

Volume 5 ▪ Issue 1 ▪ March 2011

Editor-in-Chief
Professor João Manuel R. S. Tavares

INTERNATIONAL JOURNAL OF

BIOMETRICS AND BIOINFORMATICS (IJBB)

ISSN : 1985-2347

Publication Frequency: 6 Issues / Year

CSC PUBLISHERS
<http://www.cscjournals.org>

INTERNATIONAL JOURNAL OF BIOMETRICS AND BIOINFORMATICS (IJBB)

VOLUME 5, ISSUE 1, 2011

**EDITED BY
DR. NABEEL TAHIR**

ISSN (Online): 1985-2347

International Journal of Biometrics and Bioinformatics (IJBB) is published both in traditional paper form and in Internet. This journal is published at the website <http://www.cscjournals.org>, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJBB Journal is a part of CSC Publishers

Computer Science Journals

<http://www.cscjournals.org>

INTERNATIONAL JOURNAL OF BIOMETRICS AND BIOINFORMATICS (IJBB)

Book: Volume 5, Issue 1, March 2011

Publishing Date: 04-04-2011

ISSN (Online): 1985-2347

This work is subjected to copyright. All rights are reserved whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication of parts thereof is permitted only under the provision of the copyright law 1965, in its current version, and permission of use must always be obtained from CSC Publishers.

IJBB Journal is a part of CSC Publishers

<http://www.cscjournals.org>

© IJBB Journal

Published in Malaysia

Typesetting: Camera-ready by author, data conversion by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2011

EDITORIAL PREFACE

This is the first issue of volume five of International Journal of Biometric and Bioinformatics (IJBB). The Journal is published bi-monthly, with papers being peer reviewed to high international standards. The International Journal of Biometric and Bioinformatics is not limited to a specific aspect of Biology but it is devoted to the publication of high quality papers on all division of Bio in general. IJBB intends to disseminate knowledge in the various disciplines of the Biometric field from theoretical, practical and analytical research to physical implications and theoretical or quantitative discussion intended for academic and industrial progress. In order to position IJBB as one of the good journal on Bio-sciences, a group of highly valuable scholars are serving on the editorial board. The International Editorial Board ensures that significant developments in Biometrics from around the world are reflected in the Journal. Some important topics covers by journal are Bio-grid, biomedical image processing (fusion), Computational structural biology, Molecular sequence analysis, Genetic algorithms etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 5, 2011, IJBB appears in more focused issues. Besides normal publications, IJBB intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

The coverage of the journal includes all new theoretical and experimental findings in the fields of Biometrics which enhance the knowledge of scientist, industrials, researchers and all those persons who are coupled with Bioscience field. IJBB objective is to publish articles that are not only technically proficient but also contains information and ideas of fresh interest for International readership. IJBB aims to handle submissions courteously and promptly. IJBB objectives are to promote and extend the use of all methods in the principal disciplines of Bioscience.

IJBB editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJBB. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJBB provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members

International Journal of Biometric and Bioinformatics (IJBB)

EDITORIAL BOARD

EDITOR-in-CHIEF (EiC)

Professor João Manuel R. S. Tavares
University of Porto (Portugal)

ASSOCIATE EDITORS (AEiCs)

Assistant Professor. Yongjie Jessica Zhang

Mellon University
United States of America

Professor. Jimmy Thomas Efirid

University of North Carolina
United States of America

Professor. H. Fai Poon

Sigma-Aldrich Inc
United States of America

Professor. Fadiel Ahmed

Tennessee State University
United States of America

Mr. Somnath Tagore (AEiC - Marketing)

Dr. D.Y. Patil University
India

Professor. Yu Xue

Huazhong University of Science and Technology
China

Associate Professor Chang-Tsun Li

University of Warwick
United Kingdom

Professor. Calvin Yu-Chian Chen

China Medical university
Taiwan

EDITORIAL BOARD MEMBERS (EBMs)

Dr. Wichian Sittiprapaporn

Maharakham University
Thailand

Assistant Professor. M. Emre Celebi

Louisiana State University
United States of America

Dr. Ganesan Pugalenti
Genome Institute of Singapore
Singapore

Dr. Vijayaraj Nagarajan
National Institutes of Health
United States of America

Dr. Paola Lecca
University of Trento
Italy

Associate Professor. Renato Natal Jorge
University of Porto
Portugal

Assistant Professor. Daniela Iacoviello
Sapienza University of Rome
Italy

Professor. Christos E. Constantinou
Stanford University School of Medicine
United States of America

