Huffman Code Function and Mahalanobis Distance-base Face Recognition

Babatunde R. S.
Kwara State University
Malete, Nigeria
Department of Computer Science

Ajao. J. F.
Kwara State University
Malete, Nigeria
Department of Computer Science

Balogun B. F.
Kwara State University
Malete, Nigeria
Department of Computer Science

ABSTRACT
Human facial appearance is affected by lots of environmental and personal factors. The human face is a very challenging pattern to recognize because of its rigid anatomy. The main problem of face recognition is large variability of the recorded images due to pose, illumination conditions, facial expression, cosmetics, different hair styles and presence of glasses amongst others. Another major issue is the ability to project facial faces into a low sub space due to the non-linear manifold nature of face, resulting in high features, affecting the automated recognition of face. Due to the aforementioned problems this research develops an improved face recognition system using Huffman Encoding method for selecting optimal features from the high dimensional face image. Recognition of faces was carried out by obtaining the Mahalanobis distance of the test image and all the training images. The experimental results obtained showed that the method employed gave comparable recognition accuracy to existing literature.

General Terms
Pattern Recognition

Keywords
Huffman code, Mahalanobis distance, high dimension and face recognition

1. INTRODUCTION
Face recognition has been considered as an important subject of research work over the last fifteen years. This subject has gained as much importance as the areas of image analysis, pattern recognition and more precisely biometrics (Delac, 2004), because it has become one of the identification methods to be used in e-passports and identification of candidates appearing in various national and international academic examinations.

In the past few years, face recognition has received a significant attention and regarded as one of the most successful applications in the field of image analysis (Zhao, Chellappa, Phillips and Rosenfeld, 2003)). The human faces represent complex, multidimensional, meaningful visual stimulant. Developing a computational model for face recognition is difficult.

The resolution or the size of the image plays an important role in the face recognition. However, the image compression effects on the face recognition system are not given as much importance as is given to other aspects of image analysis (Vijay and Nagabhushana, 2013).

Furthermore, the human face is not a unique, rigid object. Indeed, there are numerous factors that cause the appearance of the face to vary. The sources of variation in the facial appearance can be categorized into two groups: intrinsic factors and extrinsic ones. (Srinivas and Bibhas, 2009)

i) Intrinsic factors are due purely to the physical nature of the face and are independent of the observer. These factors can be further divided into two classes: intrapersonal and interpersonal (Delac, 2004). Intrapersonal factors are responsible for varying the facial appearance of the same person, some examples being age, facial expression and facial paraphernalia (facial hair, glasses, cosmetics, etc.). Interpersonal factors, however, are responsible for the differences in the facial appearance of different people, some examples being ethnicity and gender (Vijay and Nagabhushana, 2013).

ii) Extrinsic factors cause the appearance of the face to alter via the interaction of light with the face and the observer. These factors include illumination, pose, scale and imaging parameters (e.g. resolution, focus, imaging, noise, etc.).

Images are compressed for different reasons like storing the images in a small memory like mobile devices or low capacity devices, for transmitting the large data over network, or storing large number of images in databases for experimentation or research purpose. This is essential due to the reason that compressed images occupy less memory space or it can be transmitted faster due its small size (Georgia, Stefanos, Maja and Lijun, 2012)

Due to this reason, the effects of image compression on face recognition started gaining importance and have become one of the important areas of research work in other biometric approaches as well like iris recognition and fingerprint recognition. In addition to paying importance to standard compression methods in recognition, researchers have focused in developing special purpose compression algorithms, e.g. a recent low bit-rate compression of face images (Bryt and Elad, 2008).

Huffman coding is a popular method for compressing data with variable-length codes. Huffman constructs a code tree from the bottom up (and the bits of each codeword are constructed from right to left). Huffman codes have been widely used for source coding and have shown high efficiency in exploiting the source redundancy (Huffman, 1952). Huffman codes along with run-length codes have been widely used in most international multimedia standards (e.g. MPEG and ISO standards).Huffman decoding can be implemented with a lookup-table (LUT) or multiple lookup-tables (Anissa, 2011).

Mahalanobis distance is a distance measure based on correlations between variables by which different patterns can be identified and analyzed. It is a useful way of determining similarity of an unknown sample set to a known one. It differs from Euclidean distance in that it takes into account the correlations of the data set and is scale-invariant, i.e. not dependent on the scale of measurements.

2. RELATED WORK

Sahoolizadeh (2008) developed a hybrid approach to face recognition based on combined Gabor wavelet with ANN classifier. The Gabor wavelet was used to represent face image. The representation of face images using Gabor wavelets was effective for facial action recognition and face identification. The result obtained shows good performance of 93% recognition rate on ORL data set.

