VILNIUS UNIVERSITY

Justas Kranauskas

PERSON IDENTIFICATION BY FACE AND IRIS

Doctoral Dissertation

Physical Sciences, Informatics (09 P)

Vilnius, 2010

Doctoral dissertation was prepared at Vilnius University in 2005-2009.

Scientific Supervisor

Assoc. Prof. Dr. Algirdas Bastys (Vilnius University, Physical Sciences, Informatics - 09 P)

Abstract

In this thesis, person identification by combining automatic face and iris recognition is analyzed. Person identification by his face is one of the most intuitive from all biometric measures. We are used to recognizing familiar faces and confirming identity by a short glance at one's id card which contains image of the face. We are also used to being observed by surveillance cameras, which can perform biometric authentication without even being noticed. However, facial biometrics is one of most unstable metrics because the face gets noticeably older in several years and can frequently change depending on the mood of its owner. The core algorithm for facial recognition presented in this work is based on Gabor features. Deep analysis of each step helped to develop the method with better or similar accuracy to the best published results received on the same datasets, while being simple and fast.

On the other hand, person identification by his iris is one of the most sophisticated, stable and accurate biometrics. The core algorithm for iris recognition presented in this work is based on a novel iris texture representation by local extremum points of multiscale Taylor expansion. The proposed irises comparison method is very different from the classic phase-based methods, but is also fast and accurate. Combining it with our implementation of phase-based method results in superior recognition accuracy which is comparable or better than any published results received on the same datasets.

A combination of aforementioned algorithms was implemented and successfully tested in a recent Multiple Biometrics Grand Challenge Version 2 Portal Challenge experiments, where iris and face videos were captured simultaneously. As expected, recognition accuracy was significantly better when both biometrics were combined.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Research Area

Automatic person identification using biometric characteristics has huge potential and thousands of applications starting from everyday personal computer login applications and finishing by high security demanding systems like national level border cross control or an instrument to safeguard democratic elections. The most attractive technologies are those which do not need a contact to perform the biometric measurements and those which can capture person's biometrics without any interaction with the user.

In the last decade, facial recognition was thought to be a holly grail in biometrics because of huge amount of already installed surveillance cameras in the world and possibility to perform person identification even without informing him about that. While society was debating about the privacy concerns and countries were fighting terrorists by installing facial screening systems in all the major airports, scientists improved facial recognition accuracy by an order of magnitude [87]. However, major improvements were done mostly for facial images captured in controlled environment at high resolution which is not the case of typical surveillance cameras. That didn't reduced interest in person identification by face because new imaging technologies enabled getting face images at high enough resolutions. It's only the matter of time when these technologies will be installed everywhere. On the other hand, there are plenty of other applications which can benefit from face recognition.

Iris recognition was always marketed as most stable and accurate biometric measurement which can be used for person identification (except perhaps DNA). Additionally, similar to face, it can also be "measured" without a contact and interaction with a person.

The research area of this work is person identification by the biometric measurements (that can be done at a distance and on the move) - face and iris images, independently as unimodal and combined as multimodal biometric systems.

1.2 Problem Relevance

Face recognition is one of the most natural biometrics because humans are extremely good at recognizing persons by their faces. Recent studies showed that some face recognition algorithms already surpass humans at matching faces [79]. Nevertheless, high number of applications always demand better accuracy and robustness to such factors as changing illumination, expression, pose, aging, facial hair, glasses, etc. Emerging new imaging technologies suggested to use better image capturing devices which enabled capturing additional biometrics together with face - iris texture. Although, recognition algorithms for each of these biometrics are improved constantly, fusing them together results in even more significant boost in recognition accuracy.

Recently, several commercial systems which use this concept became available [3], [51], [93], [97] and several publications were also published [8], [38], [71], [119], mostly presenting previously mentioned commercial systems. However, they focus on iris recognition which is more stable and accurate than face recognition. Additionally, usually two irises are captured simultaneously, which further improves accuracy of iris only recognition systems. On the other hand, capturing irises at the distance and on the move must introduce additional noise, which is not present in typical iris recognition systems which capture irises in more constrained conditions. This results in lower than expected accuracy from the most stable and

accurate biometrics. Thus, we believe that in such case adding face modality will further improve accuracy of iris only recognition systems.

1.3 Research Object

This thesis research object is person identification at a distance using combination of face and iris biometric modalities which complement each other by being the least intrusive (face recognition part) and most accurate (iris recognition part) biometrics.

1.4 The Objectives and Tasks of the Research

The main objective of this thesis is to analyze current state-of-the-art face and iris recognition algorithms, and propose the improvements to increase recognition accuracy and speed. Furthermore, to implement core ready-to-use person identification algorithms for a multi-modal biometrics system utilizing the strengths of both biometric modalities.

In order to achieve these objectives the following tasks were formulated:

1. To analyze existing face and iris recognition methods.
2. To propose an improvement to a baseline face recognition from single image algorithm based on Gabor features both in speed and accuracy.
3. To evaluate impact of each face recognition step to identification and verification accuracy.
4. To create a real-time iris segmentation method.
5. To propose a new representation of iris texture based on local features.
6. To propose a new iris matching method which could complement traditional binary xor-based matching of phase-based iris texture representation.
7. To combine face and iris biometric modalities in one person identification system and evaluate its performance.

1.5 Scientific Novelty

The proposed face recognition method is built on traditional face identification from single image scheme: real-time face detection, followed by extraction of Gabor features and classic vectors' similarity measure for comparison of two features sets. Although, the smallest components of the algorithm are ordinary and they are often used in facial recognition, a unique combination with deep analysis of each component is created in this thesis:

- Geometric normalization - optimal resolution and cropping parameters for the proposed features were found.
- Photometric normalization - three different methods were analyzed. The first method does not influence face recognition accuracy if the proposed features are used. The other two are comparable and both significantly improve face recognition accuracy. It was shown that it is an essential step in facial recognition.
- Number of used features - an optimal density of regular grid (which controls number of used features) was found.

The proposed accelerated calculation of Gabor features enables real-time conversion of face image to facial features set. It is always faster than direct calculation and faster than convolution in frequency domain when image dimensions are not convenient for fast Fourier transform.

Gabor features similarity score was modified to use the same responses of Gabor filters in two complementary ways: by forming vectors along different directions and scales. The proposed modification to Gabor features similarity calculation can be straightforward extended to other object recognition tasks.

And finally, the proposed method is more accurate than best of the published algorithms, which do not require training step, by an order of magnitude, and even more, it is still more accurate than those algorithms, which are based on training step.

The proposed iris recognition method is built on proposed novel representation of texture - a set of local extrema of multiscale Taylor expansion.

Completely different from traditional ones iris matching method is proposed. It is designed to cope with inaccuracies of iris segmentation step. Combined with a real-time iris segmentation step it gives similar recognition accuracy as current state-of-the-art iris recognition algorithms.

Additionally, we showed how the proposed iris texture representation can be converted into traditional phase-based representation. If both methods are used simultaneously (this does not introduce additional complexity in feature extraction step), iris recognition accuracy is improved significantly and becomes higher than any published results on the tested recently created public iris datasets of different quality.

Finally, we showed how both algorithms can be easily fused at a score level and successfully used together in a complete person identification system to further improve recognition accuracy.

1.6 Practical Significance of the Work

The proposed face and iris recognition methods are ready to be deployed as core algorithms of person identification system either as single modality, or a combination of both biometric modalities. A modern personal computer can perform person identification based on the proposed methods in real time which is an important criteria for real-world applications.

1.7 Defended Propositions

1. Improved Gabor features extraction from facial image speed by an order of magnitude without any loss in the accuracy of Gabor filters responses. This enables to use Gabor features based recognition methods in real time.

2. Analyzed and proposed better parameters for each step of typical face recognition from single image method. This allowed to achieve identi-

fication and verification accuracy, which is comparable with accuracy of the methods which additionally incorporate a training step and require to have a representative training set of facial images.

3. Proposed Gabor features comparison method which reuses responses of Gabor filters along directions and scales. Thus, without any additional features improves identification and verification accuracy to the better than any known published results.

4. Proposed iris texture representation by a set of multiscale Taylor expansions. The proposed representation can be encoded in two different and complementary ways - traditional binary phase-based and novel extrema-based, thus, creating two different iris texture representations.

5. Proposed similarity metrics to compare the proposed iris texture representations encoded as a sets of two-dimensional extrema points. The proposes similarity metric can cope with minor errors in iris segmentation step and gives similar accuracy as existing state-of-the-art iris recognition methods.

6. Proposed a way to combine two different iris texture representations to further improve recognition accuracy.

7. Proposed a way to combine face and iris biometric modalities to improve person identification accuracy when iris images are captured at a distance because getting a face image at the same time is a cheap operation. The proposed face and iris recognition algorithms are similar in nature - they both compare texture, thus resulting similarities scores can be easily used in fusion of different biometric modalities at scores level. It was evaluated in Multiple Biometrics Grand Challenge Version 2.

1.8 Approval of Research Results

The results of the proposed methods were presented in the following workshops and conferences:

1. "The Multiple Biometrics Grand Challenge Workshop III", 4 December 2009, Washington (USA).

2. "The 3^{rd} IAPR/IEEE International Conference on Biometrics", 2-5 June 2009, Alghero (Italy).

This thesis is an essential continuation of previous scientific work performed by the author:

1. Justas Kranauskas (2005), "The Attributes of Faces Recognition", Master Thesis in the Faculty of Mathematics and Informatics, Vilnius University.

1.9 Publications

1. Justas Kranauskas: Accelerated Calculation of Gabor Features in Spatial Domain. Electronics and Electrical Engineering, no. 1 (97), pp. 39-44, 2010.

2. Algirdas Bastys, Justas Kranauskas, Rokas Masiulis: Iris Matching by Local Extremum Points of Multiscale Taylor Expansion. Advances in Biometrics, Third International Conference, ICB 2009, Alghero, Italy, June 2-5, Proceedings, vol. 5558, pp. 1070-1079, 2009.

3. Algirdas Bastys, Justas Kranauskas, Rokas Masiulis: Iris Recognition by Local Extremum Points of Multiscale Taylor Expansion. Pattern Recognition, vol. 42, no. 9, pp. 1869-1877, 2009.

4. Andrej Kisel, Alexej Kochetkov, Justas Kranauskas: Fingerprint Minutiae Matching without Global Alignment Using Local Structures. Informatica (Lithuanian Academy of Sciences), vol. 19, no. 1, pp. 31-44, 2008.

1.10 Thesis Structure

This thesis is organized as follows. Chapter 2 reviews face recognition from single still image problem. A survey is followed by an introduction to method of accelerated calculation of Gabor features in spatial domain. Gabor features make the foundation for a baseline face recognition algorithm, which is further significantly improved by deep analysis of each face recognition step - geometric normalization, photometric normalization, number and position of used Gabor features, and proposed similarity metric. The accuracy is evaluated on FERET - most popular face recognition protocol and dataset.

Chapter 3 reviews iris recognition from single still image problem. A survey is followed by the description of the proposed real-time iris segmentation method. Classic iris texture geometric normalization by rubber sheet method is presented. It is followed by introduction to proposed iris texture representation by local features as local extrema of multiscale Taylor expansion and a similarity metric for robust comparison of created features sets. Further, a way to convert the proposed iris texture representation into classic phase-based method is presented. Developed iris recognition algorithms are evaluated on largest publicly available iris datasets - CASIA and ICE. Finally, evaluation results are presented for all datasets.

Chapter 4 presents evaluation results for previously introduced face and iris recognition algorithms on the Multiple Biometrics Grand Challenge Version 2 Portal Challenge experiment. Iris localization method is extended to work with specific very high resolution NIR face images, which contain part of the face and usually one or two irises visible. This experiment requires fusion of biometric modalities. The results of fusion at score level are presented.

Chapter 2

Face Recognition

2.1 Introduction

Roughly, face recognition problem can be divided into two categories: the *one sample problem* and the *multiple samples problem* [104]. The first assumes that only one image per person is available in the database and the goal of facial recognition system is to identify a person later in any time, any pose, any illumination conditions from that single image. While the second problem assumes that many images of every person are available in the database and that these images can be used to train the face recognition system. We will focus on the *one sample problem* in this thesis.

Most of the previous works in 2D facial recognition from a single still image rely on the same scheme and we will not make an exception here:

1. Face detection - localization of face in the image, usually combined with detection of reference points in the localized face.

2. Geometric normalization - face image transformation according to predefined reference points into a reference shape.

3. Photometric normalization - a series of image processing operations for eliminating different illumination conditions.

4. Features extraction - an extraction of person's face representative features set, usually combined with training step.

5. Features matching - a comparison of previously extracted features sets.

2.1.1 Face Detection

We will not give much attention to face detection step in this thesis. After a real breakthrough in face localization [108] even a cell phone is capable of detecting faces in images. When an open source implementation [64], [63] appeared as a part of OpenCV library [44] no significant progress in this field was observed (except plenty of modifications which slightly improve the performance). The main features of this approach can be summarized as:

- Highly accurate - our implementation [53] is able to detect $\sim$99.6% of all near frontal face images on five publicly available datasets (BANCA [7], BioID [49], FERET [88], XM2VTS [72], and FRGC [86]) totaling in more than 50000 face images.
- Real-time - our implementation [53] is capable of processing more than 30 images of 640×480 pixels per second on a single core of *Intel Core 2 Duo CPU P8400 @ 2.26 Ghz*.

2.1.2 Geometric Normalization

Since the early methods of human face parametrization [102] and recognition [106] using Principal Component Analysis (PCA), there was a need to precisely align face images for getting acceptable recognition accuracy. The easiest and yet very effective method is a two-dimensional affine face image transform for mapping several reference points (most often centers of eyes) to predefined locations, see Fig. 2.1. Similar methods using additional reference points like center of mouth or nose tip usually did not improved the recognition of near-frontal faces accuracy. However, they help in compensating the in-depth rotation of faces. Geometric normalization approaches can be divided into the following major categories:

1. Affine image transform by the centers of eyes.
2. Image warping into a shape-free texture - active shape models (ASM)

(a) (b)

Figure 2.1: Face alignment and cropping (a) by the centers of eyes in original image (b).

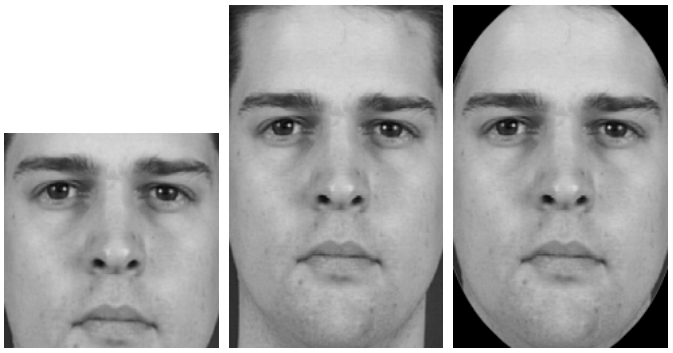
[57], active appearance models (AAM) [20], 3D morphable models [13], etc.

3. Local features alignment to the reference points - elastic bunch graph matching (EBGM) [116], scale invariant feature transform (SIFT) [12], etc.

We will focus on near-frontal face recognition, thus the first and simplest geometric normalization method will be used further.

Although affine image transform by the centers of eyes is a simple image processing operation, several unknowns still need to be explored. Firstly, face image resolution can greatly affect an algorithm's performance [68]. It is concluded that face image of 32×32 pixels is the lower and almost optimal limit for most typical holistic methods. Some suggest that face image of 6×7 pixels still can be used for recognition [113], but this is definitely an overtraining issue. On the other hand, modern algorithms which took part in FRVT2006 competition organized by US National Institute of Standards and Technology (NIST) made advantage of higher resolution images [89].

Secondly, it is not clear how tightly face should be cropped from the original image. Should we crop a tight square, a rectangle, or even mask out the background from face oval additionally, see Fig. 2.2. The tight square



(a) (b) (c)

Figure 2.2: Several types of face cropping: a tight square (a), a rectangle (b), and rectangle with masked-out background (c)

crop seems reasonable, because we are working on face recognition but not neck or haircut recognition (which are visible in the rectangular crop). The looser crop reveals more details on forehead and chin, but can result in adaptation to the training database because more background is present in face image. Masking out face oval with ellipse is kind of compromise in this case. It hides non-facial details, which also could become a reason of overtraining. On the other hand, the mask could influence local features to generate very similar values on artificially created border between the face image and mask. Some [116] recommend blurring out the face image near image borders (or mask).

2.1.3 Photometric Normalization

One of the main weaknesses of all facial recognition systems is poor recognition performance when face illumination varies. Several approaches to deal with this problem are suggested. The first group of them suggest choosing illumination invariant features - edge maps, Laplacian of a Gaussian features, image intensity derivatives, or 2D Gabor-like features [1], local binary patterns [2]. Authors of [126] also claim that no illumination normalization is needed after geometric normalization, if local Gabor features are used to encode facial features. The second group suggest modeling illumination variation in some subspace or manifold that is closest to the analyzed face

image [107], [36], [37]. The third group suggest normalizing face image to the canonical form by image transformation [98] or synthesis [35]. Any of these techniques can significantly increase face recognition accuracy when dealing with different illumination. We will focus on the last approach, also known as photometric face image normalization.

The first attempts of photometric normalization were global image processing operations, known as contrast stretching, histogram equalizations, and histogram fitting [90], [123]. The global methods were followed by their local versions with the same operations applied within several different regions of geometrically normalized face image [98] or even within the local neighborhood of each pixel [41]. Recently, a combined approach of global and local photometric normalization methods has been published [117].

Five different photometric normalization methods were compared in [101]: principal component analysis method, multiscale retinex method, homomorphic filtering, isotropic smoothing, and anisotropic smoothing. They were tested on three databases: Yale B [36], XM2VTS [72] and BANCA [7]. The authors conclude that anisotropic smoothing is superior to other tested methods in dealing with illumination variation. However, all the methods require training which adds complexity and clumsiness to them. A set of less complex normalizations were tested in [98]: global histogram equalization (HE), region-based histogram equalization (RHE), global gamma intensity correction (GIC), region based GIC (RGC), quotient illumination relighting, and combinations of those methods. They conclude with quotient illumination relighting as superior method, which, on the other hand, makes assumption that lighting modes of the images are known or can be estimated.

None of the mentioned photometric normalization methods is revolutionary in dealing with illumination variation. Some of them work better in one conditions (or on one database) others are better elsewhere. But generally, they improve face recognition accuracy.

2.1.4 Features Extraction and Matching

We will not separate features extraction from matching step because unlike the fingerprint minutiae [52], there exist no standard facial features which could be reliably extracted and reused by completely different extraction and matching algorithms. Most often the facial features matching step is specific for the extraction step and cannot be separated.

2.1.4.1 Holistic Methods

A classic eigenfaces [106] technique was one of the first holistic methods. It was extended by Linear Discriminative Analysis (LDA) based algorithms [9], [124], [66], Probablistic-based method [75], Support Vector Machine (SVM) based method [84], Laplacianfaces [40], Correlation-based methods [54], etc. All of them are superior to the PCA, but require multiple face images of the same person and, usually, a length training procedure. On the other hand, PCA was extended into methods which work on a single image per person with generating more samples from that image by adding various type of noise [78], [31], [67]. But most of them improved the recognition accuracy only marginally.

2.1.4.2 Local Features Methods

Local features methods often achieved higher accuracy compared to holistic ones. Among these were Elastic Bunch Graph Matching (EBGM) [116] which evolved from Dynamic Link Matching [56]. Additionally, well known techniques were applied: Convolution Neural Networks (CNN) [58], Hidden Markov Models (HMM) [59], Gabor jets [126]. Furthermore, Discrete Cosine Transform (DCT) [32], modified LDA [42] and other techniques were applied to the components of the face instead of the whole face at once. A successful Local Binary Patterns (LBP) [2] application to face recognition was extended in many ways [122], [105] and significantly improved the upper bound of face recognition from single image accuracy. In conclusion, most

of the recent face recognition technologies rely on local features to cope with the most challenging conditions in facial recognition: pose, lighting, aging, occlusions, etc.

2.2 Proposed Face Recognition Algorithm

We started developing the proposed face recognition algorithm by choosing potential set of features and then adapting all the other steps. From a literature review Gabor features seemed to be the most promising ones. They have a property of illumination invariance (which we found rather limited in our experiments), they are closely related to processes in primary visual cortex in human brains [29], and systems based on Gabor features consistently show one of the best results in many computer vision tasks including face recognition. To use Gabor features in real-time face recognition system we needed a new computationally efficient Gabor features calculation method.

2.2.1 Accelerated Calculation of Gabor Features in Spatial Domain

Gabor filters have been widely used in constructing various Gabor features for different computer vision tasks: competitive texture classification [19], segmentation [15], [48] and synthesis [16], fast and accurate object detection and tracking [76], [46], one of most precise biometrics - iris recognition [24] and other. Gabor features are proved to perform very well because of their properties like rotation, scale, translation and uniform lighting semi-invariance [55]. On the other hand, computational complexity still limits their application in practice. We will focus on most widely used Gabor features - convolution with multi-resolution structure of Gabor filters of several frequencies and orientations.

A straight forward implementation of Gabor features extraction would be an image convolution with Gabor filters in spatial domain. It can be improved by an order of magnitude using the separability property of 2D

filters [4] or symmetry/anti-symmetry/wavelet characteristics [92] for special cases of Gabor filters orientations and frequencies. Several schemes of calculating Gabor features more effectively by approximations were presented: effective area of filters and Laplacian pyramid [43], recursive Gabor [120], decomposition into Gaussians [10]. However, the most effective Gabor features extraction at every location in the image is done by using Fast Fourier Transform (FFT) for image convolution with Gabor filters in frequency domain. Several works in pattern recognition [109], [114] use Gabor features that are calculated at some regular grid but not every pixel of the image. Motivated by that we will explore how a structure of regular grid and generalized separability of Gabor filter can be exploited to speed up the calculation of Gabor features almost to the speed of FFT without loss of precision.

2.2.1.1 Gabor Features

2.2.1.1.1 Gabor Filter Following [4], we assume complex-valued 2D Gabor filter as a product of isotropic Gaussian and complex exponential plane wave, that is,

$$\begin{aligned} G(x, y, \theta, f) &= e^{-\frac{x_\theta^2 + y_\theta^2}{2\sigma^2}} \cos(2\pi f x_\theta + \phi), \\ x_\theta &= x\cos\theta + y\sin\theta, \\ y_\theta &= -x\sin\theta + y\cos\theta, \end{aligned} \tag{2.1}$$

where $\theta \in [0, \pi)$ is filter orientation, f is filter frequency, σ is a standard deviation of Gaussian function, and $\phi \in \{0, \frac{\pi}{2}\}$ corresponds to real and imaginary parts of Gabor filter. We will show the Gabor filter of any orientation θ and any frequency f becomes linearly-multi-separable (formal definition will be presented later). Later, we will show how this and the filter symmetry properties can be exploited for accelerated calculation of Gabor features.

