

A New Method for Iris Segmentation and Feature Extraction in Spatial Domain

Dujan B. Taha

Shamil Q. Ibraheem

dujan_taha@uomosul.edu.iq

College of Computer Sciences and Mathematics

University of Mosul

Received on: 16/03/2011

Accepted on: 16/05/2011

ABSTRACT

In this paper, a new proposed method is developed, this method contains two algorithms, one for human iris recognition, the other is for extracting the features of the recognized iris.

Many studies tried to extract the iris from images, most of those studies succeeded but after using very complex processing and filtering operations in addition to transforming the image to other domains (such as frequency domain) to achieve the required operations. The proposed method deduct those complexities operations to a minimum and requires only wavelet transform in a small part of the second algorithm, all other operations will be applied directly to the spatial domain. Experimental results shows the efficiency of the method as its being applied to the images of the Chinese Academy of Sciences – Institute of Automation “CASIA” iris database which contains a lot of deformations. Finally, Matlab R2010a (Ver. 7.10.0) was used to implement the algorithms presented in this paper because it facilitates handling images, arrays, and filters.

Keywords: Iris Segmentation, Feature Extraction, wavelet transform

طريقة جديدة لتجزئة القرنية واستخلاص الصفات في المجال المكاني

شامل قيس ابراهيم

دجان بشير طه

كلية علوم الحاسوب والرياضيات، جامعة الموصل

تاريخ قبول البحث: 2011/5/16

تاريخ استلام البحث: 2011/3/16

المخلص

تم في هذا البحث تطوير طريقة مقترحة جديدة تحتوي على خوارزميتين، الأولى تستخدم من اجل التعرف على قرنية العين والثانية من اجل استخلاص الخواص من القرنية التي تم التعرف عليها. تم إجراء عدة بحوث في هذا المجال، تكمل معظم هذه الدراسات بالنجاح بعد استخدام طرق معالجة معقدة جدا تتضمن استخدام خوارزميات تعتمد على التحول إلى حيز آخر (مثل الحيز الترددي) لكي يتم إجراء العمليات المطلوبة. الطريقة المقترحة تقلل من المعالجات والحسابات المعقدة إلى اقل قدر ممكن ولا تستخدم أي حيز ثاني، أي أن العمليات ستتم في الحيز المكاني بشكل مباشر. أثبتت التجارب العملية نجاح الطريقة المقترحة حتى عند تنفيذها على قاعدة بيانات "CASIA" التي تحوي عددا كبيرا من الصور التي تعتبر متباينة ومتنوعة أو مشوهة. أخيرا، تم استخدام لغة (Matlab R2010a) لتمثيل الخوارزميات المقدمة في هذا البحث وذلك لأنها تقدم تسهيلات في التعامل مع الصور والمصفوفات، والمرشحات.

الكلمات المفتاحية: تجزئة القرنية، استخلاص الصفات، التحويل المويجي

1. Introduction

Biometrics comprises methods for uniquely recognizing humans based upon one or more intrinsic physical or behavioral traits[4]. In computer science, in particular, biometrics are used as a form of identity access management and access control. The

word “Biometric” is used to refer to the human body parts that can be used to identify it. Biometric characteristics can be divided in two main classes:

- Physiological are related to the shape of the body. Examples include - but are not limited to - fingerprint, face recognition, DNA, Palm print, hand geometry, iris recognition, which has largely replaced retina, and odor/scent [4,9].
- Behavioral are related to the behavior of a person. Examples include - but are not limited to - typing rhythm, gait, and voice [14,9].

Biometrics are used to identify a single human based on the unique characteristics of the used biometric, since those characteristics are almost impossible to be created or initiated by other people they are used to create systems that are called “Biometric Systems”, those biometric systems use one or more human biometric in order to identify people and give them access to different levels of security and applications [3,12].

The iris is a part of the human eye, it's covered with the cornea (which covers both the iris and the pupil) and it's shaped like a circle, an inner circle to the iris is the pupil area, it normally appears darker than the iris. It is believed that the iris texture is determined randomly during the fetal development of the eye. Meanwhile, iris is believed to be different between one person to another, and between the left and right eye of the same person [2], The iris is considered one of the best biometrics to be used because it's always unique and different in features [12].

The term iris recognition refers to identifying an unknown iris or eye image by using computational algorithms and image segmentation, sometimes even pattern recognition is used and in extreme cases image compression is used too, iris recognition includes eye imaging, iris segmentation, iris verification, and so on [11].

When iris segmentations is applied, the iris image will be segmented into regions each of which corresponds to a portion in that image, those regions will be used to identify the iris and transformed mainly to numbers by extracting their features, this approach is mainly used to create secret keys that are used with encryption algorithms such as the Advanced Encryption Standard (AES) which needs a large key size that can easily be obtained through iris segmentation [2].

If iris segmentation was used for pattern recognition the iris image's extracted features will be stored in a database and a comparison or a similarity algorithm will be used on the stored values and a newly calculated feature values taken from a provided image to decide whether the identity of the person who supplied the new image is correct or not, this is mainly used for identity authentication in security systems [8].

