared to LRPCA.They were only concentrated the face specific feature such as illumination and expressions. But they were omitted the image specific feature and the face specific feature such as aging and pose.

Yui Man Lui et al [3] characterized quality as an interaction between compared pairs of biometric samples (Multiple biometric Evaluation)while showing the theoretical equivalence between perfect matching and perfect quality analysis.They presented the Cohort Linear Discriminant Analysis FR algorithm and it has respectable performance on the GBU Challenge Problem. CohortLDA performs illumination compensation from color spaces and normalization from the training set.The CohortLDAprovides a new baseline on the GBU where performance can be compared with other algorithms.They observed the GBU evaluation protocol where the subject identities do not overlap between the training and the test sets.In particular kernel methods and local regions may be promising areas and it needs to be improved further.

In addition to the other factors discussed, GauravAggarwalet al[4],considered image specific characters (image sharpness,hue,saturation) and face specific characters (expressions) included into account for predicting the performance. The above said image specific factors are also contributed impact on the performance of face recognition.So Partial Least Squares (PLS) based regression is used to perform the prediction task. In all other GBU related works the image and facial specific characteristics are ignored or significantly reduced by the data collection and experimental protocol.

Pedro Tome et al [5],presented a general variability compensation scheme based on the Nuisance Attribute Projection (NAP) [6,7] that can be applied to compensate any kind of variability factors that affects the face recognition performance. They used the segmented datasets provided by Multiple Biometric Grand Challenge(MBGC) [8] compressed to 20KB with 120 pixels between the centers of the eyes. The faces were normalized following the ISO norm described in [9], from a size of 408×528 to size 168×192 pixels. They used the sparse representation for classification (SRC) as a base line algorithm.

Givens et al[2], completed the performance evaluation for seven different algorithms such as fusion of algorithms from face recognition vendor test 2006, cohort linear discriminant analysis, Gabor wavelets, kernel generalized discriminant analysis, local region Principal Component Analysis, Local Binary Pattern(LBP) and Elastic bunch graph matching with the GBU datasets. They fitted a generalized linear mixed model (GLMM) using the residual mean pseudo-likelihood estimation approach[10, 11]. They considered two sources of over dispersion modeled with random effects. First their models included a subject-specific random effect. Second , the location variables contributed random effects.Their results indicated (see table 1) that subject and location specific variation were both very important factors affecting algorithm [2].

Table 1 : Magnitude of random effects on log odds scale, expressed as the ratio of random effect standard error to mean log odds of verifications

Covariates	Ugly	Bad	Good
Subject	0.72	0.72	0.92
Location	0.59	0.60	0.15

The different GBU based face recognition techniques are discussed and the comparison is shown in the Table 2.

Table 2: comparison of GBU based FR techniques

Study	Scenario	Database	Technique	Remarks
P. Jonathon Phillips et al [1]	Still frontal image	FRVT	Ocular LRPCA & LRPCA	Face-specific characters such as illumination and expression are considered but image specific characters are not taken into consideration
Yui Man Lui et al [3]	Still frontal image	FRVT & FRGC	CohortLDA	Face-specific characters such as illumination and expression , image specific character like space color are considered but other image specific characters like Sharpness, hue, saturation are not bothered
Gaurav Aggarwal et al [4]	Still frontal image	FRVT	PLS-regression	Face-specific characters such as illumination is considered but image specific characters are not taken into consideration
Pedro Tome et al [5]	Still frontal image	MBGC	SRC & NAP	Face-specific characters such as illumination is considered but image specific characters are not taken into consideration
Givens et al [2]	Still frontal image	FRVT & FRGC	GLMM with Random Effects	The subject specific covariates, environment specific covariates and image specific covariates are considered

3 GENERATION OF GBU PARTITIONS

The data set is divided into three partitions called the good, the bad and the ugly. The Good partition consists of pairs of face images of the same person that are easy to match; the Bad partition contains pairs of face images of a person that have average matching difficulty; and the Ugly partition concentrates on difficult to match face pairs. The performance of GBU is based on fusing the result from three partitions

Each GBU partition contains two sets of images: a target set and a query set. For each partition,computes a similarity score between all pairs of images in that partition's target and query sets. A similarity score is a measure of the similarity between two faces. Higher similarity scores imply greater likelihood that the face images are of the same person. If an algorithm reports a distance measure, then a smaller distance measure implies greater likelihood that the face images are of the same person. Distances are converted to similarity scores by multiplying by negative one. The set of all similarity scores between a target and a query set is called a similarity matrix. A pair of face images of the same person is called a match pair, and a pair of face images of different people is called a non-match pair.