2.2.1.1.2 Effective Filter Envelope Effective filter envelope corresponds to the filter area with significant coefficients [43]. Filter coefficients

outside that area can be discarded depending on what accuracy and speed ratio is needed. Although speed of convolution in frequency domain is not affected by smaller filter size[1], it can significantly reduce computational complexity when filtering in spatial domain and memory consumptions for storage of filters. Effective filter envelope of Gabor filter can be calculated directly from standard deviation of approximately normally distributed data. For all further experiments doubled standard deviation will be used as the radius of effective filter envelope retaining approximately 95% energy of the filter, which is defined as sum of absolute Gabor filter coefficients.

2.2.1.1.3 Convolution Direct convolution of a linear $M \times M$ (where $M = 2m + 1$) 2D filter C and image I in spatial domain is defined as

$$H(r,c) = \sum_{i=-m}^{m} \sum_{j=-m}^{m} C(i,j)\, I(r-i, c-j), \tag{2.2}$$

which requires $O\left(M^2\right)$ (operations) calculations to calculate convolution at one point of the image and $O\left(M^2N^2\right)$ (operations) calculations for convolution with the whole $N \times N$ image. Linear 2D filter is said to be separable if it can be decomposed as a product of two one-dimensional signals filters. Convolution of whole image with separable 2D filter can be speeded up by convolving each row of the image with the horizontal projection of filter, resulting in the intermediate image. Then, convolving each column of the intermediate image with the vertical projection of filter. The resulting image is identical to direct convolution, no matter which step (horizontal or vertical) is performed first, and requires $O\left(MN^2\right)$ calculations. We will generalize notion of separable 2D filter that will be applicable for our complex valued Gabor filter.

Definition It is said that complex-valued 2D filter $C(x,y)$ is *linearly multi-*

[1]Here we neglect the fact that convolution by FFT without a signal wraparound requires complementing of the longer signal by half of the shorter signal with zeros.

separable if for some finite K_1 and K_2

$$C(x,y) = \sum_{k=1}^{K_1} a_k^1(x)\, b_k^1(y) + \sqrt{-1} \sum_{k=1}^{K_2} a_k^2(x)\, b_k^2(y), \tag{2.3}$$

where $a_k^j(x)$ and $b_k^j(y)$ are real-valued functions. Sum $K_1 + K_2$ is referred as *order* of multi-separability.

Note, that $N \times N$ image filtration with linearly multi-separable filter would require $2(K_1 + K_2)MN^2 + (K_1 + K_2 - 2)N^2$ arithmetic operations. We will show that the Gabor filter 2.1 is multi-separable of order 4. The real part of the Gabor filter can be decomposed in

$$\begin{aligned}
\Re G(x,y,\theta,f) &= e^{-\frac{(x\cos\theta+y\sin\theta)^2+(-x\sin\theta+y\cos\theta)^2}{2\sigma^2}} \cdot \\
&\quad \cos\left(2\pi f\left(x\cos\theta + y\sin\theta\right)\right) \\
&= e^{-\frac{x^2+y^2}{2\sigma^2}} \cos\left(2\pi f x\cos\theta + 2\pi f y\sin\theta\right) \\
&= e^{-\frac{x^2+y^2}{2\sigma^2}} \cos\left(2\pi f x\cos\theta\right)\cos\left(2\pi f y\sin\theta\right) - \\
&\quad e^{-\frac{x^2+y^2}{2\sigma^2}} \sin\left(2\pi f x\cos\theta\right)\sin\left(2\pi f y\sin\theta\right) \\
&= G_{h1}(x,\theta,f)\, G_{v1}(y,\theta,f) - \\
&\quad G_{h2}(x,\theta,f)\, G_{v2}(y,\theta,f).
\end{aligned} \tag{2.4}$$

Here,

$$\begin{aligned}
G_{h1}(x,\theta,f) &= e^{-\frac{x^2}{2\sigma^2}} \cos\left(2\pi f x\cos\theta\right), \\
G_{h2}(x,\theta,f) &= e^{-\frac{x^2}{2\sigma^2}} \sin\left(2\pi f x\cos\theta\right), \\
G_{v1}(y,\theta,f) &= e^{-\frac{y^2}{2\sigma^2}} \cos\left(2\pi f y\sin\theta\right), \\
G_{v2}(y,\theta,f) &= e^{-\frac{y^2}{2\sigma^2}} \sin\left(2\pi f y\sin\theta\right).
\end{aligned} \tag{2.5}$$

Similarly, the imaginary part of Gabor filter can be decomposed in

$$\Im G(x,y,\theta,f) = G_{h2}(x,\theta,f)\, G_{v1}(y,\theta,f) + G_{h1}(x,\theta,f)\, G_{v2}(y,\theta,f). \tag{2.6}$$

Combining 2.4 and 2.6 we have that complex Gabor filter is linearly multi-separable of order 4. Using additional benefits of symmetry of 2.4 and 2.6 we can reduce the filtration complexity $8MN^2 + 6N^2$ to $6MN^2 + 2N^2$. The last significantly improves complexity of direct convolution which requires $4M^2N^2$ arithmetic operations.

Convolution H in frequency domain is done by converting image I to frequency domain with FFT, multiplying by a converted to frequency domain filter C and converting back to the spatial domain with IFFT:

$$H = IFFT(FFT(I) \cdot FFT(F)). \tag{2.7}$$

This approach gives a periodic version of convolution. If one likes to obtain non-periodic convolution, the approach of filtering in frequency domain requires a modification involving an additional complexity. For the simplicity of analysis we will restrict ourselves on periodic version of convolution. Computational complexity of FFT, as well as IFFT, is $O(N^2 \log N)$, however the lowest bound of the exact count of arithmetic operations of 1D FFT (split-radix FFT algorithm [118]) is $4N \log_2 N - 6N + 8$ real additions and multiplications[2] and applies only for N a power of two greater than 1. Since complexity of convolution in frequency domain does not depend on filter size (if filter is smaller than the whole image) and complex Gabor filter can be computed directly in frequency domain by

$$\Gamma(u, v, \theta, f) = 2\pi\sigma^2 e^{-(2\pi\sigma)^2\left((u - f\cos\theta)^2 + (v - f\sin\theta)^2\right)}. \tag{2.8}$$

Arithmetic complexity of convolution with complex Gabor filter is $4N^2 \log_2 N - 4N^2 + 8N$ (one IFFT of 2D signal plus one complex multiplication in frequency domain, if image and filter are already in frequency domain[3]).

2.2.1.1.4 Multi-resolution and Multi-orientation The most attractive property of Gabor feature - orientation and scale semi-invariance - is achieved by using Gabor filters of many different orientations and scales which describe local structure of the image. In [43], Laplacian pyramid of images for faster calculation of multi-resolution Gabor feature is suggested. However, speed improvement comes with several drawbacks, and once again works only for very special case of Gabor features:

[2] Recently it was improved to $\approx \frac{34}{9} N \log_2 N$ [50].

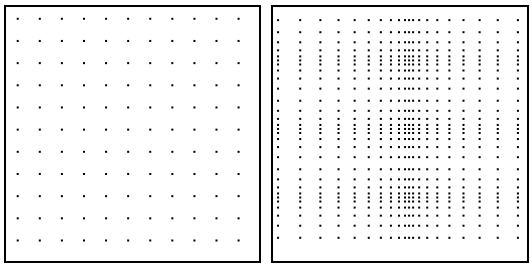
[3] The complexity of image conversion to frequency domain is not added here because further we analyze a Gabor feature (which is composed of several Gabor filters) calculation time.

1. Pyramid of images can be effectively constructed only for integer scaling factors, downscaling by real factors creates aliasing effects and should be avoided.

2. Responses of Gabor filters are approximated and, additionally, special care should be taken to upscale the responses of low frequency filters back to the higher resolution.

We will focus on calculation of exact values of Gabor filters of equally distributed directions (covering the $[0 \dots \pi]$ range with a constant step) which are not limited to particular scales that are convenient for constructing Laplacian pyramid of images. Authors of [92] try to calculate convolution with several Gabor filters at once, but is tuned to orientations with $\theta \in \left\{0, \frac{\pi}{4}, \frac{2\pi}{4}, \frac{3\pi}{4}\right\}$ and 3 scales (starting from a 3×3 discrete filter and scaling it by a factor of 2). All of the rest deal with each Gabor filter independently. One of the main contributions of the proposed method is calculating the whole Gabor feature at once in spatial domain, including convolution with Gabor filters of all equally distributed directions and any number of scales.

2.2.1.1.5 Regular Grid Proposed optimizations are most effective if adjacent points where Gabor features will be calculated are closer than the half of the largest Gabor filter. The simplest distribution of points in 2D is a regular grid where each point is in the same distance from the adjacent points. A regular grid is used in some pattern recognition tasks like [109], [114]. Furthermore, the proposed method is not limited to regular grids defined as above but filter separability property can be exploited to speed up calculations at a grid which is also separable, i. e. can be separated to one-dimensional vectors of rows and columns. Several examples of such grid can be seen in Fig. 2.3.



(a) (b)

Figure 2.3: Examples of regular grids on a 128×128 image.

2.2.1.2 Implementation

A classical example of Gabor feature is a vector of responses got from image convolution with Gabor filters of 8 orientations and 5 frequencies homogeneously distributed in a frequency band [116]. Number of frequencies is not limited by the proposed implementation, however number of orientations should be even to use full ensemble of optimizations. On the other hand, this limitation is not exceptional because almost every application of Gabor features in the literature uses even number of orientations. In section 2.2.1.1.3, we showed that every Gabor filter with isotropic Gaussian part is multi-separable of order 4 and this can be exploited to speed up the convolution in spatial domain. Further we will show how symmetry and anti-symmetry of Gabor filter as well as Gabor feature can be used for speeding up the convolution at any location in the image up to four times.

2.2.1.2.1 Symmetry (Anti-symmetry) of Gabor Filter Real and imaginary parts of Gabor filters are symmetric (Fig. 2.4(a), Fig. 2.4(c)) and anti-symmetric (Fig. 2.4(b), Fig. 2.4(d)).

Direct convolution of 1D signal I and filter $F = (f_0, \ldots, f_M)$, where $f_i = f_{M-i}$, at location x (which requires $M + 1$ multiplications and M additions)

$$\begin{aligned} H(x) &= f_0 I(x-m) + f_1 I(x-m+1) + \ldots + \\ &\quad f_{M-1} I(x+m-1) + f_M I(x+m), \end{aligned} \tag{2.9}$$

where $M = 2m + 1$, can be replaced by symmetric (or anti-symmetric if

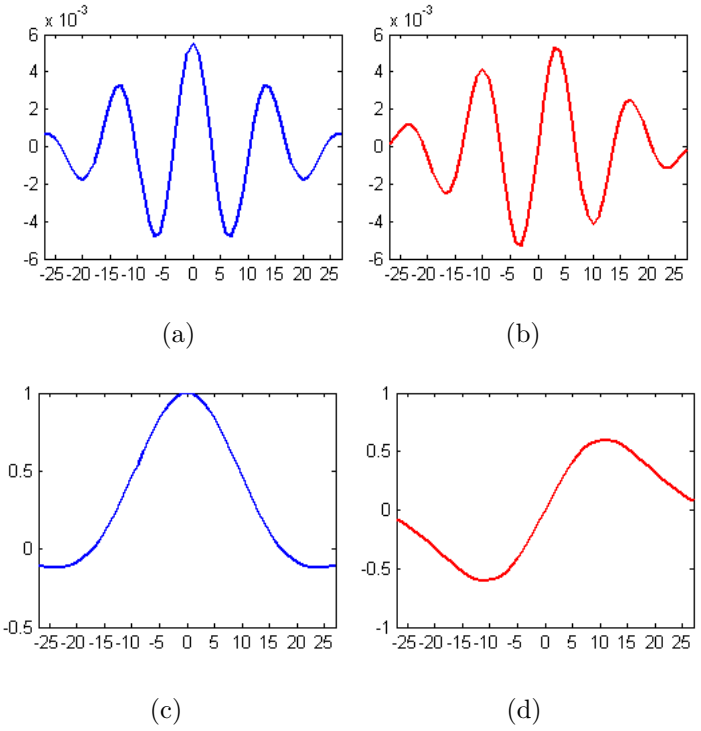


Figure 2.4: G_{h1} (a), G_{h2} (b), G_{v1} (c) and G_{v2} (d) parts of Gabor filter with orientation $\theta = \frac{\pi}{16}$.

sums of signal values will be replaced by differences) version (which requires $m+1$ multiplications and M additions):

$$\begin{aligned} H_s(x) &= f_{m+1} I(x) + f_m (I(x-1) + I(x+1)) + \ldots + \\ &\quad f_0 (I(x-m) + I(x+m)). \end{aligned} \tag{2.10}$$

Although in [43] authors state that this will not lead to any improvement on modern computers because multiplication is not an expensive operation, symmetric convolution requires 25% less arithmetic operations which can reduce convolution time with one filter by one fourth. Additionally, when symmetric convolution is used to calculate Gabor feature, each Gabor filter can use the same sums (or differences) of signal and they can be precalculated only once for the largest filter. Another 25% of arithmetic operations can be saved for all but the largest filter.

2.2.1.2.2 Symmetry (Anti-symmetry) of Gabor Feature One more symmetry (anti-symmetry) exists in Gabor feature between filters

of same scale but different orientations. If D orientations are used in Gabor feature, full convolution must be performed only with filters of $\frac{D}{2}+1$ orientations that fall in the range $\left[0,\ldots,\frac{\pi}{4}\right]$. The remaining filters are symmetric (anti-symmetric) and their responses can be calculated by reusing previously calculated filters responses as shown below for G_{h1} filter part:

$$\begin{aligned} G_{h1}\left(x,\pi-\theta,f_0\right) &= e^{-\frac{x^2}{2\sigma^2}}\cos\left(2\pi f_0 x\cos\left(\pi-\theta\right)\right) \\ &= e^{-\frac{x^2}{2\sigma^2}}\cos\left(-2\pi f_0 x\cos\left(\theta\right)\right) \\ &= e^{-\frac{x^2}{2\sigma^2}}\cos\left(2\pi f_0 x\cos\left(\theta\right)\right) \\ &= G_{h1}\left(x,\theta,f_0\right). \end{aligned} \tag{2.11}$$

Similarly,

$$\begin{aligned} G_{h2}\left(x,\pi-\theta,f_0\right) &= -G_{h2}\left(x,\theta,f_0\right), \\ G_{v1}\left(y,\pi-\theta,f_0\right) &= G_{v1}\left(y,\theta,f_0\right), \\ G_{v2}\left(y,\pi-\theta,f_0\right) &= G_{v2}\left(y,\theta,f_0\right). \end{aligned} \tag{2.12}$$

Responses of corresponding Gabor filters can be calculated by:

$$\begin{aligned} G\left(x,y,\pi-\theta,f_0,0\right) &= G_{h1}\left(x,\theta,f_0\right)G_{v1}\left(y,\theta,f_0\right)+ \\ &\quad G_{h2}\left(x,\theta,f_0\right)G_{v2}\left(y,\theta,f_0\right), \\ G\left(x,y,\pi-\theta,f_0,\tfrac{\pi}{2}\right) &= -G_{h2}\left(x,\theta,f_0\right)G_{v1}\left(y,\theta,f_0\right)+ \\ &\quad G_{h1}\left(x,\theta,f_0\right)G_{v2}\left(y,\theta,f_0\right). \end{aligned} \tag{2.13}$$

Using symmetry along orientations $\frac{D}{2}-1$ of orientations won't be recalculated and will save almost 50% of arithmetic operations if more than two orientations will be used.

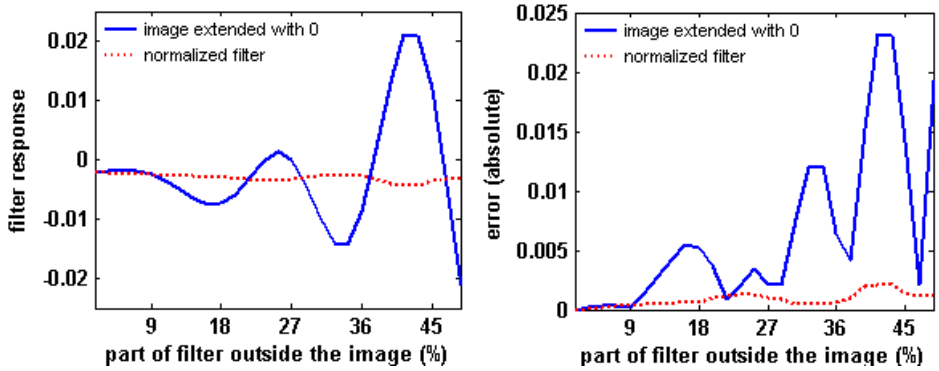
2.2.1.2.3 Filtering at Regular Grid Image convolution with a linear separable 2D filter can be optimized by an order of magnitude exploiting the filter multi-separability property. Actually, convolution of the whole image is the same as convolution at the dense regular grid which has a distance of one pixel between the adjacent grid positions. Same optimizations are possible if the distance between adjacent grid positions is greater than one pixel (but not greater than the length of the filter). Often, there is no need

to have Gabor response at every location in the image but calculating direct convolution is time consuming and FFT must be used. Direct convolution, FFT convolution and the proposed implementation will be evaluated in respect of speed when calculating Gabor feature responses at grids of different dense in section 2.2.1.3.

2.2.1.2.4 Filtering Near Image Boundary Several practices of filtering near image boundaries when part of the filter slips outside the image come from image processing:

1. Extend the image with a constant (possibly zero) intensity value.
2. Extend the image periodically or by mirroring it at the boundaries.
3. Normalize the response of convolution by sum of values from filter part which does not slip outside the image.

Discrete Gabor filters are constructed to have the DC free property, i. e. sum of filter coefficients equals to zero. When part of the filter slips outside the image Gabor filter loses the DC free property and its response can change unacceptably as it is shown in Fig. 2.5. Filter response normal-



(a) (b)

Figure 2.5: One specific Gabor filter response (a) and absolute error (b) when filter slips outside the image without and with filter response normalization.

ization is necessary and can be done by subtracting a DC free component:

$$H_{DCfree}(r,c) = H(r,c) - g_{avg} \sum_{x,y} I(r+y, c+x), \tag{2.14}$$

where $H(r,c)$ is Gabor filter response without normalization at row r and column c, g_{avg} is an average value of filter part which does not slip outside the image, and $(r+y, c+x)$ runs through the image part which is under the filter. The DC free component can be calculated rapidly in a constant time by using intermediate integral image and integral filter (also called "summed area tables" in [21]):

1. Integral image is the image representation where location (r,c) contains the sum of all the pixels above and to the left of (r,c) inclusive:

$$II(r,c) = \sum_{r' \le r, c' \le c} I(r', c'). \tag{2.15}$$

2. Similarly integral filter is defined:

$$IG(r,c) = \sum_{y' \le y, x' \le x} G(y', x'). \tag{2.16}$$

3. DC free component can be calculated by accessing integral image at four locations as it is shown in Fig. 2.6 (to get a sum of intensity values under the filter) and integral filter at another four locations (to get a sum of filter values that do not slip outside the image).

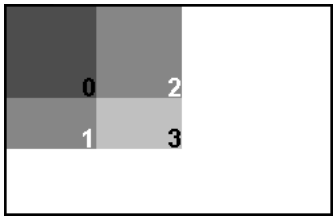


Figure 2.6: Sum of values in light gray rectangle S can be calculated by accessing integral representation at 4 locations: $S = 0 + 3 - 1 - 2$.

As can be seen in Fig. 2.5, compensating for the DC property of the Gabor filter results in much more stable filter response even when up to 45% of the filter is outside the image.

2.2.1.3 Evaluation

First of all, arithmetic complexity of the proposed method is compared to direct and FFT based methods by calculating number of arithmetic operations (see section 2.2.1.1.3) required to perform Gabor features extraction at regular grids of different sizes. Regular grids were chosen in the following order: each point, every second point, every third point, ..., one center point in the image. The proposed method is expected to be faster than direct calculation of Gabor features and to be faster than convolution in frequency domain when Gabor features are needed only on every second point of the image (see Fig. 2.7). Complexity of convolution in frequency domain was

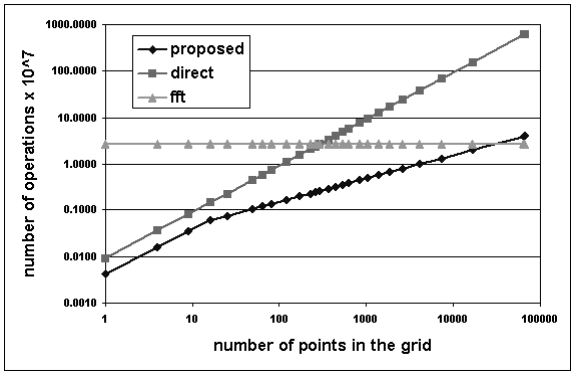


Figure 2.7: Number of arithmetic operations required for the evaluated methods to calculate Gabor features (9 scales, 16 orientations) at regular grids of different sizes on 256×256 image.

calculated according to the split-radix FFT method (see section2.2.1.1.3) that is applicable only to the signals of power-of-two length. In practice its complexity heavily depends on effectiveness of implementation. For further experiments one of the most efficient publicly available FFT software - FFTW library [34] will be used. Results of practical experiment with the same Gabor features, image size and regular girds that were used in theoretical evaluation can be seen in Fig. 2.8. Different timings for the same number of points for the direct Gabor features calculations appear from the regular grids where the same points are situated further or closer to image boundary. Gabor filters from the features that are calculated closer to the

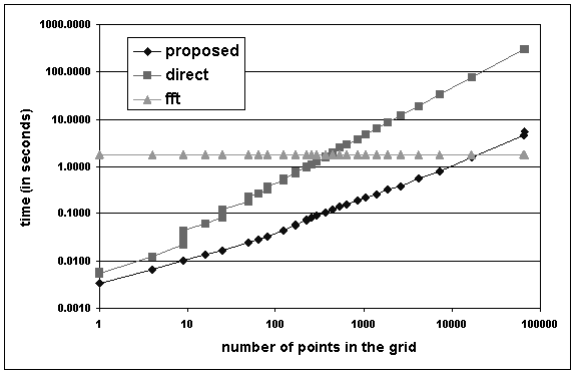


Figure 2.8: Time (in seconds) required for the evaluated methods to calculate Gabor features (9 scales, 16 orientations) at regular grids of different sizes on 256×256 image.

image boundary slip outside the image and are calculated faster because parts of them are not used in convolution. However this does not affect the speed of the proposed method because the problem of image boundary is solved at the precalculation of signal sums (and differences) step (see Section 2.2.1.2.1). One more difference from theoretical evaluation of complexity is the form of the proposed method curve. This can be explained by the fact that arithmetic complexity was calculated without taking into account (actually, taking the worst case) how Gabor filters overlap in the regular grid, i. e. how close are adjacent points in the regular grid.