Srinivas and Bibhas (2009) presented an approach, which can efficiently extract the facial landmarks like eyes, nose and lips from a face image. Statistical and geometrical features extracted from these facial features are used as a descriptor for a face. The results of experiments indicated that 90% recognition rate was recorded when the method was tested of ORL dataset.

Jun, Jing and Xing (2015) developed a novel random facial variation modeling system for sparse representation in face recognition. The work challenged the single-sample face recognition problem with intra-class differences of variation in a facial image model based on random projection and sparse representation. The facial variation modeling systems developed composed only of various facial variations. A novel facial random noise dictionary learning method that is invariant to different faces was employed. The experimental results on the AR, Yale B, Extended Yale B, MIT and FEI databases validate that the method leads to substantial improvements, particularly in single-sample face recognition problems.

Vijay and Nagabhushana (2013) examined the effects of feature selection and feature normalization on the face recognition scheme. From the local features that are extracted using block-base discrete cosine transform (DCT), three feature sets were derived. These local feature vectors were normalized in two different ways by making them unit norm and by dividing each coefficient by its standard deviation that is learned from the training set. This was done to mitigate the effects of expression, illumination and occlusion variations by performing local analysis and by fusing the outputs of extracted local features at the decision level. It was discovered that the DCT out performs the Karhunen-Loeve Transform (KLT).

Yongjiao, Chuan and Lei (2015) applied sparse representation theory on face recognition. Their intention was to find the best sparse coefficients from a variety of combinations of solutions and obtain the minimum value. The effectiveness of the algorithm was evaluated by comparing face recognition based sparse representation (SR) with the common methods such as nearest neighbor (NN), linear support vector machine (SVM) and nearest subspace (NS). Experimental results show that sparse representation method obtains better performance than the other methods.

Facial expression recognition using PCA (Principal component Analysis with SVD (Singular Value Decomposition) is superior to PCA in terms of recognition rate as reported in (Rajeev, 2011). The universally accepted five principal emotions considered are: angry, happy, sad, disgust and surprise along with neutral. SVD provides a robust method of storing large images as smaller and manageable images using matrix of the image.

3. METHODOLOGY

The developed system comprises of five stages: Data acquisition, image preprocessing, feature extraction and classification. Detailed of this is presented with the flow diagram in Figure 1.

3.1 Data Gathering

Face images of selected students of Computer Science Department of Kwara State University were captured. The subjects were made to appear in different head pose so as to capture the faces with high degree of within class variation. Samples of these are shown in Figure 3.1 These faces were used to construct a database for the system.

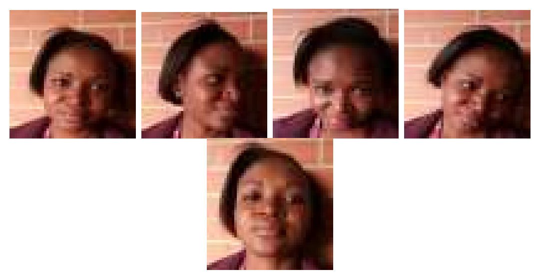

Figure 3.1. Five samples of a subject.

3.2 Image Cropping and Resizing

Face images were cropped using Adobe Photoshop. The cropping was essential so as to concentrate on the facial features alone in the entire images and ignore hair, ear and forehead region. Each subject of all the class was cropped to a localized view of the facial discriminant areas (eye, nose, mouth, and chin). To ensure uniformity of image sizes, the entire cropped image will formulate the input database that will be passed into Matlab for resizing using the image resize command in Matlab environment. The Facial Images were resized arbitrarily to the following dimensions: 70 X 70 and 80 X 80 dimensions.