Professor. Fiorella SGALLARI
University of Bologna
Italy

Professor. George Perry
University of Texas at San Antonio
United States of America

Assistant Professor. Giuseppe Placidi
Università dell'Aquila
Italy

Assistant Professor. Sae Hwang
University of Illinois
United States of America

Associate Professor Quan Wen
University of Electronic Science and Technology
China

Dr. Paula Moreira
University of Coimbra
Portugal

Dr. Riadh Hammami
Laval University
Canada

IRIS Based Human Recognition System

Mansi Jhamb

*USIT, Guru Gobind Singh Indraprastha
University, Delhi, India*

mansi.jhamb@gmail.com

Vinod Kumar Khara

*Guru Tegh Bahadur Institute of
Technology Guru Gobind Singh
Indraprastha University, Delhi, India*

vinodkhara@gmail.com

Abstract

The paper explores iris recognition for personal identification and verification. In this paper a new iris recognition technique is proposed using (Scale Invariant Feature Transform) SIFT. Image-processing algorithms have been validated on noised real iris image database. The proposed innovative technique is computationally effective as well as reliable in terms of recognition rates.

Keywords: Iris Recognition, Hough Transform, SIFT, Key-Points

1 INTRODUCTION

Today, biometric recognition is a common and reliable way to authenticate the identity of a living person based on physiological or behavioral characteristics. A physiological characteristic is relatively stable physical characteristics, such as fingerprint, iris pattern, facial feature, hand silhouette, etc. This kind of measurement is basically unchanging and unalterable without significant duress. A behavioral characteristic is more a reflection of an individual's psychological makeup as signature, speech pattern, or how one types at a keyboard. The degree of intra-personal variation in a physical characteristic is smaller than a behavioral characteristic. For examples, a signature is influenced by both controllable actions and less psychological factors, and speech pattern is influenced by current emotional state, whereas fingerprint template is independent. Nevertheless all physiology-based biometrics don't offer satisfactory recognition rates (false acceptance and/or false reject rates, respectively referenced as FAR and FRR). The automated personal identity authentication systems based on iris recognition are reputed to be the most reliable among all biometric methods: we consider that the probability of finding two people with identical iris pattern is almost zero [1]. That's why iris recognition technology is becoming an important biometric solution for people identification in access control as networked access to computer application [2]. Compared to fingerprint, iris is protected from the external environment behind the cornea and the eyelid. No subject to deleterious effects of aging, the small-scale radial features of the iris remain stable and fixed from about one year of age throughout life. This paper is divided into 4 main parts. The *Section 1* introduces what is the position of iris technology in personal authentication. In the *Section 2*, we sum up the state of the art in the domain of iris recognition. The more widely known iris recognition system developed by *J.Daugman* [4] is taken as reference for comparison. The *Section 3* presents in details our approach, and discusses the different issues we chose. At last a conclusion is done in *Section 4*, which tasks about the next considerations for the improvement of the proposed solution.

2. LITERATURE SURVEY

The French ophthalmologist *Alphonse Bertillon* seems to be the first to propose the use of iris pattern (color) as a basis for personal identification [3]. In 1981, after reading many scientific reports describing the iris great variation, *Flom* and San Francisco ophthalmologist *Aran Safir* suggested also using the iris as the basis for a biometric. In 1987, they began collaborating with computer scientist *John Daugman* of *Cambridge University* in England to develop iris identification software who published his first promising results in 1992 [4]. Later on a little similar