Feature extraction is defined as locating pixels or areas in an image that contains distinctive characteristics. Iris images depends mainly on feature extraction algorithms because no matter how is the method of using an iris segment, features must always be extracted. For instance if an iris image was used as a key to an encryption algorithm then the features are first extracted and then used a key after they are processed[10].

The current biometric systems have two factors that measure their performance, the **first** factor is the False Accept Rate(FAR) which represents the probability that the system incorrectly matches the input pattern to a non-matching template in the database, It measures the percent of invalid inputs which are incorrectly accepted[4]. the **Second** factor is the False Reject Rate(FRR) which represents the probability that the system fails to detect a match between the input pattern and a matching template in the database, It measures the percent of valid inputs which are incorrectly rejected [7].

Many papers and researches regarding iris segmentation and feature extraction were published before. In (1997) Wildes presented a method for automated iris recognition as a biometrically based technology for personal identification and verification [15].

In (1998) Lester proposed a method that showed the ability to transform a two dimensional images into three dimensional by segmenting the image and extracting its feature, then using those features to create the third dimension [10].

In (2004) Cui et. al., presented a new method for identifying iris existence in an image by extracting four types of features (contrast, symmetric, isotropy, and disconnected features), then use those features to build a classifier [6].

In (2010) Annapoorani et.al., proposed a new method for accurate iris segmentation in noisy images, using spectral removal then localizing the pupil, iris, and eyelid using reflection map computation [3].

In (2010) too, Manikandan and Sundararajan proposed a method to use the two dimensional discrete Wavelet transform in feature extraction from the iris image and using those features to built a secure biometric system [12].

In (2011) Al-Hijaili and Abdulaziz proposed a new method for iris recognition and fusing the iris recognition algorithm with a facial recognition algorithm to built an Electronic Medical Record (EMR) which can be used in any medical facility [1].

The proposed method differs from the past work in three ways ; **first** is the processing domain, all the past work was done in domains other than the spatial, the proposed method is the first to be done in the spatial domain. **Second** is the amount of work for iris segmentation, where other algorithms succeeded in recognizing the iris through very complex processing, the proposed method succeeded while using simpler processing. **Third** is the feature extraction, in the past work the features are always extracted using complex mathematical algorithms and not in the spatial domain.

2. The Proposed Method

The proposed method consists generally of two algorithms :

- A. Recognition of the iris using the proposed new iris segmentation algorithm.
- B. Extracting the features from the segmented iris using the proposed feature extraction algorithm.

The following sections represent the algorithms with their detailed steps.

2.1 Iris Segmentation

The iris segmentation algorithm is applied entirely in spatial domain. The algorithm includes the following steps:

Step1: Image preprocessing. The algorithm begins with reading the original iris image and turning it into gray if it's a colored image. The reason for this operations (as almost all biometric images are turned to gray before further processing) is that the colors of the image is of no importance. For instance; the color of the hand have no significance over a palm print or a finger print whereas the difference lies exclusively in the details of the print. The same goes for the iris, colors are not important, only the details in the iris are. Notice that the CASIA images are already gray [5].

Step2: Zeroing (blackening) the pupil area. Because of its physical build, the pupil reflects less light than other parts of the eye [2] as shown in figure (1) [13]. This makes the pupil looks darker in the image and is used to identify the pupil area in a grey image, practice shows that in most times the pupil pixels have the intensity of (30) to (60), a filter will be applied to darken the pupil area by zeroing any pixel that have the intensity

of (60) or lower. This will produce a picture with a completely black pupil, without affecting the original image because normally the iris pixel values are greater than (60).

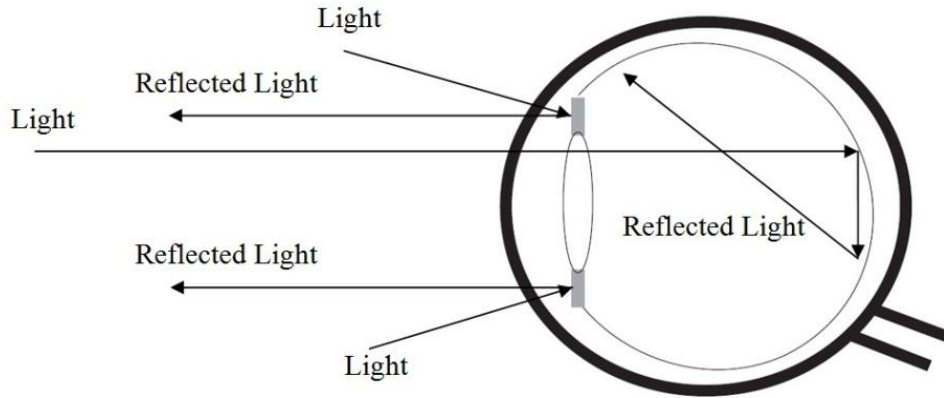


Figure (1). The human eye reflecting an incoming light

Step3: Locating the pupil. This is done by looking for a circle that has a certain radius and is filled with only zero intensity pixels (the radius is calculated as multiplying (10) by the division of the number of columns by the number of rows), this circle will only be found inside the pupil, the reason for choosing this technique is that after step2 the pupil will be the only place where this circle can be found because even if other areas had some dark pixels there will be non-zero pixels that prevent the circle from being filled.