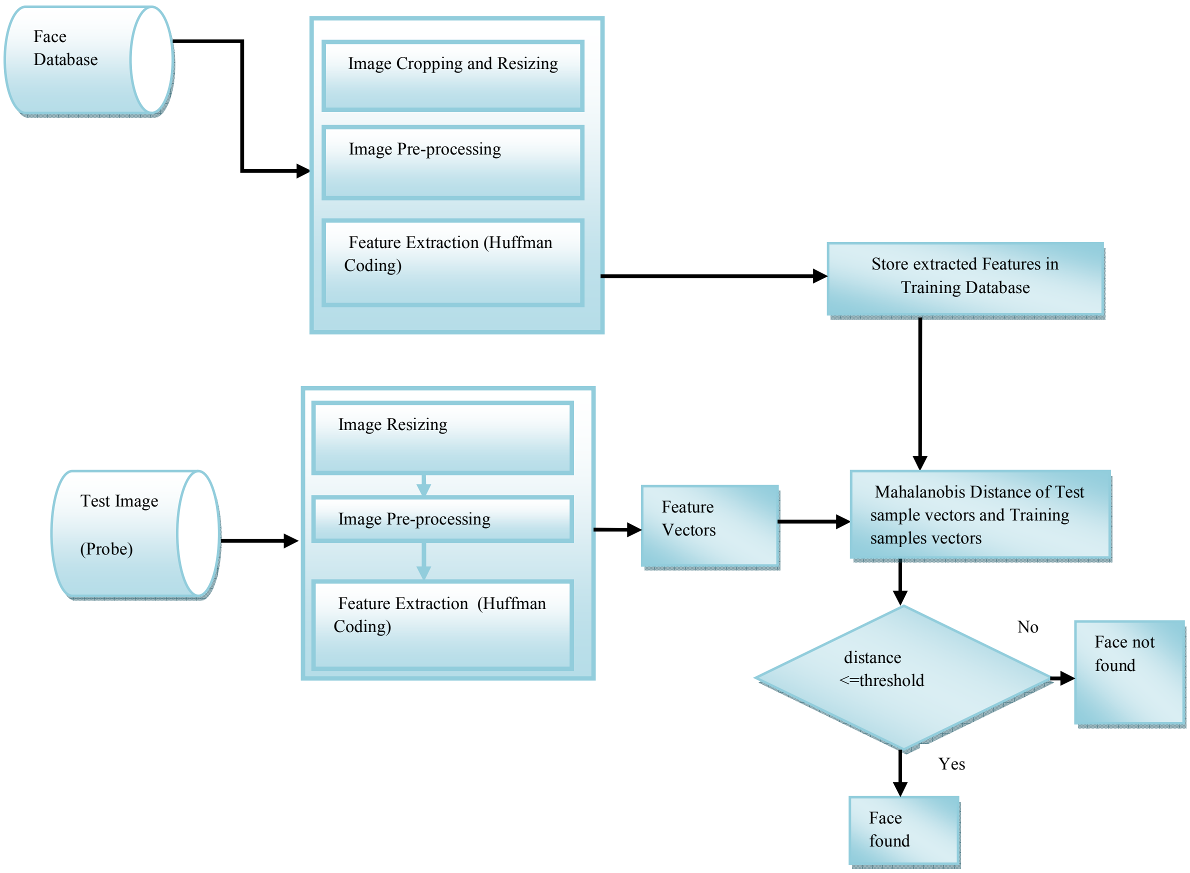

Figure 1: The flow diagram for the developed system

3.3 Image Preprocessing

The cropped and resized faces were converted to gray scale leveling and were further preprocessed using histogram equalization technique for illumination compensation and contrast improvement. Equalization implies mapping one distribution (the given histogram) to another distribution of wider and more uniform distribution of intensity values so the intensity values are spread over the whole range.

3.4 Feature Extraction using Huffman Coding

Feature extraction is a process used to obtain discriminant features from the face images of each subject so as to obtain minimal and most importantly salient information from the face. Feature extraction involves transformation of features leading to a lower dimensional feature space. In this research, Huffman code was used for feature extraction. Huffman code was used for compressing the faces and only code words of the images were considered to create discriminant features of faces

Huffman coding is an efficient source-coding algorithm for source symbols that are not equally probable such as face pixel values. The source symbol probabilities are P (xi), i=1, 2… L. The average number of bits required to represent the source symbol is minimum provided the prefix condition is met i.e. item 5. The steps of Huffman coding algorithm are given thus:

1. Arrange the source symbols in increasing order of their probabilities.
2. Take the bottom two symbols & tie them together. Add the probabilities of the two symbols & write it on the combined node. Label the two branches with a ‘1‘ & a ‘0‘
3. Treat this sum of probabilities as a new probability associated with a new symbol. Again pick the two smallest probabilities tie them together to form a new probability. Each time we perform the combination of two symbols we reduce the total number of symbols by one. Whenever we tie together two probabilities (nodes) we label the two branches with a ‘0‘ & ‘1‘.
4. Continue the procedure until only one procedure is left. Note that it should be one to indicate correctness of addition. This completes construction of the Huffman tree and the result indicates the extracted features.
5. To find out the prefix code word for any symbol, follow the branches from the final node back to the symbol. While tracing back the route, read out the labels on the branches. This is the codeword for the symbol.