works have been investigated, such as *R.Wildes'* [5], *W.Boles'* [6] and *R.Sanchez- Reillo's* [7] systems, which differ both in the iris features representation (iris signature) and pattern matching algorithms. *R.Wildes'* solution includes (i) a *Hough* transform for iris localization, (ii) *Laplacian* pyramid(multi-scale decomposition) to represent distinctive spatial characteristics of the human iris, and (iii) modified normalized correlation for matching process. *W.Boles'* prototype operates in building (j) a one dimensional representation of the gray level profiles of the iris followed by obtaining the wavelet transform zero-crossings of the resulting representation, and (jj) original dissimilarity functions that enable pertinent information selection for efficient matching computation. To finish *J.Daugman's* and *R.Sanchez-Reillo's* systems are implemented exploiting (I) *integrodifferential* operators to detect iris inner and outer boundaries, (II) *Gabor* filters to extract unique binary vectors constituting *iriscodexTM*, and (III) a statistical matcher (logical exclusive OR operator) that analyses basically the average *Hamming* distance between two codes (bit to bit test agreement). Because of unified reference database of iris images does not exist, a classic performance comparison of the described systems is not trivial. However in terms of recognition rates (FAR, FRR), the commercial success of the patented *Daugman's* system speak in his favor. Indeed *Daugman's* mathematical algorithms have been contributing to a commercial solution patented by *IriScan Inc.* This biometric identification platform processes iris recognition through (i) a specific optical unit that enables noninvasive acquisition of iris images, and (ii) a data processing unit. Although capturing a well-defined image of the iris while not interacting actively with the device seems to be one the major challenge we encountered for iris recognition system design, our research focus on the second block both in charge of (j) the enrolment process, and (jj) the matching which quantifies the similitude between two biometric templates.

3. PROPOSED APPROACH

Previous work on iris recognition, derived from the information found in the open literature, led us to suggest a few possible improvements. For justification of these new concepts we implemented in *Matlab/C* .The algorithm used is as follows:

- Image Acquisition
- Iris Localization.
- Find the darkest point of image (referred as black hole) in the global image analysis.
- Determine a range of darkness (based on 1) designated as the threshold value (t) for identification of black holes.
- Determine the number of black holes and their coordinates according to the predefined threshold. Calculate the centre of mass of these black holes.
- Construct a $L \times L$ region centred at the estimated centroid.
- Repeat step 3 to improve the estimation of actual centroid of pupil.
- Find key points using SIFT.
- Match the key points of the input image with the key points of images in database.

The algorithm is beautifully explained by following algorithmic flow chart ,figure 1

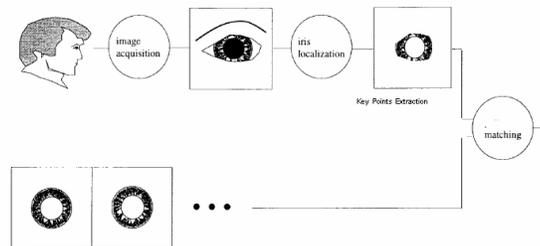


FIGURE 1: Iris Recognition: The Process

3.1 IMAGE ACQUISITION

One of the major challenges of automated iris recognition is to capture a high-quality image of the iris while remaining noninvasive to the human operator. Given that the iris is a relatively small (typically about 1 cm in diameter), dark object and that human operators are very sensitive about

their eyes, this matter requires careful engineering. Several points are of particular concern. First, it is desirable to acquire images of the iris with sufficient resolution and sharpness to support recognition. Second, it is important to have good contrast in the interior iris pattern without resorting to a level of illumination that annoys the operator, i.e., adequate intensity of source (W/cm) constrained by operator comfort with brightness (W/sr-cm). Third, these images must be well framed (i.e., centered) without unduly constraining the operator (i.e., preferably without requiring the operator to employ an eye piece, chin rest, or other contact positioning that would be invasive). Further, as an integral part of this process, artifacts in the acquired images (e.g., due to specular reflections, optical aberrations, etc.) should be eliminated as much as possible. Schematic diagrams of two image-acquisition rigs that have been developed in response to these challenges. The acquired Image is as shown in figure 2 below:

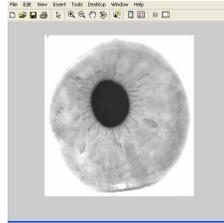


FIGURE 2: Acquired Image

3.2 IRIS LOCALIZATION

Without placing undue constraints on the human operator, image acquisition of the iris cannot be expected to yield an image containing only the iris. Rather, image acquisition will capture the iris as part of a larger image that also contains data derived from the immediately surrounding eye region. Therefore, prior to performing iris pattern matching, it is important to localize that portion of the acquired image that corresponds to an iris. In particular, it is necessary to localize that portion of the image derived from inside the limbus (the border between the sclera and the iris) and outside the pupil. Further, if the eyelids are occluding part of the iris, then only that portion of the image below the upper eyelid and above the lower eyelid should be included. Typically, the limbic boundary is imaged with high contrast, owing to the sharp change in eye pigmentation that it marks. The upper and lower portions of this boundary, however, can be occluded by the eyelids. The papillary boundary can be far less well defined. The image contrast between a heavily pigmented iris and its pupil can be quite small. Further, while the pupil typically is darker than the iris, the reverse relationship can hold in cases of cataract: the clouded lens leads to a significant amount of backscattered light. Like the pupillary boundary, eyelid contrast can be quite variable depending on the relative pigmentation in the skin and the iris. The eyelid boundary also can be irregular due to the presence of eyelashes. Taken in tandem, these observations suggest that iris localization must be sensitive to a wide range of edge contrasts, robust to irregular borders, and capable of dealing with variable occlusion. The systems differ mostly in the way that they search their parameter spaces to fit the contour models to the image information. To understand how these searches proceed, let $I(x,y)$ represent the image intensity value at location (x,y) and let circular contours (for the limbic and papillary boundaries) be parameterized by center location (x_c,y_c) and radius r . The Daugman system fits the circular contours via gradient ascent on the parameters (x_c,y_c,r) so as to maximize

$$\left| \frac{\partial}{\partial r} G(r) * \oint_{r=x_c, y_c} \frac{I(x,y)}{2\pi r} ds \right|$$

Where $G(r) = (1/\sqrt{2\sigma\pi})\sigma^{-(r-r_0)^2/2\sigma^2}$ is a radial Gaussian with center r_0 and standard deviation σ that smooths the image to select the spatial scale of edges under consideration $*$, symbolizes convolution, ds is an element of circular arc, and division by $2\pi r$ serves to normalize the integral. In order to incorporate directional tuning of the image derivative, the arc of integration ds is restricted to the left and right quadrants (i.e., near vertical edges) when fitting the limbic boundary. This arc is considered over a fuller range when fitting the pupillary boundary; however,

the lower quadrant of the image is still omitted due to the artifact of the specular reflection of the illuminant in that region (see Section II-A). In implementation, the contour fitting procedure is discretized, with finite differences serving for derivatives and summation used to instantiate integrals and convolutions. More generally, fitting contours to images via this type of optimization formulation is a standard machine vision technique, often referred to as active contour modeling. The Wildes *et al.* system performs its contour fitting in two steps. First, the image intensity information is converted into a binary edge-map. Second, the edge points vote to instantiate particular contour parameter values. The edgemap is recovered via gradient-based edge detection [2], [44]. This operation consists of thresholding the magnitude of the image intensity gradient, i.e.,

$$\nabla G(x, y) * I(x, y) \text{ where}$$

$$\nabla \equiv (\partial / \partial x, \partial / \partial y) \text{ while}$$

$$G(x, y) = 1 / 2\pi\sigma^2 e^{-((x-x_0)^2 + (y-y_0)^2) / 2\sigma^2}$$

is a two-dimensional Gaussian with center (x_0, y_0) and σ is standard deviation that smooths the image to select the spatial scale of edges under consideration. In order to incorporate directional tuning, the image intensity derivatives are weighted to favor certain ranges of orientation prior to taking the magnitude. For example, prior to contributing to the fit of the limbic boundary contour, the derivatives are weighted to be selective for vertical edges. The voting procedure is realized via Hough transforms [27], [28] on parametric definitions of the iris boundary contours. In particular, for the circular limbic or pupillary boundaries and a set of recovered edge points $(x_j, y_j) \ j = 1 \dots n$. Hough transform is defined as

$$H(x_c, y_c, r) = \sum_{j=1}^n h(x_j, y_j, x_c, y_c, r)$$

where

$$h(x_j, y_j, x_c, y_c, r) = \begin{cases} 1, & \text{if } g(x_j, y_j, x_c, y_c, r) = 0 \\ 0, & \text{otherwise} \end{cases}$$