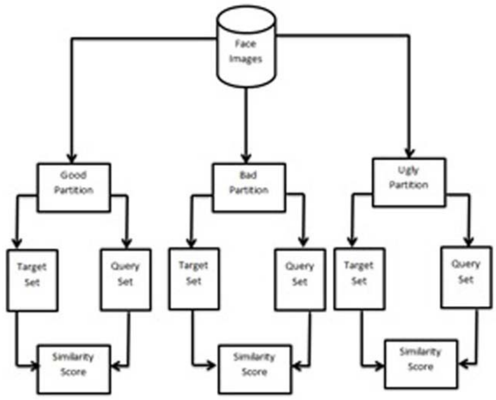


Figure 1: Generation of GBU

The below table 3 shows the comparison between face recognition with GBU and without GBU.

Table 3: comparison of face recognition with and without GBU

Face Recognition without GBU	Face Recognition with GBU
Data set consist of single image for a person	Data set consist of multiple image for a person
Data set is not divided into partitions	Data set is divided into three partitions such as good, Bad and Ugly
False Reject rate is high compared to GBU based face recognition	False reject Rate is low compared to normal face recognition

4 FUNDAMENTALS OF ALGORITHM

A face recognition algorithm is defined as a mapping from a pair of faces to a real number S representing some measure of the similarity score between those images. The first image is from a set of target images ‘t’and the second image is from a set of query images ‘q’. A similarity score S(t, q)is said to be a match score when the same person is pictured in images q and t and a non-match score when the pictures are of different people.

All face recognition algorithms first establishes a spatial correspondence between the faces from two images before they can proceed to measure the similarity. It is commonly done by first locating the eyes in both images. Then the image may be adjusted by positioning, scaling and rotating the face and then obtain a new smaller image called a face image chip. The eyes always fall at exactly the same position. Thisis called as localization.

In all FR algorithms, the two major and distinct tasks are identification and verification of faces. In an identification task, a query image q ∈ Q is compared to a set of target images in a gallery T. The gallery is then sorted by similarity with respect to q and either the most similar or a small set of the most similar target images are returned as a presumed match or match ranking for q.In a verification task, a person presents themselves to a system that already has a stored (target) image of face. The system acquires a new(query) image of the person and compares the resulting similarity S(t, q) to an acceptance threshold. The acceptance threshold is chosen using a training dataset where true identities are known If the similarity exceeds the threshold then the system confirms that the person claims identity; otherwise the person’s claimed identity is rejected. The threshold is usually set to achieve a pre-determined False Verification Rate(FVR). The rate of correct verification will depends on the extent to which the distribution of match and non-match scores overlap.

Most of the major face recognitionalgorithms carried out areconcentrated on the verification task. One reason is that outcome of a verification test depends only on the similarity score and the threshold for acceptance. In contrast, identification performance depends upon the other people and images in the gallery.

