



## Technical Section

## A multiresolution approach to iris synthesis

L. Wecker<sup>a,b,\*</sup>, F. Samavati<sup>b</sup>, M. Gavrilova<sup>b</sup><sup>a</sup> 422 19th ST NW, Calgary, Alberta, Canada T2N 2J1<sup>b</sup> Department of Computer Science, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4

## ARTICLE INFO

## Article history:

Received 12 June 2009

Received in revised form

22 April 2010

Accepted 26 May 2010

## Keywords:

Biometrics

Iris synthesis

Multiresolution

Reverse subdivision

Image generation

Verification

## ABSTRACT

Databases of human iris images are created and distributed for the purposes of testing iris identification algorithms. For logistical and privacy reasons, these databases are often too small to fulfill their potential applications.

In this work we develop a novel multiresolution approach to augment iris image databases. First, using a multiresolution obtained from reverse subdivision we decompose the example iris images into a set of lower resolution components. The components are a complete representation of the original image and consist of a low resolution approximation and a set of characteristics. To generate synthetic iris images we combine a set of components chosen from the original images. To ensure a unique, yet realistic, iris image each component of a synthetic image is chosen from different iris images. We quantitatively validate our approach by employing a classical iris recognition algorithm to compare our synthetic images with those that were used to create them. The results demonstrate that our approach is effective at augmenting iris image databases with iris images that are unique, yet exhibit both visually and statistically realistic characteristics.

© 2010 Elsevier Ltd. All rights reserved.

## 1. Introduction

Digital models can be created by gathering a set of measurements from geometric objects. For various reasons, these models may be incomplete representations of the objects. For example, there may be too few sample objects to produce a sufficiently large database.

The development of new iris identification algorithms can be hindered by the lack of test databases with enough sample iris images [36]. At the time of this research there were only three, freely available, public databases that could be used for testing iris recognition algorithms [5,19,8]. This motivated our interest in the development of a new technique for iris synthesis.

Although free databases of real iris images may become more readily available in the future, our method will continue to be relevant. Given high-quality sample iris images we can produce new high-quality iris images thereby augmenting whichever input database that is chosen. The sample irises will inherit properties from the original set of irises allowing for investigation into what kind and how many irises we can have if we start from an initial small set.

The goal of our approach is to augment existing iris image databases. We use a novel iris image synthesis technique that has

two distinct components. First, a number of preprocessing steps based on iris recognition methods are used to rid the image of extraneous (non-iris) information. For the synthesis stage, we capture and combine the characteristics of real images to compose new images. Generally speaking, this is similar to genetic inheritances, in which the irises of the children resemble those of the parents. For the purposes of combination, we explore a classification of iris images that increases the compatibility of the selected components.

To verify that our augmented database exhibits the same characteristics as the original database we use a well-known iris recognition algorithm on both the original database and the database augmented with our synthetic images. Although this alone validates our primary goal of database augmentation, our synthetic results also demonstrate a high level of realism.

The paper is organized as follows. Section 2 provides the necessary background information and related work. Section 3 describes methodology we developed for creating iris databases from the sample iris images. Section 4 presents experimental results, followed by conclusions and future work.

## 2. Background

Multiresolution representation (MR) and wavelets are powerful tools that have been recently applied in biometric and other image processing areas. They have been used for the purpose of biometric identification [7,12,32,37], fingerprint edge detection

\* Corresponding author at: 422 19th ST NW, Calgary, Alberta, Canada T2N 2J1.  
E-mail addresses: [lakin@structuredabstraction.com](mailto:lakin@structuredabstraction.com) (L. Wecker),  
[samavati@cpsc.ucalgary.ca](mailto:samavati@cpsc.ucalgary.ca) (F. Samavati), [marina@cpsc.ucalgary.ca](mailto:marina@cpsc.ucalgary.ca) (M. Gavrilova).

and matching [29,28], document contour extraction [34], as well as handwritten numeral recognition [20]. Recently, reverse subdivision (RS) filters have been proposed as alternatives to Haar wavelets for biometric applications [24]. RS refers to a kind of multiresolution that is directly created from operations on discrete datasets and therefore leads to more efficient and compact results than conventional wavelets [22]. Other techniques can be used for data synthesis purposes. See [24] for a discussion of MR techniques for data synthesis and see [31] for a discussion of non-MR techniques for data synthesis. Energy minimization MR can create details that are better suited to synthesis; however, their construction is more complex [21]. In this work we use Chaikin RS filters as they are very simple, and improve upon traditional Haar wavelets by providing continuous scaling and wavelet functions.

RS methods have been used successfully in a number of areas; such as polygonal silhouette error correction [9], and visualization of clinical volume data [26]. More closely related to this work, RS methods have been used in a synthesis capacity. Brosz et al. modeled terrain, by example, using MR to match the resolution of the terrains and provide the finer details needed for synthetic terrains of higher resolution [1]. In order to extract characteristics of line drawings, Brunn et al. used MR to decompose the drawings which allowed them to carry the characteristics to other line drawings [2]. In this paper, we present the novel application of reverse subdivision methods to generate synthetic irises and thus augment biometric databases.

Data synthesis refers to the creation of new data to meet some intended purposes and includes areas such as texture synthesis, domain specific rendering and biometric synthesis. Biometrics conventionally involves the analysis of biometric data for identification purposes. Due to logistical and privacy issues with collecting and organizing large amounts of biometric data, a new direction of biometric research concentrates on the synthesis of biometric information. One of the primary goals of the synthesis of biometric data is to provide databases on which the classical biometric algorithms can be tested [36,11]. There is very little work directly related to iris synthesis. In order to provide a better background, we review some works that are generally related to the area of biometric synthesis.

Cappelli et al. describe a complete system for human fingerprint image generation [3]. Their system uses fingerprint classifications to generate the global shape of each fingerprint. This global shape then undergoes a series of deformations and random variations in order to achieve the master ridge patterns. Once the master ridge patterns are known, their system can generate multiple samples of the original image using combinations of deformations throughout the synthesis process. They validate their results by exploiting the ubiquity of fingerprint identification and the subsequent availability of fingerprint experts and competitions.

Luo and Gavrilova propose an approach for facial synthesis and expression modeling based on the underlying mesh modification. Selection of control points in their method is guided by the 3D Voronoi diagram [15]. A general overview of geometric algorithms in facial expression modeling can be found in a plenary lecture delivered by Gavrilova [10].

Cui et al. propose a method to synthesize iris images, based on principal component analysis (PCA) and super-resolution [6]. Irises are grouped into five classes and 75 dimensional PCA global feature vectors are generated by applying eigenvector-based multivariate analysis to gray-level irises. The synthesis method uses a feature vector of the same size, and the problem is constrained to a search in a limited high dimensional space. Their method is computationally expensive, and uses a process of fine data decomposition while our method has a linear time

implementation and uses low resolution color components to produce synthetic images.

A different technique to synthesize human irises for use in computer graphics applications was recently developed by Lefohn [14]. The method makes use of the domain knowledge of ophthalmologists to obtain results that take on a high level of realism. Their approach uses 30–70 layers of painted textures, scanned into the program using a conventional flat bed scanner. An alternative approach uses a similar assembling technique to form an iris image from iris layers, such as collarette or stroma [35]. The set of iris layers is taken from a collection of synthetic and original elementary patterns of the iris. A superposition of these so-called iris primitives is used to assemble the final iris image. In addition to original elementary patterns, this approach synthesizes the collarette's outer boundary and the stroma using Bezier curves and a Voronoi transform, respectively. However, both approaches are highly computationally involved and complicated.