Images with dimensions of power-of-two are very convenient for the FFT. To show the efficiency of the proposed method the same experiment was performed with image of 191×191 pixels, results can be seen in 2.9. Now the proposed method outperforms convolution in frequency domain by 10% even at calculating Gabor features at each point of the image (though, the difference is only marginal in logarithmic scale).

Similar experiments were performed with different number of scales $(3, 5, 9)$ and orientations $(4, 8, 16, 32)$ in Gabor features and different sizes of images $(128 \times 128, 191 \times 191, 256 \times 256)$. The results are almost identical to those that were presented above.

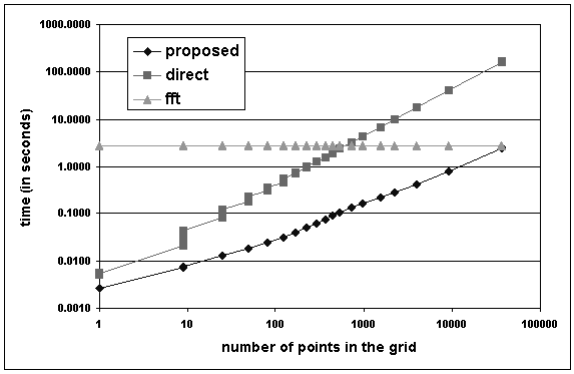


Figure 2.9: Time (in seconds) required for the evaluated methods to calculate Gabor features (9 scales, 16 orientations) at regular grids of different sizes on 191×191 image.

2.2.1.4 Conclusion

The regular grid together with Gabor filters symmetry (anti-symmetry) and Gabor features symmetry (anti-symmetry) along directions were successfully used to improve Gabor features calculation time. The problem of filtering near image boundary was also addressed and efficient solution was proposed. After the detailed comparison of the proposed method with direct and FFT based calculation of Gabor features the following conclusions can be drawn from the evaluation results:

1. Proposed method is always faster than the direct convolution.

2. Proposed method is faster than convolution in frequency domain if Gabor features are required at every second (sometimes every third) point of the image and image dimensions are convenient for FFT.

3. Proposed method is always faster than convolution in frequency domain if image dimensions are not convenient for FFT.

Several approaches of optimizations were left out of the scope of this research. Firstly, it is exploiting Gabor wavelet property, i. e. different Gabor scales are generated from one mother wavelet. Secondly, modern processors are able to do several arithmetic operations in parallel and this

can be used for parallel version of the proposed method. And lastly, various approximations of Gabor filters could be used to speed up features extraction with small errors in Gabor responses. If the errors are too small to influence the quality of further steps in the algorithms, approximations should be definitely used.

2.2.2 Baseline Face Recognition Algorithm

Instead of using a common baseline algorithm based on PCA [106] we have chosen to start from a more sophisticated algorithm as a baseline for our research. Similar to [116], we will build our algorithm on Gabor features. In a previous section we showed how Gabor features can be efficiently calculated on a fixed regular grid which is overlaid on the image (which contains a human face in our case). While [116] suggests creating a human face model to cope with head rotations, we focus on near-frontal face recognition, thus, a regular grid is a simple but potential approximation of the same idea.

The starting point for our further research is the following algorithm:

1. Find centers of eyes as reference points for geometric face image normalization.

2. Geometrically normalize face image to a constant dimensions of 128×128 pixels with left and right eyes at fixed positions $(32, 32)$ and $(96, 32)$ respectively. See Fig. 2.1.

3. Overlay a regular 10×10 grid over the geometrically normalized face image.

4. Calculate a classic configuration of Gabor feature (5 scales, 8 orientations) at each grid's position.

5. Convert the responses of complex Gabor filters $g_c = g_{c,real} + i g_{c,imaginary}$ to their magnitudes by $g_{c,magnitude} = \sqrt{g_{c,real}^2 + g_{c,imaginary}^2}$.

6. Similar to [126], save facial features representation as 100×5 8-dimensional vectors for each grid position and frequency.

7. Compare two sets of vectors generated from different face images by calculating a normalized dot product between the corresponding pairs of vectors. The mean of received values is a similarity between two representations of facial images.

One can find some similarities of the proposed baseline algorithm to the work of [126]. The main difference is the proposed regular grid which is constant for all scales of Gabor filters. And the main strength of our method is low computational cost as we showed in section 2.2.1.

2.2.3 Evaluation

To evaluate all variations of the proposed algorithm the most widely publicly available face recognition protocol was chosen. FERET database along with the FERET protocol [88] was assembled to support government monitored testing and evaluation of face recognition algorithms using standardized tests and procedures. The whole image set is divided into the following subsets:

1. Fa - serves as a set of 1196 gallery images for simulating persons that are enrolled into facial recognition system.

2. Fb - serves as a set of 1195 probe images, where images were taken at the same day as Fa images, with the same camera and illumination conditions, but changed facial expression.

3. Fc - serves as a set of 194 probe images, where images were taken at the same day as Fa images, but with different cameras and illumination conditions.

4. Dup1 - serves as a set of 722 probe images, where images were taken on different days than Fa images, but within one year.

5. Dup2 - serves as a set of 234 probe images, where images were taken on different days than Fa images, but at least one year later.

6. Training - serves as a set of 736 training images, used to train algorithms which require training.

We will further identify our experiments by the probe images subsets identifiers: experiment Fb, experiment Fc, experiment Dup1, and experiment Dup2.

As we didn't want our face recognition algorithm evaluation to be dependent on face detection part, we chose to use manually labeled eyes positions (which are provided together with FERET database) as reference points for geometric normalization in all experiments.

FERET protocol evaluates face recognition algorithms in the following tasks:

1. Verification - by calculating a Detection Error Tradeoff (DET) curve of the algorithm.

2. Identification - by calculating a $Rank1$ identification rate. $Rank1$ identification rate stands for percentage of probe images which were more similar to their corresponding gallery image than any other gallery images.

We will skip the easiest Fb experiment and focus on Dup1 experiment in our research. Fig. 2.10 presents a DET curve for our baseline algorithm. The proposed baseline algorithm is already better than a top performer [116] of FERET test performed by NIST until the last testing in 1997. While typical $Rank1$ identification rates for Dup1 experiment are $50\% - 60\%$, our baseline algorithm achieves $Rank1 = 65.24\%$. It is a good performance compared to twelve years old algorithms but recent publications declare 85% [126] or even 90% [105] $Rank1$ identification rate in the same experiment.

2.2.4 Improvements to Baseline Algorithm

2.2.4.1 Geometric Normalization

We fixed Gabor feature as a combination of 40 Gabor filters of 5 scales and 8 orientations. Additionally, we fixed the number of pixels between each

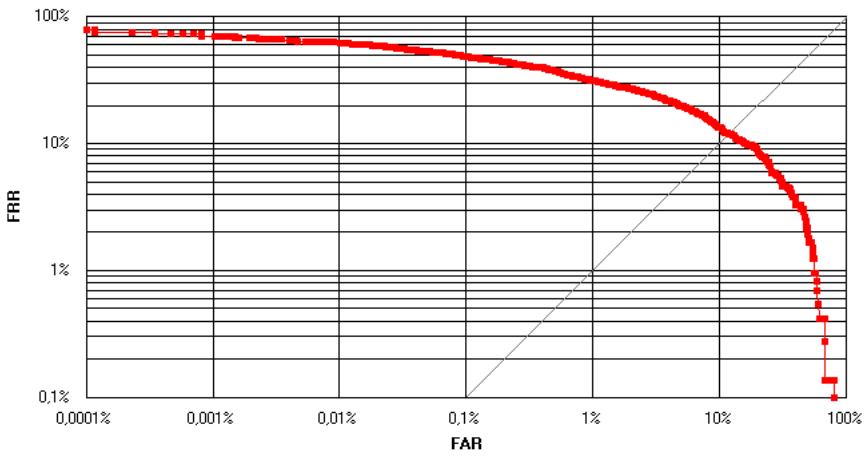


Figure 2.10: Verification performance of our baseline algorithm (with identification rate $Rank1 = 65.24\%$).

Gabor feature in the regular grid in three combinations, which generated regular grids of three different densities: 8×8, 10×10 and 12×12 for the square geometrically normalized images 2.1(a). Further, we performed the following experiments to determine how geometric normalization should be performed:

1. Resolution experiment - to find out, how number of pixels between the eyes after geometric normalization influence face recognition accuracy.

2. Vertical face alignment - to find out, how vertical alignment of the face image influence face recognition accuracy.

3. Cropping tightness experiment - to find out, how much of face width and height should be cropped from the original face image.

2.2.4.1.1 Resolution Experiment We measured identification accuracy for the following inter-ocular distances (IOD) in pixels of geometrically normalized square face images: 32, 48, 64, 80, 96, 112, 128. We could not fit a 10×10 density grid in the smallest resolutions (32 and 48), but other densities showed significantly lower identifications rates in these resolutions 2.11. So, we dropped them from all further experiments. Other resolutions gave almost identical accuracy with $IOD = 96$ resolution being slightly

better in average. Additionally, we noticed that grid of rarest density 8×8 performed worst, while other two tested densities gave even accuracy. So, we also dropped this grid density from all further experiments. Best identification rate achieved in this experiment is $Rank1 = 72.30\%$.

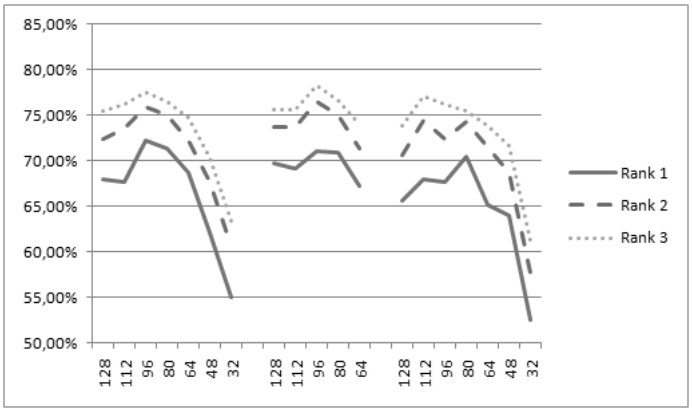


Figure 2.11: Identification accuracy dependency on number of pixels between the centers of eyes in geometrically normalized face. Three groups of rankings are shown (from left to right) by grid density: 12×12, 10×10 and 8×8.

2.2.4.1.2 Cropping Tightness and Vertical Face Alignment Experiments Geometric face image normalization into square image seems reasonable, however, some publications suggest using rectangular face cropping. We tested this in two steps. Firstly, we fixed width of face crop to twice the inter-ocular distance and adjusted height of face crop by a factor ranging from 1.00 (height equals to width) to 1.55 (height is larger than width by this factor), see Fig. 2.12. Every tested factor improved identification accuracy with $1.25 - 1.45$ range of factors giving the best results across all the resolutions and grid densities. This was not unexpected, because by expanding the face image and not changing spacing between Gabor features in the regular grid, we increased number of used Gabor features. Best identification rate achieved in this experiment is $Rank1 = 76.59\%$.

When geometrically normalized face image was expanded vertically, it was a good time to find out the vertical face alignment which could im-

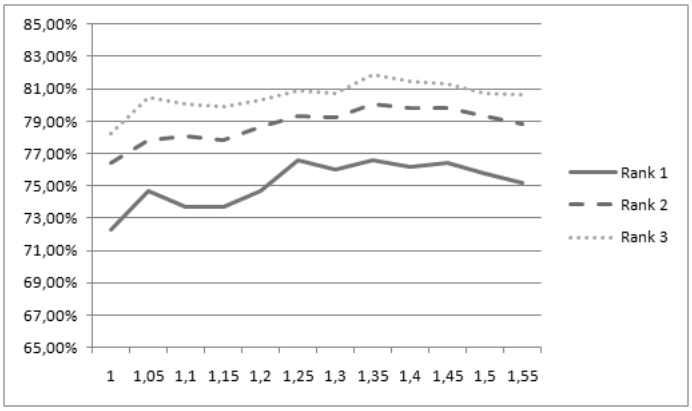


Figure 2.12: Identification accuracy dependency on rectangular cropping, when face crop is expanded by some factor vertically.

prove identification rate further. The lower face part was fully visible in all geometrically normalized images after we expanded images vertically Fig. 2.13(a), but often a lot of non-face details were visible. We tried to to avoid non-face details by gradually lowering positions of the eyes, see Fig. 2.13. We started from the baseline eyes positions and lowered them by 5% of inter-ocular distance in the following test.

(a) (b)

Figure 2.13: The highest (a) and the lowest (b) tested positions of the eyes. Showed on the average geometrically normalized face image.

This time identification rate was also improved when lowering eyes positions by any percentage of IOD, see Fig. 2.14. Best identification rate achieved in this experiment is $Rank1 = 79.92\%$.

Finally, we fixed the most successful factor for the height of face crop in average 1.35 and most successful percentage for lowering eyes positions in

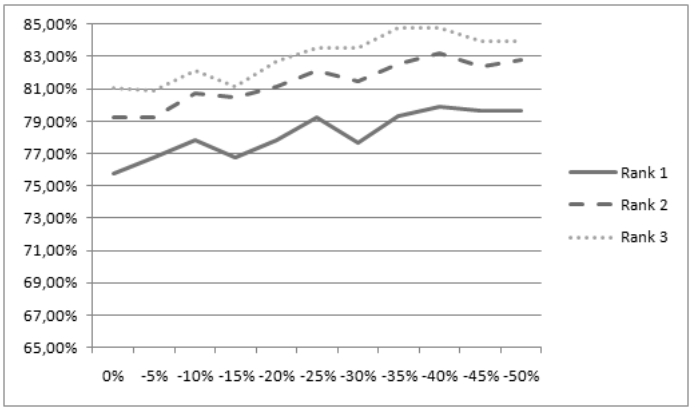


Figure 2.14: Identification accuracy dependency on vertical face alignment in vertically expanded face image.

average -40% from the previous experiments and further adjusted the factor for the width of face crop ranging from 0.85 to 1.15 of original face crop. As can be seen in Fig. 2.15, this almost did not change identification rate. Best identification rate achieved in this experiment is $Rank1 = 80.33\%$.

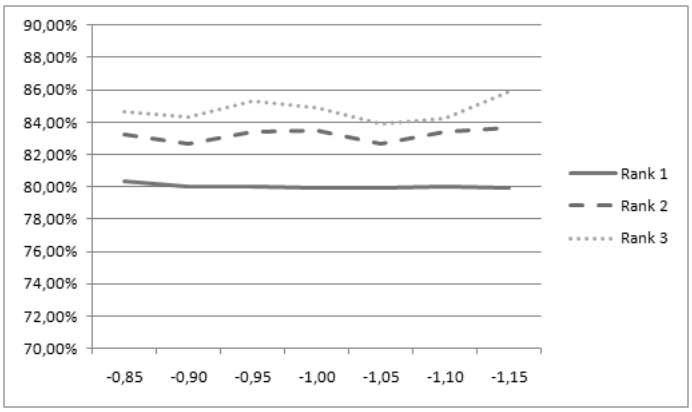


Figure 2.15: Identification accuracy dependency on rectangular cropping, when face crop is expanded by some factor horizontally.

2.2.4.1.3 Summary Only by changing parameters of geometrical face image normalization we were able to improve identification rate $Rank1$ from 65.24% to 80.33%. Verification performance change is presented in Fig. 2.16. Although our parameters were adapted to classic Gabor features, the achieved results suggest that choosing the right parameters for geometrical face image normalization can significantly improve face recognition accuracy

for any other feature set.

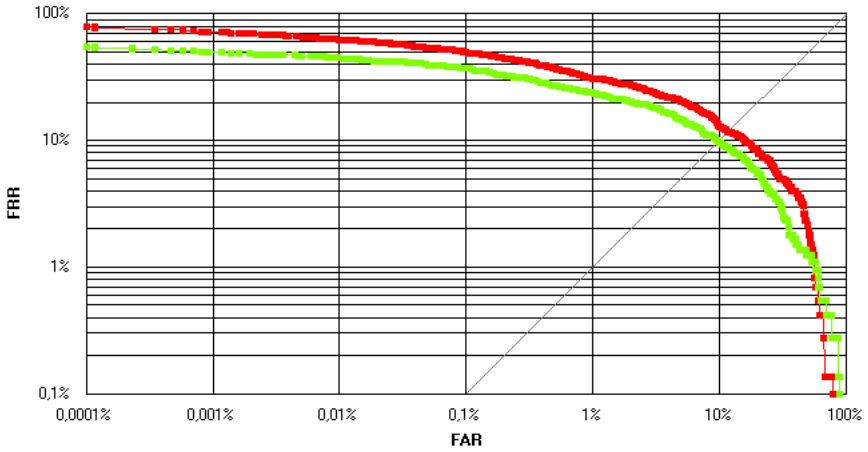


Figure 2.16: Verification accuracy improvement from the baseline algorithm (red curve) to final geometric face image normalization parameters (green curve).

2.2.4.2 Photometric Normalization

From the previous experiments we fixed classic Gabor features and geometric face image normalization. Although sometimes it is claimed that Gabor features are illumination invariant, it is wrong assumption for a complex texture of human face. To deal with illumination changes we tested the following photometric normalization methods:

1. Self-quotient image [110].
2. Local mean and variance normalization.
3. Local histogram equalization.

Many more methods exist, but we have chosen to test a representative simple subset because our goal is an algorithm oriented towards speed.

2.2.4.2.1 Self-Quotient Image Self-quotient image (SQI) is an extension to quotient image [100]. It has similar lighting invariant properties to quotient image but can be calculated from one image, does not require

neither additional empirical learning, nor assumption about face images [110].The algorithm can be summarized in the following steps:

1. Select several smoothing kernels $G_1, G_2, \ldots, G_n$ and calculate corresponding weights $W_1, W_2, \ldots, W_n$ according to image I, and than smooth it by each of the weighted anisotropic filter WG_i:

$$I'_k = I \oplus \frac{1}{N} WG_k, k = 1, 2, \ldots, n, \tag{2.17}$$

   where weights W_k are constructed for each smoothing kernel at each image pixel's neighborhood (i, j) as

$$W(i,j) = \begin{cases} 0, & I(i,j) < \tau; \\ 1, & I(i,j) \geq \tau. \end{cases} \tag{2.18}$$

   where $\tau = meanI(i,j)$.

2. Calculate self-quotient image between input image and its each smoothed version:

$$Q_k = \frac{I}{I'_k}, k = 1, 2, \ldots, n. \tag{2.19}$$

3. Transfer self-quotient image with nonlinear function (logarithm, arctangent, or sigmoid):

$$D_k = T(Q_k), k = 1, 2, \ldots, n. \tag{2.20}$$

4. Sum all nonlinear transformation results:

$$Q = \sum_{k=1}^{n} m_k D_k, k = 1, 2, \ldots, n, \tag{2.21}$$

We used 5 Gaussian smoothing kernels with $\sigma_k = \sqrt{2}^k$ and, as suggested by the authors, $m_k = 1$, for each $k = 1, \ldots, 5$. Typical normalization results are presented in Fig. 2.17.

Despite the fact that SQI normalization generates visually better face images, neither identification, nor verification accuracy of our enhanced with optimized geometric normalization baseline algorithm was improved.

(a) (b) (c) (d)

Figure 2.17: Face images before (a), (c) and after SQI photometric normalization (b), (d).

Although the authors [110] show higher face recognition algorithm performance after SQI normalization applied to geometrically normalized face images, they tested that only with the holistic methods, which are very sensitive to even very small illumination changes.

2.2.4.2.2 Local Mean and Variance Normalization This local normalization is a simple operation for each image pixel. It is computed by subtracting an estimation of a local mean in neighborhood with radius r from the original pixel value, and dividing the got value by an estimation of a local standard deviation in the same neighborhood:

$$\begin{aligned} I'(x,y) &= \frac{I(x,y)-m_I(x,y,r)}{\sigma_I(x,y,r)}, \\ m_I(x,y,r) &= \frac{\sum_{y_0=-r}^{r}\sum_{x_0=-r}^{r} I(x+x_0,y+y_0)}{(2r+1)^2}, \\ \sigma_I(x,y,r) &= \sqrt{\frac{\sum_{y_0=-r}^{r}\sum_{x_0=-r}^{r} I(x+x_0,y+y_0)^2}{(2r+1)^2} - m_I(x,y,r)^2}. \end{aligned} \tag{2.22}$$

Typical normalization results are presented in Fig. 2.18.

We tested how identification rate is influenced by the radius of local neighborhood 2.19. Best identification rate achieved in this experiment is $Rank1 = 88.23\%$ using neighborhood limited by radius 5.

2.2.4.2.3 Local Histogram Equalization This local normalization is applied for each image pixel by equalizing the histogram in its local neigh-

(a) (b) (c) (d)

Figure 2.18: Face images before (a), (c) and after local mean and variance normalization with local neighborhood radius $r = 7$ (b), (d).

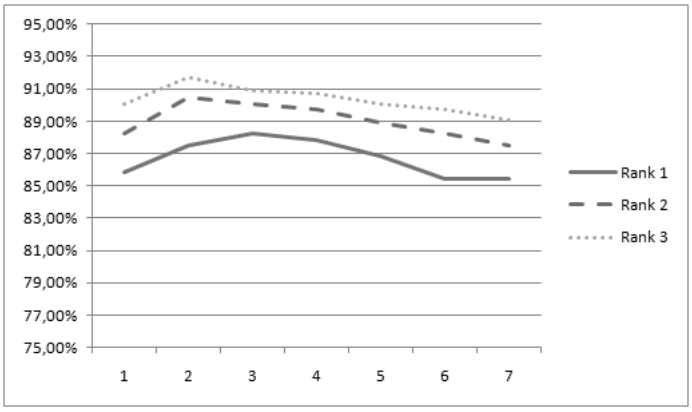


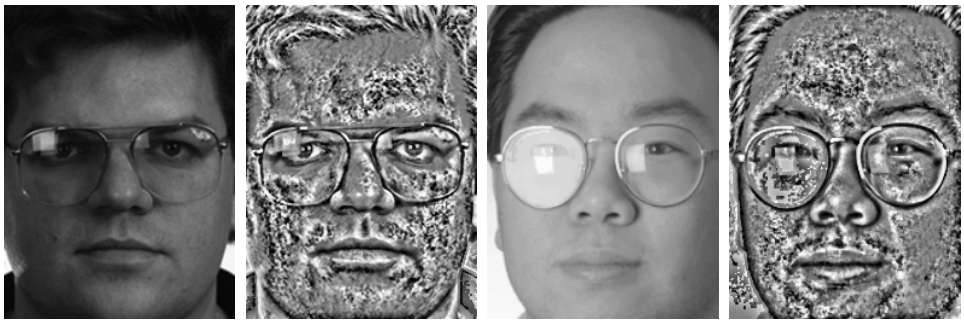
Figure 2.19: Identification accuracy dependency on local mean and variance normalization neighborhood radius.

borhood limited by radius r:

$$I'(x,y) = \frac{1}{(2r+1)^2}\left(\sum_{\substack{y_0=-r\ldots r, x_0=-r\ldots r \\ I(x+x_0,y+y_0)<I(x,y)}} 1 + \sum_{\substack{y_0=-r\ldots r, x_0=-r\ldots r \\ I(x+x_0,y+y_0)=I(x,y)}} 0.5\right). \quad (2.23)$$

We tested how identification rate is influenced by the radius of local neighborhood 2.21. Best identification rate achieved in this experiment is $Rank1 = 87.67\%$ using neighborhood limited by radius 5.

