**Research on a Face recognition system using image processing**

Uditanshu Khatri

Image Processing

New Jersey Institute of Technology

Newark, NJ 07102

**Abstract**

The automatic face recognition has been a challenging task that has caught researchers' attention in the recent past for its vast applications in diverse areas. The concept of face recognition is among those challenges that have not been fully defined to address all situations. This research will present a new innovative method for face recognition. The technique involves using an image-based approach that leans towards the foothold of artificial intelligence by continuously excluding data from images through various methods such as image compression. Image compression involves the use of a two-dimensional cosine transform denoted as 2D-DCT. (Kim, Lee, & Sohn, 2014) The DCT is responsible for extracting properties from images based on skin tone and color. The DCT coefficients are used to create feature vectors. Research shows that for the DCT-based feature vectors into various categories and determine whether the element in the input image is "present" or "not present" in the database, which is a Self-organizing map utilizing an unobserved learning method is implemented. (Yan & Liu, 2020) Face recognition with SOM is accomplished by categorizing grayscale image intensity levels into diverse groups. The evaluation was carried out in MATLAB with a database of several face images, with distinct facial expressions for each subject. The system attained an 81.36 percent recognition rate for ten consecutive trials after training for roughly 850 epochs. In terms of both speed and computational requirements, the real benefit of this technology is its unrivaled fast processing capabilities and relatively low requirements for computation.

**Keyword**s- face recognition, artificial intelligence, self-organizing map, discrete cosine transform.

**INTRODUCTION**

The concept of face recognition has recently become a topic of interest of research. This is due to rising security concerns in commercial and law enforcement applications. (Yan & Liu, 2020) surveillance, human-computer interaction (HCI), and content-based coding of videos and images, biometric analysis, etc., have made significant progress in this area a few years back. While face recognition is a simple activity for the normal human, it has proven to be incredibly challenging to replicate it artificially because of several factors; for instance, while there are similarities between faces, and they differ in terms of gender, age, and skin color. Different image quality, facial expressions, facial tone, background, and lighting situations further contribute to the problem. (Yan & Liu, 2020)

This research work proposes a novel face recognition strategy based on an idea proposed by Hjelms and Low. Their study explains a preprocessing phase that tries to separate pixels connected with skin from those associated with facial features. In comparison to earlier methods, this strategy requires significantly less computing power. In addition, because human skin color varies by individuals, research suggests that intensity, not chrominance, is the most differentiating feature. (Yan & Liu, 2020) For subsequent processing, the recognition stage commonly uses an intensity referred to as the grayscale re-presentation of images compressed by the 2D-DCT. The intensity values for skin pixels are then included in this grayscale version.

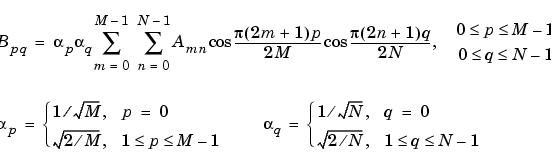
The 2D-DCT for every image is computed first, and the discrete cosine transform (DCT) coefficients are used to create feature vectors. (Kim, Lee, & Sohn, 2014) The second stage makes use of the self-organizing map and the unobserved learning method to group vectors and determine whether the individual in the input picture element is "present" or "not present" in the database. If the individual is identified as a present, the best matching image discovered in the database is given as the output result; if the subject is not classified as a present, the result states that the subject was not traced or identified in the database.

The following is a breakdown of the paper's structure: The computation of DCT on face photos is discussed in Section II. (Kim, Lee, & Sohn, 2014) The design and architecture are described in Section III. Section IV presents the outcomes of the experiments and analyzes prospective system adjustments and enhancements. Finally, in Section V, there are some closing observations.

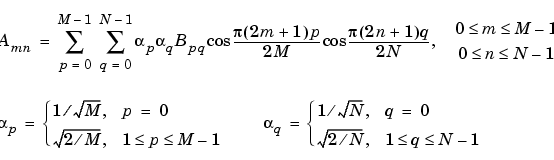
1. **OVERVIEW OF THE SYSTEM**

This Face recognition system is essentially based on discrete cosine transform algorithm, which is a commonly used method in various applications. The DCT is most commonly used for compression of data, as it is the foundation for the JPEG loss image compression technique, which is a worldwide standard. (Kim, Lee, & Sohn, 2014) The DCT has the virtue that the most visually meaningful information about a picture is concentrated in only a few coefficients for a typical image. DCT co-efficient extracted maybe implemented as a form of signature for major recognition functions like facial recognition.