The features obtained from applying the above algorithm to the face image at the feature extraction stage create an encoding value for each image of each class that will enhance the facial recognition process.

The feature extraction process uses a binary method for extracting salient information from the preprocessed and normalized face images so as to generate reduce features of the image. The Huffman encoding takes a "greedy" approach to obtain reduced pixels by converting the feature (pixels) from each face image into a character encoding from which a dictionary was constructed in conjunction with the probabilities with which each encoded character (pixel) occurs. The Huffman dictionary code serve as the extracted feature vector of each facial image as shown in Figure 3.2. This results to the bit/pixel binary encoding which is shown in the Figure 3.3.

Figure 3.2. Huffman Encoding

0000000 1111111 0000000 111111 0000000

Figure 3.3. Sample Encoding Pattern

3.5 Training

The database was divided into Training and Testing sets. Four images per subject were selected randomly to construct the training dataset while two images per subject constitute the testing set. The training dataset was subjected to the Huffman algorithm to obtain the pattern of the faces as well as optimized feature vectors. The covariances of the vectors were obtained which is to be used by the Mahalanobis distance metric for testing of the probe face.

3.6 Recognition using Mahalanobis Distance

The Mahalanobis distance is a very useful way of determining the similarity between a set of values from an unknown sample to a set of values measured from a collection of "known" samples. One of the main reasons the Mahalanobis distance method was used is that it is very sensitive to inter-variable changes in the training data. In addition, since the Mahalanobis distance is measured in terms of standard deviations from the mean of the training samples, the reported matching values is expected to give a statistical measure of how well the spectrum of the unknown sample matches (or does not match) the original training spectra. The Mahalanobis distance between a test sample x and training images μ is given by:

$$\Delta_{ik}^2 = (x_i - \mu_k)^T \Sigma^{-1} (x_i - \mu_k), \qquad (1)$$

Where μ_k are the mean and x_i is the input vector of attributes, $\sum$ is the covariance matrix given by

$$\sum = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{1L} \\ \sigma_{21} & \sigma_{22} & \sigma_{2L} \\ \sigma_{L1} & \sigma_{L2} & \sigma_{LL} \end{bmatrix} \qquad (2)$$

and the individual covariance values of $\sum$ are computed from the outer product sum given by

$$\Sigma = \frac{1}{N} \sum_{i=1}^{n} (xi - \mu j)(xi - \mu j)T \qquad (3)$$

Thus Mahalanobis distance can be seen as the generalization of Euclidean distance, and can be computed for each cluster if the covariances of the cluster are known (Brian, 2009).

Mahalanobis distance was computed between a probe image (test image) and all the face images available in the database. The Mahalanobis metrics computes the Mahalanobis distance (in squared units) of each probe in test database from the reference point of all the training samples. A threshold value was set heuristically to obtain matching value between the testing and training images. The system was evaluated based on the training time, testing time and the classifier accuracy.

The face recognition system was designed and implemented in MATLAB is a high-performance language for technical computing created by The MathWorks in 1984. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. Typical uses include: Algorithm development, Data acquisition, Modeling, simulation, and prototyping. The experimental setup was carried out using Matlab programming language (MATLAB 2015A). Various functions were developed and linked to a graphic user interface for user interactivity and responsiveness.

4. RESULTS AND DISCUSSION

A total number of one hundred and twenty 120 face images were used in this experiment to test the effectiveness and compatibility of the developed system. Table 4.1 shows the information of the database used to evaluate the system.

Table 4.1: Database information

No of Samples/Individual	**6**
No. of Individual	20
Total No. of samples	120
Total No of samples/Training set	80
Total No of samples/Testing set	45

It was observed from the experiment that the varying dimension of the images had a little effect on the recognition accuracy but an increase in training time was discovered when the face resolution of 80x80 was used compared to 70x70 .

Face images that were not used to train, i.e. whose class sample was not represented in the training database (knowledge base) are not supposed to have an equivalent match, indicating True Negative.

The training time obtained for the varying face resolution is shown in Table 4.2. This is the time taken to adapt the face images in the training set to the developed system. It was observed that the 70 X 70 resolution had a lower training time than the 80 X 80 resolution. This is due to the variance in the image pixel area allocation. The total number of images trained are 80.

Table 4.2 Training time result

Resolution	Total training time	Average training time
70 X 70	109.66 secs	1.370 secs
80 X 80	122.0 secs	1.525 secs

The recognition time obtained for the varying resolution is shown in Table 4.2. This is the time taken to recognize a probe image by the developed system. It was observed that the 80 X 80 resolution gave the overall best timing model, however at large it can be said that the developed system gave a fast recognition time for both 70 x 70 and 80 x 80 resolutions. 40 number of faces were used to test the developed system.