In [13], a model-based method is presented which generates iris images in five steps. To start, 3D continuous fibers are projected into a 2D polar space. A top layer with an irregular edge is modeled using cosine functions. Then the top of the layer is blurred and a smooth Gaussian noise layer is added to make the area bumpy. The final step adds eyelids using two low frequency cosine curves.

Model-based methods are important and interesting approaches for synthesizing objects. However, for complex objects such as the iris, the model and its parameters (including noise functions) become very complicated which, in turn, requires major simplifications and consequently rigorous verification. However, our approach uses a by-example strategy which starts with examples of real iris images. Since the synthetic irises inherit their characteristics from the real examples they also become realistic. Furthermore, adding new features to the synthetic irises produced by our method is a simple task. To obtain new features, one must simply add an iris image that contains the specific feature into the input dataset. Conversely, adding a new feature to a model-based method is usually a challenging task.

In this paper, we propose a radically different approach to augmenting iris databases with new and unique iris images that are easily obtainable through multiresolution technique. Multiresolution representation has been used for iris synthesis in [30]. In this work we take this idea and extend it to complete a systematic approach to iris synthesis with an emphasis on extensive experimentation and validation.

### 3. Methodology: iris synthesis

We now present a method that describes how multiresolution based synthesis can be applied to the problem of iris synthesis in the area of biometrics. Biometrics is the science of using biological properties to identify individuals. The purpose of biometric identification algorithms is to determine if a given biometric identifies anyone within a database of known individuals. Images of an individual's irises are a biometric that can be reliably used for identification [32,7]. However, the development of new iris identification algorithms can be hindered by the lack of test databases with enough sample iris images [36]. This motivated our interest in the development of a new technique for iris synthesis. This section describes our new multiresolution approach to synthesize new iris images.

We use multiresolution methods to decompose existing iris images into several components. We then reconstruct new iris images by combining components from different irises. By using components of real irises in our method we are able to synthesize

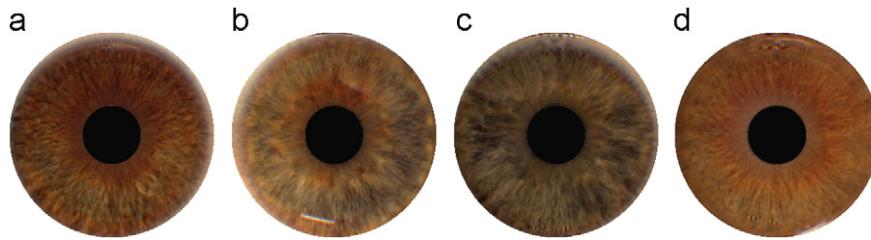


Fig. 1. Selected synthetic irises generated with our algorithm.

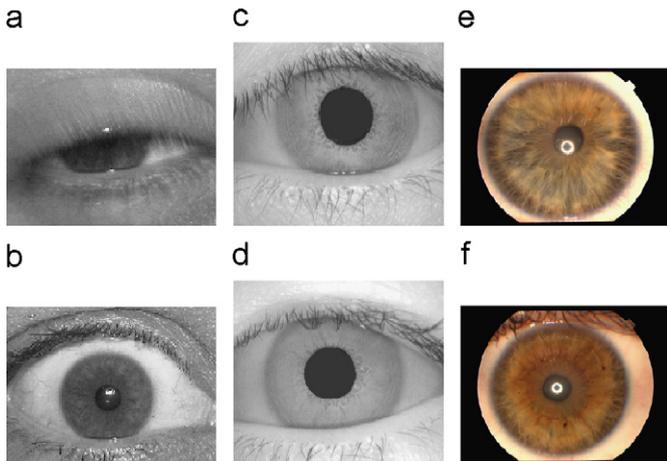


Fig. 2. (a,b) Example UBIRIS iris images. (c,d) Example CASIA iris images. (e,f) Example UPOL images.

iris images that are very realistic. Each unique combination of the extracted components creates a new and unique iris. Therefore, all comparisons using our synthesized irises will be inter-class comparisons (i.e. comparisons between images of different irises).

### 3.1. Image databases

Our synthesis technique requires a pre-existing set of iris images. There are currently three publicly available databases of iris images: CASIA [5], UPOL [8], and UBIRIS [19] (see Fig. 2). Each of the three databases has distinct characteristics: CASIA contains black and white images of medium resolution, UPOL contains color images of high resolution, and UBIRIS contains color images of high resolution as well as noisy images. Our method uses the existing images in the database to synthesize new images. The synthesized images will be of the same color space, quality and resolution of the input iris images which affects our decision of which iris database to use. Although UBIRIS is the largest publicly available iris database, it contains many noisy iris images for testing purposes. Extracting characteristics from noisy iris images produces noisy characteristics which is unsuitable for generating high quality synthetic iris images. The CASIA database images are black and white and as such, will result in black and white synthetic iris images. The UPOL database contains high-quality, high-resolution colored iris images which have minimal artifacts. We chose UPOL as the starting database because it allows our method to produce high quality, high resolution images. The resulting images have minimal artifacts that can be attributed to the input images thereby allowing us to easily identify any artifacts introduced by our method.

### 3.2. Preprocessing: isolating the iris

The images in the UPOL database may also contain the surrounding eye, eyelids and the pupil (see Fig. 3). In order to rid the image of this extraneous information, iris recognition algorithms use a number of preprocessing steps. Any of the preprocessing steps that result in the identification and removal of non-iris information contained in the image may be used [33,32,7]. There are two main requirements that our algorithm imposes on these preprocessing steps. The first is that they must fully remove the eyelid, and any highlights contained in the pupil leaving only the iris itself as in Fig. 3(b). The second is that they must produce an image that is a suitable resolution for our multiresolution step. We use an image editing tool, such as photoshop to remove the lighting reflections from the pupil and removing most of the eyelid and surrounding eye resulting in an image such as Fig. 3(b). UPOL input images are  $768 \times 576$  pixels in size. After we removed the eyelid and eyeball from the images, the images had  $< 512$  pixels of height. The RS filters we use are best employed on images that have a width and height that is a power of two. To satisfy this, the resulting images were resized such that the largest dimension was 256 pixels and extra rows or columns of black pixels were added to result in an image of size  $256 \times 256$ .

The iris itself is a circular, pigmented portion of the eye that functions to regulate the amount of light allowed to pass through to the retina. The stroma is a fibrovascular tissue that is visible as the lines connecting the outer edge of the pupil with the outer edge of the iris. Additionally, irises have circular rings that are called the collarette. The stroma and collarette give the iris some structure. We use a preprocessing step that unwraps the iris image into polar coordinate system such that the stroma and the collarette are aligned with the rows and the columns of the image.

The circular nature of the iris lends itself to an alternate coordinate system. In the first step, we use a polar coordinate transform to *unwrap* the iris image into a rectangular shape, resulting in the image in Fig. 3(c). The polar coordinate system is defined in terms of  $(r, \theta)$  where  $\theta$  is an angle from a polar axis and  $r$  is the distance from the origin. This effectively unwraps the circular iris information into columns and rows. Next, the pupil is removed such that only the iris information is left in the image, a necessary requirement of our synthesis method. Once a new iris image has been synthesized, we re-insert the pupil information and return the iris to its original coordinate system. The sequential organization of the preprocessing stages and subsequent synthesis framework are provided in Fig. 4. We explain each of the preprocessing stages in detail in the following sections.

