Distinguishing Between Identical Twins Using Optimized Feature Extraction and Edge Detection Techniques

Raghav Ravishankar
Department of Computer and Information Science and Engineering
University of Florida
Gainesville, United States of America
raghav29@ufl.edu

***Abstract*— In the current scheme of things, reliable and exact verification of people has become very necessary in many fields such as business, crime investigation, forgery etc. Since identical twins have similar features, distinguishing between them has always been a big challenge in the field of image processing. The innumerable methods which have been employed till date, to distinguish between identical twins, rely mainly on features that are unique to an individual such as the position of iris, ears and facial marks like moles, scars etc. This project aims to achieve twin recognition solely through the process of facial recognition. It comprises of two phases. The first phase deals with the implementation of Hybrid-Discrete Artificial Bee Colony algorithm. The second phase consists of edge-length computation and entropy comparison. Both the phases make use of only facial recognition techniques and find extensive applications in the fields of forgery, mistaken identity etc.**

Keywords*— *face recognition, artificial bee colony algorithm, edge detection, entropy comparison.

I. Introduction

Distinguishing between identical twins is a serious issue in applications ranging from security, identity theft etc. There have been several cases where identical twins have taken advantage of their similar facial features to exchange criminal charges, benefits etc. In the early 1930's, Alsman was mistaken for America's most wanted fugitive - John Dillinger because of their similar facial features and as a result, was arrested more than 17 times. Such instances, and others which involve forgery and mistaken identity are completely uncalled for, besides being a moral and legal violation. The technique proposed in this paper tries to eliminate the possibility of such mishaps from occurring altogether owing to its accuracy and the fact that it can distinguish between identical twins based solely on their facial images. Most of the techniques which have been employed till date make use of various datasets consisting of facial gestures, facial muscles etc. to identify the minutiae differences between individuals. They also take several factors such as lighting, gender and age into consideration to determine the recognition rate for an individual. This project, however, aims to distinguish between identical twins based only on facial recognition techniques.

Identical twins are formed when an egg has been fertilized by a single sperm and the zygote splits into two. This process occurs in the earliest stages of development, when the zygote is just a cluster of cells. This implies that, in the initial stages of development, identical twins share the same genetic information. However, as the cells undergo mitosis, differences between the identical twins become more visible and these differences increase with age. Thus, older twins are more easily distinguishable than their younger counterparts.

Fig 1: Identical Twins sharing similar facial features

The method used in this paper primarily comprises of two stages. The first stage deals with the implementation of Hybrid-Discrete Artificial Bee Colony algorithm (henceforth, H-DABC), a robust swarm intelligence algorithm, which enhances the feature extraction and selection mechanisms. The H-DABC algorithm majorly uses Discrete Cosine Transforms(DCT). In certain cases, Discrete Wavelet Transforms are also employed. This algorithm uses optimized feature extraction and selection mechanisms by the participant onlooker, scouts and employed bees. This algorithm is apt for distinguishing between twins due to its high accuracy and ability to exploit the infinitesimally minor differences that exist between identical twins.

From a database consisting of the frontal facial images of both the twins, the H-DABC algorithm is initially trained with the images of twinA. Due to its high sensitivity, significant differences are obtained in the recognition rates, when the algorithm is tested upon the images of twinB.

The second stage involves the process of morphing the images of identical twins followed by edge length computation and entropy comparison. The images of the twins are symmetrically divided in half, cutting across the image longitudinally. The left half of twinA and right half of the twin to be recognized are morphed together generating a new image. The edge length of the morphed image along with the entropy of the sum of edge lengths are compared with the original edge length. If the difference obtained is negligible, the person depicted in both the images can be inferred to be the same, else, otherwise. The above experiments have been tested rigorously and have been found to be highly efficient.

II Motivation

Positive identification of individuals is a very basic social requirement today. In today's large complex society, the need for accurate personal authentication has become very important in many applications such as banking, business, security etc. Though the old run of the mill methods like passwords and pins are easy to implement, they are not completely secure. It involves the risk of being stolen or being obtained by someone else. The introduction of biometric evaluation techniques has significantly solved the identification problems today. However, the need for more reliable identification of individuals is increasing day by day. In applications like surveillance and monitoring of a public place, traditional age old biometric techniques like magnetic swipe cards, passwords and pins will fail due to the huge volume of people involved in the process. Further, identification of people with similar facial features is one of the most challenging tasks in the field of image processing till date. This paper proposes a system which, in a sense, is similar to the human eye and provides a novel approach towards the identification of individuals based solely on their facial features.