5 GBU BASED FACE RECOGNITION TECHNIQUES

5.1 Local Region Principle Component Analysis(LRPCA)

Local Region PCA (LRPCA) [1] is a refined implementation of the standard method based on Principle Component Analysis (PCA) face recognition algorithm, also known as Eigenfaces [12]. It first extracts a cropped and geometrically normalized face region from an original face image. The original image is assumed to be a still image whose pose of the face is close to frontal. The face region in the original is scaled, rotated and cropped to a specified size and the centers of the eyes are horizontally aligned and placed on standard pixel locations. In the baseline algorithm, the face chip is 128 by 128 pixels with the centers of the eyes spaced 64 pixels apart. The PCA representation is conducted on thirteen local regions cropped out of a normalized face image and the complete face chip. The local regions are centered relative to the average location of the eyes, eyebrows, nose and mouth. Figure 2 shows a cropped face and the thirteen local regions. All the 14 face regions are normalized to attenuate variation in illumination. First, self-quotient normalization is independently applied to each of the 14 regions [15]. The self-quotient normalization procedure first smooths each region by convolving it with a two-dimensional Gaussian kernel and then divides the original region by the smoothed region. The influence of illumination can be reduced by self-quotient normalization. A final normalization step adjusts the pixel values in each region to have a sample mean of zero and a sample standard deviation of one.

During training, PCA is computed for each of the 14 regions and the 3rd through 252th eigenvectors are retained to represent the face. As a result, a face is encoded by concatenating the 250 coefficients for each of the 14 regions into a new vector of length 3500. The representation is whitened by scaling each dimension to have a sample standard deviation of one on the training set. Then the weight on each dimension is further adjusted based on Fishers criterion. The Fishers criterion weight emphasizes the dimensions along which images of different people are spread apart according to the training set. During testing, coefficients computed from PCA projection of each of the 14 regions are concatenated as a vector. Each image corresponds to one vector. Similarity between pairs of faces is measured by computing the Pearsons correlation coefficient between pairs of these vectors.

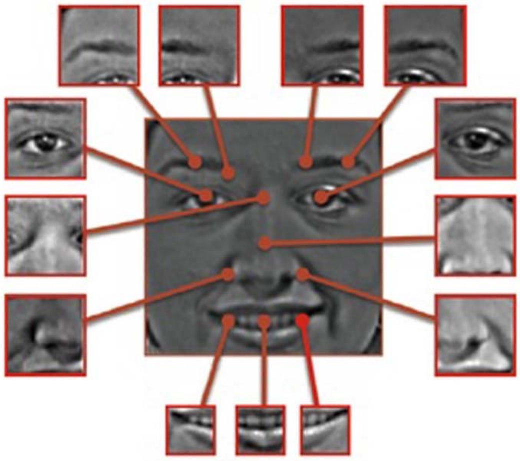


Figure 2: A cropped face and the thirteen local regions.

5.2 Cohort Linear Discriminant Analysis

Cohort Linear Discriminant Analysis (CohortLDA) [3] is a Linear Discriminant Analysis (LDA) algorithm with color spaces and cohort normalization. The main differences between CohortLDA and standard LDA are two-fold. One is the preprocessing step and the other is that CohortLDA introduces a cohort set to normalize the score. Specifically, CohortLDA uses both the R channel from RGB color space and the I channel from YIQ color space to conserve the structure of the face and reduce the influence of strong illumination. Since the red channel is similar to the gray-scale image, it usually does not work well with large lighting variation. Therefore, logarithm transformation and z-normalization are applied after extracting the R channel. During training, it seeks a projection that maximizes the ratio of between-class scatter and within-class scatter in order to make the data belonging to the same cluster more similar and the data belonging to different clusters more different. Figure 3 presents the LDA faces computed from the red channel and I chrominance.

In face verification, a decision threshold is needed to determine whether a pair of faces is a match or not. Since some images are harder than others, a fixed threshold may not adapt well from image to image. As a result, a set of images called the cohort set is adopted to adjust the match distance. During face verification, a subset of cohort images is selected

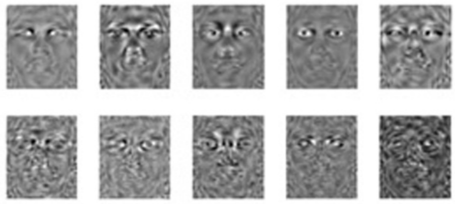


Figure 3: Top row: LDA faces acquired from the red channel after log and z-norm. Bottom row: LDA faces obtained from the I chrominance after z-norm.

for each query image and target image using the k-nearest neighbor rule. Then their average distance to the k-neighbors is computed as an indication of the difficulty of the query/target image. The distance of the query and target image is calculated as their original distance subtracted by their difficulty. The final distance is obtained based on the red channel and I chrominance images using a simple sum rule.