Each input image is a uniform sampling of the light reflecting off of an iris. The image resulting from the polar coordinate transform represents a uniform sampling in the angle and radius axes. A standard polar coordinate transform produces an output image of the same width and height as the input image. However, the Cartesian and polar sampling rates are not equivalent. The

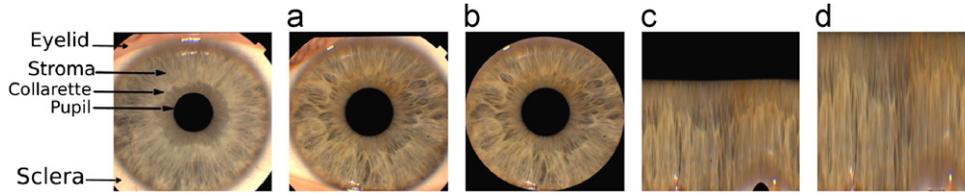


Fig. 3. Iris image preprocessing steps: (a) the original image, (b) after removing the eyelids, eyeball and pupil highlights, (c) after being unwrapped with the polar coordinate transform, and (d) after being scaled using piecewise linear scaling algorithm.

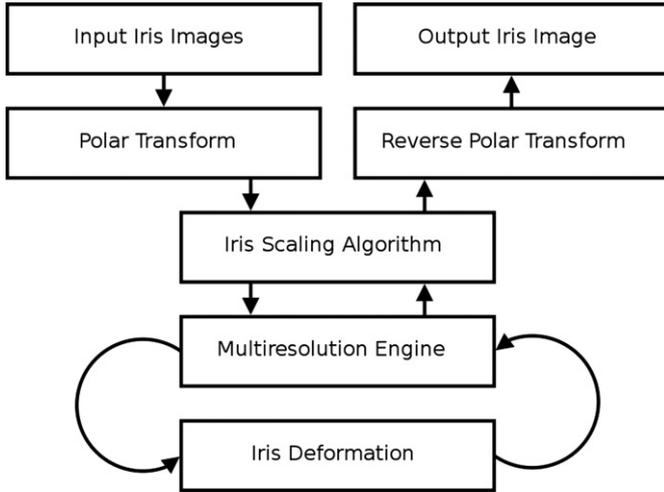


Fig. 4. The preprocessing stages, and subsequent multiresolution engine organization.

outer edge of the iris is a higher resolution sampling of the iris in the angle dimension, than is the inner edge. In our case, a polar coordinate transform that does not change image resolution will result in loss of information from the outer edge of the iris. We address this issue by using an output image that has a resolution sufficiently large enough to contain the unwrapped circumference of the iris. Analysis, discussed below, demonstrates that a moderate increase in quality can be achieved by using this approach instead of a standard resolution transform.

The circumference of any circle is  $\pi$  multiplied by the diameter. For our database, the diameter of the iris is approximately the width of the image. Therefore, the resulting image size must be at least  $\pi w_i$ , where  $w_i$  is the width of the input image. Digital images use integer sizes, hence we use the ceiling of  $\pi = 4$  which results in a polar coordinate image with dimensions  $4w_i \times w_i$ . The inverse transform reverses this process, producing an iris image in the Cartesian coordinate system that has the same size as the original iris image.

To test this high-resolution transform, we used the standard measure of peak signal to noise ratio (PSNR) [17]. PSNR compares an original signal with a reconstructed version and provides a measure of the similarity of the original signal to the reconstructed signal. We use the following definition of PSNR:

$$20.0 \log_{10} \left( \frac{255.0}{\sqrt{\frac{\sum [f(i,j) - F(i,j)]^2}{p^2}}} \right) \quad (1)$$

where  $p$  is the number of pixels in the image,  $f$  is the original iris image and  $F$  is the image obtained by applying a polar coordinate transform followed by the inverse transform. We use our high resolution polar transform and compare its PSNR value with that of a standard resolution transform. Fig. 5 demonstrates that the

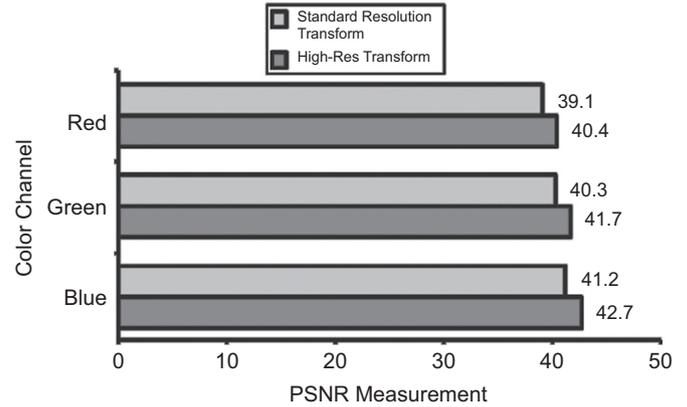


Fig. 5. Chart of the PSNR measurements used to test the two polar coordinate transforms. *Standard transform* refers to a transform that results in an image size of  $w_i \times w_i$ . *High-res transform* denotes our high resolution transform that results in an image size of  $4w_i \times w_i$ .

high resolution polar coordinate transform provides a moderate improvement in the quality of the resulting image.

### 3.2.1. Removing the pupil

The pupil is the central portion of the iris, where the light passes through to the retina. Its size is dependant upon the amount of light available. In addition, the pupil may not be centered in the iris image. Both of these conditions produce abnormalities in the polar coordinate images, see Fig. 3(c). The variability of the pupil size causes a black rectangular area in the top of the image. There may also exist small parabolas along the bottom of the image due to non-centered pupils. In order to isolate the iris for our synthesis algorithms we should remove these parabolas.

To solve these issues, we employ a technique which scales the iris image information to fill each column. This effectively normalizes the length of each column in the iris image and also removes all empty values. We use a piecewise linear scaling algorithm [27] to perform the scaling.

The first step is to isolate area within each column that contains the iris information and nothing else. Edges of the iris are found in each column by detecting the first pixel from the top and the first pixel from the bottom which is not black. The detection may be sensitive to noise in the pupil although in our cases such abnormalities were removed when we removed the eyeball, eyelid and pupil highlights in the first step. Although it is not the focus of this work, automation of this method might be achieved through noise removal and smoothing filters with edge detection algorithms [33,7]. These edges denote a new, smaller column of values that will be scaled to fit the original column size.

We use a piecewise linear scaling algorithm [27] to scale the iris.  $C_o$  denotes the original column and  $C_i$  denotes the subset of that column that contains the iris information. The scaling

algorithm maps each pixel in  $C_o$  into the domain of  $C_i$ . The mapping is not one to one, so each pixel in  $C_o$  is a linear interpolation of two adjacent values from  $C_i$  effectively resampling the iris image in a uniform manner across the entire image. The resulting image is similar to the one in Fig. 3(d).

Once the synthesis algorithm is finished, the new iris image will also appear similar to Fig. 3(d). It is then post-processed to return it to the original shape and size of a real iris (see Fig. 4). During the post-processing each column is scaled to a fraction of its original height. The size used during the post-processing step can be varied to achieve varying sized pupils. This allows our synthesis algorithm to produce images of the same iris with different pupil sizes.

### 3.3. Iris synthesis

Our goal is to augment existing iris image databases through our image synthesis process. The synthesized images should display similar characteristics to real irises, yet be unique. To achieve the goal of realism, we introduce a method that uses characteristics of real iris images. For the goal of uniqueness, our method uses a set of unique characteristics extracted from multiple differing irises.

