THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

# *Image Processing and Computer Vision*

Virginia Negri

Biometrics Research Report

May 31, 2019

# Contents

# 1 Introduction

This work is focused on how to achieve a robust biometrics system by combining several cutting edge technologies.
Biometrics is a science and technology used for identification and access control based on human physical or behavioural characteristics [7]. A biometric system is built in two phases: enrolment and authentication. The first phase consists in generating a database of templates, that are compressed representations of each individual's identifying features. A database never collects a whole image or fingerprint of a user, but adopts an efficient representation that captures all the essential features necessary to guarantee an accurate identification. The second phase consists in collecting new data from a user wanting to access the system and matching it to a template in the database. The most common and accurate systems implemented today are those based on face, iris and fingerprint recognition. Precisely, iris and face recognition are considered to be the state of the art technologies in biometrics as their accuracy and precision rates have no comparison with other systems. While face recognition is the most natural and easily verifiable method for identification, iris recognition can reach even higher levels of precision [3]. However, while both methodologies succeed in reaching high performances in recognition, both have several issues related to the measurement environment, the distinction of twins and the computational expenses.
Instead, as it is highlighted in [9], by combining together multiple methodologies it is possible to overcome unimodal issues such as noise, non-universality and spoof attacks.
Several works have been performed to solve these issues by combining these technologies at different levels. The addressed systems are so-called multibiometric systems. The sources of biometric information to establish the identity of an individual are multiple and are combined at different levels such as sensor-level, feature-level, score-level, rank-level or decision-level [6].

This report will thus focus on how it is possible to build a robust multimodal system that can outperform state of the art unimodal ones and overcome their critical issues. As explained in [5] multimodal systems "are expected to be more reliable due to the presence of multiple, independent pieces of evidence". To tackle this problem the report will be organised as follows. First of all an analysis on unimodal recognition systems based on the face and iris will take place. Then different methodologies for multimodal systems will be proposed, precisely these will be simple and hierarchical bimodal systems and a combination of multimodal biometrics and cryptography. Finally there will be a discussion on the results of these methods and what can be argued about the proposed systems.

# 2 Issues

## 2.1 Face recognition

A face recognition system is a biometric technology able to recognise different individuals based on the features of their face. It is one of the most preferable recognition systems as it has pretty low false acceptance rates, it is human verifiable and can work in non cooperative situations. However, as mentioned in [9] the most critical issues concern image acquisition and the change of appearance over time, of posture and expression that can significantly decrease the system's reliability. Furthermore, while being quite accurate with high definition pictures such as those in passports and used in Smart Gates, problems arise with low quality images such as those coming from cctv. Finally, face recognition has issues in distinguishing twins from one another. This problem can in fact be solved only through fingerprints or iris recognitions as they are randomly generated and do not depend on the genome. Iris especially is the most accurate at twins distinction. As C. N. Devi explains in [2] "the iris patterns are different even in genetically similar sources such as the right and left eyes of a single person and the eyes of twins". Thus an impostor can immediately be recognised and rejected during authentication.

## 2.2 Iris recognition

Iris recognition is considered to be the most reliable and accurate biometric technology. It works by locating and segmenting the iris from a human face and producing an appropriate mathematical encoding of the pattern captured through infrared illumination. From this pattern it is possible to extract the most relevant features that distinguish that particular iris. While guaranteeing the lowest false acceptance rates, being able to work in different lighting conditions and able to tell twins apart, it still comes with different issues. Among these the most relevant are the need for user cooperation and the cost of both image acquisition and model training. As mentioned in [9] a full collaboration of users is needed during image acquisition and this involves no movement, background fixes and predefined acquisition angles and illumination conditions. Furthermore the whole equipment is expensive.