Step4: Finding the pupil's boundary. After step3 the pupil's location will be known. Finding the center, top, left, right, and bottom boundaries pixels of the pupil is this step's goal, which is done by taking any pixel in the dark circle that was found in step3 and drawing an imaginary line from the first non-zero pixel on the right of the selected pixel to the first non-zero pixel on the left, this will create a horizontal straight line between two pixels on the boundary of the pupil as shown in figure (2a).

The center of the imaginary line is located and the same operation is repeated to draw a vertical line. This second line is drawn from top to bottom around the first line's center as shown in figure (2b). The second line's top and bottom pixels will represent the actual top and bottom pixels of the pupil's boundary.

The same method used for drawing the first line is repeated to draw another horizontal line around the center of the vertical as shown in figure (2c). The left and right most pixels in this line will represent the left and right most pupil boundary pixels. The center of the pupil will be the point where the vertical and the second horizontal line cross.

Step5: Finding the iris boundaries. This step is considered crucial in the algorithm because it is where the actual iris is found. This step is done by applying the Canny filter.

The **threshold** is a two-element vector in which the first element is the low threshold, and the second element is the high threshold. When only one value is provided for the threshold it will be used for the high threshold and (0.4 x high threshold value) is used for the low threshold. **Sigma** is used as the standard deviation of the Gaussian filter and is used to find if the weak edges are connected to strong ones.

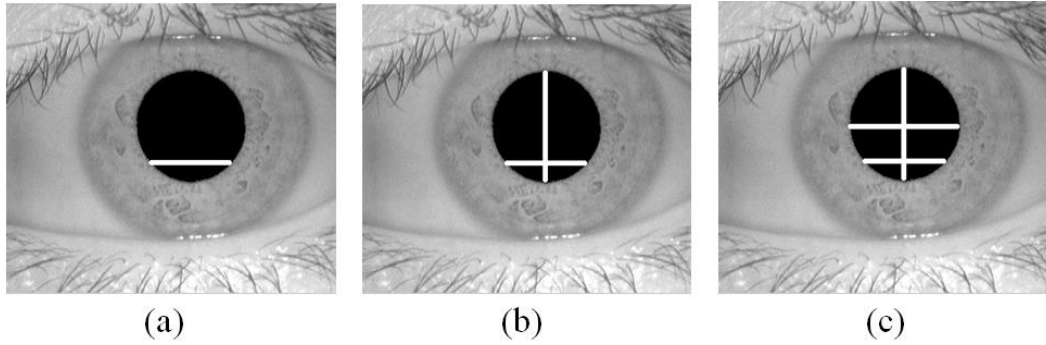


Figure (2). Finding the pupil's boundary (a) The first horizontal line (b) Finding the top and bottom pixels (c) Finding the left, right, and center pixels

Applying this filter requires a wise selection for the values for threshold and sigma. Provided dynamically rather than statically, a number of threshold and sigma values are manipulated until the correct ones are found.

Since the Canny filter produces skeleton images, the result will be a binary image, and since the pupil boundaries are already known, the correct values for sigma and threshold can be found using the distance equation. This is applied by starting from the left most pixel in the pupil's boundary and moving left until a non-zero pixel is reached or the image boundary is reached. If the boundary is reached, the values of threshold and sigma will be changed, if not, then from those two points the left distance is calculated. The same process is repeated for the right side.

Next the calculated left and right distances are compared, the comparison result will be called the "accepted difference", if the accepted difference is within (5) pixel margin then the result is accepted and the iris is taken, otherwise the results are refused and the values for threshold and sigma are changed again.

Step6: Cropping the iris image. After the iris boundary points are found in step5 the iris will be cropped out of the image but since Matlab have no function to crop a circle, a new method was developed for this purpose using the circle equation, distance equation, the structure function, image dilation function, and image filling. This allowed for the cropping of a circle shape and blacking everything other than the iris. Blacking will be very useful in the feature extraction phase.

Step7: Removing deficiencies. It is mainly done by cutting around 10% from the top and 5% from the bottom of the segmented iris image to eliminate the eyelashes because the CASIA data base images contains a lot of deformities that are not processed . The chosen values (5% and 10%) were acquired by practice from observing all the (756) images of CASIA database.

Finally, the result can be spotted when a comparison is made between the original iris image shown if figure (3a), the cropped iris image shown in figure (3b), and the final segmented iris image after removing the deficiencies shown in (3c).

Note that the selected accepted difference is less than (10) pixels for the segmented iris shown in figure (3), this value can be static because it is not relevant to the size of the image. a flowchart of the proposed iris segmentation algorithm steps is shown in figure (4).