5.3 Partial Least Square Regression(PLS-Regression)

Partial Least Square regression (PLS-Regression) is a recent technique [16] that generalizes and combines features from Principal Component Analysis (PCA) and multiple regressions. It is particularly useful when we need to predict a set of dependent variables from a (very) large set of independent variables (i.e., predictors).

The I observations described by K dependent variables are stored in a IxK matrix denoted by Y, the values of J predictors collected on these I observations are collected in the $I \times J$ matrix denoted by X.PLS regression finds components from X that are also relevant for Y. Specifically, PLS regression searches for a set of components (called latent vectors) that performs a simultaneous decomposition of X and Y with the constraint that these components explain as much as possible of the covariance between X and Y. This step generalizes PCA. It is followed by a regression step where the decomposition of X is used to predict Y.

PLS regression decomposes both X and Y as a product of a common set of orthogonal factors and a set of specific loadings. So, the independent variables are decomposed as $X = TP^T$ with $T^TT = I$ with I being the identity matrix (some variations of the technique do not require T to have unit norms). By analogy with PCA T is called the score matrix, and P the loading matrix (in PLS regression the loadings are not orthogonal). Likewise, Y is estimated as $Y_1 = TBC^T$ where B is a diagonal matrix with the "regression weights" as diagonal elements (see below for more details on these weights). The columns of T are the latent vectors. When their number is equal to the rank of X, they perform an exact decomposition of X. Note, however, that they only estimate Y.

The latent vectors could be chosen in a lot of different ways. In fact in the previous formulation, any set of orthogonal vectors spanning the column space of X could be used to play the role of T. In order to specify T, additional conditions are required. For PLS regression this amounts to finding two sets of weights w and c in order to create a linear combination of the columns of X and Y(respectively) such that their covariance is maximum. Specifically, the goal is to obtain a first pair of vectors $t = Xw$ and $u = Yc$ with the constraints that $w^Tw = 1$, $t^Tt = 1$ and t^Tu be maximal. When the first latent vector is found, it is subtracted from both X and Y and the procedure is re-iterated until X becomes a null matrix

The properties of PLS regression can be analyzed from a sketch of the original algorithm. The first step is to create two matrices: E = X and F = Y. These matrices are then column centered and normalized (i.e., transformed into Z-scores). The sums of squares of these matrices are denoted SS_X and SS_Y. Before starting the iteration process, the vector u is initialized with random values. (in what follows the symbol ∝ means "to normalize the result of the operation").

Step 1.$w \propto E^Tu$ (estimate X weights).

Step 2.$t \propto Ew$ (estimate X factor scores).

Step 3.$c \propto F^Tt$ (estimate Y weights).

Step 4. $u = Fc$ (estimate Y scores).

If t has not converged, then go to Step 1, if t has converged, then compute the value of b which is used to predict Y from t as $b = t^Tu$, and compute the factor loadings for X as $p = E^Tt$. Now subtract (i.e., partial out) the effect of t from both E and F as follows $E = E-tp^T$ and $F = F-btc^T$. The vectors t, u, w, c, and p are then stored in the corresponding matrices, and the scalar b is stored as a diagonal element of B. The sum of squares of X (respectively Y) explained by the latent vector is computed as p^Tp (respectively b^2), and the proportion of variance explained is obtained by dividing the explained sum of squares by the corresponding total sum of squares (i.e., SS_X and SS_Y). If E is a null matrix, then the whole set of latent vectors has been found, otherwise the procedure can be re-iterated from Step 1 onwards

5.4 Sparse Representation based Classification (SRC)

A system based on sparse representation for classification purposes (SRC) [13, 14] has been adopted as base line algorithm. Essentially, this kind of systems span a face subspace using all known training face images, and for an unknown face image they try to reconstruct the image sparsely. The motivation of this model is that given sufficient training samples of each person, any new test sample for this same person will approximately lie in the linear span of the training samples associated with the person.