2.2.4.2.4 Summary The tested photometric normalizations can reduce lighting variation between face images. Although SQI gives the most visually acceptable results, it does not improve recognition accuracy. This



(a) (b) (c) (d)

Figure 2.20: Face images before (a), (c) and after local histogram equalization with local neighborhood radius $r = 7$ (b), (d).

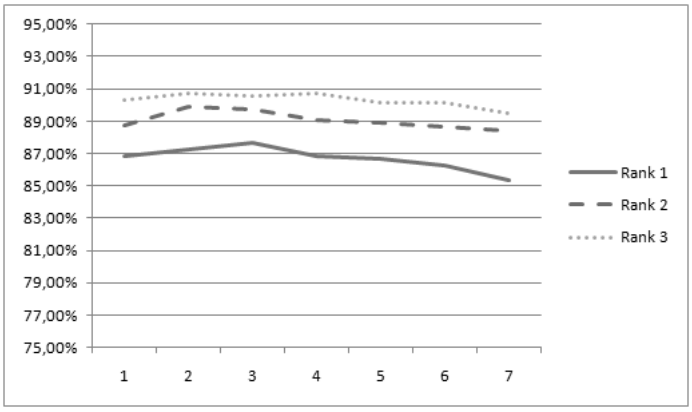


Figure 2.21: Identification accuracy dependency on local histogram equalization neighborhood radius.

happens because SQI is optimized for holistic methods and is not suitable for local features. In this case, Gabor features are already resistant to such type of lighting variation. The other two - LMVN and LHE - improve face recognition accuracy in the similar manner by emphasizing the noise (actually, face skin texture) in the regions which seem almost constant in the original face images. A total improvement of $Rank1$ from 80.33% to 88.23% was achieved by simple local photometric normalization method. Verification performance change is presented in Fig. 2.22.

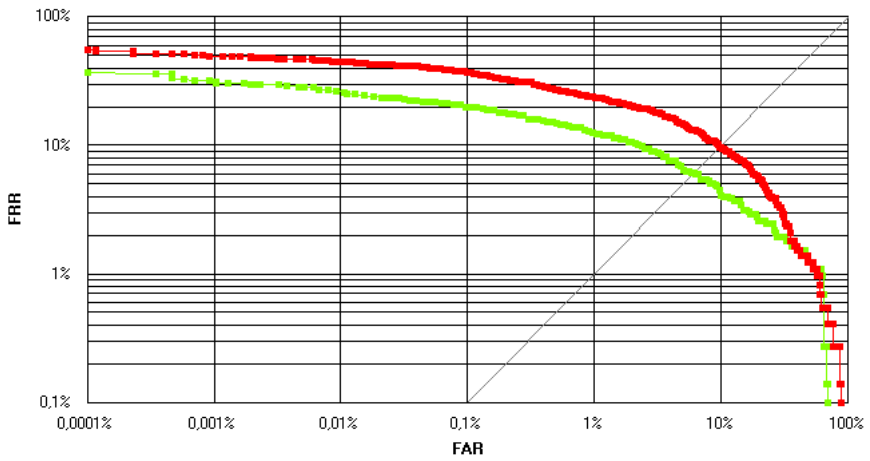


Figure 2.22: Verification accuracy improvement from the final geometric face image normalization parameters (red curve) to the final photometric face image normalization (green curve).

2.2.4.3 Regular Grid Density and Gabor Feature Extension

From the previous experiments we fixed geometric and photometric normalizations and further analyzed how identification rate depends on regular grid density. Fig. 2.23 shows that there are only marginal changes after the grid is more dense than 13 pixels between adjacent grid positions. 13 pixels correspond to $\frac{1}{7}$ of inter-ocular distance, and generates 12×20 grid with 240 positions. That is a vector of 9600 Gabor responses.

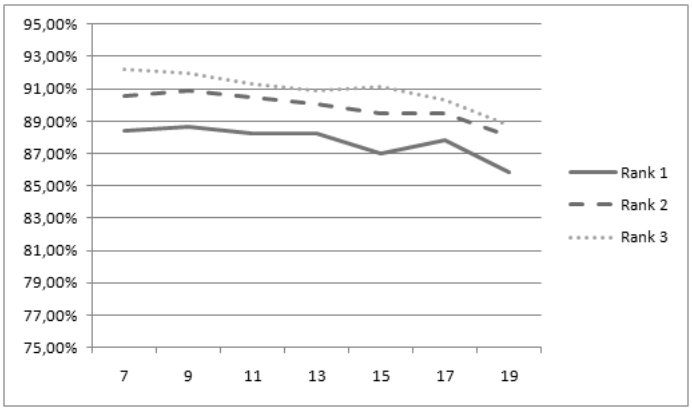


Figure 2.23: Identification accuracy dependency on regular grid density (number of pixels between adjacent grid positions).

Decreasing density to 9 pixels between adjacent grid position can im-

prove identification rate to 88.64%, see Fig. 2.23. However, this increases number of Gabor responses in features vector more than twice to 20160. On the other hand, we could extend Gabor feature by increasing number of used scales or directions. A natural extension of 5 scales to 9 (4 additional intermediate scales) gives identification rate of 88.09% (17280 Gabor responses). While adding 8 additional intermediate directions gives identification rate of 89.64% (19200 Gabor responses). Further, if both, scales and directions, are extended, identification rate will jump to 90.17% (34560 Gabor responses). If similar number of Gabor responses is generated by making even more dense Gabor gird but using classic Gabor feature, identification rate is not improving anymore, see Fig. 2.23. A total improvement of $Rank1$ from 88.23% to 90.17% is achieved when number of Gabor responses was increased by a factor of 3.6. Features extraction time is not an issue here because of the suggested accelerated calculation of Gabor features in spatial domain 2.2.1. But, features vector length is huge and that's an open problem in our current algorithm. On the other hand, 33376 of Gabor responses are used in [126] with significantly lower $Rank1 = 85.00\%$ identification rate.

2.2.4.4 Gabor Feature Similarity

The baseline similarity function between two face images computes the average of all Gabor jets similarities, where Gabor jet is a set of Gabor responses calculated at one specific point of one specific frequency but 16 different directions. When in previous section we expanded Gabor feature to 9 frequencies, it made sense to form the Gabor jet in a different manner, i.e. use all 9 frequencies of one specific direction as a Gabor jet. However, Gabor filters of higher frequency tend to generate responses with higher values than those of lower frequency. We illustrated an average Gabor filter response for different frequencies in Fig. 2.24. The difference between the highest and lowest average response is almost 3 times. Additionally, all the vectors took similar form of descending values, which reduces recognition

accuracy of our similarity function.

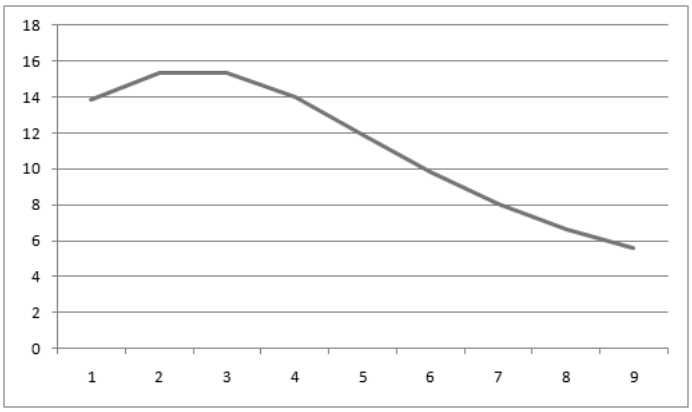


Figure 2.24: Average Gabor response value dependency on frequency $\{1, \ldots, 9\}$.

To overcome this problem we introduce Gabor responses normalization. We used the gallery dataset 2.2.3 to calculate an average Gabor responses for each used Gabor filter. Further, we normalize each calculated Gabor response by dividing it by previously found average value. After the normalization is performed we used suggested new form of Gabor jet to calculate similarity between any pair of face images. Identification rate jumped from 90.17% to 91%. An increased verification accuracy is demonstrated in Fig. 2.25.

Furthermore, using the same feature set we can calculate similarity in two different ways simultaneously. This does not introduce additional computational heaviness but produces almost twice number of similarities between different Gabor jets. By averaging all the similarities to produce final similarity score, we increase identification accuracy one more time to $Rank1 = 92.94\%$.

2.2.5 Further Evaluation

We performed a greedy optimization of different steps from the proposed baseline algorithm and achieved high face recognition accuracy on one standard experiment. To validate these results we tested the proposed optimizations in the rest of FERET experiments 2.2.3. We showed that the proposed

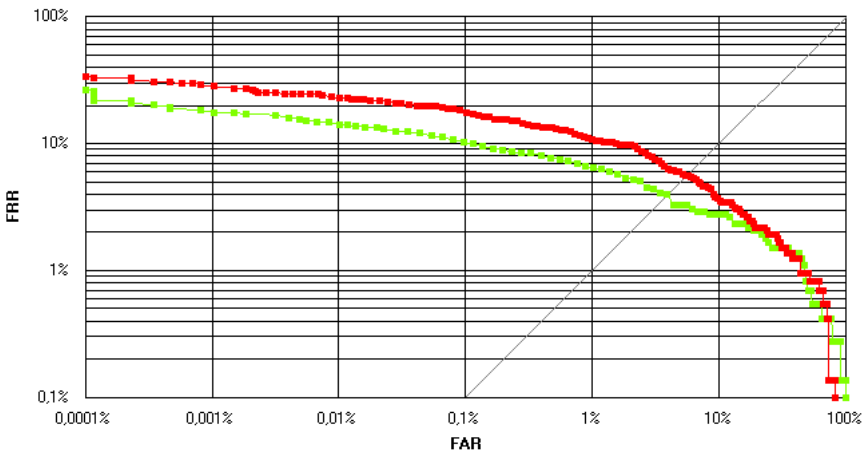


Figure 2.25: Verification accuracy improvement if Gabor jet formed from different directions (red curve) is changed to Gabor jet formed from different frequencies (green curve).

baseline algorithm is highly configurable and the final chosen parameters set is not always the best choice in real-world applications. Thus, we report the results for several examples of the parameters sets, which we found during this research, see Table. 2.1:

1. $Proposed_{BASELINE}$ - the proposed baseline parameters set, as described in section 2.2.2.

2. $Proposed_{FINAL}$ - this is the final parameters set found during our greedy optimization, i.e. geometric normalization to 164×260 pixels with fixed eyes positions at $(36, 120)$ and $(128, 120)$, local mean and variance normalization with $radius = 7$ local neighborhood, Gabor features of 9 frequencies and 16 directions calculated at every 13 pixel in the image resulting in a regular grid of 12×20 points, each Gabor response normalization by the average response of gallery dataset, and calculating similarity of two different forms of Gabor jets simultaneously.

3. $Proposed_{LHE+F\_ONLY}$ - same as $Proposed_{FINAL}$ but using local histogram equalization for photometric normalization, and calculating similarity only between frequency-based Gabor jets.

4. $Proposed_{GN\_PHN\_REV}$ - same as $Proposed_{FINAL}$ but geometric and photometric normalizations were revisited, i.e. geometric normalization to 128×192 pixels with fixed eyes positions at $(30, 70)$ and $(98, 70)$, local mean and variance normalization with $radius = 5$ local neighborhood, and regular grid of 12×18 points (at every 11 pixel in the image).

5. $Proposed_{LOW\_DIM}$ - same as $Proposed_{GN\_PHN\_REV}$ but using classic Gabor features of 5 frequencies and 8 directions.

Final verification accuracy achieved by using best parameters sets in two most common FERET experiments - Fb and Dup1 - is reported in Fig. 2.26.

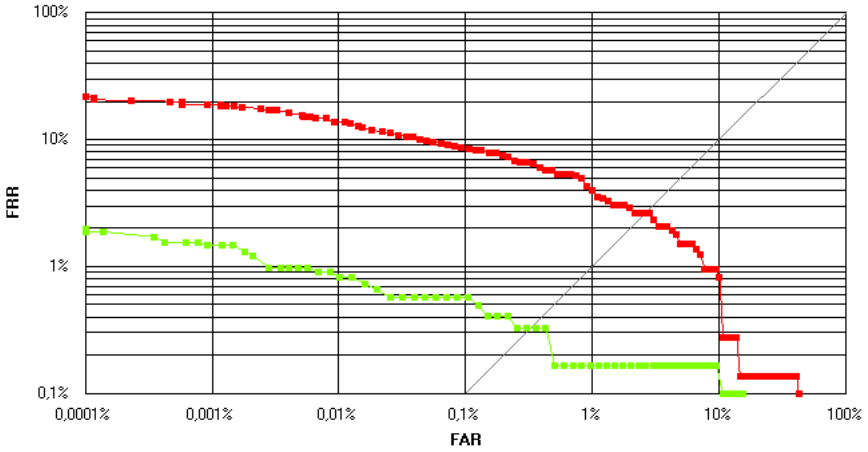


Figure 2.26: Final verification accuracy measured in Dup1 experiment using $Proposed_{GN\_PHN\_REV}$ parameters set (red curve) and Fb experiment using $Proposed_{LHE+F\_ONLY}$ parameters set (green curve).

Table 2.1: Identification accuracy as *Rank*1 (in percents) for different face recognition methods in different FERET experiments. Additional columns indicate whether the method needs training, number of dimensions in features vector, extraction speed on a single core of *Intel Core 2 Duo CPU P8400 @ 2.26 Ghz*.

Method	Experiment				Trained	Dimensions	Extraction Speed (ms)
	Fb	Fc	Dup1	Dup2			
$EBGM$ [116]	95.00	82.00	59.10	52.10	yes		
$LBP - Weighted$ [2]	97.00	79.00	66.00	64.00	yes	2891	
$Log - GaborPCA$ [82]	97.99	90.21	72.44	65.81	yes	900	
$LGBPHS - Weighted$ [122]	98.00	97.00	74.00	71.00	yes	74000	
$PGFC$ [103]	99.00	97.00	87.00	82.00	yes	$\sim$3000	
$Gabor - LBP - KDCV$ [105]	98.00	98.00	90.00	85.00	yes		
$FIGHT - RF$ [61]	99.58	99.00	91.90	88.89	yes	5000	
$EPFDA - LGBP$ [99]	99.58	99.00	92.00	88.89	yes	11000	
$GRM - Local$ [67]	97.50	97.90	79.50	83.80	no		
$GaborJets, BordaCount$ [126]	99.50	99.50	85.00	79.50	no	33376	
$Proposed_{BASELINE}$	91.05	93.81	66.90	61.97	no	4000	$\sim$3
$Proposed_{FINAL}$	99.92	**100.00**	92.94	88.89	no	34560	$\sim$30
$Proposed_{LHE+F\_ONLY}$	**100.00**	**100.00**	91.55	87.18	no	34560	$\sim$30
$Proposed_{GN\_PHN\_REV}$	99.75	**100.00**	**94.60**	**93.59**	no	31104	$\sim$20
$Proposed_{LOW\_DIM}$	99.75	**100.00**	92.38	88.89	no	8640	$\sim$8

2.3 Summary

We showed how the proposed baseline face recognition algorithm can be evolved into the state-of-the-art method, which, to the best of our knowledge, achieves the highest accuracy of all currently published methods, see Table 2.1. Each face recognition algorithm step was greedily optimized, which led to very high identification rate, but in the same time generalized well to other FERET datasets[4]. We would like to emphasize here that the proposed method doesn't require any training step. It achieves better accuracy in comparison with methods which require training. It is also an order of magnitude more accurate in the most difficult Dup1 experiment than other methods which do not use training. Additionally, identification step is not overburden with Borda Count method, as was done in [126].

The proposed rapid calculation of Gabor feature at a regular grid allows to use Gabor features together with real-time face detection 2.1.1 in real-time face recognition applications. The proposed method with $Proposed_{GN\_PHN\_REV}$ parameters set can perform facial features extraction at 50 faces per second rate. Slightly less accurate $Proposed_{LOW\_DIM}$ parameters set will increase the processing speed to 125 faces per second.

Chosen geometric face image normalization is not novel, yet very simple and effective. Furthermore, chosen photometric face image normalization could be replaced by more sophisticated method which in no doubt can further increase face recognition accuracy, but it was left out of this thesis scope. However, the proposed Gabor feature similarity metric is not met in the literature and is a key factor of achieving high recognition accuracy.

One more thing, which was intentionally skipped in this thesis, is face recognition accuracy degradation when automatic face detection is used to locate faces in images. We used manually labeled coordinates of eyes,

[4]We should mention here that we fixed several mistakes in original grayscale FERET protocol which assigned several different unique identifiers to the same person. These mistakes were reported and fixed while constructing color FERET database. If someone still works with original grayscale FERET protocol (and makes a huge mistake by trying to identify different photos of the same person as belonging to different persons), our results using such faulty protocol can be found in appendix A

however, more detailed analysis showed that manual labeling is not very accurate by itself. Thus, we expect small degradation in recognition accuracy if automatic face detection will be used. On the other hand, that degradation could be coped with by using Gabor feature ability to estimate displacement vector [116] in a more flexible matching function.

Chapter 3

Iris Recognition

3.1 Introduction

For a long time iris recognition was monopolized by one patented algorithm [30] supported by another patent of the same idea of person identification by iris [33]. Patent expiration date created enlarged interest in developing new methods of iris recognition which will surely outperform the fifteen years old algorithm. Recent studies [85] showed that already more accurate algorithms exist.

Iris texture features provides a unique high dimensional information that explains why iris recognition based verification has the lowest false acceptance rate among all types of biometric verification systems [28], [23]. A transparent and simple distribution of calculated distances between iris pairs allows to make a fine prediction of verification quality for a chosen distance threshold [30]. Large-scale application of iris recognition for border-crossing control showed up robustness of the technology even in case of more than half million different iris images spanning about 150 nationalities [25]. Appearance of new technologies that enable to capture iris images even at distance of 20 meters significantly reduced requirements for subject cooperation with the automated identification system [3].

Iris is an annular part of the eye image with the inner boundary that coincides with the pupils outer boundary and the outer boundary that borders with the sclera. The iris image has a complex texture originating from many particles of different size. The particles comes from freckles, furrows,

stripes, coronas, superimposed eyelashes, etc. Distribution of the particles is random from eye to eye and that is the key why iris image is magnificent biometric source for verification and identification. If one would have an robust algorithm for detection and classification of iris texture particles, the iris recognition would be rather similar to fingerprint recognition: local features, their type, position, and other attributes of the shape, replacing fingerprints minutiae type, position and orientation information. Such iris recognition scenario is non-utopic and in parallel to phase-based [23], [24], [26], [30], [73], and zero-crossing iris representations [14], [77], [96], more intuitive on iris texture based descriptors, are used for iris recognition [60], [81], [115], [121], [125]. Using image processing terminology the particles may be renamed to noisy dark and bright blobs having different shapes and sizes. It seams that particles or noisy blobs and their attributes can be detected with essentially bigger uncertainty than the fingerprints minutiae and their descriptors. The good news are that frequently all iris image is formed from a lot of particles and one can expect much larger amount of local features extracted on basis of noisy local blobs than in the case of fingerprints.

Despite seemingly differences in phase, zero-crossing, and texture based iris description all approaches are pretty similar. It is well known that multiscale phase and zero-crossing information is often sufficient for initial signal reconstruction [65], [70]. For example 256×256 image can be reconstructed [22] from the binarized Fourier transform phase with probability $p > 1 - 2^{-2000}$. Thus, all three iris descriptions with big probability preserve information about the original image texture. On the other hand, it is well known that distortion of image complex Fourier phase leads to major transmutations in reconstructed image and parallel distortions of the Fourier modulus with preservation of its phase gives still recognizable reconstructed images. The phase keeps the information "where are the details" and this is the most significant information that differentiates one iris pattern from another. We think that's why phase-based iris recognition

approach is so efficient and popular.

Typical iris recognition system is comprised of the following steps:

1. Segmentation - localization of iris inner and outer boundaries by circles [30], ellipses [62], active contours [95], etc. Additionally, localization of occlusions by eyelids/eyelashes and illumination noise, usually over-exposed parts of the image.

2. Geometric normalization - warping of iris image into polar coordinate system.

3. Features extraction - extraction of iris texture encoding features.

4. Features matching - comparison of previously extracted features sets.

3.2 Proposed Iris Recognition Algorithm

3.2.1 Segmentation

Iris segmentation consists of iris inner and outer boundaries localization, optional detection of superimposed upper and lower eyelids, and detection/removal of reflections from the cornea or eyeglasses. Our segmentation algorithm is a sequence of the following basic steps. The input of each step is the output of the previous algorithm step:

1. Deinterlace the original image.

2. Detect and remove the reflections.

3. Reduce the noise.

4. Estimate approximate pupil center.

5. Estimate second order image derivatives along radii coming from the pupil center.

6. Group zero crossings of the second order derivatives to connected contours.

7. Calculate distributions of the connected contours distances to the pupil center.

8. Group connected contours to iris inner and outer boundaries.

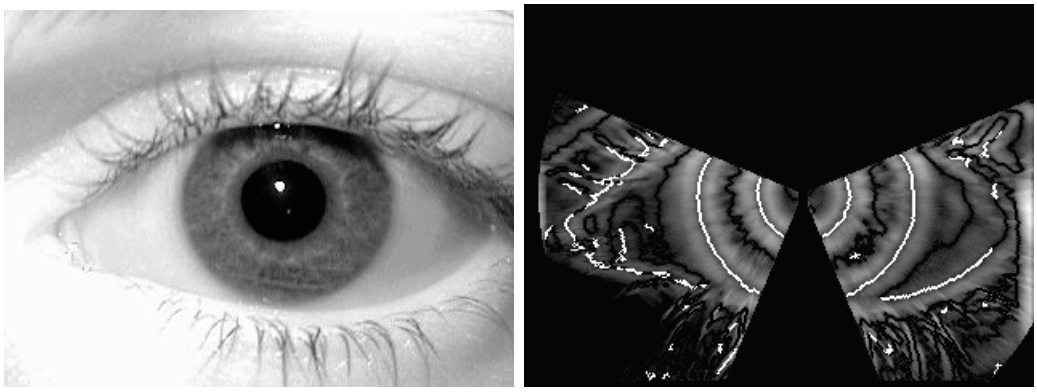
9. Describe the iris inner and outer boundaries by short Fourier series.