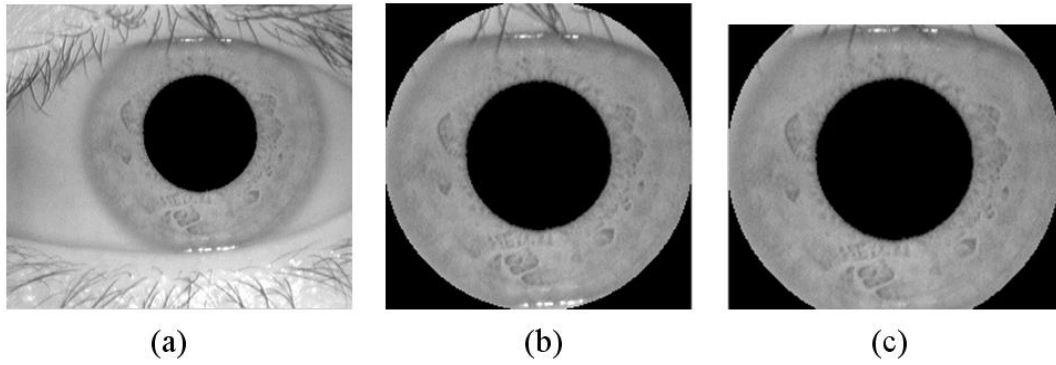


Figure (3). (a) The original iris image (b) The cropped segmented image (c) The final result after deficiency removal