It is important to note Face images feature a lot of redundant and correlated information, which puts a strain on processing speed and storage. The DCT, defined as a technique for converting image elements from the spatial to the frequency domain, discards high-frequency coefficients and is responsible for quantizing the remaining coefficients because apparently low frequencies are identified as more visually relevant in an image than high frequencies. (Kim, Lee, & Sohn, 2014) This saves space while maintaining image quality structure. The 2D-DCT of an M by N Matrix A, for example, can be written as:



The Bpq values represent the DCT coefficients. The 2D-IDCT (2D Inverse-DCT) is an invertible transform that is defined as follows:

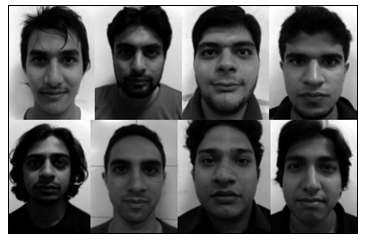


In this research, the proposed method uses the MATLAB Image Processing Toolbox's DCT transform matrix. This method works well with small square inputs, such as eight by 8-pixel image blocks.

1. **IMAGE PREPROCESSING**

Consider photographs of several applicants with various facial expressions captured with a digital camera with a resolution of 1200 x 1600 pixels, or 2.0 megapixels. The following characteristics can be found in all of the images:

1. Consistent lighting situations
2. Light color background
3. Upright and frontal faces
4. Tilting and rotation tolerance up to 20 degrees



*An example of an image representation of candidates*

In this case, the images are preprocessed in the Adobe Photoshop application. The preprocessing includes the following steps;

1. Adjusting brightness and contrast to a predetermined scale
2. Auto-adjusting hue and saturation levels
3. Converting a 24-bit RGB color to an 8-bit grayscale
4. Image resizing to 512 x 512 pixels
5. Image saving in jpeg format

RGB to greyscale conversion in Adobe

Face image from the 2-megapixel camera

Image preprocessing in Adobe

2D-DCT

Image adjusted in mat lab (64\*1 feature vector

Image resize in MATLAB

Resized with nearest-neighbor interpolation

Image adjustments

SOM neural network

*Face image fabrication process ((ZHANG, ZHONG, GUO, & PENG, 2012)*

**2D-DCT image compression technique**

The MATLAB Image Processing Toolbox is used to downsize most pre-processed images for instance from the 512by512 pixels to blocks of size 8by8 pixels using nearest-neighbor interpolation. The proposed design technique uses '8' out of the 64 DCT coefficients to calculate the 2D-DCT of picture blocks of size 8by8 pixels for masking. (Kim, Lee, & Sohn, 2014) The remaining 56 coefficients are thrown out. The image is then reconstructed using the DCT transform matrix computation technique to compute the 2D-IDCT of each block. Finally, a set of arrays is produced. Each array represents a single image and is eight by 8 pixels in size. The most important values are found in the uppermost left corner of every 2D-DCT matrix, which corresponds to a low-frequency component inside the preprocessed picture block.

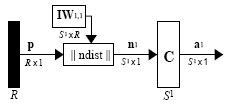
1. **SELF ORGANIZING MAPS**

The self-organizing map, commonly referred to as the Kohonen Map, is a neural artificial network that can play a significant role in facial recognition. It is a defined unsupervised learning approach that learns the different distribution patterns without prior knowledge of the class. It has the property of maintaining topology. (Raja, 2012) The neurons compete for the opportunity to be engaged or fired. As a result, one neuron wins the competition gets triggered, and this neuron is referred to as the winner.

A Self Organizing Map network uses the same process as a competitive layer to identify a winning neuron. Using the Kohonen Rule, neurons within a specified neighboring block of the winning neuron are upgraded and updated, rather than just the winning neuron. (Raja, 2012) The Kohonen rule is vital in recognition applications because it permits the weights of the specified neuron to learn the input vector. As a result, a SOM is used in the system to categorize DCT-based vectors into various groups to determine if the individual in the input image is "present" or "not present" in the database.

**The architecture of the network**

The one-dimensional, 2-dimensional, and multidimensional maps are all examples of Self organizing maps. Determining the input number of connections in a Self-organizing map network structure is determined by the number of categorization properties applied. (Raja, 2012)



*The basic architectural design of a simple SOM*

The DCT compressed image's row of pixels is represented by the input vector p, as portrayed in the architecture. The ||dist.|| box takes p as an input vector and IW1,1 as the input matrix weight and outputs a vector with S1 elements. (Raja, 2012) The elements are the negative distance figures between the input vector as well as the vectors iIW1,1 produced from various rows for the input weight matrix. The ||dist|| box finds the (Raja, 2012) Euclidean distance between input p and the respective weight vectors to compute the competitive layer's input n1.

Furthermore, with the exception of the winning neuron, the next neuron with a positive character input say input n1 and the defined competitive transfer function say C takes in an input vector for a particular layer and returns 0’s for all neurons leaving just the winning neuron. The output of the winning neuron is given as 1. The neuron with the closest weight vector to that of the input vector contains the smallest negative input and hence wins the competition to give a1. As a result, the defined competitive transfer function C returns a1 for the output element a1 I which corresponds to I \*, which is the "winner." In a 1, and the rest of the output subjects are given as 0.