Once a new test image y is acquired, it can be represented using samples from the database by the linear equation Y=AX, where matrix A defines our training data and X represents the sparse solution.

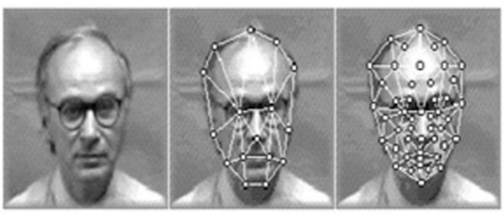
According to the assumption that images from a given subject are sufficient to represent themselves, the solution X in the linear equation=AX should be very sparse. This can be approximately recovered by solving the following noise-aware l1-minimization problem:

$$x_1 = \mathrm{argmin}||x||_2 \mathrm{subjecttoAx} - y \in \varepsilon$$

To recognize a probe test image, the SRC algorithm identifies the class by computing the minimum among the residuals reconstructed per class.

5.5 Elastic Bunch Graph Matching

In this technique(Elastic Bunch Graph Matching)[17], a data structure called a bunch graph [18] is used. It recognizes the faces in three stages,viz., face finding, landmark finding and recognition by comparison. The first two stages serve to create a scale invariant model of the face in an input image. In the last stage, face models are compared to achieve recognition. The first stage serves to find a face in an image and determine its size. This is accomplished by a set of matches to bunch graphs of appropriate pose and of three different sizes. The detailed schedule of this match is described in [19]. The best matching bunch graph determines the size and position of the face. They were next place a square frame around the face so that the face occupies about a quarter of the area of the frame. The resulting image is warped to a standard size (currently 128 x 128 pixels) and a new wavelet transform is computed, thus defining the image frame. The image frame is passed to the next module, the Landmark Finder.See Figure 4(b) for the graph placed over the facial image during face finding.The purpose of second stage is to find facial landmarks with high positional accuracy and reliability and to encode the information contained in the image as accurately as possible. This step is equally crucial since a node not correctly placed over its landmark will lead to distorted similarity values during the comparison stage.Figure 4(c) shows a typical result of this stage. This model graph represents all the information extracted from an image. For a face in frontal pose it contains 48 nodes, compared to the 16 nodes used during face finding.The graph comparison model produced as the result of the landmark finding step are compared pair wise to compute a similarity value. This value is computed as the sum of jet similarities between pairs of corresponding nodes divided by the number of pairs. A jet describes a small patch of grey values in an image around a given pixel. Since model graphs for different poses differ in structure, a little conversion table was used to identify correspondence between nodes referring to the same landmark. The result of the graph comparison step is a complete comparison score, containing for each of the face entries in the gallery provided by Army Research Laboratory (ARL), an ordered list of all other entries in descending order of similarity



(a) (b) (c)

Figure 4: Graph representation of a facial image. (a): Input image. (b): Face-finding graph. (c): Model graph as defined for landmark-finding. The image frame determined by face finding is used in each case

5.6 Face Description with Local Binary Patterns[LBP]

The original LBP operator, introduced by Ojala et al. [21], is a powerful means of texture description. The operator labels the pixels of an image by thresholding the 3x3-neighbourhood of each pixel with the center value and considering the result as a binary number. Then the histogram of the labels can be used as a texture descriptor. See Figure 5 for an illustration of the basic LBP operator[22]. Later the operator was extended to use neighborhoods of different sizes [23]. Using circular neighborhoods and bilinearly interpolating the pixel values allow any radius and number of pixels in the neighborhood.

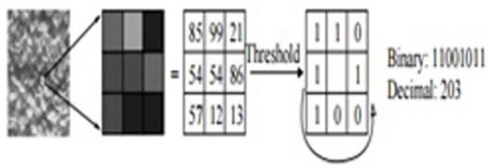


Figure 5: The basic LBP operator

For neighborhoods we will use the notation (P,R) which means P sampling points on a circle of radius of R. See Figure 2 for an example of the circular (8,2) neighborhood. Another extension to the original operator uses so called uniform patterns [24]. A Local Binary Pattern is called uniform if it contains at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered circular. Forexample,00000000,00011110and10000011areuniformpatterns.Ojalaet al. noticed that in their experiments with texture images, uniform patterns account for a bit less than 90 % of all patterns when using the (8,1) neighborhood and for around 70 % in the (16,2) neighborhood. They used the following notation for the LBP operator:$LBP^{u2}_{P,R}$. The subscript represents using the operator in a (P,R) neighborhood. Superscript u2 stands for using only uniform patterns and labeling all remaining patterns with a single label.