Iris image characteristics can be extracted from the original iris image by using a suitable multiresolution method. Haar wavelets are traditionally used in biometric applications, however, we use Chaikin reverse subdivision filters which have been proposed as alternatives [23]. Chaikin reverse subdivision filters improve upon Haar wavelets by providing continuous scaling and wavelet functions. In addition the filters improve upon conventional B-Spline wavelets [25] by providing the banded regular structure with shifted filter values necessary for efficient implementation [23].

A common approach to multiresolution notation is to use a matrix notation where  $A$ ,  $B$ ,  $P$ , and  $Q$  represent the multiresolution operations as follows. The iris image  $I^n$  can be decomposed to a coarser approximation  $I^{n-1}$  and a vector of details  $D^{n-1}$  by using

$$I^{n-1} = \mathbf{A}^n I^n \quad (2)$$

and

$$D^{n-1} = \mathbf{B}^n I^n \quad (3)$$

In addition, the original image  $I^n$  can be easily reconstructed from  $I^{n-1}$  and  $D^{n-1}$  using

$$I^n = \mathbf{P}^n I^{n-1} + \mathbf{Q}^n D^{n-1} \quad (4)$$

The process can be iteratively applied to the coarse approximations for further decomposition  $I^{n-l}, D^{n-l}, D^{n-l+1}, \dots, D^{n-1}$ . The core of our synthesis algorithm is built upon an interesting possibility: any detail component,  $D^{n-1}$  can be replaced with the details from another iris image,  $D^{n-1}$ . The reconstructed image based on these two components can be thought of as a new unique iris image  $I_{new}^n$ . This leads us to have many options for combining components of different irises (Fig. 6).

Looking at any iris image, it becomes obvious that most of its characteristics are made of high resolution data and consequently the value of  $l$  should be small. Experimental results, shown in Fig. 7, demonstrate that only four levels ( $l=4$ ) of decomposition are enough to effectively capture all of the details from  $256 \times 256$  iris image. The low resolution approximation which is left after four levels of decomposition contains only the global color scheme of the given iris. To demonstrate this observation, we decomposed an iris image and replaced its low resolution approximation with a solid grey image. We then restored it to its original resolution, resulting in the grey iris in Fig. 7(f). The new iris is simply a grey-scale version of the original iris image. We can therefore conclude that the details extracted during four

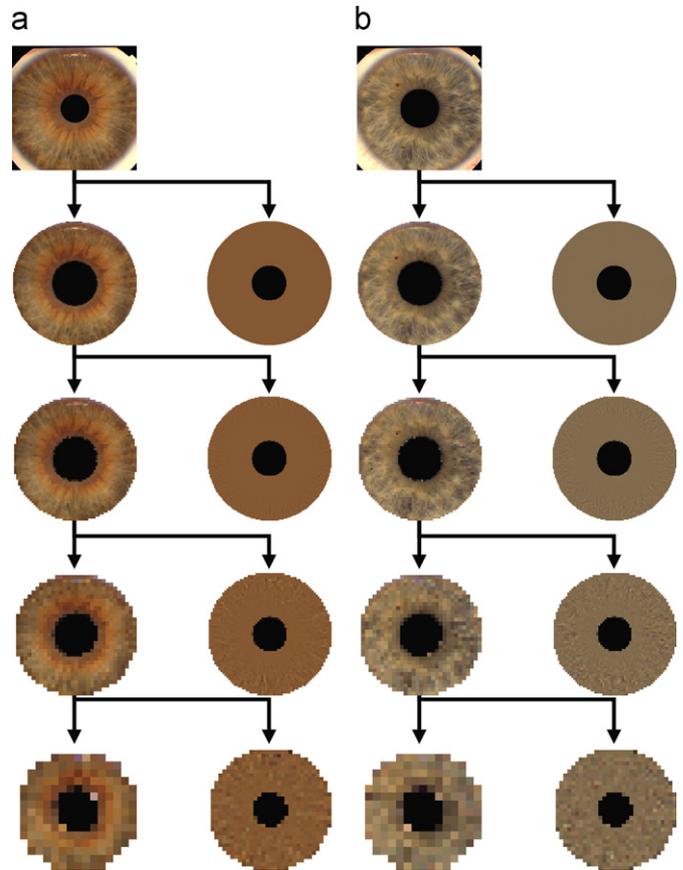


Fig. 6. Iris image decomposition. The original iris image is on the top row. Each subsequent row contains a coarse representation after one level of decomposition and the details extracted during the process.

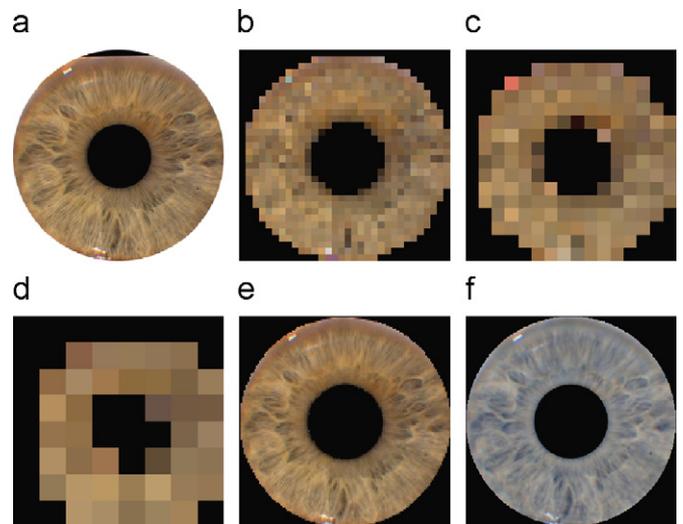


Fig. 7. (a) Original iris. Coarse approximation after (b) two, (c) three, (d) and four levels of reverse subdivision. The third level still contains pattern information, while the fourth level contains only color information. (e) Fully reconstructed iris image, compared with details (f) captured from four levels of reverse subdivision. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

levels of decomposition completely capture the characteristics of an iris image; an observation that is central to the combination algorithm. This observation allows us to completely decompose

an iris image into the following five components  $I^{n-4}$ ,  $D^{n-4}$ ,  $D^{n-3}$ ,  $D^{n-2}$ , and  $D^{n-1}$ .

The number of levels required to completely capture the characteristics of the iris is dependant on the input image resolution. Given a database which contains larger iris images, more levels of decomposition may be required to capture all characteristics. In a conceptual level, our approach can also be used when the level of the decomposition is higher than four and indeed this can increase a larger set of synthesized irises. However, our specific experimentation and implementation are based on ‘four’ levels of decomposition which is a consequence of selection of UPOL as the original database.

### 3.4. Selecting combinations of iris images

Recall that any  $256 \times 256$  iris image can be decomposed into five new components. The reverse is also true, a new  $256 \times 256$  iris image can be created from five components, due to the properties of multiresolution. For each newly synthesized iris image, we select the five necessary components from those available to us in the original database. As a way of synthesizing new unique iris images, each of the components can be selected from a different iris. Given a database of  $N$  input images, we can decompose each of these images into their five components. Therefore, when synthesizing an iris image, we have  $N$  choices available for each of the necessary components which allows us to create a total of  $N^5$  possible combinations. The original  $N$  iris images are included in these new combinations, as they are recreated when each of the selected components is from the same iris image. Clearly, given even a reasonable small database to start with, our method can generate an exponential increase in size.