When an input vector p is provided, the winning neuron's weights and those of its nearest neighbors move relatively toward p. As a result, after several exposures, nearby neurons learn vectors similar to one another. As a result, the Self organizing map network learns and adopts to classify the input vectors it encounters.

The Self organizing map network in this example has N nodes organized in a 2-dimensional lattice-like architecture. Each node has two or four neighbors in these circumstances, depending on the case. The learning phase and training phase as well as the testing phase are the three phases in a SOM's life cycle.

**Analysis of unsupervised learning**

The neuron that has weight nearest to the input vector is defined as the wining neuron during the learning phase. (Raja, 2012) The weights of all these neurons in the winning neuron's neighborhood are then changed by a certain amount that is inversely proportionate to the Euclidean distance and is based on a defined set of characteristics employed it groups and categorizes the data set based on the set of characteristics.

The Self organizing network plus matrix weight of every image saved in the training database are used to map training images into a lower dimension. It is during recognition that training test images are reconstructed by use of weight matrices, and recognition is done using Euclidean distance as the similarity measure on untrained test images. (Raja, 2012) The MATLAB Neural Network Toolbox was used for our system's training and testing.

**Training phase**

The self-organizing map is supplied with labeled DCT-vectors one by one during this phase. The number of "wins" and the defining attribute of the input sample is stored for each node. As explained in the learning phase, the weight vectors for all the nodes are respectively updated. (Yan & Liu, 2020) At the end of this step, each SOM node has two distinct values including the total number of successful times for subjects in the image databases and the number of subjects winning times not located in the database.

**Testing phase**

Every input vector is put in compared to all Self organizing map nodes during this phase, with the best matching image determined by the shortest Euclidean distance. The system's final output, depending on recognition, shows whether the image under testing is "present" or "not present" in the database.

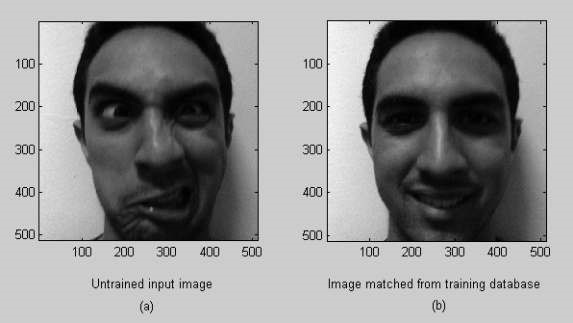
1. **EXPERIMENTAL RESULTS OF STUDY**

To study this face recognition system, an image database was developed. The database is divided into two halves, one for training and the other for basic testing. Numerous images are utilized during SOM training, each containing several subjects and each subject having their own set of images with various facial expressions.

**The validation technique**

In MATLAB, the preprocessed images (greyscale images) all of size 8by8 pixels are reshaped into a 64 by 1 array, for every image. This procedure is applied to the test photos to create the input images for the recognition system to be tested. Similarly, the picture database for the training uses photos to create a matrix containing 64 rows by n columns. The SOM's input vectors are dispersed throughout a 2D input space range, representing the grayscale pixel intensity levels. (Yan & Liu, 2020) These are utilized to train and teach the Self-organizing group training with sample dimensions [64 2], where each image sample represents 64 minimum and a 64 maximum integers of pixel intensities. With these values, a single layer of feed is forwarded to the Self organizing map containing 128 weights, and a respective competitive transfer function is produced. This network's weight function is the inverse of the Euclidean distance.

In addition, all subsequent tests use this SOM network. For the trials, up to half a dozen test images might be used with the database. There was no overlap between the training and testing sets in this study.



*Sample result of the system*

This simulation's outcome reveals that the element in the input image is "present" in the database. The best matching image in the example result demonstrates that persons in the image database with various facial expressions may be easily identified. (Yan & Liu, 2020)

**DCT Block size representation**

The first experiment is meant to investigate the impact of DCT block size on the system's recognition rate, with every DCT coefficient employed in the feature vector.

RECOGNITION RATES VS. DCT BLOCK SIZE

The maximum recognition rate is obtained at (8 x 8) block.