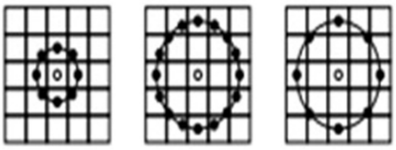


Figure 6: The circular (8,1), (16,2) and (8,2) neighborhoods. The pixel values are bilinearly interpolated whenever the sampling point is not in the center of a pixel.[20]

A histogram of the labeled image f(x,y) can be defined as

$$H_i = \sum_{x,y} I\{f(x,y) = i\}, i = 0, \dots n-1$$

in which n is the number of different labels produced by the LBP operator and

$$I(A) = \begin{cases} 1, \text{Aistrue} \\ 0, \text{Aisfalse} \end{cases}$$

This histogram contains information about the distribution of the local micro patterns, such as edges, spots and flat areas, over the whole image. For efficient face representation, one should retain also spatial information. For this purpose, the image is divided into regions R_0, R_1,...R_{m-1} and the spatially enhanced histogram is defined as

$$H_{i,j} = \sum_{x,y} I\{f(x,y) = i\} I\{(x,y) \in R_j\}, i = 0, \dots, n-1, j = 0, \dots., m-1$$

In this histogram, they effectively had a description of the face on three different levels of locality: the labels for the histogram contain information about the patterns on a pixel-level, the labels are summed over a small region to produce information on a regional level and the regional histograms are concatenated to build a global description of the face.

From the pattern classification point of view, an usual problem in face recognition is having a plethora of classes and only a few, possibly only one, training sample(s) per class. For this reason, more sophisticated classifiers are not needed but a nearest-neighbor classifier is used. Several possible dissimilarity measures have been proposed for histograms:

– Histogram intersection:

$$D(S,M) = \sum_{i,j} \min(S_{i,j}, M_{i,j})$$

– Log-likelihood statistic:

$$L(S,M) = -\sum_i S_{i,j} \log M_{i,j}$$

– Chi square statistic ($\chi 2$):

$$\chi 2(S,M) = \sum_{i,j} \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}}$$

In these equations S and M are the input and model images.All of these measures can be extended to the spatially enhanced histogram by simply summing over i and j. When the image has been divided into regions, it can be expected that some of the regions contain more useful information than others in terms of distinguishing between people. For example, eyes seem to be an important cue in human face recognition [25,26]. To take advantage of this, a weight can be set for each region based on the importance of the information it contains. For example, the weighted $\chi 2$ statistic becomes in which w_j is the weight for region j.

$$\boldsymbol{\chi^2_w(S,M) = \sum_{i,j} w_j \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}}}$$

5.7 kernel Generalized Discriminant analysis(GDA)

The classic LDA has been generalized to its kernel version, namely GDA [24]. Let $\emptyset : z \in R^N \to \emptyset(z) \in F$ be a nonlinear mapping from the input space to a high-dimensional feature space F, where different classes of objects are supposed to be linearly separable. The idea behind GDA is to perform a classic LDA in the feature space F instead of the input space R^N. Let S_b and $S_w$be the between- and within-class scatter matrices in the feature space , respectively, expressed as follows

$$S_b = \frac{1}{L} \sum_{i=1}^{C} (\bar{\emptyset}_ı - \bar{\emptyset})(\bar{\emptyset}_ı - \bar{\emptyset})^T$$

$$S_w = \frac{1}{L} \sum_{i=1}^{C} \sum_{j=1}^{C_i} (\emptyset_{ij} - \bar{\emptyset}_ı)(\emptyset_{ij} - \bar{\emptyset}_ı)^T$$