### 3.5. Classifying iris images

During validation of our initial results, it became clear that not all of the  $N^5$  combinations result in synthetic irises with a high level of realism. Irises are quite varied and the characteristics extracted from varied irises may not be compatible. This fact is also valid for fingerprints. However, fingerprint classifications allow synthesis algorithms to more effectively generate images. In a similar fashion, simple classifications for irises also help to separate the irises into compatible groups and subsequently improve the realism of the results. Based on our observations, there are two distinct types of iris images. Some iris images have two areas where the thickness and position of the visible lines have a subtle difference (Fig. 8(c) and (d)). In the second set of irises, the lines are uniform throughout the iris, although the color of the iris may change slightly (Fig. 8(c) and (d)). Therefore, we classify irises into two types: single area or double areas. Although, our framework allows us to combine components from both kinds of irises, the resulting synthesized iris may become blurry as shown in Fig. 9. To reduce the number of blurry irises, only components from compatible irises (from the same group) are combined.

Grouping databases into sets of single area or double area irises can be done either manually (when the original size of the datasets is small) or automatically using image segmentation techniques [4,18]. The query is simple: whether each iris belongs to the single or double area set. In addition, the error in segmentation is not very critical and may only increase the blurriness of some irises. Our selected database was small and therefore a manual classification was employed.

Once one set of iris images have been classified, the possible combinations of irises that can be made from this small set is very large. If, for example, only eight iris images are available, a

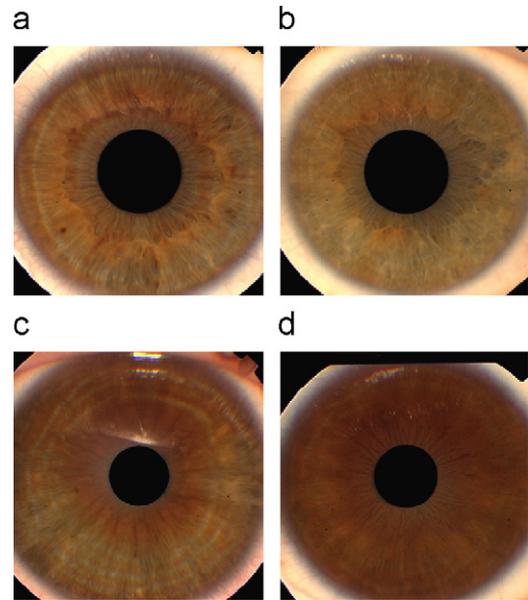


Fig. 8. Classifications of iris images, (a,b) with two rings of similar frequency characteristics and (c,d) with one ring of similar frequency characteristics.



Fig. 9. Example iris which is blurry due to incompatible components.

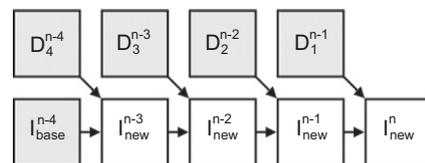
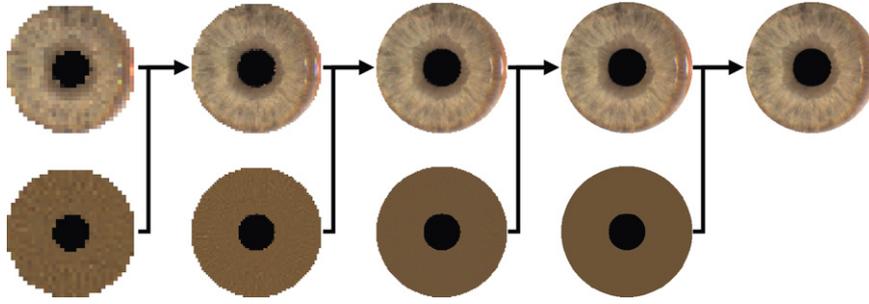


Fig. 10. Structure of the combination algorithm. Items with darker background represent components extracted from real iris images, components with white background represent synthetic iris images at varying levels of resolution.

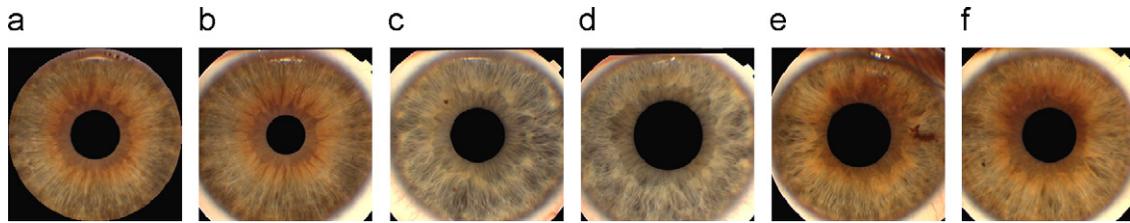
possible  $8^5 = 32768$  images can be created. Although it would be possible to automate the classifications, the combinatorial possibilities of this method make this automation a secondary goal.

### 3.6. Combining the details to form a new iris

Our combination algorithm requires five images from the database as input; using the subscript notation to identify each of the five images we have:  $I_0$ ,  $I_1$ ,  $I_2$ ,  $I_3$ , and  $I_4$ . From these images, we extract the five necessary components to build a complete iris image (e.g. for  $I_1$  the hierarchy is  $I_1^{n-1}$ ,  $I_1^{n-2}$ ,  $I_1^{n-3}$  and  $I_1^{n-4}$ ). First, the base color information for the new iris image  $I_{new}^{n-4}$ , is directly



**Fig. 11.** The components used to synthesize an iris and the final output iris. The left-most column is the initial set of details and coarse approximation extracted from the original irises. The subsequent columns have the results of the reconstruction process on the top and another set of details on the bottom. The right-most column is the final output iris image.



**Fig. 12.** Synthesized iris beside real iris samples used to create it. (a) The synthesized image, (b) the base iris image,  $I_0$ , (c–f) the irises used for details,  $D_1^{n-1}$ ,  $D_2^{n-2}$ ,  $D_3^{n-3}$ , and  $D_4^{n-4}$ , respectively.

extracted from  $I_0$  by using the low resolution approximation of  $I_0$ :  $I_0^{n-4}$ . Then the characteristics,  $D_4^{n-4}$ ,  $D_3^{n-3}$ ,  $D_2^{n-2}$ ,  $D_1^{n-1}$  are each extracted from one of the other four images. The process starts by combining the base color information,  $I_{new}^{n-4}$ , with the lowest resolution set of details,  $D_4^{n-4}$  that results in a new, intermediate resolution iris image,  $I_{new}^{n-3}$ . Similarly, we obtain the higher resolution images,  $I_{new}^{n-2}$ ,  $I_{new}^{n-1}$ , and  $I_{new}^n$  by iteratively applying

$$I_{new}^j = \mathbf{P}^j I_{new}^{j-1} + \mathbf{Q}^j D_{n+1-j}^{j-1} \quad j = n-2, n-1, n \quad (5)$$

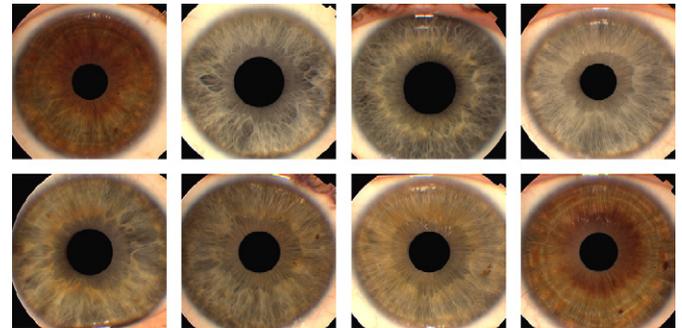
**Fig. 10** provides a visual overview of the composition of the new iris image, and **Fig. 11** shows the construction of an actual iris from synthetic components. The use of details from multiple real iris images provides the synthesized iris with both a unique and a realistic combination of characteristics as contained in the real irises.