# 3 Methods

## 3.1 Simple multimodal biometric

One first approach to build an efficient multimodal biometrics system is proposed by Y. Bouzouina et al. in [9]. The presented system requires first using two distinct unimodal systems, precisely face and iris recognition, and finally merging the two previous results to obtain a unique score. An effective face recognition system can be achieved by combining both global and local feature extractors such as PCA and DCT as show in [9]. Similarly iris features can be extracted by using both global and local methods such as respectively 1D log-Gabor filter and Zernike moment. Finally all extracted features, after being properly normalised in a same representation space, are used to compute a score to test against the database templates. The following image shows the general workflow of the system:

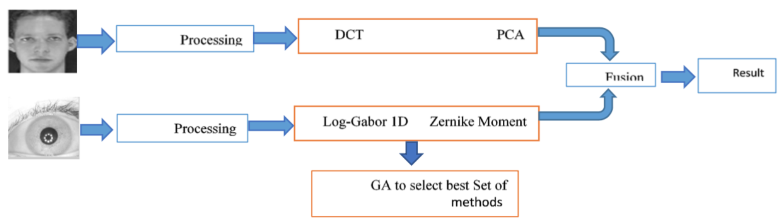



Figure 1: Block diagram of the proposed scheme in [9]

First of all face and iris recognition are performed separately. An accurate face recognition system must integrate both global and local feature extractors. The first are mostly statistical based and used to represent a face in a space of lower dimensionality. Local features are extracted by the face's morphology and point to singular details. Y. Bouzouina et al. in [9] suggest using for the two respectively PCA and DCT. Principal Component analysis is a machine learning technique that uses a statistical approach to reduce the number of variables and represent only relevant information. As explained in [4] "every image in the training set is represented as a linear combination of weighted eigenvectors called eigenfaces". Eigenfaces are obtained by the covariance matrix of the training data while the weights are computed after selecting the most relevant features. After having stored all training images in the newly compressed format a new image can be tested against the database by projecting it onto the subspace spanned by the eigenfaces [4]. In a unimodal system the best match could already be computed as suggested by [4] using Euclidean distance as a metrics. The following script provides a brief example of how PCA works:

```matlab
%% Initialize variables
I = [im1,im2,im3,im4,im5,im6];
N = 6;
[rows cols]=size(im1);


%% Compute Average image
average_face = zeros(size(im1));
for i=1:6
    average_face = average_face + I(:,(i*128 -127):(i*128));
end
average_face = average_face/6;

%% Normalize training images
normalized_images = [];

for i=1:N
    norm_img = I(:,(i*128 -127):(i*128)) - average_face;
    normalized_images = [normalized_images, norm_img];
end

%% Reduce images to column vectors
D = [];
for i=1:N
    d = [];
    img = normalized_images(:,(i*128 -127):(i*128));
    for row=1:rows
        d = [d,img(row,:)];
    end
    % Get normalized images as column vector
    d = transpose(d);
    D = [D,d];
end

%% Covariance matrix and SVD
```

```matlab
%C = D*transpose(D);
C_prime = transpose(D)*D;

% U and V containg singular vectos
% S contains singular values
[U,S,V] = svd(D);


%% Calculate eigenvalues of C'
% Square of singular values = eigenvalues in first n rows/columns
eig_values_C_matrix = S*transpose(S);

% List of eigenvalues of C
eig_values_C = [];
for i=1:N
    eig_values_C = [eig_values_C,eig_values_C_matrix(i,i)];
end

%% Calculate eigenvectors of C
% List of eigenvectors of C
eigenvectors = [];

for i=1:6
    ui = (1/sqrt(eig_values_C(1,i)))*(D*V(:,i));
    % Add eigenface as column vector to matrix of eigenfaces
    eigenvectors = [eigenvectors,ui];
end
```

```matlab
%% Load new image to classify
% The input image should already be in grayscale
imX = inputImage;

%% Normalize image as column vector and get average face column vector

% Subtract average face
imX = imX - average_face;

% Get column vector of normalized and average images
normalised_imgX = [];
average_face_vector = [];
for row=1:rows
    normalised_imgX = [normalised_imgX,imX(row,:)];
```

```matlab
    average_face_vector = [average_face_vector,average_face(row,:)];
end
normalised_imgX = transpose(normalised_imgX);
average_face_vector = transpose(average_face_vector);

%% Project input image onto eigenvector space
p = [];
eigs=size(eigenvectors,2);
for i = 1:eigs
    projection = dot(normalised_imgX,eigenvectors(:,i));
    p = [p; projection];
end
```

```
% Principal components of the training set
principal_components = transpose(eigenvectors)*D;
% Input image score
score = transpose(eigenvectors)*normalised_imgX;

% Distances of input image from principal components
distances = [];
minDist = 0;
minDistIndex = 1;
for i=1:N
    column = principal_components(:,i);
    d = 0;
    for j=1:N
        d = d + (column(j,1)-score(j,1))^2;
    end
    d = sqrt(d);
    distances = [distances, d];
    % Get class with minimum distance
    if(i==1)
        minDist = d;
    else
        if(d<minDist)
            minDist = d;
            minDistIndex = i;
        end
    end
end

class = minDistIndex;
outputImage = I(:,(class*128 -127):(class*128));

end
```