Where $\emptyset_{ij} = \emptyset(Z_{ij}), \bar{\emptyset}_ı = (\frac{1}{C_i}) \sum_{j=1}^{c_i} \emptyset(Z_{ij})$ is the mean of class Z_i . $\bar{\emptyset} = \frac{1}{L} \sum_{i=1}^{C} \sum_{j=1}^{C_i} \emptyset(Z_{ij})$ is the average of the ensemble and is the element number in , which leads to

$L = \sum_{i=1}^{c} C_i$. LDA determines a set of optimal discriminant basis vectors, denoted by $\{\varphi_k\}_{k=1}^{M}$, so that the ratio of the between- and within-class scatters is maximized [27]. Assuming $\varphi = \{\varphi_1, \dots \varphi_M\}$, the maximization can be achieved by solving the following eigenvalue problem:

$$\varphi = \operatorname{argmax}_{\varphi} \frac{|(\varphi^T S_b \varphi)|}{|(\varphi^T S_w \varphi)|}$$

The feature space could be considered as a "linearization space" [28], however, its dimensionality could be arbitrarily large and possibly infinite. Fortunately, the exact $\emptyset(z)$ is not needed and the feature space can become implicit by using kernel methods, where dot products in are replaced with a kernel function in the input R^N space so that the nonlinear mapping is performed implicitly in R^N [29].

6 CONCLUSION

The GBU based Face recognition is a challenging problem in the field of image processing and computer vision. Because of lots of application in different fields the face recognition has received great attention. In this paper, we are discussed various technique used for face recognition based on good, bad and ugly data sets.

Based on our review, we conclude that the consideration of ageing and pose variation, taking face still from videos, mobile environment images and real time face recognition may be scoped for the further researchers in GBU based face recognition.

REFERENCES

[1] P. J. Phillips, J. R. Beveridge, B. A. Draper, G. Givens, A. J.O'Toole, D. Bolme, J. Dunlop, Y. M. Lui, H. Sahizada, and S. Wiemer, "An introduction to the good, the bad, and the ugly challenge problem," in IEEE International Conference on Automatic Face and Gesture Recognition, 2011, pp. 346–353. 1, 2, 3

[2] G.H.Givns, J.R.Beveridge, P.J.Phillips, B.Draper, Y.M.Lui, D.Bolme,"Introduction to face recognition and evalution of algorithym performance",Elsevier, computational statistics and Data Analysis 67(2013), pg: 236-247

[3] Yui Man Lui, D Bolme, PJ Phillips, JR Beveridge, and BA Draper. Preliminary studies on the good, the bad, and the ugly face recognition challenge problem. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on, pages 9-16. IEEE, 2012.

[4] Gaurav Aggarwal; Soma Biswas; Patrick J. Flynn; Kevin W. Bowyer, "Predicting good, bad and ugly match Pairs", Applications of Computer Vision (WACV), 2012 IEEE Workshop on Year: 2012

[5] Pedro Tome, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-García "Variability Compensation Using NAP for Unconstrained Face Recognition" Distributed Computing and Artificial Intelligence, AISC 151, pp. 129–139, Springer-Verlag Berlin Heidelberg 2012

[6] Solomonoff, A., Campbell, W.M., Quillen, C.: Nuisance attribute projection. In: Speech Communication. Elsevier Science BV, Amsterdam (2007)

[7] Solomonoff, A., Quillen, C., Campbell, W.M.: Channel compensation for svm speaker recognition. In: Proceedings on Odyssey: The Speaker and Language RecognitionWorkshop, Toledo, Spain, pp. 41–44 (2004)

[8] Phillips, P.J., Flynn, P.J., Beveridge, J.R., Scruggs, W.T., O'Toole, A.J., Bolme, D., Bowyer, K.W., Draper, B.A., Givens, G.H., Lui, Y.M., Sahibzada, H., Scallan III, J.A., Weimer, S.: Overview of the Multiple Biometrics Grand Challenge. In: Tistarelli, M.,Nixon, M.S. (eds.) ICB 2009. LNCS, vol. 5558, pp. 705–714. Springer, Heidelberg (2009).