DCT block size

6\*6

10\*10

16\*16

Recognition

rate (%)

77.43

73.31

78.82

75.82

74.64

73.64

8\*8

12\*12

4\*4

*Example of a Sample representation for recognition rate of 8\*8 case study block*

**Reducing DCT-Vector Size**

This experiment deals with the computational effort imposed by huge DCT-feature vectors. This particular test aims to see if a reduced feature vector can be made from a set of defined DCT coefficients without affecting systems performance too much. Only 8 of the 64 DCT coefficients are used for computation with the present DCT block size of 8by8 pixels. (Dharavath, 2014) It is feasible to discover which coefficients contribute the most to the classifier's final choice using statistical analysis by analyzing the respective variance of every bit of the defined 64 dimensions belonging to the feature space. (Dharavath, 2014)

The statistical variances for the DCT-feature vectors defined for both the training and testing databases reveal two distinct areas of high variability. This shows that these characteristics play a significant impact in categorization. As a result, a smaller feature space constructed based on two high-variance DCT coefficients and removed the rest. The collected findings demonstrated that the new DCT-feature vectors only have four DCT coefficients. (Dharavath, 2014) This experiment shows that optimal face recognition performance may be achieved even using the feature vectors that are significantly smaller than those used in DCT-based analysis.

**Enhancing Processing Time on reference to Epochs**

This experiment is affiliated with the total system's processing time. Processing time accounts for most of the time required to train the SOM network. The length of training is determined by the number of epochs used. (Shah, Parah, Rashid, & Elhoseny, 2020) This test aims to explore techniques that can be used to reduce the training time while keeping the previously computed rate of recognition for potentially reduced DCT-feature vectors from experiment two. For the scenario of 850 training epochs, this experiment consists of attaining the best rate of recognition with the minimum amount of processing time. The results of the recognition rate simulations are the average of the defined number of simulations.

**CONCLUSION**

This research presents an innovative face recognition method that combines DCT coefficient-derived features with the SOM-based classifier. The system is primarily designed and implemented in MATLAB with a database of several images, a few people, and images with distinct facial expressions for each subject. The system should be designed to attain a high percent recognition rate for several consecutive trials after training for roughly 850 epochs. Compared to traditional DC feature extraction methods, a smaller feature space, as outlined for experiment 2, dramatically reduces the processing needs of the procedure. As a result, the system is suited for implementing low-cost and real-time hardware. There is currently no commercial usability of this technology. However, a viable Self organizing map-referenced face recognition system may be better developed in the future for commercial purposes.

References

DAI, Q., CHEN, S.-C., & WANG, Z. (2010). Hybrid Neural Network Architecture Based on Self-Organizing Feature Maps. *Journal of Software*, *20*(5), 1329–1336. https://doi.org/10.3724/sp.j.1001.2009.00563

Dharavath, K. (2014). Improving Face Recognition Rate with Image Preprocessing. *Indian Journal of Science and Technology*, *7*(8), 1170–1175. https://doi.org/10.17485/ijst/2014/v7i8.26

Forbes. (2020). Facial Recognition Software: Masks, Ethics & The Future | Forbes [YouTube Video]. Retrieved from https://www.youtube.com/watch?v=O7nb-8BtMl0

Kim, D.-J., Lee, S.-H., & Sohn, M.-K. (2014). Feature Extraction Method of 2D-DCT for Facial Expression Recognition. *KIPS Transactions on Software and Data Engineering*, *3*(3), 135–138. https://doi.org/10.3745/ktsde.2014.3.3.135

Lander, K., & Poyarekar, S. (2015). Famous face recognition, face matching, and extraversion. *Quarterly Journal of Experimental Psychology (2006)*, *68*(9), 1769–1776. https://doi.org/10.1080/17470218.2014.988737

Raja, A. S. (2012). Neural Network-based Supervised Self Organizing Maps for Face Recognition. *International Journal on Soft Computing*, *3*(3), 31–39. https://doi.org/10.5121/ijsc.2012.3303

Rossion, B. (2018). Humans Are Visual Experts at Unfamiliar Face Recognition. *Trends in Cognitive Sciences*, *22*(6), 471–472. https://doi.org/10.1016/j.tics.2018.03.002

Shah, A. A., Parah, S. A., Rashid, M., & Elhoseny, M. (2020). Efficient image encryption scheme based on the generalized logistic map for real-time image processing. *Journal of Real-Time Image Processing*, *17*(6), 2139–2151. https://doi.org/10.1007/s11554-020-01008-4

Solomon, C., & Tobreckon. (2011). *Fundamentals of digital image processing: a practical approach with examples in Matlab*. Hoboken: Wiley.

Wechsler, H. (2010). *Reliable face recognition methods: system design, implementation, and evaluation*. New York: Springer.

Yan, X., & Liu, G. (2020). Realization and optimization of face recognition system based on MATLAB. *Journal of Physics: Conference Series*, *1650*(3), 032091. https://doi.org/10.1088/1742-6596/1650/3/032091

Zhang, M. (2022). Unsupervised Context-Aware Feature Aggregation for Video Face Recognition. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4002170

ZHANG, P., ZHONG, J., GUO, A., & PENG, Q. (2012). Improved image preprocessing algorithm based on FPGA. *Journal of Computer Applications*, *31*(6), 1706–1708. https://doi.org/10.3724/sp.j.1087.2011.01706