III Related Work

Distinguishing between identical twins is a very daring challenge in the facial recognition community. All the solutions that have been discovered for the above-mentioned problem involves the use of various datasets such as the position of eyes, ears, facial gestures etc. which are unique to every individual.

Ref [2] presents a study of distinctiveness of facial features in identical twins which is measured using both the entire face and each facial component such as eyes, nose, mouth etc. The method proposed in this paper employs multiple biometric evaluation and takes into consideration many factors such as lighting, age effect, etc.

Ref [3] distinguishes between two images of identical twins collected a year apart using facial recognition algorithms. The experiments in this paper were conducted using Multiple Biometric Evaluations and the performance results were measured using facial features, lighting, gender and age.

This paper aims to pick the positives from the references mentioned above and tries to solve the age-old problem of twin recognition. It presents a novel and an efficient approach involving the likes of edge-length computation and entropy comparison to produce accurate and comprehensive results.

IV Data Preprocessing

The database that is chosen for this task of distinguishing between twins is ND-TWINS. This dataset consists of 24,050 color photographs of identical twins captured in different configurations in steps of 45 degrees.

Fig 2 : Training and Testing Samples for Twin A and Twin B

Since the main aim of this paper was to exploit the discrete ABC algorithm to perform face recognition, no dedicated preprocessors were used to assist the algorithm. The images that were input into the system were subjected to Gaussian and Laplacian filters.

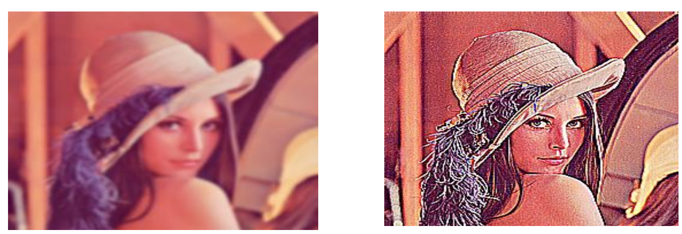

Fig 3: Gaussian and Laplacian Blurred Images

Discrete Cosine Transforms and Wavelet Transforms were applied interchangeably and contrasted at each phase to verify which of the two yielded a higher efficiency. The Discrete Wavelet Transform was also tested separately for its standalone recognition rate by changing and contrasting the wavelets Haar and Symlet. Continuous and rigorous testing has proved the self-sufficiency of the discrete algorithm employed in this paper in maintaining a swift run time without comprising on the accuracy of the recognition rates.

V. Architecture

Identification of twins has proved to be one of the most difficult tasks in the field of image processing. There is no standard technique that has been found to be efficient to distinguishing between identical twins. Most of the techniques that have been employed till date make use of datasets such as the position of iris, facial muscles, gestures etc. along with the facial images to differentiate between identical twins. This paper puts forth a novel approach which utilizes only facial image datasets. As shown in Fig 4, there are two techniques employed in this paper – implementation of a swarm optimization algorithm and employment of edge detection and entropy comparison algorithms.

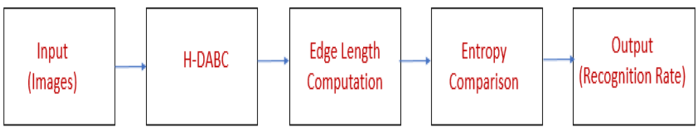

Fig 4: System Architecture

The H-DABC algorithm used in this paper is a minor variation of the standard Artificial Bee Colony algorithm used for various optimization methods. The method used in this paper involves the application of the standard Artificial Bee Colony algorithm used in particle swarm optimization techniques – hence the name hybrid-discrete artificial bee colony algorithm. The minor variation introduced in this method involves the discretization of velocity as opposed to the conventional method of discretizing the position. Tweaking the velocity is found to be more beneficial than discretizing the position for the method proposed in this paper. Furthermore, the runtime of the algorithm has also been found to be substantially low. The discretization of velocity involves the process of subtracting the present velocity of the employed bees with the previous velocity of the employed bee.