[9] ISO/IEC JTC 1/SC 37 N 504. Biometric data interchange formats part 5: Face image (2004)

[10] Wolfinger, R.,O'Cornell,M., "Generalized linear models: a pseudo-likelihood approach. Journal of Statistical computation and simulation 48, 233-243.

[11] Breslow, N.E., Clayton, D.G., 1993, Appoximate interface in generalized linear mixed models, Journal of the American Statistical Association 8, 9-25

[12] Matthew Turk and Alex Pentland. Eigenfaces for recognition. Journal of cognitive neuroscience, 3(1):71-86, 1991.

[13] Haitao Wang, Stan Z Li, Yangsheng Wang, and Jianjun Zhang. Self quotient image for face recognition. In Image Processing, 2004. ICIP'04. 2004 International Conference on, volume 2, pages 1397-1400. IEEE, 2004

[14] H. Abdi, "Partial least squares regression (PLS-regression),"in Encyclopedia for Research Methods for the Social Sciences, M. L. Beck, A. Bryman, and T. Futing, Eds., 2003, pp. 792–795. 2, 7

[15] Huang, K., Aviyente, S.: Sparse representation for signal classification, pp. 609–616 (2006)

[16] Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. 31(2), 210–227 (2009)

[17] P.J. Phillips, W.T. Scruggs, A.J. O'Toole, P.J. Flynn, K.W. Bowyer, C.L. Schott, M.Sharpe, FRVT 2006 and ICE 2006 large-scale results, IEEE Trans. Pattern Anal, Mach. Intell. 32 (2010) 831–846

[18] The Bochum/USC Face Recognition System and How it Fared in the FERET Phase III Test Kazunori Okada1, Johannes Steffens1, 2,3, Thomas Maurer2, Hai Hongl, Egor Eiagin1,3, Hartmut Neven1,3, and Christoph von der Maisburg1,2 Face Recognition © Springer-Verlag Berlin Heidelberg 1998

[19] Wiskott, L., Fellous, J. M., Kruger, N., and von der Malsburg, C. (1995): Face Recognition and Gender Determination. Proc. IntI. Workshop on Automatic Face- and Gesture- Recognition, Ziirich, 92-97

[20] Wiskott, L., Fellous, J. M., Kruger, N., and von <;ier Malsburg, C. (1997): Face Recognition by Elastic Bunch Graph Matching. IEEE Trans. on Pattern Recognition and Machine Intelligence, 19:775-779

[21] Ahonen, T., Hadid, A., Pietikainen, M., 2006. Face description with local binary patterns: application to face recognition. IEEE Transactions on Pattern AnalysisandMachineIntelligence28(12),2037–2041.

[22] Ojala, T., Pietik¨ainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition 29 (1996) 51–59

[23] Ojala, T., Pietik¨ainen, M., M¨aenp¨aa¨, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2002) 971–987

[24] Zhao, W., Chellappa, R., Rosenfeld, A., Phillips, P.J.: Face recognition: a literature survey. Technical Report CAR-TR-948, Center for Automation Research, University of Maryland (2002)

[25] Gong, S., McKenna, S.J., Psarrou, A.: Dynamic Vision, From Images to Face Recognition. Imperial College Press, London (2000)

[26] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," Neural Comput., vol. 12, pp. 2385–2404, 2000.

[27] R. A. Fisher, "The use of multiple measures in taxonomic problems," Ann. Eugenics, vol. 7, pp. 179–188, 1936.

[28] M. A. Aizerman, E. M. Braverman, and L. I. Rozonoér, "Theoretical foundations of the potential function method in pattern recognition learning," Automat. Remote Contr., vol. 25, pp. 821–837, 1964.

[29] [27] B. Scholkopf, S. Mika, C. J. C. Burges, P. Knirsch, K.-R. Muller, G. Ratsch, and A. J. Smola, "Input space versus feature space in kernelbased methods," IEEE Trans. Neural Networks, vol. 10, pp. 1000–1017, Sept. 1999.