with

$$g(x_j, y_j, x_c, y_c, r) = (x_j - x_c)^2 + (y_j - y_c)^2 - r^2.$$

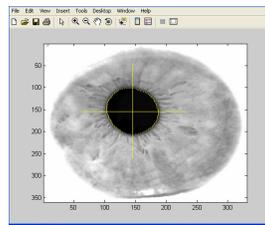


FIGURE 3: Iris and centroid detection

3.3 IRIS MATCHING

Image matching is a fundamental aspect of many problems in computer vision, including object or scene recognition, solving for 3D structure from multiple images, stereo correspondence, and motion tracking. This method describes image features that have many properties that make them suitable for matching differing images of an object or scene. The features are invariant to image scaling and rotation, and partially invariant to change in illumination and 3D camera viewpoint. They are well localized in both the spatial and frequency domains, reducing the probability of disruption by occlusion, clutter, or noise. Large numbers of features can be extracted from typical images with efficient algorithms. In addition, the features are highly distinctive, which allows a single feature to be correctly matched with high probability against a large database of features, providing a basis for object and scene recognition. The cost of

extracting these features is minimized by taking a cascade filtering approach, in which the more expensive operations are applied only at locations that pass an initial test.

Following are the major stages of computation used to generate the set of image features:

- Scale-space extrema detection:
- Key point localization:
- Orientation assignment:
- Key point descriptor:

3.3.1 Detection Of Scale Space Schema

As described in the introduction, we will detect keypoints using a cascade filtering approach that uses efficient algorithms to identify candidate locations that are then examined in further detail. The first stage of keypoint detection is to identify locations and scales that can be repeatedly assigned under differing views of the same object. Detecting locations that are invariant to scale change of the image can be accomplished by searching for stable features across all possible scales, using a continuous function of scale known as scale space (Witkin,1983).It has been shown by Koenderink (1984) and Lindeberg (1994) that under a variety of reasonable assumptions the only possible scale-space kernel is the Gaussian function. Therefore, the scale space of an image is defined as a function, $L(x, y, \sigma)$, that is produced from the convolution of a variable-scale Gaussian, $G(x, y, \sigma)$, with an input image, $I(x, y)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y),$$

Where * is the convolution operation in x and y,

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

and To efficiently detect stable keypoint locations in scale space, we have proposed (Lowe, 1999) using scale-space extrema in the difference-of-Gaussian function convolved with the image, $D(x, y, \sigma)$ which can be computed from the difference of two nearby scales separated by a constant multiplicative factor k:

$$D(x,y,\sigma)=(G(x,y,k\sigma) - G(x,y,\sigma) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma)$$

There are a number of reasons for choosing this function. First, it is a particularly efficient function to compute, as the smoothed images, L , need to be computed in any case for scale space feature description, and D can therefore be computed by simple image subtraction. In addition, the difference-of-Gaussian function provides a close approximation to the scale-normalized Laplacian of Gaussian, $\sigma^2 \nabla^2 G$, as studied by Lindeberg (1994). Lindeberg showed that the normalization of the Laplacian with the factor σ^2 is required for true scale invariance. The relationship between D and $\sigma^2 \nabla^2 G$ can be understood from the heat diffusion equation (parameterized in terms of σ rather than the more usual $t = \sigma^2$): $\partial G / \partial \sigma = \sigma \nabla^2 G$

From this, we see that $\nabla^2 G$ can be computed from the finite difference approximation to $dG/d\sigma$, using the difference of nearby scales at $k\sigma$ and σ :

$$\sigma \nabla^2 G = \partial G / \partial \sigma \approx G(x, y, k\sigma) - G(x, y, \sigma) / K\sigma - \sigma$$

And therefore

$$\sigma \nabla^2 G = \partial G / \partial \sigma \approx G(x, y, k\sigma) \approx (k - 1)\sigma^2 \nabla^2 G .$$

The factor $(k - 1)$ in the equation is a constant over all scales and therefore does not influence extrema location. An important aspect of this approach is that it generates large numbers of features that densely cover the image over the full range of scales and locations. For iris matching and recognition, SIFT features are first extracted from a set of reference images and stored in a database.