#### 4. Experimentation

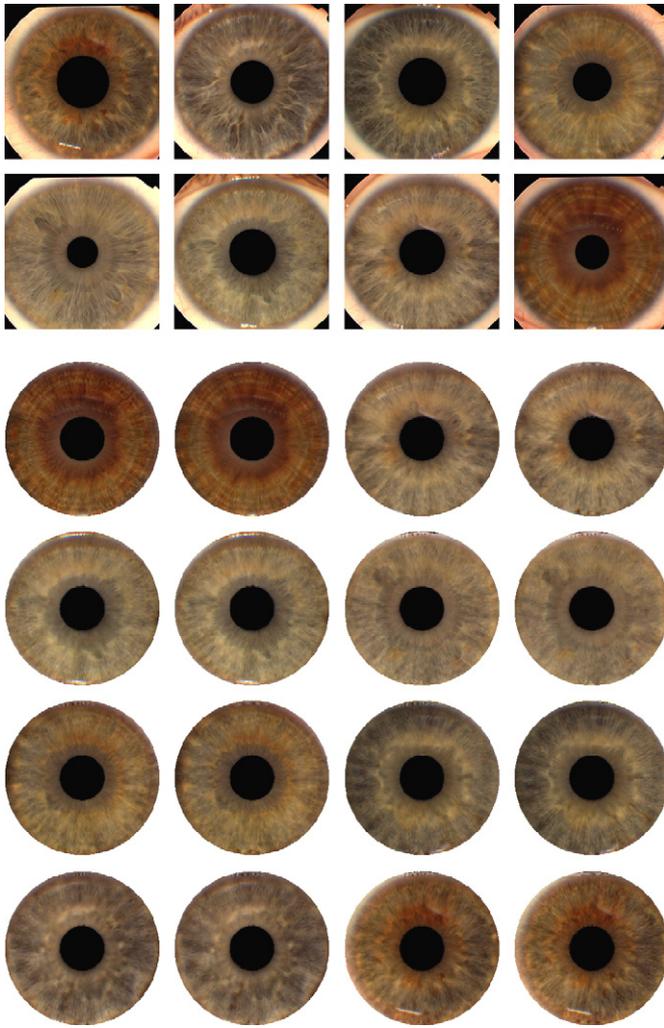
The goal of this method is to increase the size of existing iris image databases with realistic, synthetic irises. For the main validation we used an iris-recognition algorithm to test the individuality of the produced iris images when compared to the original iris images. This validation confirms that our synthetic images do not vary the characteristics of the database in a manner which allows us to determine that it contains synthetic iris images. In addition to such a validation, we highly value the realism of the synthetic irises. In order to evaluate the realism, we individually considered each synthetic image in terms of quality and uniqueness. For this approach to be tractable, we did not generate an exhaustive set of iris images. We restricted the method to a small number of irises which are a good distribution of the possible combinations.

#### 5. Realism

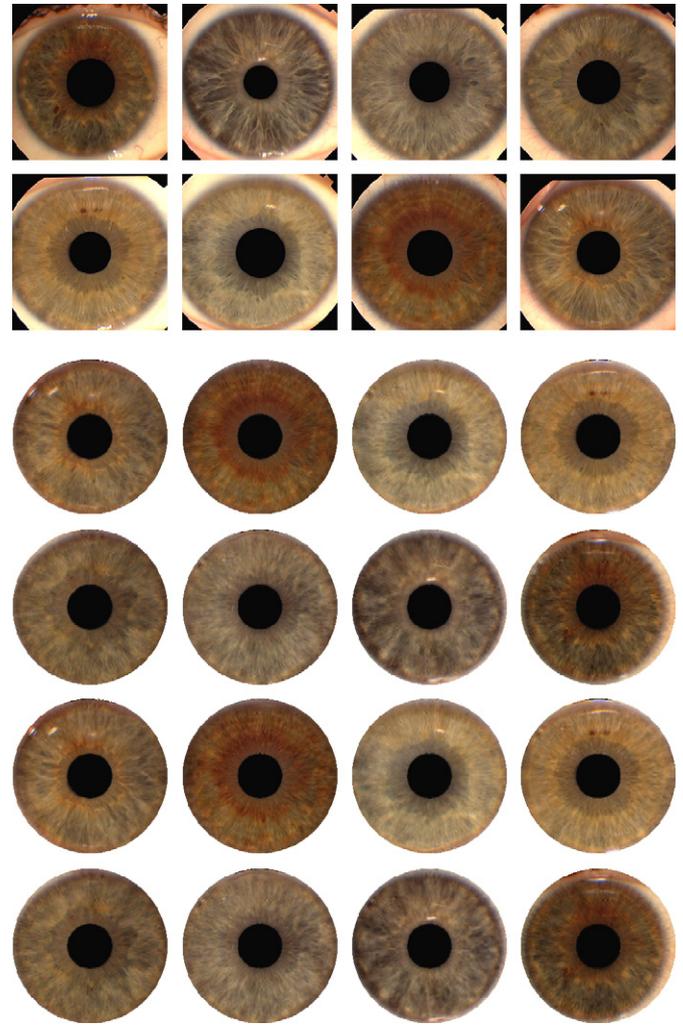
**Figs. 12–15**, each contain eight iris images from the original database followed by 16 out of the 70 synthetic irises generated from the originals. The synthetic images demonstrate the high



**Fig. 13.** Eight original irises from the UPOL [8] database. Synthetic iris images generated by our algorithm in the bottom box by using combinations of the original eight.



**Fig. 14.** Eight original irises from the UPOL [8] database. Synthetic iris images generated by our algorithm in the bottom box by using combinations of the original eight.



**Fig. 15.** Eight original irises from the UPOL [8] database. Synthetic iris images generated by our algorithm in the bottom box by using combinations of the original eight.

level of realism achievable with our method: the stroma, collarette and pupil are clearly present and the coloring of the iris has a realistic nature. The synthetic images display reflections similar to the ones found in the original irises (Fig. 1(b)). The multiresolution method extracts all characteristics from an image and cannot distinguish between the reflection and the underlying iris characteristics. The reflections add to the realistic nature of the resulting images. Future work may investigate methods to selectively include or exclude highlights such as these reflections.

### 5.1. Database augmentation

In the first test of our method, we selected two sample databases of irises from the existing UPOL database. The first database consisted of all of the UPOL images with duplicates removed; leaving a total of 128 iris images. We named this database *Complete*. The second database is a copy of the first database but is separated into two groups. The first group contains the irises that have two rings of distinctly different characteristics. The second group contains the irises that have one ring of characteristics. We called this second database *Grouped*; it also consists of 128 iris images. Our method was used to augment each of these two databases.

**Table 1**

Grouping compatible irises together increases the percentage of iris that are visually realistic.

Database	Irises created	Visually realistic irises	Ratio (%)
Complete UPOL	1120	1025	91.5
Grouped UPOL	1120	1098	98.0

To fully investigate the realism of our results, we needed to consider each synthesized iris image individually in terms of quality. We also needed to test the method's ability to significantly increase the number of images in a particular database. However, as the number of output images increases, our ability to individually consider each output image significantly decreases. Therefore, we decided not to generate the exhaustive set of possibilities. Rather, we generated a smaller subset that represents a good distribution of all possible combinations. To achieve this goal we exclude the component used as the base color from the selection process, reducing the possibilities to  $N^4$ , where  $N$  is the number of irises in the database. The component to be used for the base color information is then selected uniformly from the original set of irises that provides an even distribution of the color schemes amongst the generated

irises. Next, we use only the unique combinations of the remaining four components, rather than the ordered permutations. This further reduces the number of images to the unique combinations of four iris images chosen from  $N$ . In our case, each database will contain 128 iris images. Even with the reduced number of possibilities, using our method on a database of this size will generate approximately 10 million iris images. To reduce this to a reasonable number, we separate each database into groups of eight images, and consider each group as a separate database. With each of these databases we obtain 70 new iris images from a unique combinations of four irises from eight. Therefore, the total number of iris images synthesized is reduced to

$$\frac{N}{8} \times 70 \quad (6)$$

or in our case: 1120 synthesized irises.