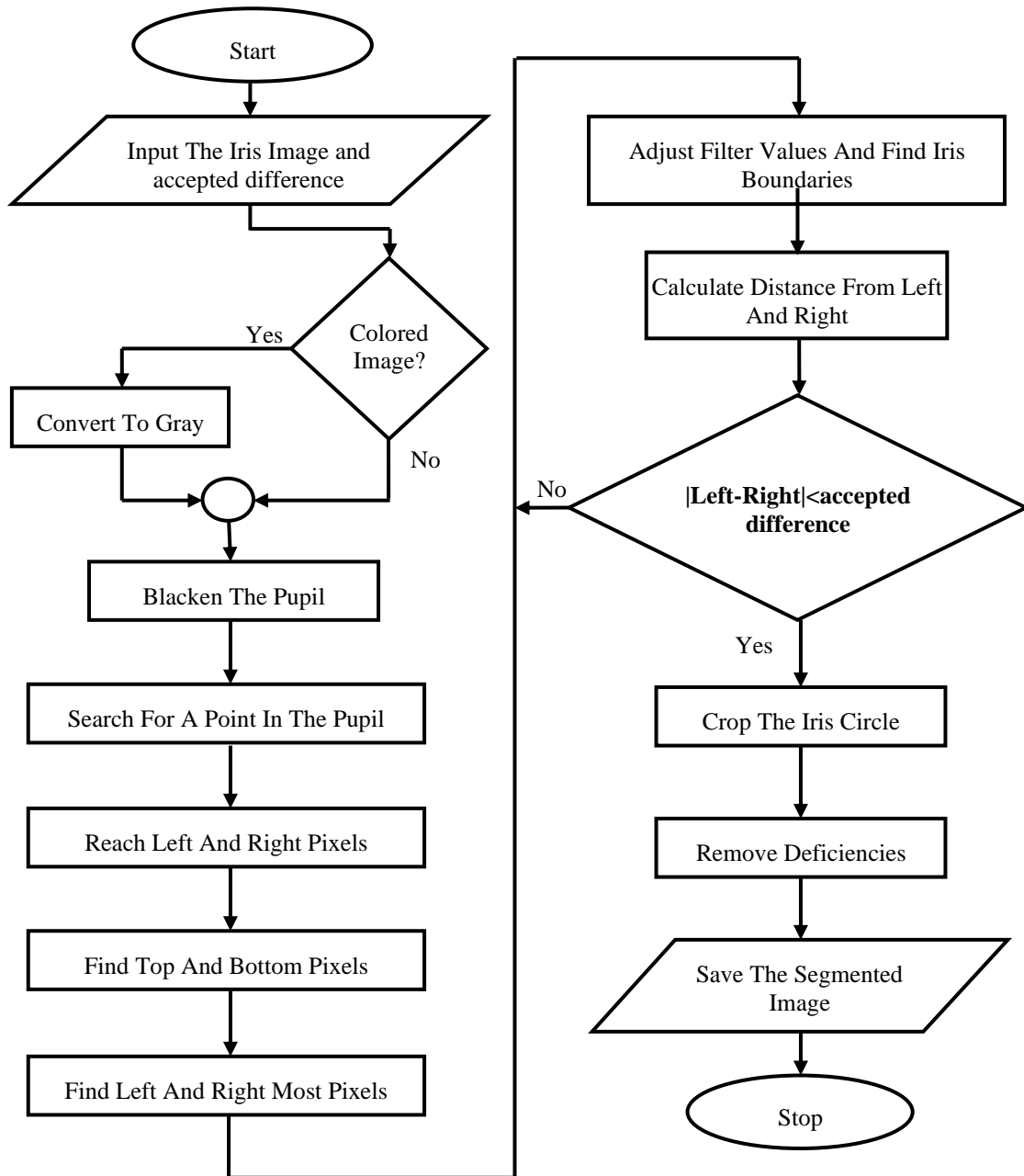


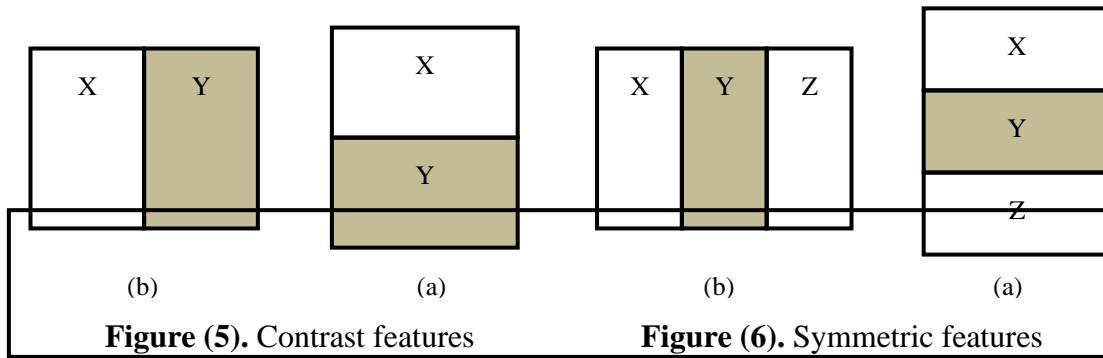
Figure (4). flowchart of the proposed iris segmentation algorithm

2.2 Feature Extraction

The main function of the feature extraction algorithm is to extract quantized features that represent iris patterns. One of the widely used methods is the Gray Level Co-occurrence Matrix (GLCM)[6]. It is mainly used as a feature for structural patterns, however, it's also a statistical feature that maintains sufficient computational complexity. In the proposed algorithm the GLCM was revised and four extraction methods were developed to extract features from the iris based on this revision.

In addition to the four extraction methods, two additional extraction methods were developed with the help of the 2D discrete wavelet transform, this will bring the count to six features.

The first four features are split into two types, first are the horizontal and vertical contrast features as shown in figures (5a) and (5b) respectively, second are the horizontal and vertical Symmetric features as shown in figures (6a) and (6b) respectively[6]. Next the six features will be summarized:



First and Second features: the idea behind calculating the features is to calculate the difference between the X and Y regions of the image, then normalizing the difference by using the variance of the image, since a gray image is being used the variance equation will be modified to accept a gray histogram instead of the two dimensional matrix that represents the colored images pixels.

After finding the variance, the feature is calculated using equation (1)[6]. Let (F) be the feature, (D) is the difference between the image's areas, and (V) is the Variance of the Image, then

$$F = D / V \quad \dots(1)$$

Third and Fourth features: the same extraction method used with the first and second features are used, only the difference calculation is changed. Instead of finding the difference between X and Y regions, the difference will be calculated using equation (2)[6]. Let $A=X+Z$, then

$$D=A - Y \quad \dots(2)$$

Fifth feature: after finding the first four features the last two will be calculated next, this will involve using the two dimensional discrete Wavelet transform. Wavelet transform has been extensively studied in various fields. The multi resolution architecture is very suitable for image analysis and feature extraction.

After reaching the fourth LL sub band, the fifth features extraction method starts by normalizing the iris values in the LL sub band. This is done by finding the average of all the pixels after normalizing them by the maximum value in the LL sub band.

Sixth feature: the sixth feature is extracted using a double standard deviation on the two dimensional level 4 LL sub band, the first standard deviation will results in a one dimensional array of values each representing the standard deviation of all the values in a column in the LL sub band, the second standard deviation will take the one dimensional array as input and outputs a single number that will be used as the sixth and final feature.

Figure (7) shows the flowchart of the proposed feature extraction algorithm, note that each extracted feature is made of one number contains 20-25 digits which means around 160-200 bits per feature. For all six feature there are 720-1200 bits. This large number of bits will provide the ability to use those features in many fields such as cryptography and pattern recognition systems.

Another thing to note is that everything was zeroed (blackened) except the iris, this is because the zeros will not appear in all the mathematical calculations to find the features, since all those calculations are based on the summing rather than multiplication of the pixels values, which means any non-iris pixel will be ignored by the calculations.