3.3.2 Accurate Key-point Localization

Once a keypoint candidate has been found by comparing a pixel to its neighbors, the next step is to perform a detailed fit to the nearby data for location, scale, and ratio of principal curvatures. This information allows points to be rejected that have low contrast (and are therefore sensitive to noise) or are poorly localized along an edge. The initial implementation of this approach (Lowe,

1999) simply located keypoints at the location and scale of the central sample point. However, recently Brown has developed a method (Brown and Lowe, 2002) for fitting a 3D quadratic function to the local sample points to determine the interpolated location of the maximum, and his experiments showed that this provides a substantial improvement to matching and stability. His approach uses the Taylor expansion (up to the quadratic terms) of the scale-space function, $D(x, y, \sigma)$, shifted so that the origin is at the sample point.

$$D(\mathbf{x}) = D + \frac{\partial D}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x}$$

Where D and its derivatives are evaluated at the sample point and $\mathbf{x} = (x, y, \sigma)^T$ is the offset, from this point. The location of the extremes, $\hat{\mathbf{x}}$, is determined by taking the derivative of this function with respect to \mathbf{x} and setting it to zero, giving

$$\hat{\mathbf{x}} = -\frac{\partial^2 D}{\partial \mathbf{x}^2}^{-1} \frac{\partial D}{\partial \mathbf{x}}$$

As suggested by Brown, the Hessian and derivative of D are approximated by using differences of neighboring sample points. The resulting 3×3 linear system can be solved with minimal cost. If the offset $\hat{\mathbf{x}}$ is larger than 0.5 in any dimension, then it means that the extreme lies closer to a different sample point. In this case, the sample point is changed and the interpolation performed instead about that point. The final offset $\hat{\mathbf{x}}$ is added to the location of its sample point to get the interpolated estimate for the location of the extremum. The function value at the extremum, $D(\hat{\mathbf{x}})$, is useful for rejecting unstable extrema with low contrast. For the experiments, all extrema with a value of $|D(\hat{\mathbf{x}})|$ less than 0.03 were discarded (as before, we assume image pixel values in the range $[0, 1]$). The key point selection is shown in figure 4

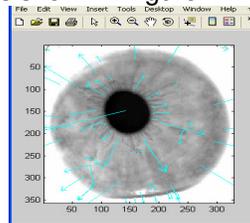


FIGURE 4: Key points detection

By assigning a consistent orientation to each keypoint based on local image properties, the keypoint descriptor can be represented relative to this orientation and therefore achieve invariance to image rotation.

3.3.3 Key-Point Matching

The best candidate match for each key point is found by identifying its nearest neighbor in the database of key points from training images. The nearest neighbor is defined as the keypoint with minimum Euclidean distance for the invariant descriptor. However, many features from an image will not have any correct match in the training database because they arise from background clutter or were not detected in the training images. Therefore, it would be useful to have a way to discard features that do not have any good match to the database. A global threshold on distance to the closest feature does not perform well, as some descriptors are much more discriminative than others. A more effective measure is obtained by comparing the distance of the closest neighbor to that of the second-closest neighbor. If there are multiple training images of the same object, then we define the second-closest neighbor as being the closest neighbor that is known to come from a different object than the first, such as by only using images known to contain different objects. This measure performs well because correct matches need to have the closest neighbor significantly closer than the closest incorrect match to achieve reliable matching. For false matches, there will likely be a number of other false matches within similar distances due to the high dimensionality of the feature space. We can think of the second-closest match as providing an estimate of the density of false matches within this portion of the feature space and at the same time identifying specific instances of feature ambiguity.

3.3.4 CLUSTERING With Hough Transform

Each of our keypoints specifies 4 parameters: 2D location, scale, and orientation, and each matched keypoint in the database has a record of the keypoint's parameters relative to the training image in which it was found. Therefore, we can create a Hough transform entry predicting the model location, orientation, and scale from the match hypothesis. This prediction has large error bounds, as the similarity transform implied by these 4 parameters is only an approximation to the full 6 degree of- freedom pose space for a 3D object and also does not account for any non rigid deformations. Therefore, we use broad bin sizes of 30 degrees for orientation, a factor of 2 for scale, and 0.25 times the maximum projected training image dimension (using the predicted scale) for location. To avoid the problem of boundary effects in bin assignment, each keypoint match votes for the 2 closest bins in each dimension, giving a total of 16 entries for each hypothesis and further broadening the pose range.