Local features instead can be extracted using DCT algorithm that consists in transforming an image from spatial to frequency representation. As explained in [9] "the low-frequency part is the most visually significant region, and the higher frequencies represent the finer details".

Iris recognition requires first of all iris segmentation and secondarily feature extraction and encoding. Y. Bouzouina et al. in [9] suggest using active contours for iris segmentation. Active contours, also called Snakes algorithm, consists in identifying an object's boundary that is shaped by two predominant forces: an internal and an external one. In [9] the initial boundary is found after an application of the Hough Circle transform to locate the pupil's circle. Snakes is then used to find both contours of pupil and iris. An active contour is defined by its energy through the following equation:

$$E_{\text{snake}} = \int_0^1 E_{int}\left(v(s)\right)ds + \int_0^1 E_{ext}\left(v(s)\right)ds$$

Where the following hold:

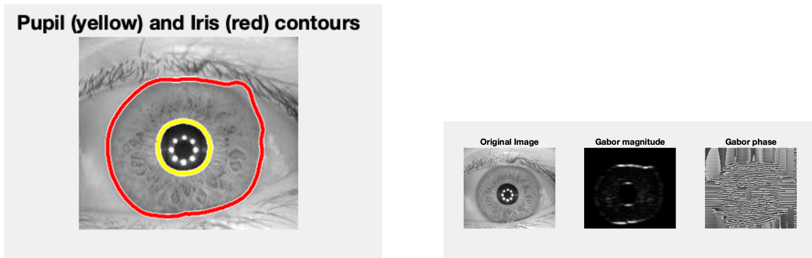
$$E_{\text{int}}(\text{v(s)}) = \frac{1}{2}\ \alpha(s)|v(s)|^2 + \frac{1}{2}\ \beta(s)|v(s)|^2$$

$$\text{E}_{\text{ext}}(\text{v(s)}) = |\nabla(G\sigma(\text{x,y})*I(x,y))|^2$$

C. N. Devi in [2] suggests instead using an integrodifferential operator to obtain the iris's boundaries by looking at the maximum variation in pixels. This is more effective if preceded by histogram equalisation to enhance the image's contrast.