Original images of the irises are often captured by devices producing interlacing artifact. Our approach in solving "interlacing noise" problem is trivial, we downsample original iris image in horizontal and vertical dimensions by a factor of two. Such deinterlacing entirely excludes interlacing artifact. Thus, we get four times increased processing speed and decreased memory consumptions in further segmentation procedures. Additionally, iris segmentation quality is also increased.

Reflections from the cornea or eyeglasses are removed by contrast stretching followed by detection of small connected areas having brightness greater than a threshold and replacing them with an average of neighboring pixel values that are lower than a threshold. Additionally, some authors use parallel procedure of removing eyelashes. Our experiments showed that our implementation of eyelash removal leads to a slight degradation of the following iris recognition quality, thus, we skipped this procedure from the segmentation. Our segmentation procedure also skips detection of upper and lower eyelids and their boundaries.

Dark area of the pupil is detected by convolving original image with isotropic two-dimensional Gaussian filter followed by additional blurring with a cross form averaging filter to suppress thin dark eyelashes and possible small bright areas of reflection that were not correctly detected at previous stage. Area of the pupil is defined by global minimum point of the blurred image and its surrounding pixels of similar values (also controlled by a threshold). The boundary of pupil area is approximated by a short Fourier series calculated in respect of point A (center of the pupil) that, together with the Fourier series coefficients, is free parameter in minimization of the approximation error.

Further, directional second order image derivatives are calculated. Their directions goes along radii outgoing from the point A. The derivatives are calculated only along directions that form angle ϕ with horizontal line such that $-3\pi/8 \leq \phi \leq \pi/8$ or $7\pi/8 \leq \phi \leq 11\pi/8$ radians and limited by the distance which is sufficiently big to cover up irises of the biggest size. An example of original iris image with its directional second order derivatives are showed in the Fig. 3.1. Upper and lower black sectors correspond to skipped angles motivated by a high probability of eyelids and eyelashes.



(a) (b)

Figure 3.1: (a) Original 242116.tiff image of the NIST ICE2005 iris data set. (b) Directional second order derivatives estimated along radii outgoing from the point A. White dots mark the zero crossing.

Zero crossings of the directional second order derivative give possible locations of iris inner and outer boundaries. A histogram of distances to point A is calculated for each connected contour formed from zero crossings (as showed in Fig. 3.2). For the particular iris image two histograms with big maximums around 125 and 135 pixels correspond to left and right iris outer boundary and two histograms with clear maximum at 55 pixels correspond to iris left and right inner boundary.

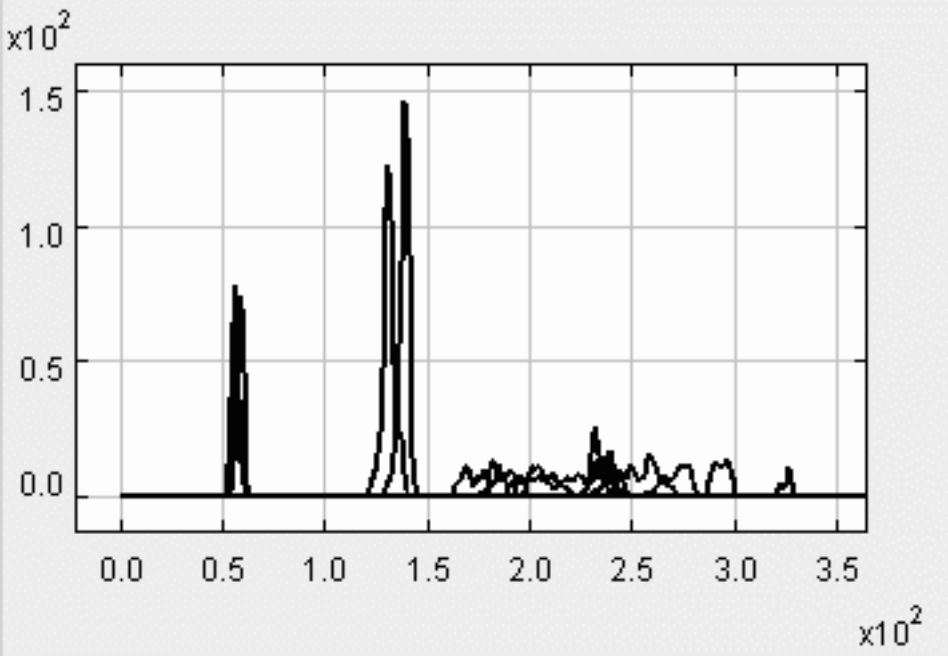


Figure 3.2: Distributions of zero-crossing points distances to pupil center. Connected contours of zero-crossings points correspond to different distributions.

The detected iris inner and outer boundaries are approximated by the following Fourier series:

$$\begin{aligned} x^{\alpha}(\phi) &= a_0^{\alpha} + a_1^{\alpha}\cos(\phi) + b_1^{\alpha}\sin(\phi), \\ y^{\alpha}(\phi) &= c_0^{\alpha} + c_1^{\alpha}\cos(\phi) + d_1^{\alpha}\sin(\phi), \\ & 0 \leq \phi < 2\pi. \end{aligned} \tag{3.1}$$

Here $\alpha = i$ and $\alpha = o$ stand for inner and outer iris boundary respectively. The Fourier coefficients and polar axis (angle where $\phi = 0$) are found by minimizing the approximation error of iris inner and outer boundaries using equations 3.1. Note that any circles or ellipsis are only particular cases of boundaries for which the exact Fourier approximations can be found.

3.2.2 Geometric Normalization

Geometric normalization of iris texture is necessary because diameter of pupil relative to iris diameter is constantly changing, even under steady illumination. As suggested by [30], we used a rubber sheet method for converting iris texture from captured image to doubly dimensionless system

(see Fig. 3.3). The rubber sheet method assigns to each point on the iris,

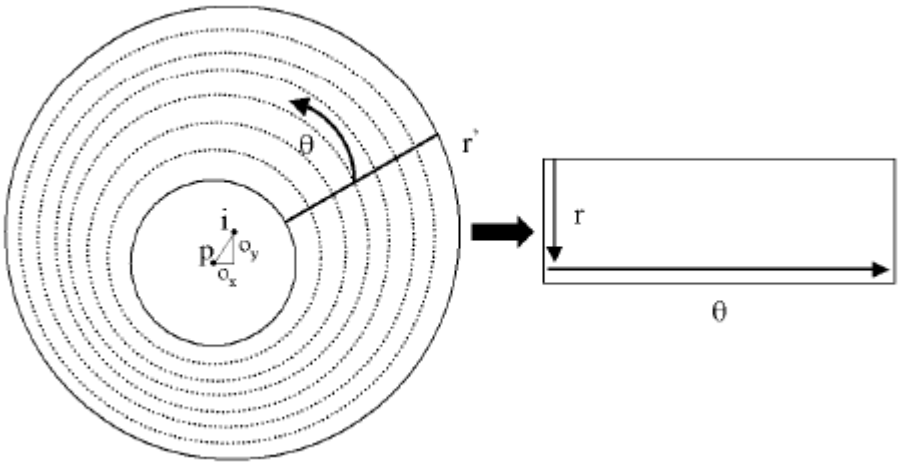


Figure 3.3: Geometric normalization of the iris image through the rubber sheet method, adapted from [91].

regardless of its size and pupillary dilation, a pair of real coordinates (r, θ), where r is on the unit interval $[0, 1]$ and θ is an angle in $[0, 2\pi]$.

[5] suggests to use linearly-guessed reference point instead of pupil center for the rubber sheet method. However, our experiments showed no improvements to recognition accuracy after such change.

3.2.3 Local Features as Local Extrema of Multiscale Taylor Expansion

We developed our scheme of iris representation by local features considering the following constraints: number of local features should be easily controlled and their attributes should be as compact and simple as possible. The main goal of our representation is to control memory consumption and iris matching speed, thus, leading to the fast and robust iris-based verification and identification. The proposed extraction of compact iris representation (which could be further compared to other iris representations created by the proposed method) can be summarized by the following algorithm:

1. Choose the local descriptors of the iris texture.

2. Calculate the chosen descriptors at each pixel of the iris texture.

3. Leave only the most significant descriptors.

4. Divide iris texture into a fixed number of sectors.

5. Limit the number descriptors in each sector.

3.2.3.1 Local Descriptors

Let us consider one-dimensional iris signal having a fixed pseudopolar radius $r = const$ as an analytic one-dimensional signal $u = u(x)$, where x is pseudopolar angle. It is well known that any signal can be expressed in classic Taylor series expansion centered on any fixed $x = x_i$. The zero coefficient of the expansion equals to $u(x_i)$ and is strongly influenced by iris lighting conditions, thus cannot be a robust iris descriptor. The second two coefficients equal to $\frac{u'(x_i)}{1!}$ and $\frac{u''(x_i)}{2!}$. The first derivative $u'(x)$ estimates transition rate of the signal around point x, i.e. the big positive value of $u'(x)$ means a fast increasing transition of the signal around point x and a big negative$u'(x)$ means a fast decay around the same point. The sign of the second derivative $u''(x)$ provides information about the type of concavity of graph of the transition around x, i.e. if $u''(x) > 0$ then graph of transition of the signal is concave upward and $u''(x) < 0$ gives that graph of the signal that is concave downward around point x. If transition of x values corresponds to transition of $u = u(x)$ from left to right through a dark or bright blob one can expect the following transitions of the signs: Left and right

Table 3.1: Transitions of signs.

Blob	Derivative	Transition of signs
Dark	$u'(x)$	$-+$
	$u''(x)$	$-+-$
Bright	$u'(x)$	$+-$
	$u''(x)$	$+-+$

boundary of a dark blob can be marked by local negative minimum and positive maximum of the first derivative. The second derivative will have zero crossings around these local extrema. If we expand such features to two-dimensional case, only a fraction of$u_{xx}(x, y)$ zero crossings will correspond to $u_x(x, y)$ local extrema. Therefore, local extrema of both two-dimensional functions $u_x(x, y)$ and $u_{xx}(x, y)$ were used as primary source for derivation of iris local features. Notice that $u'(x)$ and $u''(x)$ derivatives can be estimated only numerically in the context of digital iris data and the result of approximations significantly depends on scale at which increments of the signal are estimated. $u''(x)$ approximation will have a positive maximum close to a dark blob center and a negative minimum close to a bright blob center if the chosen scale resonates with blob size.

We neglect Taylor series expansion zero term $u(x_i)$ which sign do not provide a meaningful information and is strongly conditioned by background illumination if no sophisticated normalization or thresholding procedure is applied. The first two terms of Taylor series expansion defines local linear and quadratic trend around the center point of the expansion. Extrema of the first derivative indicate points where signal $u = u(x)$ has greatest local asymmetry. Similarly, extrema of second derivative indicate points where $u = u(x)$ has greatest local symmetry. Similar meaning can be provided to the extrema of higher order derivatives.

The well known Taylor series expansion

$$u(x) = \sum_{n=0}^{\infty} \frac{u^{(n)}(a)}{n!}(x - a)^n \tag{3.2}$$

tells us that sufficiently regular function can be reconstructed from its Taylor coefficients $\frac{u^{(n)}(a)}{n!}$. In theory Taylor coefficients can be derived from $u(x)$ values known at any narrow surrounding of the center point $x = a$. Let us consider some details of estimation of local extrema of the first two derivatives of iris texture and conditions under which a particular extremum point is included in the set of iris local features. Estimation technique of derivatives at different scales is similar to the wavelet multiresolution anal-

ysis. Let us fix any symmetric (even) smooth filter $p = p(x) \geq 0$ having maximum at center$x = 0$ and L_1 energy equal to 1:

$$\int_{-\infty}^{\infty} p(x)\, dx = 1. \tag{3.3}$$

Let us fix some scale $\sigma > 0$ and consider convolutions

$$(u' * D_\sigma p)(x) = 1/\sigma \int_{-\infty}^{\infty} p(\xi/\sigma) u'(\xi - x)\, d\xi \tag{3.4}$$

and

$$(u'' * D_\sigma p)(x) = 1/\sigma \int_{-\infty}^{\infty} p(\xi/\sigma) u''(\xi - x)\, d\xi. \tag{3.5}$$

In the bigger scale σ the more blurred version of $u'(x)$ and $u''(x)$ represents $(u' * D_\sigma p)(x)$ and $(u'' * D_\sigma p)(x)$. One the other hand, if σ approaches zero then $(u' * D_\sigma p)(x)$ and $(u'' * D_\sigma p)(x)$ approaches to $u'(x)$ and $u''(x)$ respectively.

Filter $p = p(x)$ defines rule of derivatives averaging. Function $\mathrm{sech}(x) = 2/(e^x + e^{-x})$ meets our requirements of symmetry, exponentially tends to zero at infinity and its L_1 norm equals to π. Therefore we chose the following filter function

$$p(x) = \frac{\mathrm{sech}(x)}{\pi}. \tag{3.6}$$

We have discrete periodic sequence $u_i = u(i)$ of length NX in case of iris texture. Defining $u(x)$ by linear interpolation $u(x) = u_i + (x - i)(u_{i+1} - u_i)$, $x \in [i, i + 1)$, we will have

$$u''(x) = \sum_{i=-\infty}^{\infty} (u_{i+1} - 2u_i + u_{i-1})\delta(x - i) \tag{3.7}$$

and

$$\begin{aligned}
(u'' * D_\sigma p)(x) &= \frac{\sum_{i=-\infty}^{\infty} (u_{i+1} - 2u_i + u_{i-1}) p\left(\frac{x+i}{\sigma}\right)}{\sigma} && (3.8)\\
&= \sum_{i=0}^{NX-1} \Delta u_i p_\sigma^{NX}(x + i) && (3.9)\\
&= \sum_{i=0}^{NX-1} u_i \Delta p_\sigma^{NX}(x + i). && (3.10)
\end{aligned}$$

Here $\delta(x)$ is Dirac's delta function,

$$\Delta u_i = u(i + 1) - 2u(i) + u(i - 1), \tag{3.11}$$

and

$$p_{\sigma}^{NX}(x) = \frac{\sum_{i=-\infty}^{\infty} p(\frac{x+i*NX}{\sigma})}{\sigma}. \tag{3.12}$$

Similarly, for the first derivative we have

$$(u' * D_{\sigma}p)(j) = -\sum_{i=0}^{NX-1} \Delta u_i q^{NX}\sigma(j+i), \tag{3.13}$$

$$q_{\sigma}^{NX}(x) = \sum_{i=-\infty}^{\infty} 2\,\mathrm{arctg}\left(\frac{x+i*NX}{\sigma}\right). \tag{3.14}$$

From symmetry of the filter function 3.6 follows that 3.8 defines a filtration of one-dimensional data $\{u_i\}$ by an even-symmetric filter and 3.13 defines an odd-symmetric filter. In wavelet terminology the even and odd filters has two and one vanishing moments, respectively. The proposed filters has more freedom in choosing basic function $p^{\sigma} = p^{\sigma}(x)$ in comparison with dyadic wavelet filtration and do not have dyadic restrictions for the scale σ. To regularize the filtration results that at different scales σ approximate second (even filters) or first (odd filters) order derivatives we used some smoothing in orthogonal direction by applying the same filter $p^{\sigma} = p^{\sigma}(y)$ with smaller scale value σ.

Table 3.2 presents scales, type of symmetry and direction which was used to calculate differences. The differences calculated along x and y directions correspond, respectively, to horizontal and vertical type filters. Derivatives along x direction were estimated at three different scales while derivatives along y direction were estimated only at one scale.

3.2.3.2 Significant Local Descriptors

Blurred derivative may be considered as the coefficients of Taylor series expansion beside $(x-a)/1!$ and $(x-a)^2/2!$ terms estimated at different scales or coefficients of wavelet expansion calculated at different scales with asymmetric or symmetric mother wavelets that have one or two vanishing moments. It is well known that coefficients with highest absolute values give the main contribution in approximation of original signal by the expansions. Thus, extrema or more exactly, positive local maximums and negative local

Table 3.2: Derivatives and averaging parameters used in derivation of multiscale Taylor expansion and local features.

Scale and filter type index s	Blurred derivative	$\sigma_x = \sigma$ of horizontal direction	$\sigma_y = \sigma$ of vertical direction
0	u_x	$\frac{6}{8}$	$\frac{1}{2}$
1	u_x	$\frac{9}{8}$	$\frac{1}{2}$
2	u_x	$\frac{15}{8}$	$\frac{1}{2}$
3	u_{xx}	$\frac{9}{8}$	$\frac{1}{2}$
4	u_{xx}	$\frac{15}{8}$	$\frac{1}{2}$
5	u_{xx}	$\frac{24}{8}$	$\frac{1}{2}$
6	u_y	$\frac{1}{2}$	$\frac{9}{8}$
7	u_{yy}	$\frac{1}{2}$	$\frac{15}{8}$

minimums of blurred derivatives, are good candidates for local descriptors of iris texture. Additionally, we required that the local extremum at particular scale σ_0 exceeded the same derivative at the same position but with rougher scales $\sigma > \sigma_0$ (that conditioned more stable extrema). And the very last condition for including the extremum point to the set of local iris features was its magnitude. However, we decided to avoid thresholds in this place and used the following procedure for choosing extrema points (see Fig. 3.4):

1. Divide iris texture annual along polar angle direction into 16 equal sectors.

2. Enumerate sectors clockwise starting at "0" hour.

3. Eliminate local extrema which are found in upper or lower lids:

   (a) Find a convex domain that includes only iris with pupil and excludes lids.

   (b) Ignore sectors 0, 1, 2, 13, 14, and 15 because iris texture in these sectors will likely be hidden by upper lid. Ignore sectors 7 and 8

because iris texture in these sectors will likely be hidden by lower lid.

4. Choose no more than K biggest local extrema from each scale and each active sector.

(a) original image (b) a convex domain that includes only iris (c) iris division into sectors

Figure 3.4: Elimination of local extrema which are found in upper or lower lids. In (c), active sectors are marked in white and inactive sectors are marked in light gray.

Fig. 3.5 illustrates local extrema points that were found in active sectors for 242116.tiff and 241643.tiff iris images of NIST ICE2005 iris data-set. These two images constitute genuine pair. The two presented derivatives u_{xx} and u_{yy} are calculated with averaging parameters defined by the 6th ($s = 5$) and 8th ($s = 7$) row of the table 3.2. White "+" and "-" mark local maximum and minimum points of the averaged Taylor expansion. The presented iris texture has no clearly expressed freckles, furrows, stripes or coronas in active sectors. However there are blurred dark and bright blobs where local maximum and minimum points indicate the blobs of the size that resonates with the averaging scale σ.

3.2.4 Similarity Metric for Comparison of Local Features Sets

Let us fix local features sets A and B formed from Taylor local extrema estimated at different scales. A particular feature $f_i = \{x_i, y_i, s_i, z_i\}$ has the following attributes:

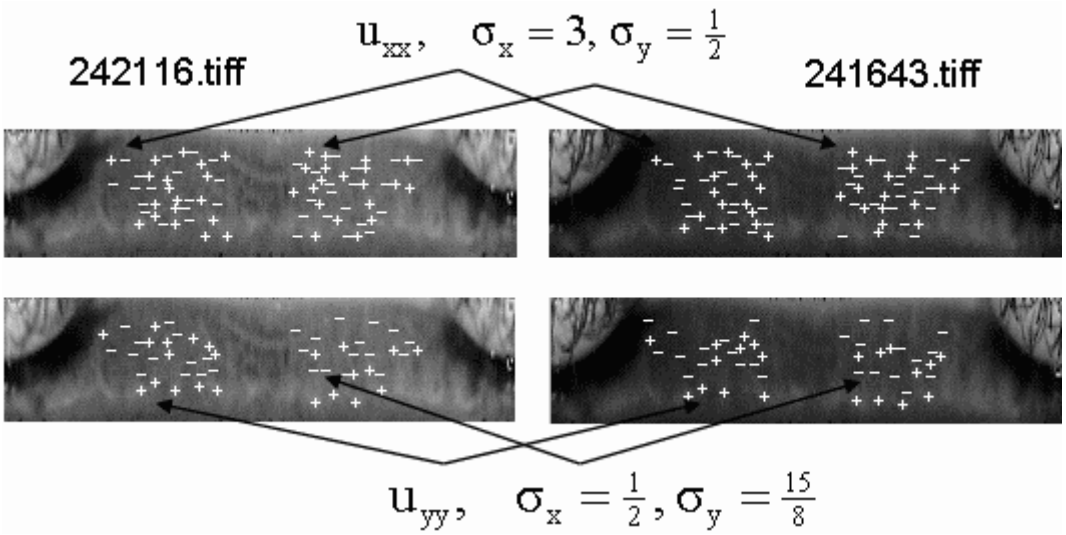


Figure 3.5: Unwarped 242116 and 241643 iris images of the NIST ICE2005 data-set. "+" and "-" mark local maximums and minimums of averaged u_{xx} and u_{yy} derivatives (u_{xx} - above, u_{yy} - below). Arrows points to the examples of extrema points that have correspondences in both images.

- (x_i, y_i) - local extremum position defined by two integer numbers x_i and y_i.
- $s_i \in \{0, 1, \cdots, 7\}$ - scale and filter type index (see table 3.2 for details).
- $z_i \in \{0, 1\}$ - extremum type marker. 0 was used for local maximum points and 1 for local minimum points.

The similarity metric between two local features sets is defined by a normalized sum of similarity scores of feature pairs. A particular pair (f_i^A, f_j^B) can gain non-zero similarity scores only if f_i^A and f_j^B correspond to the same filter and extremum type. Such features will be further called *congeneric* features. Hence the following definition can be formulated:

Definition Two local features $f_i^A = \{x_i^A, y_i^A, s_i^A, z_i^A\}$ and $f_j^B = \{x_j^B, y_j^B, s_j^B, z_j^B\}$ are *congeneric* iff they are defined by the same filter, i.e. $s_i^A = s_j^B$, and correspond to the same type of extremum, i.e. $z_i^A = z_j^B$.

Two congeneric features can gain positive similarity only if their positions are sufficiently close. Distance between two affine (x_i^A, y_i^A) and (x_j^B, y_j^B) points is estimated considering possible alignment of two irises by means of an angular rotation and radial translation. Angular rotation and radial translation correspond to some shift along x and y directions respectively[1]. With a fixed α and β shifts we use the following scores similarity expression:

$$s(f_i^A, f_j^B; \alpha, \beta) = \begin{cases} 0, \text{if } s_i^A \neq s_j^B \text{ or } z_i^A \neq z_j^B; \\ |DX - |x_i^A - x_j^B - \alpha||_+ \times \\ \times |DY - |y_i^A - y_j^B - \beta||_+, \text{otherwise.} \end{cases} \tag{3.15}$$

Here DX and DY are two positive integer parameters that define maximal acceptable distance along x and y directions and

$$|a|_+ = \begin{cases} a, & \text{if } a > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{3.16}$$

The similarity scores between A features set and α rotated and β radially shifted B set is defined as sum of the similarity scores of local features pairs, i. e.

$$score(A, B; \alpha, \beta) = \sum_{i=1}^{NA} \sum_{j=1}^{NB} s(f_i^A, f_j^B; \alpha, \beta). \tag{3.17}$$

Similarity scores between A and B features sets is defined by the following maximum:

$$score^*(A, B) = \max_{-KX \leq \alpha \leq KX, -KY \leq \beta \leq KY} score(A, B; \alpha, \beta). \tag{3.18}$$

Here KX and KY are integer parameters defining $2KX + 1$ rotations and $2KY + 1$ radial shifts under which similarity scores of shifted unwarped irises are optimized.