Where V_t - Velocity of the bees
V_n - Discretized bee velocity

The discretized velocity, a discrete entity, selects only those features which are essential for the recognition algorithm and discards the other remaining features, thereby decreasing the gallery. Thus, the number of features in the gallery is reduced drastically. Fig 5 shows the flow of the H-DABC algorithm after coupling it with the training and testing processes. The training process takes a few images as it's input. From these few selected images for the training process, optimized features are extracted and these required features are placed in the gallery. After retrieving the important features from the input images, these features are provided as input to the Euclidean Classifier. Similarly, the Euclidean Classifier obtains features from the testing stage. Based on the features obtained from both the training and testing stages, the Euclidean Classifier outputs the results.

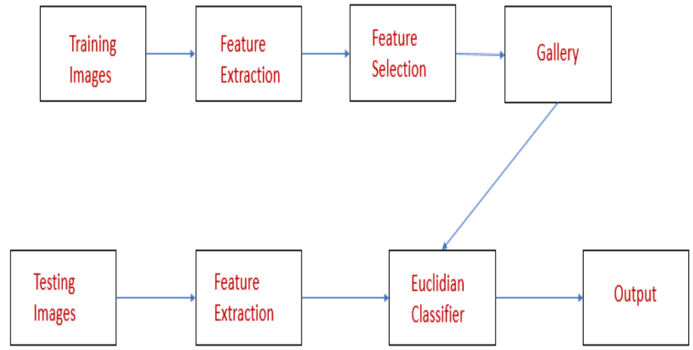

Fig 5: H-DABC Flow Algorithm

The second technique involves the process of edge length computation using Canny Edge Detector and Entropy Comparison. Canny Edge Detection involves the process of extracting the structural information of an image and dramatically reducing the amount of data that needs to be processed. Among all the edge detection techniques developed so far, Canny Edge Detector has been found to be highly reliable and popular owing to its optimality and ease of implementation. This process includes three steps:

1) Remove the noise and smoothen the image by the application of a Gaussian filter.
2) Identify the intensity gradients of the image and apply non-maximum suppression.
3) Determine the potential edges by double threshold and track them using hysteresis.

After the images undergo edge detection by the process of Canny Edge Detector, the Entropy of both the images are computed and their difference is tabulated. Results can be inferred based on the entropy difference between the images.

VI. Design Implementation

The method proposed in this paper employs two techniques to distinguish between identical twins.

The first technique involves the implementation of the H-DABC algorithm. Fig 6 shows the application of H-DABC to distinguish between twins. The first step towards the implementation of the algorithm involves the initialization of the food sources between upper and lower limits. The fitness of each food source between the limits is evaluated and the employed bees phase is initiated. The new position of the food sources is identified, fitness values are updated and the onlooker bee phase is initiated. Based on the fitness of each food source, certain food sources are identified and the position of individual bees are updated. Then, the velocity of each bee is computed by subtracting the previous position from the current position. The above process is repeated until the limit and later, the features are selected upon the application of the discretization algorithm.

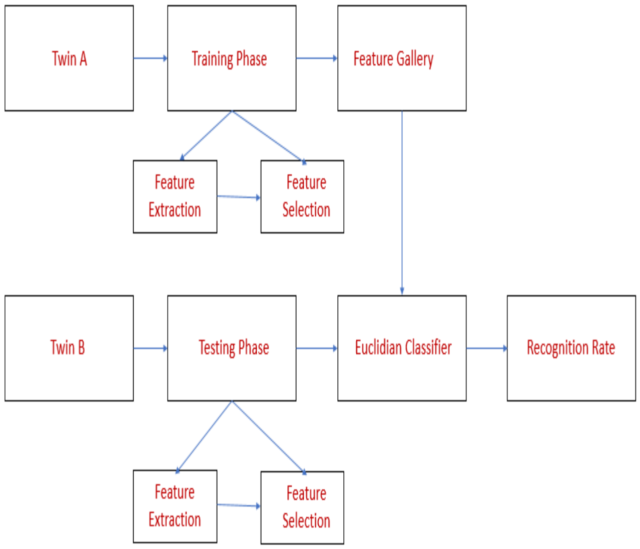

Fig 6: H-DABC to distinguish between identical twins

The original D-ABC algorithm is modified to suit the needs of our project. The H-DABC algorithm employed in this paper takes a dataset consisting of images of multiple subjects as input. A sizeable number of images are used initially to train the algorithm. Feature extraction is then achieved through certain preprocessing techniques such as Discrete Cosine Transforms and Discrete Wavelet Transforms. Once the training is completed, the remaining set of images of the subjects are tested upon by this algorithm and recognition rates are produced. In the case of twins, assume A and B are twins. The images of A are initially fed into the algorithm for training. After the necessary features of A are extracted during the training phase, the algorithm is tested upon the images of B based on the hits of the features obtained from the previous step. Since A and B are different people, the recognition rate obtained is very close to 0. The dissimilarity between the identical twins can also be obtained through this process.