Finally, having both pupil and iris contours the iris's segmentation can be performed.
Feature extraction and encoding follow segmentation. After normalising the iris into a fixes size, for example by representing each pixel of the Cartesian domain in polar coordinates, in [9] "global features are extracted by employing 1D log-Gabor filter and the local features with Zernike moment". Among the extracted features only a subset must be chosen to store only the most relevant information and to guarantee a low computational complexity. [9] proposes a fast converging technique for feature selection that has been empirically proven effective: Genetic Algorithms.
Here below is a coding example implemented in MATLAB of an iris segmentation using active contours and subsequent feature extraction through a Gabor filter. The image derives from the popular CASIA-IrisV4 database, commonly used to test iris recognition systems.

```
clear all; close all; clc;

img = imread('eye3.jpeg');
img = rgb2gray(img);

compl = imcomplement(img);

figure
imshow(img);
hold on;

%% Pupil segmentation
% Snakes algorithm
mask1 = zeros(size(img));
mask1(90:end-60,80:end-90) = 1;

bw1 = activecontour(compl, mask1, 800, 'Chan-Vese', 'SmoothFactor',8);
visboundaries(bw1,'Color','y');

%% Iris segmentation

mask2 = zeros(size(img));
mask2(50:end-20,50:end-50) = 1;

bw2 = activecontour(compl, mask2, 800, 'Chan-Vese','SmoothFactor',8);
visboundaries(bw2,'Color','r');
title('Pupil (yellow) and Iris (red) contours');

hold off;

%% Segment the iris

irisSegmentedMask = bw2-bw1;
irisSegmented = img;
for i=1:size(img,1)
    for j=1:size(img,2)
        if(irisSegmentedMask(i,j)==0)
            irisSegmented(i,j)=0;
        end
    end
end
figure, imshow(irisSegmented);

%% Extract featues with Gabor filter
wavelength = 4;
orientation = 90;
[mag,phase] = imgaborfilt(irisSegmented,wavelength,orientation);
figure
subplot(1,3,1);
imshow(img);
title('Original Image');
subplot(1,3,2);
imshow(mag,[])
title('Gabor magnitude');
subplot(1,3,3);
imshow(phase,[]);
title('Gabor phase');
```




(a) Active contours of the iris and the pupil

(b) Gabor filter for feature extraction applied to the segmented iris

Figure 2: Iris feature extraction demo

Once both face and iris feature extractors have been applied, to find the best match of a new user that wants to authenticate in the system the two previous results must be combined to obtain a unique score. A first step of normalisation is required to represent all features in the same representation space. As [9] suggests, the best match can be found by applying SVM to combine the scores. SVM works by finding the decision boundaries by performing pattern recognition using the support vectors found in the training phase [9].

## 3.2 Hierarchical multimodal biometric

While the previous method combines face and iris features at the very end and extract the most significant ones to use for template matching, hierarchical multimodal biometrics has a preliminary phase in which different biometrics are combined. This system is proposed by X. Zhang et al. in [8]. It is mainly addressed to applications that have access to low quality images due to difficulties in environmental control. The training stage consists in learning a statistical mapping at pixel level between iris and face samples using canonical correlation analysis. Testing instead begins by identifying a subset of the gallery set that at pixel level best matches the new face. Finally, [8] explains that ”ordinal representation and sparse representation are performed on these candidate samples for iris recognition and face recognition respectively”.

The first step in the hierarchical method proposed in [8] is to define a statistical mapping between face and iris to store in the gallery. This is essential in systems where training images are captured in low quality. In fact, the more the training set grows and thus the number of classes addressing different individual grows, the more the overall performance decreases. To improve the performance, the number of classes must be reduced. [8] suggests to make an initial class distinction based on the face images and then learn an iris to face mapping. The great variation between faces due to posing and different expressions can be effectively decreased by through iris mapping. The result is a subset of gallery image to test against the probe face.

Both iris and face recognition are then performed on the resulting subset. Merging the results of these two can be performed either through features selection or through their scores. [8] suggests that greater performance can be achieved through score fusion, equally to how it is done in simple multimodal biometrics. In [8] scores are computed from sparse and ordinal representation using residuals and Hamming distance respectively for face and iris recognition.

### 3.3 Multimodal biometric and Cryptography

While combining different unimodal biometrics can significantly increase robustness and accuracy in recognition, one last piece of the puzzle is contrasting potential attacks targeted especially to biometric templates in the gallery database that can significantly break down the system's robustness. One approach to solve this issue is proposed by K. Oyetola Oluwadamilola et al. in [5]. It consists in using Advanced Encryption Standard (AES) and storing all templates generated by the hybrid biometrics system in a BLOB format in MySQL databases, encrypted using MD5.
The first phases consists of a multimodal system as the ones described in the previous sections to recognise individuals. [5] suggests using a hybrid system slightly different as those above as it uses fingerprint in combination with face recognition as opposed to iris. While recognition is normally performed through enrolment and authentication, a change is seen in how the data is stored. The templates generated through enrolment can in fact be securely stored into a database by using sophisticated encryption algorithms. These can either be symmetric or asymmetric. However, both require generating keys to encrypt and decrypt the data. Encryption keys should be randomly generated and not suggest anything about the template data. In crypto biometrics systems keys can be generated from the helper data derived from the original biometric templates [5]. This type of data does not reveal anything about the templates, but is a simple and immediate way to generate keys. An overview of crypto-biometric systems' workflow can be seen in the following diagram from [5].