3.2.4.1 Warped Similarity

The similarity metric defined by equations 3.15, 3.16, and 3.18, formulas uses fixed angular and radial shifts. In [26], it is shown that iris description based on two circles model is inaccurate and significant improvement

[1]Shift along the x direction inherits angular periodicity.

of iris recognition can be obtained when an elastic adaptive grid is used for iris pixels positions. The elastic grid is extracted during the segmentation process. Our idea is to preserve simple and robust pseudopolar model with rectangular grid in segmentation and perform a sophisticated elastic matching that recovers possible incongruities in the grids. The elastic matching is based on variable shifts along angular and radial directions that are derived from principle of maximization of similarity scores. The elastic or warped similarity metric uses the following modifications of equations 3.15, 3.16, and 3.18, formulas:

$$\widetilde{score}^*(A,B) = \max_{\alpha,\beta\in T} \widetilde{score}(A,B;\alpha,\beta), \tag{3.19}$$

$$\widetilde{score}(A,B;\alpha,\beta) = \sum_{i=1}^{NA}\sum_{j=1}^{NB} \tilde{s}(f_i^A, f_j^B;\alpha,\beta), \tag{3.20}$$

$$\tilde{s}(f_i^A, f_j^B;\alpha,\beta) = \begin{cases} 0, \text{ if } s_i^A \neq s_j^B \text{ or } z_i^A \neq z_j^B; \\ |DX - |x_i^A - x_j^B - \alpha(\frac{x_i^A + x_j^B}{2})||_+ \times \\ \times |DY - |y_i^A - y_j^B - \beta(\frac{x_i^A + x_j^B}{2})||_+, \text{otherwise.} \end{cases} \tag{3.21}$$

Here $\alpha = \alpha(x)$ and $\beta = \beta(x)$ are variable shifts along iris angular and radial directions and $T = \{t = t(x), x = 0, 1, \ldots, W-1\}$ is the space of *admissible* shifts that are tabular functions defined for $x = 0, 1, \ldots, W-1$, where W is the width of rectangular iris image. The variable shifts potentially can repair minor and moderate segmentation inaccuracies and give substantial increase of genuine pairs similarity scores. On the other hand hyper freedom in shifts may lead to overscoring of impostors pairs similarities. The main content of our "admissibility" notion consists of continuity in the sense that the neighbor shifts $t(x)$ and $t(x+1)$ do not differ substantially. After some experimentations we came to the following notion:

The tabular shifts function $t = t(x), x = 0, 1, \ldots, W-1$, *is admissible iff*

- $t(x) \equiv const_k$ *on any interval* $x = k * W/16, k * W/16 + 1, \ldots, (k+1) * W/16 - 1, \quad k = 0, 1, \ldots, 15.$
- $|const_k - const_{k-1}| \leq 1 \ \forall k = 1, 2, \ldots, 15.$

In other words, rectangular iris image is divided into 16 equal slices, shifts $t = t(x)$ have fixed values on any slice, and the shifts on any two neighbor slices may not differ more than in one pixel. Note that due angular 2π periodicity it would be natural to assume that the first ($k = 0$) and the last ($k = 15$) slices are neighbors and require $|c_0 - c_{15}| \leq 1$ condition. However, such a condition would add some difficulties for constructive algorithm of optimization of similarity scores among space of all admissible shifts. Though angular periodicity condition was refused, we enumerated the slices starting from the middle of upper lid and excluded $0, 1, 2, 7, 8, 13, 14$, and 15 slices from the sum of similarity scores since these slices could be superimposed by upper or lower lid with biggest probability. Under such slices enumeration procedure the problem of refusal of periodicity condition is softened.

Additionally, we will require that $\max\limits_x |\alpha(x)| \leq KX$ and $\max\limits_x |\beta(x)| \leq KY$ in the case of angular and radial shifts respectively. Under these assumptions similarity score of the k-th slice is defined by the equation

$$score_k(A, B; \alpha_k, \beta_k) = \sum_{i,j: k*W/16 \leq \frac{x_i^A + x_j^B}{2} < (k+1)*W/16} s(f_i^A, f_j^B; \alpha_k, \beta_k) \quad (3.22)$$

and can be quickly computed using the scheme described above. However, the final similarity scores

$$\begin{aligned} \widetilde{score}^*(A, B) &= \max_{\{\alpha_k, \beta_k\}_{k=0}^{15} \in T^{16}} \sum_{k=0}^{15} score_k(A, B; \alpha_k, \beta_k), \\ T^{16} &= \left\{ \begin{array}{l} \{\alpha_k, \beta_k\}_{k=0}^{15} : \\ |\alpha_k - \alpha_{k-1}| \leq 1, k = 1, \ldots, 15, \\ |\beta_k - \beta_{k-1}| \leq 1, k = 1, \ldots, 15, \\ |\alpha_k| \leq KX, |\beta_k| \leq KY, k = 0, 1, \ldots, 15 \end{array} \right\}, \end{aligned} \quad (3.23)$$

require an efficient algorithm. The last sections scores sum maximization problem can be solved by dynamic programming: let us set initial values of $(2KX + 1) \times (2KY + 1)$ warped sum to zero, we know optimal warped sums till the $(k-1)$-th order, then the k-th order warped sum is calculated

by the rule

$$\begin{aligned} \widetilde{sum}_k(\alpha,\beta) &= score_k(\alpha,\beta)+ \\ &\max_{i,j=-1,0,1:|\alpha+i|\leq KX,|\beta+j|\leq KY} \widetilde{sum}_{k-1}(\alpha+i,\beta+j), \end{aligned} \tag{3.24}$$

and maximum of $\widetilde{sum}_{15}$ gives the final warped or elastic similarity scores, i.e.

$$\widetilde{score}^*(A,B) = \max_{|\alpha|\leq KX,|\beta|\leq KY} \widetilde{sum}_{15}(\alpha,\beta). \tag{3.25}$$

3.3 Experiments and Discussion

We used the following publicly available iris databases in all experiments: Chinese Academy of Sciences Casia 1.0, Casia 2.0 (device1), Casia 3.0 (interval) [17] and NIST "Iris Challenge Evaluation", experiment 1, (ICE-1) [85]. Their properties are summarized in table 3.3. Fully automatic iris segmentation was performed for all iris images in all databases. Firstly, average size of iris templates formed from local features is analysed. Further, verification performance of the proposed method is evaluated by generating Detection Error Tradeoff (DET) curves for the algorithm.

Table 3.3: Iris databases used in all experiments.

Database	Images	Different eyes	Impressions per eye
Casia 1.0	756	108	7
Casia 2.0 (device 1)	1200	60	20
Casia 3.0 (interval)	2655	396	1-26
ICE-1	1425	128	1-31

3.3.1 Size of Template

Since different irises has distinct number of local sharp variations the iris template size for the proposed features is variable. The average template size can be controlled by the parameter K, which defines the upper limit of *congeneric* local features in one sector (see Fig. 3.4(c)). We used only half of 16 sectors in the following experiment.

Each sector of unwarped rectangular iris image (of width $W = 256$ and height $H = 32$) has 16×32 pixels. Congeneric features of one sector are written to the template in the following order: $\{N, x'_1, y_1, x'_2, y_2, \ldots, x'_N, y_N\}$, where N is number of congeneric features, (x'_n, y_n), $n = 1, \ldots, N$ - position of local features in sector. We assumed that $K \leq 7$. Thus, $0 \leq n \leq 8$, $0 \leq x'_i \leq 16$, and $0 \leq y_i \leq 32$ can be represented in 3, 4, and 5 bits respectively. Therefore, the maximal number bytes in one template is at most $(bitsForN + (bitsForX'_i + bitsForY_i) * K) * (\#$ of used sectors $) * (\#$ of different congeneric classes$))/8) = (3 + (4 + 5) * 7) * 8 * 16)/8 = 1056$. Table 3.4 shows real average iris templates size for three different databases corresponding to maximal allowable local extrema number K in one sector. The average size of template in Casia 1.0, Casia 2.0, and ICE-1 is 592 bytes that constitutes 56% of possible maximal template size for maximum value $K = 7$. For $K = 5$ average template size will be similar to the Daugman's iriscode [30] size. We observed clusters of local extrema at the same (x'_i, y_i) positions but different scales. There is no doubt that these clusters can be used to compress iris templates even more, however analysis of such compressibility was not done.

Table 3.4: Dependency of average size of template (in bytes) for different iris databases on maximum allowable local extrema amount in one section (number K).

K	Casia 1.0	Casia 2.0	ICE-1
1	177	190	177
2	290	328	288
3	383	452	379
5	508	619	498
7	556	679	540

3.3.2 Results of Local Extrema Only Method

For Casia 2.0 (device 1) iris database ${}_{20}C_2 \times 60 = 11400$ genuine and ${}_{60}C_2 \times 20^2 = 708000$ impostor similarity scores were evaluated. It is recognized that Casia 2.0 (device 1) iris database contains iris images of lower quality than Casia 3.0 (interval) [73]. Best result that can be found in the literature $EER = 0.58\%$ is achieved by [73]. We obtained $EER = 0.13\%$, $ZeroFAR = 1.63\%$, and $ZeroFRR = 2.52\%$ with proposed local extrema of averaged Taylor expansion coefficients as features (filled diamond marked curve in Fig. 3.6).

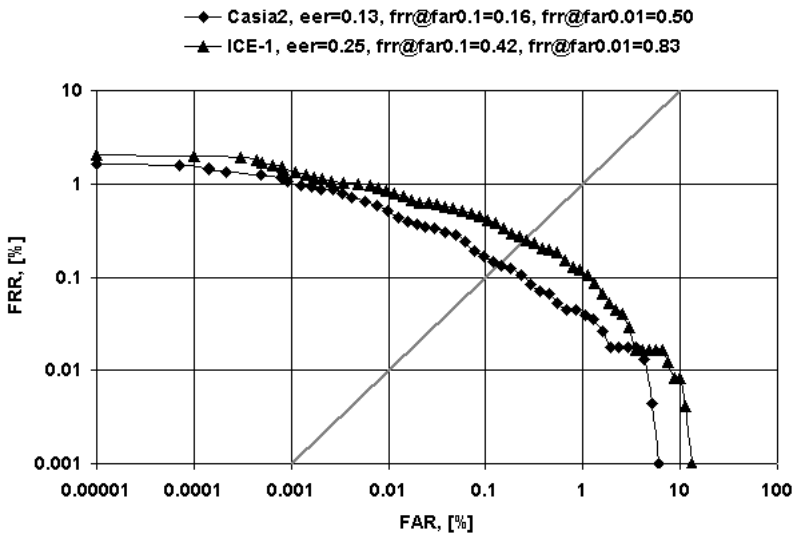


Figure 3.6: DET curves of the proposed algorithm for Casia 2.0 (device 1) and ICE-1 iris databases.

For Casia 3.0 (interval) iris database 9023 genuine and 3514162 impostor similarity scores were evaluated. Impostor and genuine pairs are completely separable that's why direct estimation of Equal Error Rate (EER) is not possible. We approximated distributions of genuine and impostor empirical similarity scores by Gaussian distributions and concluded that proposed iris verification algorithm has $EER = 4.7 \times 10^{-2}\%$. Complete separation of impostor and genuine pairs was also received on Casia 1.0 database.

For the NIST "Iris Challenge Evaluation", experiment 1, (ICE-1) iris database local extrema based verification algorithm relatively produced the

worst results. The image database contains very difficult and corrupted examples, sometimes with big eyelid occlusion, off-angle or with the iris partly outside the image frame. Poor focus of a part of the iris images especially degrade the quality of the verification that uses only information coming from local extrema. Out of focus iris images lost all subtle details of the iris texture that results in a significant decay of number of Taylor decomposition local extrema at tiny scales. For ICE-1 iris database 12214 genuine and 1002386 impostor similarity scores were evaluated. We obtained $EER = 0.25\%$, $FRR@FAR0.1 = 0.42\%$, $FRR@FAR0.01 = 0.83\%$ (filled triangle marked curve in Fig. 3.6).

3.3.3 Results of Fusion with Phase Based Method

Previously described iris recognition method was based on local extrema points and is more similar to minutiae based fingerprint recognition [52] than to conventional iris recognition algorithms. The main difference is a variable template size which depends on number of local extrema - significant local descriptors. This requires unique and completely different from previous approaches matching algorithm 3.2.4.1. On the other hand, searching for the local extrema could be replaced by a quantization step. Similar to [30], converting each negative filter response value to a bit with value 0 and converting greater than or equal to zero filter response values to bits with values 1 will result in a fixed length vector of bits, which could be compared to other vectors by calculating a Hamming distance. We implemented such extraction algorithm and applied previously described warped similarity method for comparison of such iris templates. Although, this method gave slightly worse accuracy, we fused it in the similarity scores level with previously described method by a simple sum rule and got impressive iris recognition accuracy, as can be seen in Fig. 3.7 and Fig. 3.8, respectively for Casia 2.0 (device 1) and ICE-1 databases.

A detailed comparison to other published methods for Casia 2.0 (device 1) and ICE-1 databases is available in table 3.5.

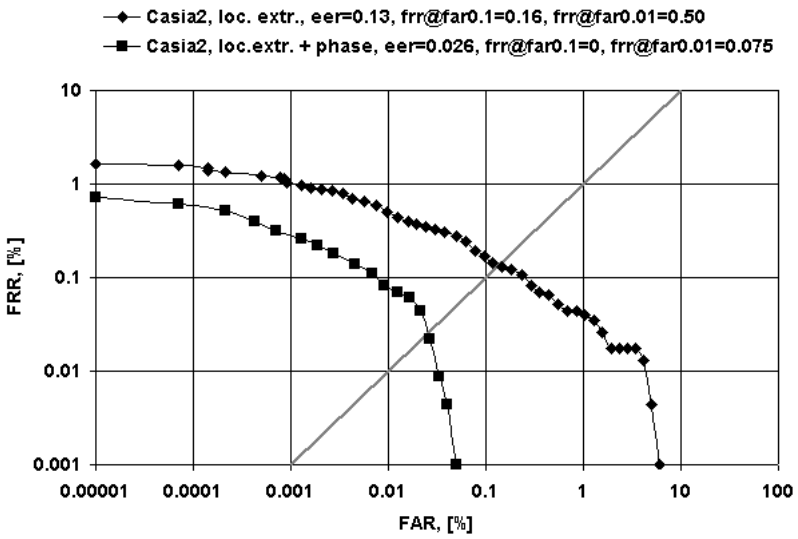


Figure 3.7: An improvement of the fused algorithm over local extrema only method for Casia 2.0 (device 1) database.

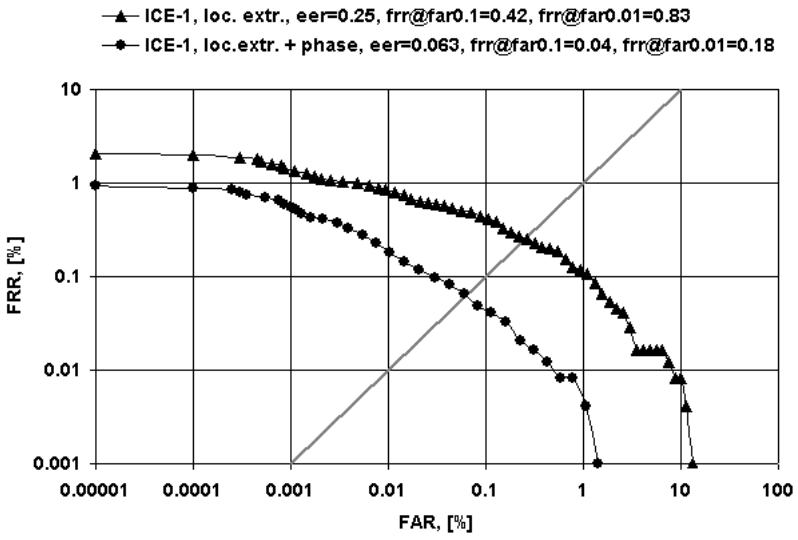


Figure 3.8: An improvement of the fused algorithm over local extrema only method for ICE-1 database.

Table 3.5: Comparison to other published results on the tested iris databases. Three numbers in each table cell presents: EER, FRR@FAR=0.1%, and FRR@FAR=0.01% in percents.

Method	Casia 2.0 (device 1)			ICE-1		
Pan [80]	3.98	–	–	–	–	–
Tohoku [73] [74]	0.58	–	–	2.64	6.94	–
Daugman (1) [26]	–	–	–	0.21	0.30	0.55
Daugman (2) [26]	–	–	–	0.11	0.12	0.30
best ICE2005 [85]	–	–	–	–	$\sim$0.10	$\sim$0.25
proposed method	0.13	0.16	0.50	0.25	0.42	0.83
proposed fused method	0.03	0.00	0.08	0.06	0.04	0.18

3.4 Summary

Each step of the proposed iris segmentation algorithm is not unique or exclusive. However, stated combination of those steps gave almost perfect performance on all tested databases. Moreover, suggested implementation worked in real-time, which together with a near zero failure to enroll rate is a crucial requirement for a real-world application.

The idea to use local extrema of wavelet [69] or other expansions [6] for iris recognition is not new either. However, after doing a good job in extracting discriminant features, they return to comparing fixed length vectors in rigid manner as suggested by the predecessor of all iris recognition algorithms [30]. Unlike the 1D local extrema, which where used in [69], we use 2D local extremum points of the multiscale expansion. Such approach reduces the size or iris template and eliminates correlation of the local extremum points along radial direction. We chose the multiscale Taylor expansion since the first two coefficients of the expansion have a transparent interpretation and are associated with computationally efficient filters. Furthermore, a novel and efficient matching algorithm is proposed.

For the proposed method iris template size varies from 452 to 751 bytes

(for $K = 7$). Blurred, out of focus or occasionally with simple texture iris images have reduced number of local extremum points. Therefore, ROC curves can have increased ZeroFAR that leads to increased slope near the ZeroFAR point in the opposite to flat ROC curves that are observed for phase based iris matching technique [27]. One of the possible solutions could be additional attributes of the local extremum points that would decrease their matching probability.

The proposed method shows comparable results with other state-of-the-art algorithms. When fused with fixed length binarized phase vector based algorithm, it achieves better accuracy than any other published algorithm on tested databases. Despite of the noticeable improvement of iris verification quality due to similarity scores fusion, we should mention that our binarized phase matrix increases template size by 2048 bytes and significantly decreases the matching speed. The decrease in matching speed is conditioned by the warped similarity calculation. It is likely, however, that more accurate iris segmentation (as in [26]) can eliminate necessity to search for optimal shifts along radial direction. On the other hand, matching speed can be significantly increased by indexing [39] or any other optimization techniques which were left out of this thesis scope.

Chapter 4

Fusion

4.1 Introduction

One could wonder why do we need to fuse these two biometric modalities into one person identification system. Iris recognition is recognized as the most accurate biometrics. Most of us have two irises with texture which is unique for each eye and their combination already improves recognition accuracy by an order of magnitude. So, why do we need even bother about the face biometric modality?

Firstly, it is cheap. To capture iris at a distance, we need to locate persons eyes, which is an easy step after the human face is detected. Or, in the opposite, if human eyes are already detected for iris recognition step by some independent method, that's everything we need to geometrically normalize face image, which can be further used for facial recognition. Secondly, capturing iris at a distance of several meters or more generates iris images which are not as good as capturing iris images with conventional iris scanners (those which work at the close range of tens of centimeters). This could result in lower iris recognition accuracy. In contrary, face can be captured at very high resolution at the same time because it is needed to get enough resolution for iris images. Thirdly, it is additional information that is available even if the person closes his eyes or accidentally blinks at the same moment when person identification system captures the images for recognition. Of course, the system could always choose to recapture iris images, but while the system is performing bunch of calculations to accept

this decision, the person could already move away or turn his face away from the cameras.

For testing fusion of our face and iris recognition algorithms we needed a database. Until recently, there were no such public databases available. Some small databases were collected in the labs internally [111], but usually, chimeric databases were composed from existing public databases of separate biometric modalities [112], [18], [94]. In December of 2007, NIST announced Multiple Biometric Grand Challenge (MBGC) with the goal "to investigate, test and improve performance of face and iris recognition technology on both still and video imagery through a series of challenge problems and evaluation" [83]. In December of 2008, together with preliminary results on MBGC datasets, the second and much larger set of images and videos was released under MBGC version 2. MBGC was divided into three challenges:

1. Still Face Challenge - to improve face recognition performance on images taken under conditions that are representative of a broad range of real-world scenarios. One set of experiments uses high resolution still face images taken in both controlled and uncontrolled illumination environments. A second set of experiments is designed to meet the International Civil Aviation Organization's (ICAO) standard for electronic passports [45] by addressing the size of the face and compression of the face image.

2. Video Challenge - to encourage the development of face recognition from video taken in hallways, atria, and outdoors under unconstrained illumination, pose and camera angle. One set of experiments examined the effect of high definition (HD) and standard definition (SD) video on performance. It was further divided into matching video sequences of walking to walking, activity to activity (usually different face pose), and walking to activity.

3. Portal Challenge - to improve iris recognition on the move, iris recog-

nition from video, and fusion of face and iris recognition algorithms.

Portal Challenge problem has the most suitable set of experiments for the algorithms proposed in chapters 2 and 3 of this thesis. We will discuss this problem in details in the following section.

4.2 Portal Challenge Problem

The design of this problem was conditioned by a number of new iris recognition systems which boasted being able to perform human iris identification at a distance and on the move. The following data was collected for this challenge:

1. Still iris images were collected with LG2200 sensor with intentionally degraded quality as described in [85]. See Fig. 4.1.

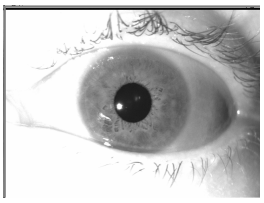


Figure 4.1: Example of still iris image (640 × 480 pixels).

2. Iris video sequences were collected with the same LG2200 sensor, digitized and transcoded to MPEG-4 format with high bit rate allowance, thus yielding near-lossless encoding. See Fig. 4.2.