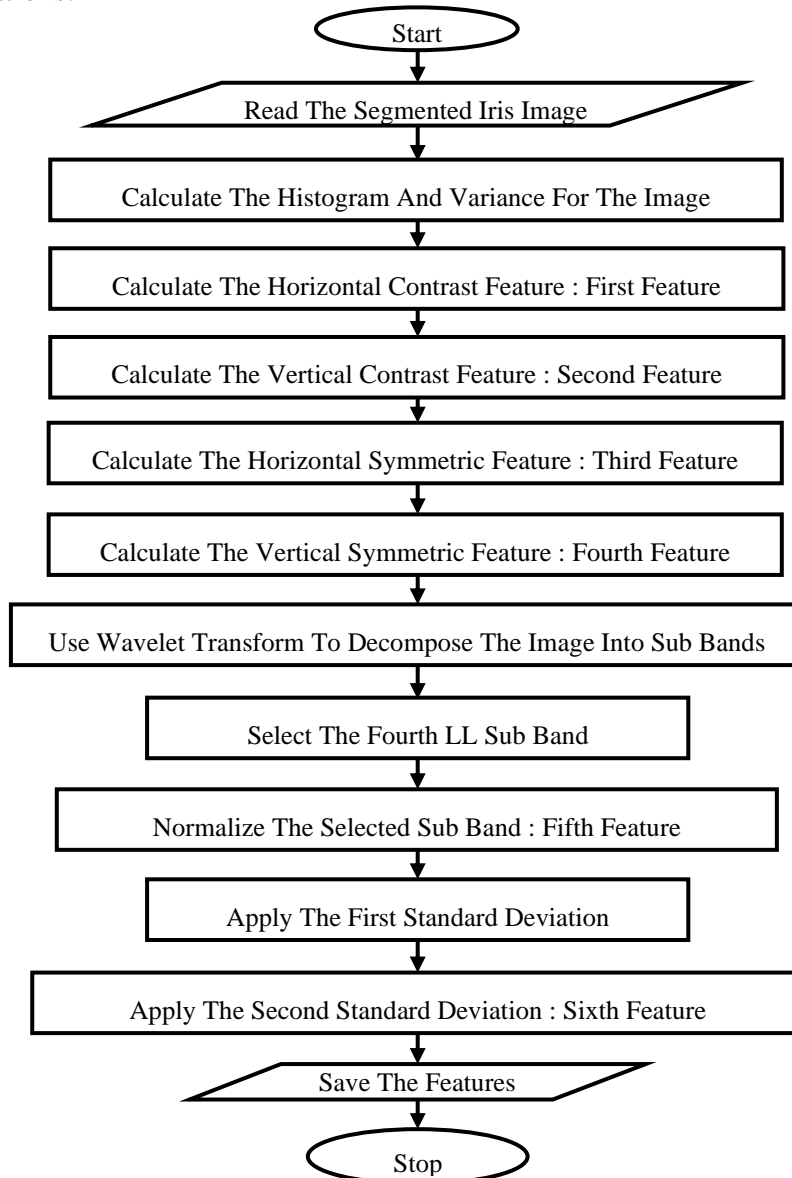


Figure (7). A flowchart of the proposed feature extraction algorithm

3. Results & Discussions

The work starts by reading the image that contains the original iris image as shown in figure (8a), then the iris segmentation algorithm is applied to darken the pupil as shown in figure (8b). The center of the pupil is then found as well as the top, bottom, left most, and right most pixels as shown in figure (8c).

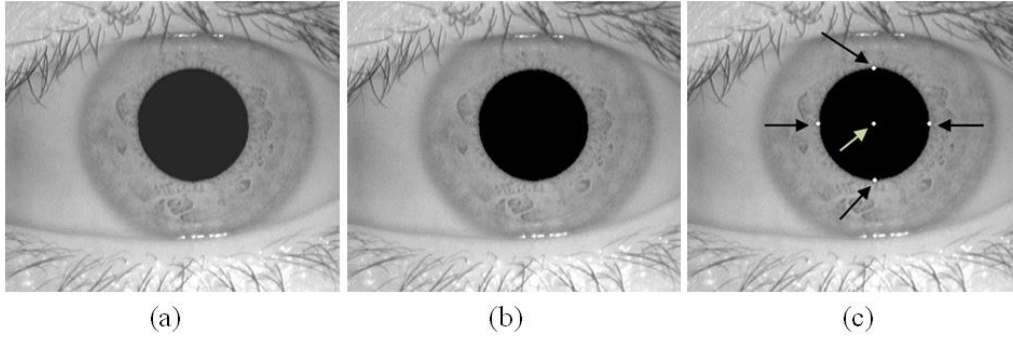


Figure (8). First stages of iris segmentation (a) Original iris image (b) Iris after blackening the pupil (c) Pupil boundaries found

Next step is choosing the right values for the sigma and threshold values of the canny filter. This is done by looping over all possible values and using the distance equation until the right values are found and the iris boundaries are known, figures (9a), (9b), (9c), and (9d) show the filter at different stages.

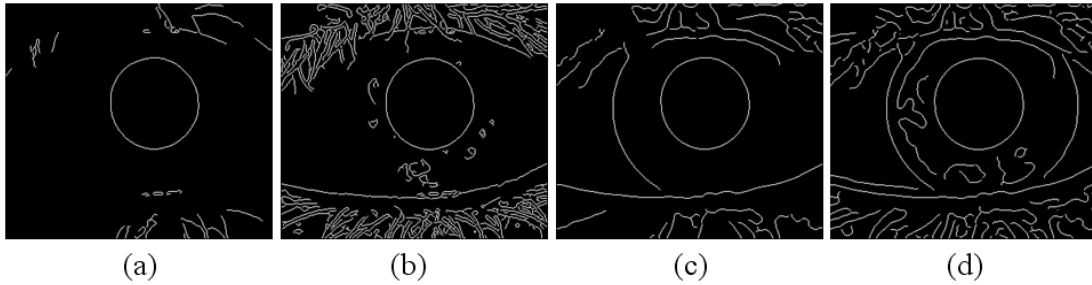


Figure (9). Examples for filter values (a) Threshold=0.5 sigma=1 (b) Threshold=0.2 sigma=1 (c) Threshold=0.2 sigma=3 (d) Threshold=0.1 sigma=3

After the correct values for the canny filter's threshold and sigma are found as shown in figure (10a) the right and left boundaries of the iris are known as shown in figure (10b), this along with the pupil center allows for the radius of the iris to be found as shown in the figure.