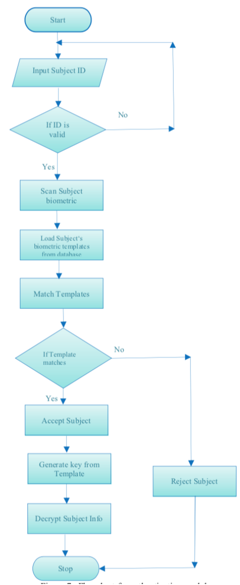



Figure 3: Flowchart for authentication module from Fig. 7 in [5]

# 4 Results

This section will focus on the results achieved by multimodal biometrics systems compared to unimodal ones.
All the systems described in the previous sections are intended to overcome several issues that can take place with unimodal systems such as noise, non-universality and issues related to specific systems. These include changes of the face appearance over time, twins distinction, change of posture and expression for face recognition systems and the requirement of user collaboration for iris recognition. A system combining these techniques can thus result in a more robust and accurate recognition.
There are several tools to effectively measure how well a multimodal system can perform with respect to a unimodal one by providing an objective mathematical result. These tools are typical of classification problems such as face template matching and are built around the values of the true positives, false acceptance and false rejection rates. As explained in [1] "a false positive is wrongly logging on an impostor and a false negative is refusing a valid user". True positives instead are the hit rate, that is the correctly classified users. One first result can be seen in the diagram below from [9]. The image below, derived from [9] plots ROC curves, that are the values of the true positives function of the false acceptance rate.

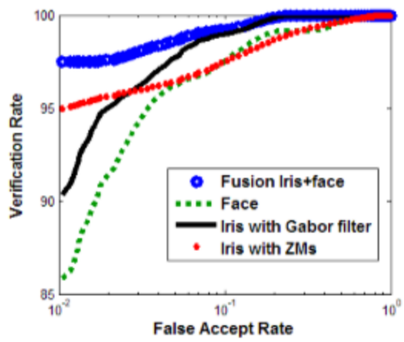



Figure 4: ROC curve of the simple bimodal system proposed in [9] compared to unimodal systems

The plot clearly shows how accuracy is significantly increased by combining iris and face methodologies. Furthermore, [9] explains how this method can "improve the stability and the performance of recognition during changes of pose, illumination and facial expression", thus overcoming some of the critical issues in unimodal systems.

Another result is shown in [8]. The proposed hierarchical method is compared to other unimodal systems again by showing the ROC curves.

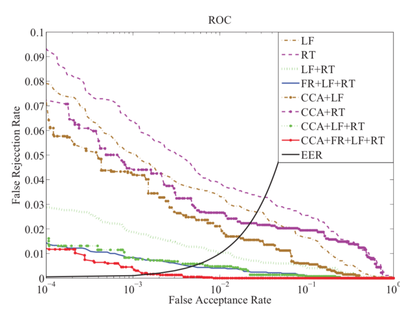



Figure 5: ROC curve of the unimodal systems compared to the hierarchical bimodal one proposed in [8]

In the plot LF and RT refer respectively to left and right iris recognition, FR to face recognition and CCA to the canonical correlation analysis, used to map iris to face at pixel level and chosen as regression model over sparse representation for its efficiency and ability to reduce variation in faces due to appearance changes. By following the ROC curves it is easy to see how the combination of iris and face allows lower false acceptance and false rejection rates compared to unimodal systems. Furthermore, using a hierarchical system allows to "learn a robust mapping from iris to face to perform coarse classification before score level fusion, which reduces the false alarm, while the other methods are all performed on the whole database" [8]. As mentioned in the previous section, this system is designed to be effective with low quality images. The combination of iris and face through a robust mapping allows in fact much greater performances with respect to unimodal systems working in uncontrolled environments.