4. RESULTS

Figures below shows the results obtained by applying SIFT

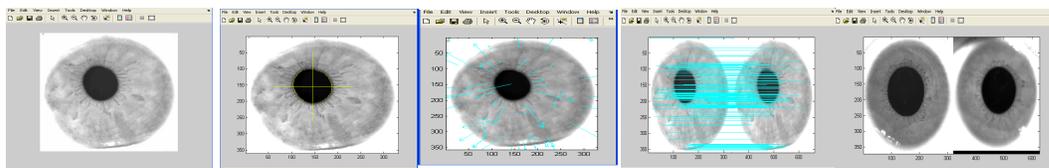


Fig. 5 : Original Image **Fig. 6** centroid detection **Fig. 7 :** Key points detection **Fig.8 :** Image matching **Fig 9 :** Match not found

5. CONCLUSION

Iris recognition system has been developed steadily with the help of MATLAB and some mathematical calculations, however limitations such as blur and dynamically taken images make it impossible to achieve perfect naturalness to combat this, we need to take images in ultraviolet environment. After getting image from the user the system will apply Hough transform detector technique to distinguish between pupillary and iris part of human eye, system applied various inbuilt MATLAB functions and mathematical calculations to encircle outer part of pupil that is inner part of iris and will mark the outer part of iris.

6. REFERENCES

- [1] Y.Belganoui, J-C.Guézel and T.Mahé, « La biométrie, sésame absolu ...», Industries et Techniques, France, n°817, July 2000.
- [2] M.M.Gifford, D.J. McCartney and C.H.Seal, « Networked biometrics systems: requirements based on iris recognition », BT Technol. Journal, Vol. 17, April 1999.
- [3] A.Bertillon, « la couleur de l'iris », Revue scientifique, France, 1885.
- [4] J.Daugman, « High confidence personal identification by rapid video analysis of iris texture », Proc. Of the IEEE, International Carnahan conf. on security technology, 1992.
- [5] R.P.Wildes, J.C. Asmuth, G.L. Green and S.C. Hsu, « A system for automated iris recognition », IEEE paper, 1994.
- [6] W.W.Boles, « a security system based on human iris identification using wavelet transform », First international conference on knowledge-based intelligent electronic systems, Adelaide, Australia. Ed, 21-23 may 1997.

- [7] R.Sanchez-Reillo, C.Sanchez-Avila and J-A.Martin- Pereda, « minimal template size for iris recognition », Proc. BMES/EMBS Conf., IEEE Publication, Atlanta, October 1999.
- [8] J.Daugman, «High confidence personal identification by rapid video analysis of iris texture », IEE Conf. Publication , European convention on security and detection, 16-18 May 1995.
- [9] D.Gabor, «Theory of communication », Journal IEE,Vol.93, no III, 1946.
- [10] J.Daugman, « Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortex filters », Journal of optical society of America, Vol. 2, 1985.
- [11] D.E.Benn, M.S.Nixon and J.N.Carter, « Robust eye extraction using H.T. », AVBPA'99.
- [12] J.P.Havlicek, J.W.Havlicek and A.C.bovik, «the analytic image », IEEE Journal, 1997.
- [13] D.Gabor, « Theory of communication », J. inst. Elect. Eng. London, Vol. 93, n° III, 1946.
- [14] A.C.Bovik, « The handbook of image processing », Ed. Bovik.
- [15] J.P.Havlicek, D.S.Harding, and A.C.Bovik, «Discrete quasi eigenfunction approximation for AM-FM image analysis », Proc. Of the IEEE Int. Conf. on Image Processing, 1996.
- [16] H. Davson, *The Physiology of the Eye*, 2nd ed. Boston, MA: Little, Brown & Co., 1963.
- [17] A. W. Drake, *Fundamentals of Applied Probability Theory*. New York: McGraw-Hill, 1986.
- [18] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
- [19] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals Eugenics*, vol. 7, no. 2, pp. 179–188, 1936.
- [20] L. Flom and A. Safir, "Iris recognition system," U.S. Patent 4 641 349, 1987.
- [21] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.*, vol. 93, pp. 429–459, 1946.
- [22] S. B. Grimes and I. Kershner, *Never Say Never Again*, Warner Brothers, 1983.
- [23] K. Hanna, R. Mandelbaum, L. Wixson, D. Mishra, and V. Paragana, "A system for nonintrusive human iris acquisition," *Machine Vision Applications*, Tokyo, Japan, 1996, pp. 200–203.
- [24] J. P. Holmes, L. J. Wright, and R. L. Maxwell, "A performance evaluation of biometric identification devices," SIL Albuquerque, NM, Tech. Rep. SAND91- 0276, 1991.
- [25] B. K. P. Horn, *Robot Vision*. Cambridge, MA: MIT Press, 1986.
- [26] P. Horowitz and W. Hill, *The Art of Electronics*, 2nd ed. New York: CUP, 1988.
- [27] P. V. C. Hough, "Method and means for recognizing complex patterns," U.S. Patent 3 069 654, 1962.
- [28] J. Illingworth and J. Kittler, "A survey of the Hough transform," *Comput. Vision, Graph. Image Processing*, vol. 44, pp. 87–116, 1988.