The second technique involves the process of edge detection using Canny Edge Detector. The Canny Edge Detector function is performed on both the images and a new set of images with highlighted edges are obtained. A new image with its own set of edges is obtained on subtracting the edges from both the images. The edges that remain on subtracting these edges determine the originality of the individual. However, these edges cannot be quantified. Hence, entropy of both the images are determined and their difference is calculated. The difference obtained points out the dissimilarity between the twins or the originality of each person depicted in the picture.

The first step towards the implementation of the Edge Detection and Entropy Comparison processes shown in Fig 7 and Fig 8 involves the application of Canny Edge detection function on the images of both the twins.

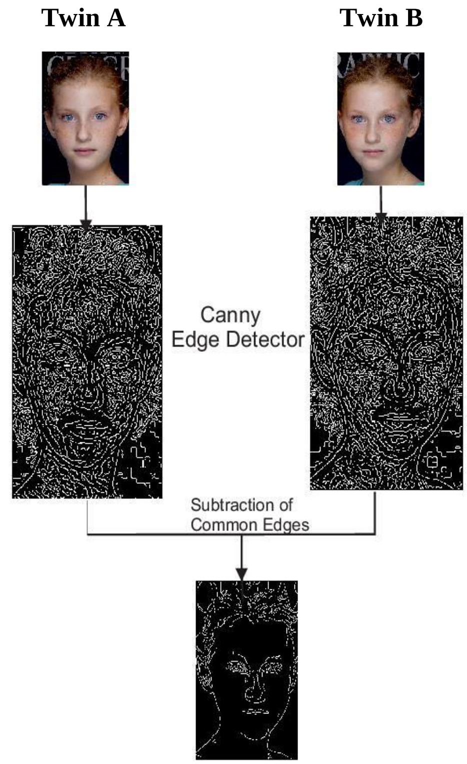

Fig 7: Edge Detection Using Canny Edge Detector

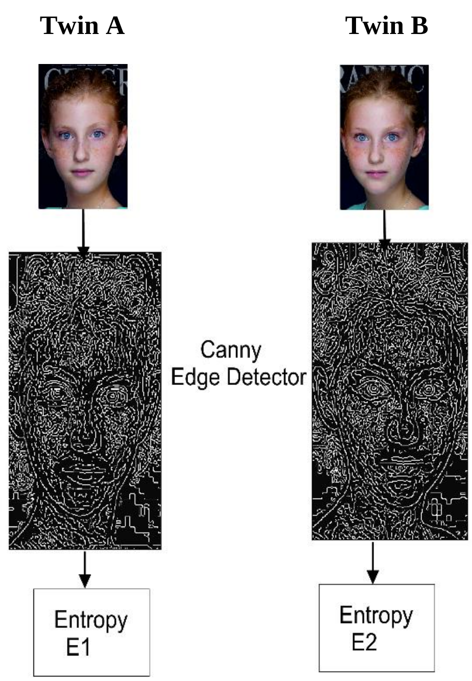

Fig 8: Entropy Computation and Comparison

Next, the entropy for the images of both the twins and their difference is calculated. The difference in the entropies obtained from the previous step give a quantitative value of originality.

The edges of the two images are then subtracted to give rise to a new image. This new image, thus formed, with a reduced number of edges is the count of edges present in A and not in B – an indication of the dissimilarity between the identical twins.

The entropies of both the images are obtained as shown in the Fig 8. Let E1 and E2 be the entropies of the two images twinA and twinB. If the entropies E1 and E2 are found to be equal, the person depicted in both the images is the same. However, if the entropies are not the same, the people depicted in the images are different.

VII. Performance Evaluation

The method proposed by this paper comprises of two stages.

1) H-DABC

The efficiency of the H-DABC algorithm is initially determined by providing a set of images of the same subject for both the Training and Testing phases. Different wavelets are employed in the pre-processing stage. It is observed that uniformity is maintained in the recognition rates and the error bandwidth does not exceed 15 % from the mean recognition rate. This recognition rate is depicted by R1 in Table 1.