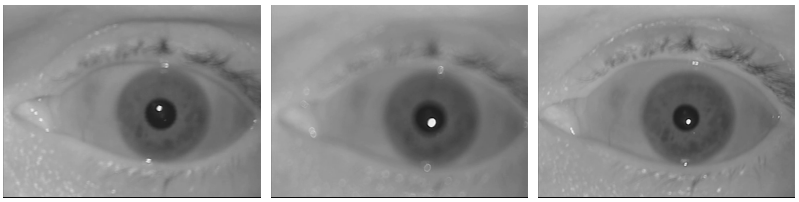
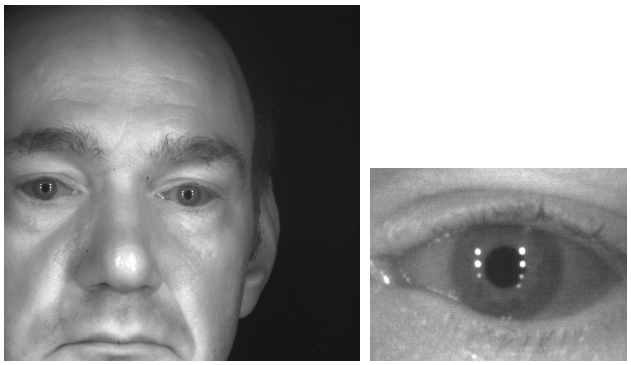


Figure 4.2: Several frames from example of iris video (640 × 480 pixels).

3. Very high resolution near-infrared (NIR) video of faces were acquired from a Sarnoff Iris on the Move (IoM) system [97]. The IoM system

was designed to capture iris imagery as a person walks through the portal. See Fig. 4.3.



(a) (b)

Figure 4.3: One frame from example of NIR video (2048 × 2048 pixels) (a) and cropped left iris from the same frame (320 × 240 pixels) (b). The diameter of the iris is approximately 140 pixels.

4. Still face images were captured with a digital camera at a medium resolution in controlled illumination and additionally in uncontrolled illumination. See Fig. 4.4.

Figure 4.4: Example of still face image (1504 × 1000 pixels).

5. High definition (HD) video sequences were captured with HD video camera which was bore-sighted beside the NIR video cameras in the IoM chassis. This camera captured persons face in visible light while the person walked through the portal. See Fig. 4.5.

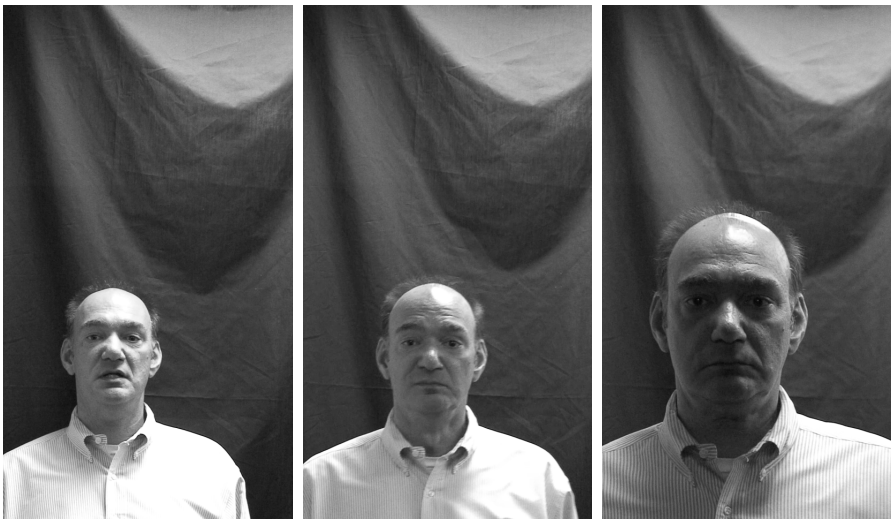


Figure 4.5: Several frames from example of HD video sequence (1080×1920 pixels).

Twelve experiments in five different categories were defined in this challenge (see Table 4.1). With algorithms described in previous chapters of this thesis, plus iris localization in very high resolution NIR videos of faces, we were able to take part in eight of them. We were not able to take part in the following experiments:

- *2C* and *2U*, because they required completely different approach of face recognition with partial matching of faces, which was not implemented in proposed face recognition algorithm.

- *4L* and *4R*, because they required another algorithm for selecting still iris images which are suitable for enrollment and identification from video sequences and such algorithm was out of the scope of this thesis.

We performed experiments *1C* and *1U* with the proposed face recognition algorithm using the MBGC evaluation framework. The results are presented in Fig. 4.6. These experiments required to compare high resolution still face image with HD video sequence in one comparison. The proposed method is able to compare two face images only, thus, we compared still face image with each frame of video sequence and chose the maximum similarity score of all comparisons as final similarity score between still image

Table 4.1: Experiment structure for MBGC Version 2 Portal Challenge.

Exp.	Gallery	Probe
1C	Still Face Controlled	HD Portal Video
1U	Still Face Uncontrolled	HD Portal Video
2C	Still Face Controlled	NIR Portal Video
2U	Still Face Uncontrolled	NIR Portal Video
3L	Left Iris Still	NIR Portal Video
3R	Right Iris Still	NIR Portal Video
4L	Left Iris Video	NIR Portal Video
4R	Right Iris Video	NIR Portal Video
5CL	Still Face Controlled + Left Iris Still	HD+NIR Portal Video
5CR	Still Face Controlled + Right Iris Still	HD+NIR Portal Video
5UL	Still Face Uncontrolled + Left Iris Still	HD+NIR Portal Video
5UR	Still Face Uncontrolled + Right Iris Still	HD+NIR Portal Video

and HD video sequence.

4.3 Localization of Irises in Very High Resolution NIR Video

A typical frame of very high resolution NIR video is illustrated in Fig. 4.3. An array of reflections from infra-red lamps are visible on each eye. Number of reflections vary from zero to eight. They are situated in two columns. When these reflections are not clearly visible, the frame is useless in iris recognition sense because iris texture is not lit well. We exploited this IoM system characteristic to locate eyes in very high resolution NIR videos by the following algorithm:

1. Smooth the image I with two different scales σ_1 and σ_2 of image-bluring filters and calculate a normalized face image by applying the following equation for each pixel:

$$I(x,y) = \frac{I_{\sigma_1}(x,y) - I_{\sigma_2}(x,y)}{I_{\sigma_1}(x,y) + I_{\sigma_2}(x,y)}. \tag{4.1}$$

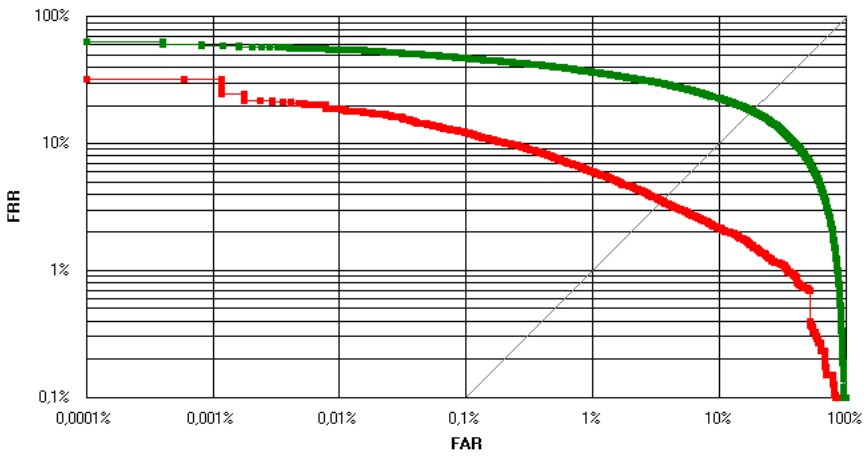


Figure 4.6: DET curves for the proposed face recognition algorithm on MBGC experiments *1C* (red curve) and *1U* (green curve).

2. Find all small areas of white pixels after thresholding the normalized image with a fixed global threshold. These spots correspond to reflections we are looking for.

3. Group the spots which are closer than some chosen distance vertically and horizontally.

4. At least three spots must be situated vertically and at least one of them must have a neighboring spot on the left or on the right. Such combination identifies reflections on the eye.

5. Declare the average position of all spots in the group as a possible location of eye.

6. If two such positions of eyes are detected (additionally, they must be situated horizontally), declare them as right and left eyes.

7. If no two positions are available, lower the required number of spots in the group vertically (defined in step 4) by one and repeat the eye selection procedure.

8. If a pair of eyes is found, use them for recognition with identified eye positions, i.e. use left eye for left iris recognition experiment and right eye for right iris recognition experiment.

9. If only one eye (group of spots) was detected, use them both for left and right iris recognition experiments.

The full data set contained 12236 video frames (images). Only 7300 of them contained at least on eye that could be identified by checking all the frames manually. Similar, only 5145 of frames contained both eyes. We were able to detect both eyes in 2878 of frames where both eyes were visible with the previously described fully automatic eye detection algorithm. Additionally, one eye with unidentified position (left or right eye) was detected in 3300 of other frames. This results in correct at least one eye localization in almost 85% of images. Although, it is not a high detection rate we should take into account that all visible eyes were selected manually, but part of them cannot be used for iris recognition because of the poor focus, high noise and little texture details. These images usually didn't have the reflections and were missed by our algorithm. But even if we were able to locate them, we wouldn't be able to use them for recognition.

Similar to experiments *1C* and *1U*, experiments *3L* and *3R* required to compare still iris images to high resolution NIR face video sequences in one comparison, but the proposed algorithm was able to compare two iris images only. Thus, we compared still image with each frame of video sequence where iris of unknown or appropriate eye position (left or right) was located. The results are presented in Fig. 4.7.

4.4 Multiple Biometric Fusion

Fusion of different biometric modalities for person identification can be performed at several different levels - sensor level, features level, scores level, and decision level. We will focus on scores level fusion in this thesis because it was recommended and implemented in MBGC evaluation framework.

Each comparison of a pair of biometric presentations of previously described recognition algorithms generates a similarity score. In identification task one probe (sometimes referred as "query") presentation is compared to

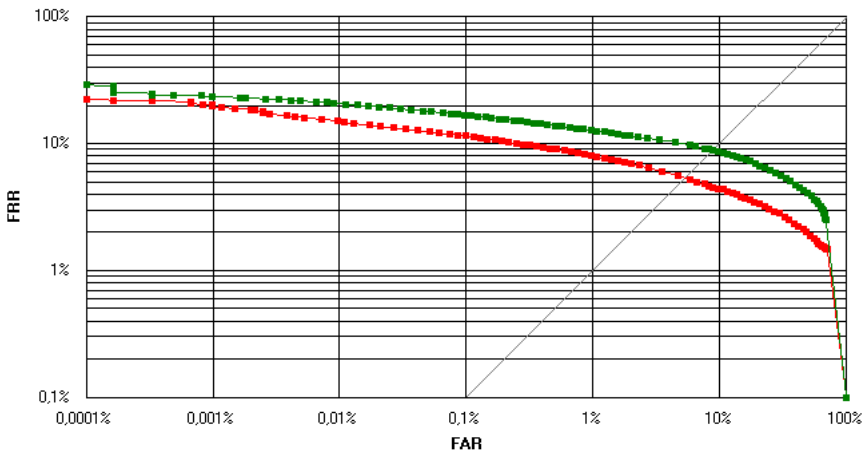


Figure 4.7: DET curves for the proposed iris recognition algorithm on MBGC experiments *3L* (red curve) and *3R* (green curve).

a set of gallery (sometimes referred as "target") presentations, which generates a set of similarity scores. Number of genuine scores (scores generated by comparing biometric presentations of the same person) is less than or equal to one, all other scores belong to impostors (are generated by comparing biometric presentation of different persons). This information can be used to normalize scores and, thus, improve recognition accuracy. Furthermore, normalized scores from different biometric modalities can be easily fused because they fall in the same range and have similar distribution parameters.

It was shown that z-score normalization is optimal when there is no information on genuine distribution and impostors distribution can be approximated by Gaussian function [47]. Luckily, this is the case for both of the proposed recognition algorithms - face and iris. z-score normalization is already applied in experiments *1C*, *1U*, *3L*, and *3R*, reported in Fig. 4.6 and Fig. 4.7 because it was default scores normalization in MBGC evaluation framework.

Normalized scores can be fused by simple rules like sum, product, min, or max. We found that sum and product rules produce almost the same results, while min and max rules perform significantly worse. After summing normalized score we applied z-score normalization once more. This

additionally slightly improved DET curves. Fusion results of the proposed face and iris recognition algorithms for the respective MBGC experiments are presented in Fig. 4.8.

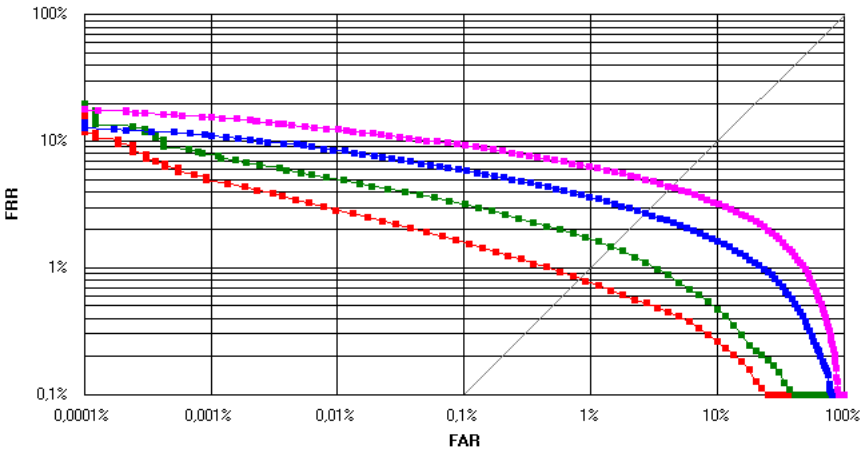


Figure 4.8: DET curves for the fusion of the proposed face and iris recognition algorithms on MBGC experiments *5CL* (red curve), *5CR* (green curve), *5UL* (blue curve), and *5UR* (magenta curve).

4.5 Summary

We showed how fusion of face an iris biometric modalities can improve person identification accuracy. For this purpose we used a recent Multiple Biometrics Grand Challenge protocol and data sets. DET curves for the face recognition experiments in controlled illumination are comparable with those of previously performed experiments on FERET dataset (see Fig. 4.6), but uncontrolled illumination experiment, when face images were captured outside, is much more challenging. Although, iris recognition is often used as a synonym for highly accurate biometric modality, MBGC Portal Challenge experiment showed that it heavily depends on iris image acquisition device. Iris recognition experiments showed terrible recognition accuracy (see Fig. 4.7) when compared to traditional twice higher resolution iris images. MBGC team did not create an experiment were both of the persons irises could be used for recognition, which is a must for real-world

applications if such low resolution iris images are being used. In spite of the fact, that iris images are of such poor quality, fusion with face recognition (in either controlled or uncontrolled illumination) give more than twice lower false rejection rates at the same false acceptance rates. That's a significant improvement over any single modality.

The obtained results are not unexpected, they are often cited in literature (i.e. fusion of several modalities usually improves overall recognition accuracy). However, fusion of the proposed face and iris recognition algorithms is attractive because the properties of their impostors distributions allows to perform it almost at no cost by applying the simplest z-score normalization and sum rule. That was confirmed by experiments performed on the most recent dataset containing multiple biometric modalities. Furthermore, a modern computer can perform person identification based on two biometric modalities - face and iris - in real time.

Chapter 5

Summary and Conclusions

In this thesis, we presented two state-of-the-art algorithms for two different biometrics modalities which can be measured at a distance - face and iris. Face detection was left out of the scope of this thesis because the method described in [108] is already suitable for our goal - it is fast and accurate enough for a real-time face localization module in the proposed system. After successful face detection three modules can start their work in parallel on a nowaday CPU: one for facial recognition and the other two for each of the detected eyes. We described the possible implementations of such modules in details. All of them can work in real-time and with really high accuracy suitable for real-world applications.

In face recognition module, described in chapter 2, Gabor features were chosen to encode facial features into a compact representation which can be later used for rapid comparison of human faces. Instead of using face model as in [116] for aligning Gabor features to facial features, we chose to normalize face image geometrically and use a regular grid to place Gabor features over the face. The benefit of such approach is threefold. Firstly, there is no need to manually locate facial features for training step, which is eliminated at all. Secondly, optimization of number and locations of used features becomes much simpler. And finally, regular grid structure can be exploited to calculate Gabor features very efficiently in spatial domain. In section 2.2.1 we showed how structure of regular grid and generalized separability of Gabor filter can be used to speed up the calculation of Gabor

features in spatial domain. It is almost as fast as calculating them in frequency domain without the loss of precision, if Gabor features are required for each pixel in the image. But if regular grid is not so dense, for example, every 10^{th} pixel in the image, the proposed method is more than 10 times faster than calculating Gabor features in frequency domain.

For geometric face image normalization a simple yet effective method of aligning faces by two reference points (centers of eyes in our case) was used. It was followed by a series of experiments to find out what resolution face images are the best for the chosen Gabor features. Additionally, the tightness of cropping face in face images was analyzed. It was shown that tight rectangular crop, which includes full forehead and usually full chin but no ears, is superior to square crop, which includes facial features from eyebrows to lower lip.

For photometric normalization, which is usually performed to eliminate lighting variations, three methods were analyzed: self-quotient image normalization, local mean and variance normalization and local histogram equalization. It was shown that self-quotient image normalization does not improve facial recognition accuracy if Gabor features are chosen as facial features descriptors. The other two complement Gabor features with additional invariance to lighting variations and perform similar both in terms of speed and accuracy.

Finally, we showed how number of Gabor filters used in Gabor features can improve face recognition accuracy. We showed that extending Gabor features by adding more frequencies and orientations is more valuable than extending number of Gabor features by using denser regular grid. Furthermore, a new similarity metric for calculating similarity between two Gabor features was proposed. A combined similarity metric gave the best face identification accuracy on the most popular face recognition accuracy evaluation FERET database.

In iris recognition module, described in chapter 3, an efficient iris segmentation method was proposed. Its main strength is the speed, which is

comparable to face detection speed. Additionally, it can describe iris inner and outer boundaries by a more complex than simple circle contours, which can significantly improve iris recognition accuracy. For geometric iris image normalization a classic rubber sheet method was implemented.

For iris texture describing features, we proposed to use local extremum points of multiscale Taylor expansion. The proposed multiscale two-dimensional features were superior to one-dimensional extrema suggested by other authors. Number of local extremum points was unacceptably large, thus, we proposed an efficient way to reduce it by choosing the most significant texture descriptors. Furthermore, a completely unique method of calculating similarity between two iris features sets was described. Two compact representations of iris texture are compared by eliminating in-plane rotation between the irises as in classic Hamming distance based methods. But additionally, iris deformations along the radius are eliminated, which not only deals with extreme pupil contractions and dilations but also copes with little inaccuracies in segmentation step. The proposed iris recognition method is comparable with other state-of-the-art methods.

Finally, we implemented a classic phase-based version of our features, which can be compared by a Hamming distance. We showed how to extend it by using warped similarity, similar to our proposed similarity metric. A simple sum of normalized similarity scores gave superior results on recent iris recognition databases by outrunning all the other published iris recognition algorithms to the best of our knowledge.

Fusion module, described in chapter 4, is constructed based on fusion at score level after z-score normalization. Literature study showed that it is optimal when impostor scores distribution is Gaussian-like and no additional information about genuine scores distribution is available. Experiments with MBGC Version 2 Portal Challenge confirmed that two biometric modalities complement each other even if their recognition accuracies differ by several times.

All declared objectives were achieved and all tasks were completed while

preparing this thesis. All results were carefully documented and presented in this text. Highly accurate and fast, multi-modal person identification by face and iris biometric modalities method was created during this research. Only several such systems (usually as proprietary prototypes) in the world were known to the author by the time this thesis was prepared. This is the main contribution of this thesis to the biometrics community.

Appendix A

Evaluation Results using Original FERET Protocol

The authors of FERET database carefully labeled more than 14 thousand images. However, when face recognition algorithms matured they automatically detected several mistakes in the labeling. In our previous experiments we fixed these mistakes because there is no use in optimizing face recognition algorithm to differentiate between several photos of the same person as belonging to different persons. The following reported mistakes were corrected:

1. *00011fa010_930831.pgm*, *00011fb010_930831.pgm* - correct subject ID is 00012.

2. *00240fa010_940128.pgm*, *00240fb010_940128.pgm*, *00240hl010_940128.pgm*, *00240hr010_940128.pgm*, *00240pl010_940128.pgm*, *00240pr010_940128.pgm* - correct subject ID is 00256.

3. *00274fa010_940422.pgm*, *00274fb010_940422.pgm*, *00274hl010_940422.pgm*, *00274hr010_940422.pgm*, *00274pl010_940422.pgm*, *00274pr010_940422.pgm* - correct subject ID is 00184.

4. *00277fa010_940422.pgm*, *00277fb010_940422.pgm*, *00277hl010_940422.pgm*, *00277hr010_940422.pgm*,

00277pl010_940422.pgm, *00277pr010_940422.pgm* - correct subject ID is 00188.

5. *00382fa010_940422.pgm*, *00382fb010_940422.pgm*, *00382hl010_940422.pgm*, *00382hr010_940422.pgm*, *00382pl010_940422.pgm*, *00382pr010_940422.pgm* - correct subject ID is 00210.

6. *00463fa010_940422.pgm*, *00463fb010_940422.pgm*, *00463hl010_940422.pgm*, *00463hr010_940422.pgm*, *00463pl010_940422.pgm*, *00463pr010_940422.pgm* - correct subject ID is 00368.

7. *00464fa010_940422.pgm*, *00464fb010_940422.pgm*, *00464hl010_940422.pgm*, *00464hr010_940422.pgm*, *00464pl010_940422.pgm*, *00464pr010_940422.pgm* - correct subject ID is 00367.

8. *00649fa010_941031.pgm*, *00649fb010_941031.pgm*, *00649hl010_941031.pgm*, *00649hr010_941031.pgm*, *00649pl010_941031.pgm*, *00649pr010_941031.pgm*, *00649ql010_941031.pgm*, *00649qr010_941031.pgm*, *00649ra010_941031.pgm*, *00649rb010_941031.pgm*, *00649rc010_941031.pgm* - correct subject ID is 00615.

9. *00696fa010_941121.pgm*, *00696fb010_941121.pgm*, *00696hl010_941121.pgm*, *00696hr010_941121.pgm*, *00696pl010_941121.pgm*, *00696pr010_941121.pgm*, *00696ql010_941121.pgm*, *00696qr010_941121.pgm*, *00696rc010_941121.pgm*, *00696rd010_941121.pgm*, *00696re010_941121.pgm* - correct subject ID is 00621.

10. *00832fa010_940307.pgm*, *00832fa010a_940307.pgm*, *00832fb010_940307.pgm*, *00832fb010a_940307.pgm*, *00832hl010_940307.pgm*, *00832hl010a_940307.pgm*,

00832hr010_940307.pgm, *00832hr010a_940307.pgm*, *00832pl010_940307.pgm*, *00832pl010a_940307.pgm*, *00832pr010_940307.pgm*, *00832pr010a_940307.pgm* - correct subject ID is 00283.