Another important result of adding iris to face recognition is gaining the ability to distinguish twins between one another. As mentioned in the previous paragraphs, twins cannot be told apart with simple face recognition, but require a unique identifier such as fingerprints or iris which have patterns generated independently of the genome. By applying the described techniques for iris recognition, such as segmentation followed by normalisation and feature extraction, recognition results can reach significantly low error rates. C.N. Devi in [2] shows how with large templates a 0% FRR in both one to one and one to many recognitions between twins can be reached.

# 5 Discussion

This section will be devoted to reviewing the previous results and extend the discussion by pointing out a few issues and possibly thinking of feasible improvements.

This work was focused on demonstrating how it is possible to build a more robust and accurate authentication system by combining different biometric technologies, particularly those that guarantee lowest false acceptance rates such as face and iris recognition. Several techniques were explored such as individual unimodal feature extraction and subsequent fusion or a hierarchical approach. Furthermore, a system can be said to be robust if it also isn't vulnerable to attacks that can drastically compromise its performance. One possible solution to this is to use encryption algorithms to protect the template gallery used for authentication.

While performance rates result to be higher than those of unimodal systems and while many of the issues discussed at the beginning of the report have been solved, there are still a few key aspects to consider.

One significant issue in multimodal biometric systems is related to their cost. Precisely, the cost involves both the training phase and the authentication. First of all the use of a multimodal biometrics requires sophisticated equipment. If these systems were to substitute normal cctv cameras these would need to be integrated with iris detection sensors that are an expensive addition. Secondarily training requires a very large database in order to have the most accurate and robust system as possible. False acceptance rates in fact significantly decrease with template size growth as explained in [2]. C. N. Devi shows in fact how passing from a 10*40 to 100*400 sized template the memory requirement from 1.6kB reaches 160kB. It is evident that one must take into account what is the price to

pay to gain in accuracy.
Overall these types of systems are to be preferred when recognition accuracy needs to be maximised and is prioritised with respect to storage capabilities. By adopting the techniques described above these can be used also in environments where there is little control and where variations in lighting and poses frequently take place.

## References


[1] Ethem Alpaydin. *Introduction to Machine Learning.* Adaptive Computation and Machine Learning. Cambridge, MA : The MIT Press, third edition edition, 2014.

[2] C.N. Devi. Automatic segmentation and recognition of iris images: With special reference to twins. *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, pages 1–5, 2017.

[3] Jun-Feng Liu Jun-Ying Gan, Jian-Hu Gao. Research on face and iris feature recognition based on 2ddct and kernel fisher discriminant analysis. *2008 International Conference on Wavelet Analysis and Pattern Recognition*, 1:401–405, 2008.

[4] A. Al Sumam L. C. Paul. Face recognition using principal component analysis method. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 1(9):135–139, November 2012.

[5] K. Oyetola Oluwadamilola ; A. Okubanjo Ayodeji ; O. Osifeko Martins ; I. Sanusi Olufunmi ; O. Abolade Rapheal. An improved authentication system using hybrid of biometrics and cryptography. *2017 IEEE 3rd International Conference on Electro-Technology for National Development (NIGERCON)*, pages 457–463, 2017.

[6] Arun Ross. *Multibiometrics*, volume Encyclopedia of Biometrics. Springer US, Boston, MA, 2009.

[7] Margaret Rouse. What is biometrics? - definition from whatis.com.

[8] T. Tan X. Zhang, Z. Sun. Hierarchical fusion of face and iris for personal identification. *2010 20th International Conference on Pattern Recognition*, pages 217–220, 2010.

[9] L. Hamami Y. Bouzouina. Multimodal biometric: Iris and face recognition based on feature selection of iris with ga and scores level fusion with svm. *2017 2nd International Conference on Bio-engineering for Smart Technologies (BioSMART)*, pages 1–7, 2017.