Hence, the iris boundaries and radius are now known, so the cropping can begin. To do this the circle equation is used as shown in figure (11a), a circle structure, image dilating, and image filling as shown in figure (11b), and finally cropping the entire iris to a zero (Black) array as shown in figure (11c).

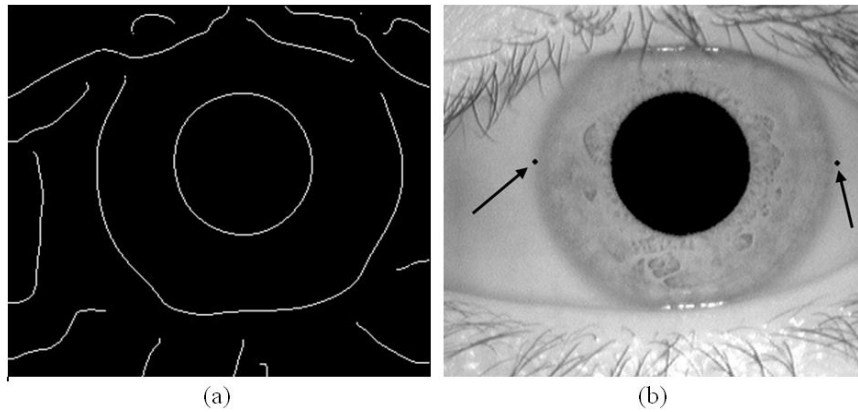


Figure (10). Correct filter values and the iris boundaries

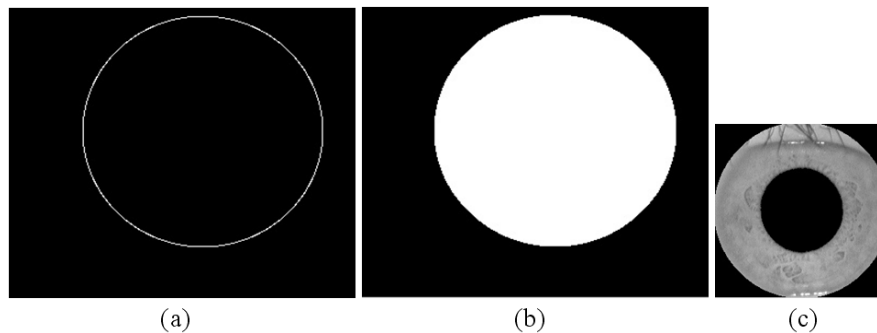


Figure (11). Final phases of the iris segmentation algorithm (a) Marking the iris boundary (b) Preparing the iris for cropping (c) The cropped iris

The segmentation algorithm results varied depending on the chosen difference threshold but the results are very promising. The algorithm was implemented for all (756) pictures in CASIA database without any exception.

When the difference threshold was less than (10) the number of successful iris segmentations was (741) making the success rate (98.02%). While the number of unsuccessful segmentations was (15) making the failure rate (1.98%). From those (15) images (13) were segmented but the iris was not complete mostly because the iris's left and right area differed in size.

To solve the non-symmetric iris size issue the difference threshold was raised to (20). This made the success rate raised to (99.6%) as (753) images out of (756) were successfully segmented. This left only (0.4%) failure rate for the (3) images from which the algorithm could not obtain the full iris.

As for the Feature extraction algorithm, obtained features give the following results(those features are extracted from the iris image shown in figure(11c)) :

First Feature	:	0.02095314479452429400
Second Feature	:	0.05980666981751550600
Third Feature	:	0.87970180644326890000
Fourth Feature	:	0.80861544497493865000
Fifth Feature	:	0.00740553396162788470
Sixth Feature	:	304.642757398440890000

The results show that the proposed algorithm provided a large number of digits and bits for each feature that is extracted, which makes it suitable for using as an encryption key for algorithms that require long sequences of bits, or for pattern recognition in databases and security systems. Since each iris is considered both unique

and random, the features extracted from it will be unique and random, which are the two most required issues in security and cryptography[14].

The usual error rate is also decreased, normally biometric systems have an error rate of (20-10%)[7]. In the proposed algorithm, since the features are extracted from the structure of the image in spatial domain, all non-iris (Blackened) pixels are ignored, this minimized the error rate to (1.98-0.4%).

The high success rate shows the efficiency of the proposed method. Comparing it to past work, the highest success rate was found by the center of biometric and security research, national laboratory of pattern recognition, institute of automation, Chinese academy of science(the institute that produced the CASIA database). By applying their work on the CAISA database they got a detection success rate of (96.5-98%). A simple comparison between their results and the results obtained from the proposed algorithm shows that the proposed algorithm produced higher success rate, and it's the highest rate found so far.

4. Conclusions

The proposed work presents a new algorithm for iris segmentation using simple calculations in the spatial domain, and an effective yet simple feature extraction method that presents effective results.