Training and Testing Time of Same person			
Wavelets	**Training Time (s)**	**Testing Time (s)**	**Recognition Rate (R1) (%)**
haar	36.44	22.87	84.56
Sym3	37.44	17.68	82.37
Sym7	34.87	12.96	68.98

Table 1: H-DABC to determine and test the uniformity among recognition rates

Training twin A and testing twin B			
Wavelets	**Testing Time (s)**	**Training Time (s)**	**Recognition Rate (R2) (%)**
haar	34.45	24.76	13.78
Sym3	40.01	18.45	15.43
Sym7	34.26	15.78	14.35

Table 2: H-DABC to distinguish between identical twins

The next step provides a set of images of different subjects during the Training and the Testing phases. Firstly, the images of twin A are used for training. During the training phase, the features of twin A are selected and extracted by the algorithm. The images of twin A are replaced by the images of twin B in the testing phase. The features obtained from the training phase are contrasted with the features obtained in the Testing phase to produce Recognition rate R2 in Table 2. Since the individuals depicted in both the images are different, theoretically the value of R2 must be zero. However, the individuals depicted in the images are identical twins and thus, share a lot of common facial features which is represented by the Recognition rate R2. The difference between the recognition rates tells us the dissimilarity between the identical twins.

Sets of Twin Images	Entropy (E1)	Entropy (E2)	Difference (|E1-E2|)
Twins Set 1	0.8165	0.1347	0.6818
Twins Set 2	0.6457	0.4893	0.1564
Twins Set 3	0.6532	0.9006	0.2474
Twins Set 4	0.5022	0.8254	0.3232
Twins Set 5	0.6310	0.4281	0.2029

Table 3: Entropy Difference Computation to Identify Twins

Subsection EDGE

Firstly, the images of the twins are read simultaneously. The Canny Edge Detection function is then, applied on the images and the entropies are calculated individually for these images. The difference between the entropies is tabulated and this value is an indication of the dissimilarity between the identical twins. A high difference between the entropies of the images is an indication that the energy contained in both in the images is sufficiently different and thus, the images do not depict the same person. The entropies of the individual images and their differences for a set of five twins are shown in Table 3.

VIII. Conclusion

This paper puts forth a simple, yet detailed technique to differentiate between identical twins. The method proposed in this paper aims to differentiate between identical twins in two simple stages. The first stage involves the implementation of the H-DABC algorithm – a slightly tweaked version of the standard Artificial Bee Colony algorithm. This algorithm optimizes the feature extraction and the selection processes enabling the identification of the minutiae differences that exist between identical twins. The second stage involves the morphing of images, edge length computation and entropy comparison. The above methods have a very high efficiency and are capable of identifying the smallest of differences that exist between identical twins. This method uses only the facial images of the subject in question. More work needs to be done towards the unification of the two stages into one to obtain a better run time and better optimization.

References

[1] N. Shu and N. Kaur, "A Literature Survey of Facial Recognition of Identical Twins," *Ijarcce*, pp. 305–307, 2015.

[2] P. J. Phillips, P. J. Flynn, K. W. Bowyer, R. W. V. Bruegge, P. J. Grother, G. W. Quinn, and M. Pruitt, "Distinguishing identical twins by face recognition," *Face and Gesture 2011*, 2011.

[3] B. Klare, A. A. Paulino, and A. K. Jain, "Analysis of facial features in identical twins," *2011 International Joint Conference on Biometrics (IJCB)*, 2011.

[4] M. T. Pruitt, J. M. Grant, J. R. Paone, P. J. Flynn, and R. W. V. Bruegge, "Facial recognition of identical twins," *2011 International Joint Conference on Biometrics (IJCB)*, 2011.

[5] B. Poornima, Y. Ramadevi, and T. Sridevi, "Threshold Based Edge Detection Algorithm," *International Journal of Engineering and Technology*, vol. 3, no. 4, pp. 400–403, 2011.

[6] D. Karaboga and C. Ozturk, "A novel clustering approach: Artificial Bee Colony (ABC) algorithm," *Applied Soft Computing*, vol. 11, no. 1, pp. 652–657, 2011.

[7] "An Approach for Canny Edge Detection Algorithm on Face Recognition," *International Journal of Science and Research (IJSR)*, vol. 4, no. 11, pp. 24–27, May 2015.

[8] Sasmita Mishra 2015, Edge Detection of Images: A Novel Approach.

[9] Jaskarandeep Kaur, Anil Kumar, 2011, Evaluating the Shortcomings of Edge Detection Operators.