11. *00881fa010_960530.pgm*, *00881fb010_960530.pgm*, *00881hl010_960530.pgm*, *00881hr010_960530.pgm*, *00881pl010_960530.pgm*, *00881pr010_960530.pgm*, *00881ql010_960530.pgm*, *00881qr010_960530.pgm*, *00881ra010_960530.pgm*, *00881rb010_960530.pgm*, *00881rc010_960530.pgm* - correct subject ID is 00807.

12. *00975fa010_960627.pgm*, *00975fb010_960627.pgm*, *00975hl010_960627.pgm*, *00975hr010_960627.pgm*, *00975pl010_960627.pgm*, *00975pr010_960627.pgm*, *00975ql010_960627.pgm*, *00975qr010_960627.pgm*, *00975ra010_960627.pgm*, *00975rb010_960627.pgm*, *00975rc010_960627.pgm* - correct subject ID is 00531.

If these mistakes are not corrected, original grayscale FERET protocol will be used. Then, final verification accuracy achieved by using best parameters sets in two most common FERET experiments - Fb and Dup1 - is reported in Fig. A.1. It is obvious that verification performance cannot be measured correctly because of the strange "tail" of the DET curve starting at false acceptance rates lower than 0.01% (compare to Fig. 2.26). Identification accuracy is reported in Table A.1.

Furthermore, several errors in eyes labeling, which is provided by NIST as a ground-truth metadata, were found for the following images:

00108fa010_960530.pgm, *00877fb010_960530.pgm*, *00885fb010_960530.pgm*, *00904fb010_960530.pgm*.

However, these mistake were corrected in CSU Face Identification Evaluation System [11]. We used ground-truth eyes coordinates from the mentioned system in all our experiments as many other researcher did.

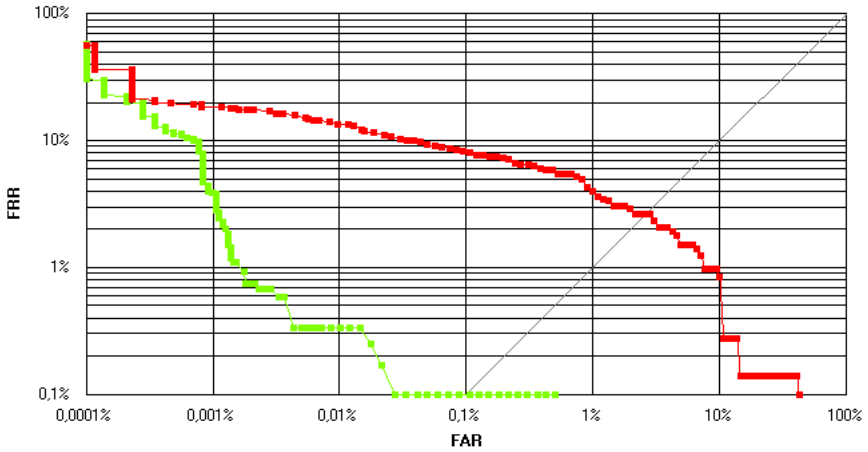


Figure A.1: Final verification accuracy measured in Dup1 experiment using $Proposed_{GN\_PHN\_REV}$ parameters set (red curve) and Fb experiment using $Proposed_{LHE+F\_ONLY}$ parameters set (green curve). Original grayscale FERET protocol is used.

Table A.1: Identification accuracy as $Rank1$ (in percents) for different face recognition methods in different FERET experiments. Original grayscale FERET protocol is used.

Method	Experiment			
	Fb	Fc	Dup1	Dup2
$Proposed_{BASELINE}$	90.79	93.81	66.76	61.97
$Proposed_{FINAL}$	99.83	**100.00**	92.80	88.89
$Proposed_{LHE+F\_ONLY}$	**99.92**	**100.00**	91.41	87.18
$Proposed_{GN\_PHN\_REV}$	99.58	**100.00**	**94.46**	**93.59**
$Proposed_{LOW\_DIM}$	99.58	**100.00**	92.24	88.89

Bibliography

[1] Y. Adini, Y. Moses, and S. Ullman. Face recognition: the problem of compensating for changes in illumination direction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:721–732, 1997.

[2] T. Ahonen, A. Hadid, and M. Pietikäinen. Face recognition with local binary patterns. pages 469–481. 2004.

[3] Aoptix. Insight. http://www.aoptix.com/biometrics.html, August 2009.

[4] V. Areekul, U. Watchareeruetai, K. Suppasriwasuseth, and S. Tantaratana. Separable gabor filter realization for fast fingerprint enhancement. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 3, pages III–253–6, March 2006.

[5] E. M. Arvacheh. A study of segmentation and normalization for iris recognition systems.

[6] K. Bae, S. Noh, and J. Kim. Iris feature extraction using independent component analysis. pages 1059–1060. 2003.

[7] E. Bailly-Bailliére, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariéthoz, J. Matas, K. Messer, V. Popovici, F. Porée, B. Ruiz, and J.-P. Thiran. The banca database and evaluation protocol. page 1057. 2003.

[8] F. Bashir, D. Usher, P. Casaverde, and M. Friedman. Video surveillance for biometrics: Long-range multi-biometric system. In *Advanced*

Video and Signal Based Surveillance, 2008. AVSS '08. IEEE Fifth International Conference on, pages 175–182, 2008.

[9] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7):711–720, August 2002.

[10] A. Bernardino and J. Santos-Victor. A real-time gabor primal sketch for visual attention. pages 335–342. 2005.

[11] J. R. Beveridge, D. Bolme, B. A. Draper, and M. Teixeira. The csu face identification evaluation system. *Machine Vision and Applications*, 16(2):128–138, February 2005.

[12] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli. On the use of sift features for face authentication. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on*, page 35, 2006.

[13] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 187–194, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[14] W. W. Boles and B. Boashash. A human identification technique using images of the iris and wavelet transform. *Signal Processing, IEEE Transactions on*, 46(4):1185–1188, 1998.

[15] A. C. Bovik, M. Clark, and W. S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(1):55–73, 1990.

[16] R. Cappelli, A. Erol, D. Maio, and D. Maltoni. Synthetic fingerprint-image generation. volume 3, pages 471–474 vol.3, August 2002.

[17] Casia. Casia-irisv3. http://www.cbsr.ia.ac.cn/IrisDatabase, 2005.

[18] C.-H. Chen and C. Te Chu. Fusion of face and iris features for multimodal biometrics. pages 571–580. 2005.

[19] D. A. Clausi and M. E. Jernigan. Designing gabor filters for optimal texture separability. *Pattern Recognition*, 33(11):1835–1849, November 2000.

[20] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. pages 484+. 1998.

[21] F. C. Crow. Summed-area tables for texture mapping. In *SIGGRAPH '84: Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, pages 207–212, New York, NY, USA, 1984. ACM.

[22] S. Curtis, J. Lim, and A. Oppenheim. Signal reconstruction from one bit of fourier transform phase. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84.*, volume 9, pages 487–490, 1984.

[23] J. Daugman. Statistical richness of visual phase information: Update on recognizing persons by iris patterns. *International Journal of Computer Vision*, 45(1):25–38, October 2001.

[24] J. Daugman. Demodulation by complex-valued wavelets for stochastic pattern recognition. *International Journal of Wavelets, Multiresolution and Information Processing*, 1:1–17, 2003.

[25] J. Daugman. Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. *Proceedings of the IEEE*, 94(11):1927–1935, 2006.

[26] J. Daugman. New methods in iris recognition. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(5):1167–1175, 2007.

[27] J. Daugman. Response to nistir-7440 and frvt/ice2006 reports, 2007.

[28] J. Daugman and C. Downing. Epigenetic randomness, complexity and singularity of human iris patterns. *Proceedings of the Royal Society (London): B. Biological Sciences*, 268:1737–1740, 2001.

[29] J. G. Daugman. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision research*, 20(10):847–856, 1980.

[30] J. G. Daugman. High confidence visual recognition of persons by a test of statistical independence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(11):1148–1161, 1993.

[31] F. De la Torre, R. Gross, S. Baker, and B. V. K. Vijaya Kumar. Representational oriented component analysis (roca) for face recognition with one sample image per training class. volume 2, pages 266–273 vol. 2, July 2005.

[32] H. K. Ekenel and R. Stiefelhagen. Local appearance based face recognition using discrete cosine transform. In *13th European Signal Processing Conference (EUSIPCO 2005*, 2005.

[33] L. Flom and A. Safir. Iris recognitions system. http://www.google.com/patents?id=4I4tAAAAEBAJ, February 1985.

[34] M. Frigo, Steven, and G. Johnson. The design and implementation of fftw3. In *Proceedings of the IEEE*, volume 93, pages 216–231, 2005.

[35] W. Gao, S. Shan, X. Chai, and X. Fu. Virtual face image generation for illumination and pose insensitive face recognition. In *ICME '03: Proceedings of the 2003 International Conference on Multimedia and Expo - Volume 3 (ICME '03)*, pages 149–152, Washington, DC, USA, 2003. IEEE Computer Society.

[36] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under

variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:643–660, 2001.

[37] R. Gross, I. Matthews, and S. Baker. Fisher light-fields for face recognition across pose and illumination. pages 481–489. 2002.

[38] G. Guo, M. Jones, and P. Beardsley. A system for automatic iris capturing. Technical report, Mitsubishi Electric Research Laboratories, 2005.

[39] F. Hao, J. Daugman, and P. Zielinski. A fast search algorithm for a large fuzzy database. *Information Forensics and Security, IEEE Transactions on*, 3(2):203–212, May 2008.

[40] X. He, S. Yan, Y. Hu, and P. Niyogi. Face recognition using laplacianfaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(3):328–340, 2005.

[41] G. Heusch, Y. Rodriguez, and S. Marcel. Local binary patterns as an image preprocessing for face authentication. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 9–14, 2006.

[42] J. Huang, P. C. Yuen, W. S. Chen, and J. H. Lai. Component-based lda method for face recognition with one training sample. In *AMFG '03: Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, pages 120+, Washington, DC, USA, 2003. IEEE Computer Society.

[43] J. Ilonen, J. K. Kamarainen, and H. K. Kälviäinen. Efficient computation of gabor features. Technical report, Lappeenranta University of Technology, Department of Information Technology, 2005.

[44] Intel. Open computer vision library. http://sourceforge.net/projects/opencvlibrary/, August 2009.

[45] International Civil Aviation Organization. Doc 9303, machine readable travel documents. Technical report, 2009.

[46] A. Jain. Object detection using gabor filters. *Pattern Recognition*, 30(2):295–309, February 1997.

[47] A. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodal biometric systemsâŸĘ. *Pattern Recognition*, 38(12):2270–2285, December 2005.

[48] A. K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recogn.*, 24(12):1167–1186, 1991.

[49] O. Jesorsky, K. Kirchberg, and R. Frischholz. Robust face detection using the hausdorff distance. pages 90–95. 2001.

[50] S. G. Johnson and M. Frigo. A modified split-radix fft with fewer arithmetic operations. *Signal Processing, IEEE Transactions on*, 55(1):111–119, December 2006.

[51] Kalovision. Fif 3.0. http://kalovision.com/products.html, August 2009.

[52] A. Kisel, A. Kochetkov, and J. Kranauskas. Fingerprint minutiae matching without global alignment using local structures. *Informatica*, 19(1):31–44, 2008.

[53] J. Kranauskas. The attributes of faces recognition. Master's thesis, Vilnius University, The Faculty of Mathematics and Informatics, 2005.

[54] B. V. K. V. Kumar, M. Savvides, and C. Xie. Correlation pattern recognition for face recognition. *Proceedings of the IEEE*, 94(11):1963–1976, January 2007.

[55] V. Kyrki. Simple gabor feature space for invariant object recognition. *Pattern Recognition Letters*, 25(3):311–318, February 2004.

[56] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wurtz, and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *Computers, IEEE Transactions on*, 42(3):300–311, August 2002.

[57] A. Lanitis, C. J. Taylor, and T. F. Cootes. Automatic tracking, coding and reconstruction of human faces, using flexible appearance models. *Electronics Letters*, 30(19):1587–1588, 1994.

[58] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back. Face recognition: a convolutional neural-network approach. *Neural Networks, IEEE Transactions on*, 8(1):98–113, August 2002.

[59] H. S. Le and H. Li. Recognizing frontal face images using hidden markov models with one training image per person. volume 1, pages 318–321 Vol.1, September 2004.

[60] K. Lee, S. Lim, K. Lee, O. Byeon, and T. Kim. Efficient iris recognition through improvement of feature vector and classifier. *ETRI Journal*, 23:61–70, 2001.

[61] J. Li, T. Wang, and Y. Zhang. Face recognition using feature of integral gabor-haar transformation. volume 4, pages IV – 505–IV – 508, November 2007.

[62] X. Li. Modeling intra-class variation for nonideal iris recognition. pages 419–427. 2005.

[63] R. Lienhart, L. Liang, and A. Kuranov. A detector tree of boosted classifiers for real-time object detection and tracking. In *Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on*, volume 2, pages II–277–80 vol.2, 2003.

[64] R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. In *Image Processing. 2002. Proceedings. 2002 International Conference on*, volume 1, pages I–900–I–903 vol.1, 2002.

[65] B. F. Logan. Information in the zero crossings of bandpass signals. *AT T Technical Journal*, 56:487–510, April 1977.

[66] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos. Face recognition using kernel direct discriminant analysis algorithms. *Neural Networks, IEEE Transactions on*, 14(1):117–126, February 2003.

[67] Y. Lui and J. Beveridge. Grassmann registration manifolds for face recognition. pages 44–57. 2008.

[68] Y. M. Lui, D. S. Bolme, B. A. Draper, J. R. Beveridge, G. H. Givens, and P. J. Phillips. A meta-analysis of face recognition covariates. In *Proceedings of IEEE Conference on Biometrics: Theory, Applications and Systems*, September 2009.

[69] L. Ma, T. Tan, Y. Wang, and D. Zhang. Efficient iris recognition by characterizing key local variations. *IEEE Trans. on Image Processing*, 13:739–750, 2004.

[70] S. Mallat. Zero-crossings of a wavelet transform. *Information Theory, IEEE Transactions on*, 37(4):1019–1033, 1991.

[71] J. R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. J. Loiacono, S. Mangru, M. Tinker, T. M. Zappia, and W. Y. Zhao. Iris on the move: Acquisition of images for iris recognition in less constrained environments. *Proceedings of the IEEE*, 94(11):1936–1947, 2006.

[72] K. Messer, J. Matas, J. Kittler, and K. Jonsson. Xm2vtsdb: The extended m2vts database. In *In Second International Conference on Audio and Video-based Biometric Person Authentication*, pages 72–77, 1999.

[73] K. Miyazawa and T. Aoki. A phase-based iris recognition algorithm. In *ICB 2006, LNCS 3832, Springer-Verlag*, volume 3832, pages 356–365, 2005.

[74] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi, and H. Nakajima. An efficient iris recognition algorithm using phase-based image matching. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 2, pages II–49–52, 2005.

[75] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7):696–710, August 2002.

[76] E. Mojaev and A. Zell. Real-time scale invariant object and face tracking using gabor wavelet templates.

[77] D. M. Monro, S. Rakshit, and D. Zhang. Dct-based iris recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(4):586–595, 2007.

[78] P. Niyogi, F. Girosi, and T. Poggio. Incorporating prior information in machine learning by creating virtual examples. *Proceedings of the IEEE*, 86(11):2196–2209, August 2002.

[79] A. J. O'Toole, P. J. Phillips, F. Jiang, J. Ayyad, N. Penard, and H. Abdi. Face recognition algorithms surpass humans matching faces over changes in illumination. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(9):1642–1646, 2007.

[80] L. Pan, M. Xie, and Z. Ma. Iris localization based on multi-resolution analysis. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4, 2008.

[81] C.-H. Park, J.-J. Lee, M. Smith, and K.-H. Park. Iris-based personal authentication using a normalized directional energy feature. page 1058. 2003.

[82] V. Perlibakas. Face recognition using principal component analysis and log-gabor filters. May 2006.

[83] P. Phillips, P. Flynn, J. Beveridge, W. Scruggs, A. O'Toole, D. Bolme, K. Bowyer, B. Draper, G. Givens, Y. Lui, H. Sahibzada, J. Scallan, and S. Weimer. Overview of the multiple biometrics grand challenge. pages 705–714. 2009.

[84] P. J. Phillips. Support vector machines applied to face recognition. In *Proceedings of the 1998 conference on Advances in neural information processing systems II*, pages 803–809, Cambridge, MA, USA, 1999. MIT Press.

[85] P. J. Phillips. Frgc and ice workshop. http://iris.nist.gov/ICE/ICE_2005_Results_30March2006.pdf, 2006.

[86] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 947–954 vol. 1, 2005.

[87] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, and W. Worek. Preliminary face recognition grand challenge results. In *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, pages 15–24, 2006.

[88] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The feret evaluation methodology for face-recognition algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(10):1090–1104, 2000.

[89] P. J. Phillips, W. T. Scruggs, A. J. O'toole, P. J. Flynn, W. Kevin, C. L. Schott, and M. Sharpe. Frvt 2006 and ice 2006 large-scale results, March 2007.

[90] P. J. Phillips and Y. Vardi. Efficient illumination normalization of facial images. *Pattern Recogn. Lett.*, 17(8):921–927, 1996.

[91] H. P. M. C. Proenca. *Toward Non-Cooperative Biometric Iris Recognition.* PhD thesis, University of Beira Interior, Department of Computer Science, October 2006.

[92] N. Ranganathan, R. Mehrotra, and K. R. Namuduri. An architecture to implement multiresolution. pages 1157–1160 vol.2, August 2002.

[93] Retica. Eagle-eyes. http://www.retica.com/products-services-eagle-eyes.htm, August 2009.

[94] S. Ribarić, I. Fratrić, and K. Kiš. A novel biometric personal verification system based on the combination of palmprints and faces. *Informatica*, 19(1):81–100, 2008.

[95] A. Ross and S. Shah. Segmenting non-ideal irises using geodesic active contours. In *Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research at the*, pages 1–6, 2006.

[96] C. Sanchez-Avila, R. Sanchez-Reillo, and D. de Martin-Roche. Iris-based biometric recognition using dyadic wavelet transform. *Aerospace and Electronic Systems Magazine, IEEE*, 17(10):3–6, 2002.

[97] Sarnoff. Iris on the move. http://www.sarnoff.com/products/iris-on-the-move, August 2009.

[98] S. Shan, W. Gao, B. Cao, and D. Zhao. Illumination normalization for robust face recognition against varying lighting conditions. In *AMFG '03: Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, pages 157+, Washington, DC, USA, 2003. IEEE Computer Society.

[99] S. Shan, W. Zhang, Y. Su, X. Chen, and W. Gao. Ensemble of piecewise fda based on spatial histograms of local (gabor) binary patterns

for face recognition. In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, pages 606–609, Washington, DC, USA, 2006. IEEE Computer Society.

[100] A. Shashua and T. Riklin-Raviv. The quotient image: class-based rerendering and recognition with varying illuminations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(2):129–139, August 2002.

[101] J. Short, J. Kittler, and K. Messer. A comparison of photometric normalisation algorithms for face verification. pages 254–259, June 2004.

[102] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America. A, Optics and image science*, 4(3):519–524, March 1987.

[103] Y. Su, S. Shan, X. Chen, and W. Gao. Patch-based gabor fisher classifier for face recognition. volume 2, pages 528–531, September 2006.

[104] X. Tan, S. Chen, Z. H. Zhou, and F. Zhang. Face recognition from a single image per person: A survey. *Pattern Recogn.*, 39(9):1725–1745, 2006.

[105] X. Tan and B. Triggs. Fusing gabor and lbp feature sets for kernel-based face recognition.

[106] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on*, pages 586–591, 1991.

[107] T. Vetter and T. Poggio. Linear object classes and image synthesis from a single example image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:733–742, 1997.

[108] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 1:511–I–518 vol.1, 2001.

[109] J. Walter, B. Arnrich, and C. Scheering. Learning fine positioning of a robot manipulator based on gabor wavelets. volume 5, pages 137–142 vol.5, August 2002.

[110] H. Wang, S. Z. Li, and Y. Wang. Face recognition under varying lighting conditions using self quotient image. pages 819–824, June 2004.

[111] Y. Wang, T. Tan, and A. Jain. Combining face and iris biometrics for identity verification. page 1060. 2003.

[112] Y. Wang, T. Tan, and A. K. Jain. Combining face and iris biometrics for identity verification, 2003.

[113] Z. Wang and Z. Miao. Scale invariant face recognition using probabilistic similarity measure. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4, 2008.

[114] G. Westphal, C. von der Malsburg, and R. Würtz. Feature-driven emergence of model graphs for object recognition and categorization. pages 155–199. 2008.

[115] R. P. Wildes, J. C. Asmuth, G. L. Green, S. C. Hsu, R. J. Kolczynski, J. R. Matey, and S. E. Mcbride. A machine-vision system for iris recognition. *Machine Vision and Applications*, 9(1):1–8, January 1996.

[116] L. Wiskott, J. M. Fellous, N. Kruger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. In *Image Processing, 1997. Proceedings., International Conference on*, volume 1, pages 129–132 vol.1, 1997.

[117] X. Xie, W.-S. Zheng, J. Lai, and P. C. Yuen. Face illumination normalization on large and small scale features. pages 1–8, August 2008.

[118] R. Yavne. An economical method for calculating the discrete fourier transform. In *AFIPS '68 (Fall, part I): Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, pages 115–125, New York, NY, USA, 1968. ACM.

[119] S. Yoon, K. Bae, K. Park, and J. Kim. Pan-tilt-zoom based iris image capturing system for unconstrained user environments at a distance. pages 653–662. 2007.

[120] I. T. Young, L. J. Van Vliet, and M. Van Ginkel. Recursive gabor filtering. volume 3, pages 338–341 vol.3, August 2002.

[121] L. Yu, K. Wang, C. Wang, and D. Zhang. Multiscale wavelet texture based iris verification. In J. P. Li, J. Zhao, M. V. Wickerhauser, Y. Y. Tang, J. Daugman, and L. Peng, editors, *WAA*, pages 200–205. World Scientific, 2003.

[122] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang. Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition. volume 1, pages 786–791 Vol. 1, December 2005.

[123] W. Zhao, R. Chellappa, and A. Krishnaswamy. Discriminant analysis of principal components for face recognition. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, pages 336–341, August 2002.

[124] W. Zhao, R. Chellappa, and P. J. Phillips. Subspace linear discriminant analysis for face recognition, 1999.

[125] Y. Zhu, T. Tan, and Y. Wang. Biometric personal identification based on iris patterns. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 2, pages 801–804 vol.2, 2000.

[126] J. Zou, Q. Ji, and G. Nagy. A comparative study of local matching approach for face recognition. *Image Processing, IEEE Transactions on*, 16(10):2617–2628, 2007.