Next, the synthesized iris images were examined and rated according to the clarity and realism of the characteristics. We also used the uniqueness of the synthetic iris when compared with those used to create it. This visual process required that we learn

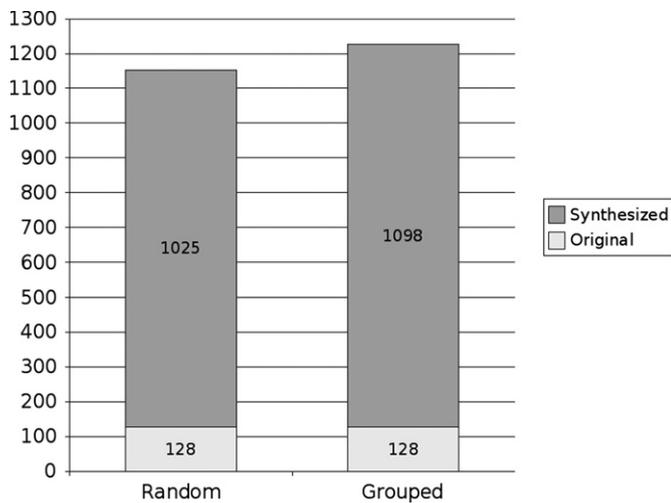


Fig. 16. Original database size and the increase provided by our augmentation method.

the patterns and characteristics found in the real irises from the database in order to distinguish the sometimes subtle differences between two similar irises. Selecting which irises to use when creating a new iris is an important step in the process as it influences the quality of the resulting images. Results show that the classification process improves the overall quality of the synthetic irises, as seen in Table 1.

Fig. 1 contains example synthesized irises images. The synthesized images display similar characteristics and color patterns to real irises resulting in very realistic iris images. In addition, each synthetic iris is significantly different from the irises used to create them and therefore are unique iris images that can be used in addition to the existing irises in the database.

Not only does the algorithm produce proper individual results but also it can significantly increase the number of unique iris images contained in the database, as the chart in Fig. 16 demonstrates. Our implementation did not generate an exhaustive set of all possible irises for logistical reasons, yet still made a significant increase to the UPOL database by providing 1098 new unique and realistic iris images.

## 5.2. Usefulness for iris recognition testing

In addition to having a realistic appearance, which is the main focus of this paper, the augmented database should have statistical properties similar to those of the original database. Each synthetic iris image is a new iris within the augmented database, therefore only inter-class comparisons can be performed. We use the Daugman iris recognition algorithm as provided by Masek and Kovesi [16] for these comparisons. The Daugman algorithm reduces each iris to a unique binary iris code and calculates a Hamming distance between two irises. Fig. 17 and Table 2 contains the resulting distribution of hamming distances as calculated within the original UPOL database and the augmented database.

As can be seen from the charts, the two distributions are close to each other. The largest differences being a 10% shift in the 0.45–0.50 range, and a 4% shift in the 0.00–0.05 range. Despite the differences on the boundaries, the results demonstrate that the augmented database maintains a satisfactory statistical similarity to the original database while significantly increasing its size.

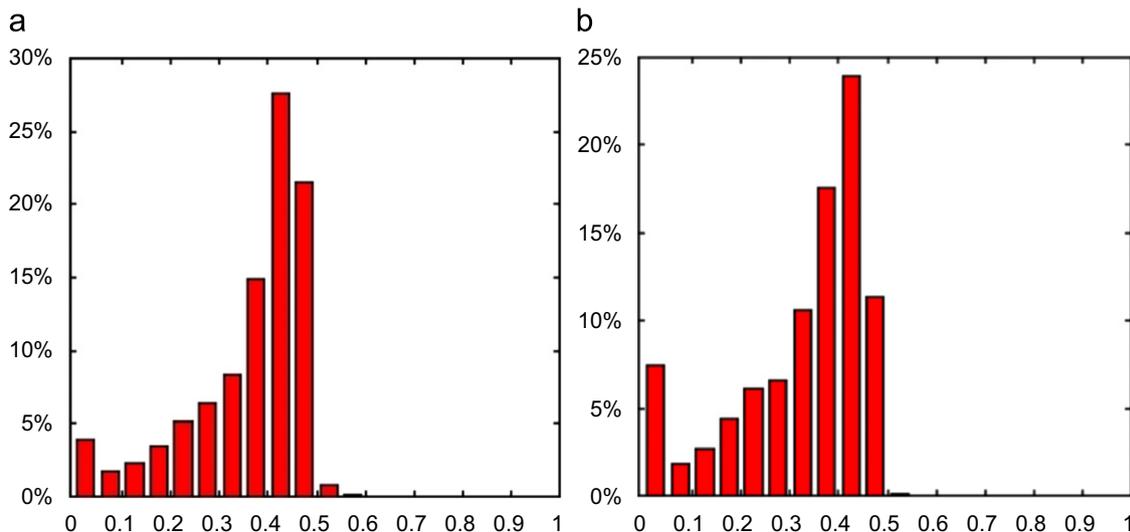


Fig. 17. Hamming distances for inter-class comparisons of (a) the original UPOL database and (b) the UPOL database augmented with our irises.

**Table 2**

A table illustrating the percentage of inter-class comparisons in our testing of the synthesis method.

Range	Original UPOL (%)	Augmented UPOL (%)
0.00–0.05	4	8
0.05–0.10	2	2
0.10–0.15	2	3
0.15–0.20	4	5
0.20–0.25	5	7
0.25–0.30	7	7
0.30–0.35	9	11
0.35–0.40	15	19
0.40–0.45	29	26
0.45–0.50	22	12
0.50–0.55	1	0
0.55–0.60	0	0
0.60–0.65	0	0
0.65–0.70	0	0
0.70–0.75	0	0
0.75–0.80	0	0
0.80–0.85	0	0
0.85–0.90	0	0
0.90–0.95	0	0
0.95–1.00	0	0

## 6. Conclusions and future work

In this paper, we presented a novel synthesis technique to augment existing iris image databases with new and unique iris images. We have shown how an iris can be isolated from surrounding anatomy by using a polar transform and normalization technique. Using reverse subdivision, we have illustrated how individual iris images can be decomposed into components that represent iris characteristics. Our technique allows characteristic components from multiple irises to be combined forming new unique iris images. We further increase the effectiveness of this technique by organizing the pre-existing irises into compatible groups resulting in synthetic images with greater realism.

Experimentation with statistical properties verified that the augmented database has a satisfactory statistical similarity to that of the original database. However, the experimentation uses only one of the many statistical properties of each database and more statistical properties may provide further insights. Future work may also investigate the effects of the number of levels of resolution and the choice of multiresolution system on the resulting statistical properties. It is particularly useful to find acceptability criteria for achieving better statistical similarity properties between the augmented and the original databases. Since the general approach proposed in this work can create a huge database of synthesized iris images, it gives enough room for filtering statistically problematic irises.