Table 4.3: Recognition time

Resolution	Total testing time	Average testing time
70 X 70	7.862secs	0.197secs
80 x 80	7.608 secs	0.191secs

To obtain the recognition rate the mathematical expression below was employed:

$$Recognition\ Rate = \frac{total\ number\ of\ faces\ correctly\ recognized}{total\ number\ of\ faces\ tested} \text{x}100\%$$

The recognition rate obtained for both resolutions is shown in Table 4.4 for tested face images.

Table 4.4: Recognition rate

	70 x 70 resolution	**80 x 80 resolution**
Faces recognized	38	37
Faces not recognized	2	3
Recognition rate	95%	92.5%

The developed system obtained good recognition rate, though it varies slightly for the resolutions used as seen in Table 4.4 above

5. CONCLUSION

The research work has implemented a face recognition system by using Huffman encoding which is an efficient binary-coding algorithm for source symbols that was used to represent individual faces. Often, its operation can be thought of as revealing the structure of the data in a way which best explains the variance in the source symbols in increasing order of heir probabilities. By implementing the Huffman Encoding for feature extraction support the reduction of number of features, minimizes the computational complexity and yielded the better recognition rates. The algorithm has been tested on locally acquired face database captured at different head pose. The algorithm delivers quite good results for dimensionality reduction. The novelty of the system is the combination of Huffman code algorithm for image feature extraction with Mahalanobis Distance classifier. The recognition performance improved with the distance metric which was used to compute a distance from a reference point of an image to all points of all the stored images. The experimental result obtained is an indicative of a system of high performance. We intend to improve on the research in the future by testing with a larger dataset, with the inclusion of identical twins. We can also extend it to Pose Invariant Face Recognition Method.

6. ACKNOWLEDGMENTS

Our thanks to the students of Kwara State University, Malete. Nigeria, for their cooperation in the processes of data capture for this research

7. REFERENCES

[1] Sahoolizadeh, A.H., Heidari, B.Z. and Dehghani, C. (2008) A New Face Recognition Method Using PCA, LDA and Neural Network. Proceedings of World Academy of Science En-gineering and Technology, Vol. 31, pp7-12.

[2] Srinivas Nagamalla and Bibhas Chandra Dhara (2009):A Novel Face Recognition Method using Facial Landmarks 2009 Seventh International Conference on Advances in Pattern Recognition. IEEE. DOI 10.1109/ICAPR.2009.59. pp.444-448

[3] Jun Cai, Jing Chen and Xing Liang (2015): "Single-Sample Face Recognition Based on Intra-Class Differences in a Variation Model". Sensors 2015, 15, 1071-1087; doi:10.3390/s150101071

[4] Zhao, W.; Chellappa, R.; Phillips, P.J.; Rosenfeld (2003): "A. Face recognition: A literature survey". ACM Comput. Surv., Vol. 35, pp34–38.

[5] Vijay B. A. and Nagabhushana Rao. (2013): "Analysis Of Face Recognition- A Case Study On Feature Selection And Feature Normalization. International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 3 No. 3 pp.971-979.

[6] Yongjiao Wang, Chuan Wang and Lei Liang. (2015): " Sparse Representation Theory and its Application for Face Recognition". International journal on smart sensing and intelligent systems Vol. 8, No. 1 pp107-124.

[7] Delac Kresimir, Grgic Mislav, Grgic Sonja. (2006). "Independent Comparative Study of PCA, ICA and LDA on the FERET Data Set". Wiley Periodicals, Inc. Vol. 15 No 5 pp.252-260.

[8] Georgia Sandbach, Stefanos Zafeiriou, Maja Pantic and Lijun Yin. (2012):`Static and Dynamic 3D Facial Expression Recognition: A comprehensive survey', Image and Vision Computing Special issue on 3D Facial Behavior Analysis, vol. 30, no. 10 pp. 683 - 697,

[9] Bryt O. and Elad M. (2008): "Compression of facial images using the k-svd algorithm," Journal of Visual Communication and Image Representation, vol. 19, no. 4, pp. 270–282.

[10] Anissa Bouzalmat, et al (2011) "Face Detection and Recognition Using Back Propagation Neural Network And Fourier Gabor Filters", Signal & Image Processing: An International Journal (SIPIJ) Vol. 2, No. 3, DOI : 10.5121/sipij.2011.2302 15