The iris segmentation algorithm proved to be very practical and have a very high success rate that ranges from (98.02%) to (99.6%). The proposed algorithm does not depend on a fixed image size, it takes all sizes of images and returns only the segmented iris image. Working in spatial domain and the simple processing operations gave the algorithm an efficient processing speed. Practice shows the processing time range to be from one to five seconds (depending on the input image).

The proposed algorithm simplifies any application that adopts it because it is fairly simple to use and needs only the input image from the user, all the needed values are calculated automatically, even filter values are generated dynamically and does not need any user intervention.

Also, the segmented iris images is saved without modification for any special algorithm. This means that the extracted image can directly be used with any feature extraction method without further processing. Another advantage of the direct saving of the iris image is that it can be used immediately in a database.

A huge advantage is the high success rate it maintained. For normal images the algorithm faced no obstacles, and even when it was used with the CASIA database the proposed algorithm found and segmented the irises for more than (98%) of the images. Images that wasn't segmented were severely deformed.

For the feature extraction algorithm, the use of the proposed methods proved to be very practical. Using the spatial domain directly in the first four extraction methods and the Wavelet and standard deviation functions provided by Matlab in the last two extraction methods provided fast feature extraction. And it is highly user friendly because it asks for only a segmented iris image and nothing more which facilitates its use.

The proposed algorithm also resulted in a large number of bits for each extracted feature, which makes it suitable for using as an encryption key or for pattern recognition in databases and security systems. But the large number will require more memory to save the database that contains the features and more time to compare those features in real time.

REFERENCES

- [1] Al-Hijaili, Shoa'a Jadallah ; AbdulAziz, Manal (2011) “**Biometrics In Health Care Security Systems, Iris-Face Fusion System**”, International Journal of Academic Research, Vol.3, No.(1), pp. 11-19.
- [2] Ali, Yasser Salem Mohamed (2009) “**Implementation Of Elliptic Curve Cryptography Using Biometric Features To Enhance Security Services**”, Unpublished MSc Thesis, University Of Malaya.
- [3] Annapoorani, G. ; Krishnamoorthi, R. ; Jeya, P. Gifty ; Petchiammal, S. (2010) “**Accurate and Fast Iris Segmentation**”, International Journal of Engineering Science and Technology, Vol. 2, No.(6), pp. 1492-1499.
- [4] Ballard, Lucas Kevin (2008) “**Robust Techniques For Evaluating Biometric Cryptographic Key Generators**”, Unpublished PhD Thesis, Johns Hopkins University.
- [5] Chinese Academy of Sciences–Institute of Automation (2003), “**Database of 756 Grayscale Eye Images**”. Cited at: <http://www.sinobiometrics.com> Version 1.0.
- [6] Cui, Jiali ; Tan, Tieniu ; Hou, Xinwen ; Wang, Yunhong ; Wei, Zhuoshi (2004) “**An Iris Detection Method Based on Structure Information**”
- [7] Hao,Feng ; Anderson, Ross ; Daugman, John (2005) “**Combining Cryptography With Biometrics Effectively**”, Technical Report Published By The University Of Cambridge.
- [8] Kekre, H.B. ; Thepade, Sudeep D. ; Jain Juhi ; and Agrawal Naman (2010) “**Iris Recognition Using Texture Features Extracted From Haarlet Pyramid**”, International Journal of Computer Applications, Vol. 11, No.(12), pp. 1-5.
- [9] Ko, Jong-Gook Ko ; Gil Youn-Hee ; Yoo Jang-Hee ;and Chung Kyo-IL (2007) “**a Novel And Efficient Feature Extraction Method For Iris Recognition**”, ETRI Journal, Vol. 29, No.(3), pp. 399-401.
- [10] Lester, Eric D. (1998) “**Feature Extraction, Image Segmentation, And Surface Fitting: The Development Of a 3D Scene Reconstruction System**”, Unpublished MSc Thesis, University Of Tennessee, Knoxville.
- [11] Liu, Wei ; Zeng, Wenjun ; Dong, Lina ; and Yao, Qiuming (2010) “**Efficient Compression of Encrypted Grayscale Images**”, IEEE Transaction On Image Processing, VOL. 19, NO.(4), pp. 1097-1102.
- [12] Manikandan, P. and Sundararajan, M. (2010) “**Discrete Wavelet Features Extraction For Iris Recognition Based Biometric Security**”, International Journal Of Electronic Engineering Research, Vol. 2, No.(2), pp. 273-241.
- [13] Nixon, Mark S. ; Aguado, Alberto S. (2002) “**Feature Extraction and Image Processing**”, Butterworth-Heinemann Publication, Linacre House - Jordan Hill - Oxford, First Edition.
- [14] Siyal, M.Y. and Ahmed, Fawad (2004) “**a Biometric-Based Scheme For Enhancing Security Of Cryptographic Keys**”, Institute of Electrical and Electronics Engineers, Vol. 8, No.(4), pp. 407-410.
- [15] Wildes, Richard P. (1997) “**Iris Recognition: An Emerging Biometric Technology**”, Proceedings of the IEEE, Vol. 85, No.(9), pp. 1348-1363.