## Acknowledgements

Authors would like to acknowledge Biometric Technologies laboratory, University of Calgary, for providing access to necessary equipment and iris recognition software for testing. Authors would also like to acknowledge support of CFI and NSERC Granting Agencies for their partial support of this project. Last but not least, the authors would like to thank Mike Simoens and Javad Sadeghi. Their editorial contributions to the final drafts are greatly appreciated.

## References

[1] Brosz J, Samavati FF, Costa Sousa M. Terrain synthesis by-example. In: GRAPP '06: international conference on computer graphics theory and applications.

INSTICC—Institute for Systems and Technologies of Information, Control and Communication. Setúbal, Portugal, 2006. p. 122–33.

[2] Brunn M, Costa Sousa M, Samavati FF. Capturing and re-using artistic styles with reverse subdivision-base multiresolution methods. *International Journal of Image and Graphics*, to appear.

[3] Cappelli R, Maio D, Maltoni D. Synthetic fingerprint-database generation. In: *Proceedings of the 16th international conference on pattern recognition*, vol. 3. Los Alamitos, CA, USA: IEEE Computer Society; 2002. p. 744–7, doi:10.1109/ICPR.2002.1048096.

[4] Carson C, Belongie S, Greenspan H, Malik J. Blobworld: image segmentation using expectation-maximization and its application to image querying. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1999;24:1026–38.

[5] CASIA. Casia iris image database, <http://www.sinobiometrics.com>; 2004.

[6] Cui J, Wang Y, Huang J, Tan T, Sun Z. An iris image synthesis method based on pca and super-resolution. In: *Proceedings of the 17th international conference on pattern recognition, ICPR 2004*. Los Alamitos, CA, USA: IEEE Computer Society; 2004. p. 471–4.

[7] Daugman J. How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology* 2004;14:21–30.

[8] Dobeš M, Machala L. Iris database, <http://www.inf.upol.cz/iris/>; 2005.

[9] Foster K, Costa Sousa M, Samavati FF, Wyvill, B. Reverse subdivision multiresolution for polygonal silhouette error correction. In: *Lecture notes in computer science*, vol. 3045, 2004. p. 247–59.

[10] Gavrilova ML. Algorithms in 3D real-time rendering and facial expression modeling. In: *3A'2006 plenary lecture, eurographics*. Limoges, France, May 2006. p. 5–8. ISBN 2-914256-08-6.

[11] Gavrilova ML. Computational geometry and image processing techniques in biometrics: on the path to convergence. In: *Image pattern recognition: synthesis and analysis in biometrics*. Hackensack, NJ, USA: World Scientific Publishers; 2007. p. 103–33.

[12] Huang Y, Luo S, Chen E. An efficient iris recognition system. In: *Proceedings of the IEEE*, 1997. p. 1348–63.

[13] Schmid NA, Zuo J, Chen X. On generation and analysis of synthetic iris images. *IEEE Transactions on Information Forensics and Security* 2007;2:77–90.

[14] Lefohn A, Caruso R, Reinhard E, Budge B, Shirley P. An ocularist's approach to human iris synthesis. *IEEE Computer Graphics and Applications* 2003;23:70–5.

[15] Luo Y, Gavrilova M. 3D facial model synthesis using voronoi approach. In: *IEEE-CS proceedings, ISVD*. Los Alamitos, CA, USA, July 2006. p. 132–7.

[16] Masek L, Kovesi P. Matlab source code for a biometric identification system based on iris patterns. *The School of Computer Science and Software Engineering, The University of Western Australia*; 2003.

[17] Netravali AN, Haskell GG. *Digital pictures: representation, compression, and standards*. 2nd ed. New York: Plenum Press; 1995.

[18] Pal SK, Pal NR. A review on image segmentation techniques. *Pattern Recognition* September 1993;26:1277–94. ISSN 0031-3203.

[19] Proena H, Alexandre LA. Ubriris: a noisy iris image database. In: *Proceedings of ICIAP 2005—international conference on image analysis and processing*, vol. 1. Berlin, Heidelberg: Springer; 2005. p. 970–7. ISBN 3-540-28869-4.

[20] Sabourin R, Correia SEN, de Carvalho JM. On the performance of wavelets for handwritten numerals recognition. In: *Proceedings of the 16th international conference on pattern recognition*, vol. 3. Los Alamitos, CA, USA: IEEE Computer Society; 2002. p. 127–30.

[21] Sadeghi J, Samavati F. Smooth reverse subdivision. *Computer & Graphics* 2009;33(3):217–25.

[22] Samavati FF, Bartels RH. Multiresolution curve and surface representation by reversing subdivision rules. *Computer Graphics Forum* 1999;18(2):97–120.

[23] Samavati FF, Bartels RH. Local filters of B-spline wavelets. In: *International workshop on biometric technologies*, 2004. p. 105–10.

[24] Samavati FF, Bartels RH, Olsen L. Local B-spline multiresolution with examples in iris synthesis and volumetric rendering. In: *Synthesis and analysis in biometrics*. Hackensack, NJ, USA: World Scientific Publishing; 2007. p. 65–101.

[25] Stollnitz E, Deroose T, Salesin D. *Wavelets for computer graphics*. San Francisco: Morgan Kaufmann; 1996.

[26] Taerum T, Costa Sousa M, Samavati FF, Chan S, Mitchel R. Real-time super resolution contextual close-up of clinical volumetric data. In: *Proceedings of eurographics/IEEE-VGTC symposium on visualization*. IEEE; 2006. p. 347–54.

[27] Van Verth JM, Bishop LM. *Essential mathematics for games and interactive applications*. San Francisco: Morgan Kaufmann; 2004.

[28] Wang C, Luo Y, Gavrilova ML, Rokne J. Fingerprint image matching using a hierarchical approach. In: *Computational intelligence in information assurance and security*. Springer SCI Series; 2007. p. 175–98.

[29] Wayman J, Jain A, Maltoni D, Maio D. *Biometric systems: technology, design and performance evaluation*. Springer; 2005.

[30] Wecker L, Samavati FF, Gavrilova M. Iris synthesis: a multiresolution approach. In: *GRAPHITE '05: Proceedings of the third international conference on computer graphics and interactive techniques in Australasia and South East Asia*. New York, NY, USA: ACM; 2005. ISBN 1-59593-201-1.

[31] Wecker L, Samavati FF, Gavrilova M. Contextual void patching for digital elevation model. *The Visual Computer* 2007;23(9–11):881–90.

[32] Wildes R, Asmuth J. A system for automated iris recognition. In: *Proceedings of the second IEEE workshop on application of computer vision*. Los Alamitos, CA, USA, 1994. p. 121–28.

[33] Wildes RP. Iris recognition: an emerging biometric technology. In: *Proceedings of the IEEE*, vol. 85, 1997. p. 1348–63.

- [34] Yang LH, Feng L, Tang YY. A wavelet approach to extracting contours of document images. In: Proceedings of the fifth international conference on document analysis and recognition. Los Alamitos, CA, USA: IEEE Computer Society; 1999. p. 71–4.
- [35] Yanushkevich S, Stoica A, Shmerko V, Popel D. Book chapter. In: Biometric inverse problems. Boca Raton, FL, USA: Taylor & Francis/CRC Press; 2005.
- [36] Yanushkevich S, Wang P, Gavrilova M, Srihari S. Image pattern recognition: synthesis and analysis in biometrics. World Scientific Publishers; 2007.
- [37] Zhu Y, Tan T, Wang Y. Biometric personal identification based on iris patterns. In: Proceedings of 15th international conference on pattern recognition, vol. 2. Los Alamitos, CA, USA: IEEE; 2000. p. 801–4.