- [29] B. Jahne, *Digital Image Processing*, 2nd ed. Berlin: Springer-Verlag, 1993.
- [30] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: PH-1984.
- [31] F. A. Jenkins and H. E. White, *Fundamentals of Optics*. New York: McMillan, 1976.
- [32] R. G. Johnson, "Can iris patterns be used to identify people," Los Alamos National Laboratory, CA, Chemical and Laser Sciences Division, Rep. LA-12331-PR, 1991.
- [33] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," in *Proc. Int. Conf. Computer Vision*, London, England, 1987, pp. 259–268.
- [34] A. L. Kroeber, *Anthropology*. New York: Harcourt Brace Jovanovich, 1948.
- [35] P. C. Kronfeld, "The gross anatomy and embryology of the eye," in *The Eye*, vol. 1, H. Davson, Ed. London: Academic, 1968, pp. 1–66.
- [36] N. A. Macmillan and C. D. Creelman, *Detection Theory: A User's Guide*. Cambridge: Cambridge Univ. Press, 1991.
- [37] A. Malickas, personal communication, 1994.
- [38] I. Mann, *The Development of the Human Eye*. New York: Grune and Stratton, 1950.
- [39] D. Marr, *Vision*. New York: Freeman, 1982.
- [40] B. Miller, "Vital signs of identity," *IEEE Spectrum*, vol. 31, pp. 22–30, Feb. 1994.
- [41] D. Miller, *Ophthalmology*. Boston, MA: Houghton Mifflin, 1979.
- [42] F. W. Newell, *Ophthalmology Principles and Practice*, 7th ed. St. Louis, MO: Mosby, 1991.
- [43] G. Olivier, *Practical Anthropology*. Springfield, IL: Charles C. Thomas, 1969.
- [44] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1978.
- [45] A. Samal and P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey," *Pattern Recognit.*, vol. 25, pp. 65–77, 1992.
- [46] J. E. Siedlarz, "Iris: More detailed than a fingerprint," *IEEE Spectrum*, vol. 31, p. 27, Feb. 1994.
- [47] P. Sinha, "A head mounted display," Bachelor's thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge

Microarray Data Classification Using Support Vector Machine

Seeja.K.R.

*Department of Computer Science
Jamia Hamdard University
New Delhi, India*

seeja@jamiahamdard.ac.in

Shweta

*Department of Computer Science
Jamia Hamdard University
New Delhi, India*

rajput.laksh@yahoo.co.in

Abstract

DNA microarrays allow biologist to measure the expression of thousands of genes simultaneously on a small chip. These microarrays generate huge amount of data and new methods are needed to analyse them. In this paper, a new classification method based on support vector machine is proposed. The proposed method is used to classify gene expression data recorded on DNA microarrays. The proposed method is tested by using benchmark datasets and it is found that the proposed method is faster than neural network and the classification performance is not less than neural network.

Keywords: Support Vector Machines, Microarray, Classification

1. INTRODUCTION

When a normal tissue becomes cancerous, the expression levels of many genes change. By identifying these changes in gene expression, the tissues can be classified as cancerous and normal. Microarray technology is a hybridization technique which allows monitoring the expression of thousands of genes in a single experiment on a small chip. The output of these microarray experiments are the expression levels of different genes and these data are publicly available. These datasets include a large number of gene expression values and need to have a good data mining method to extract knowledge from these microarray gene expression datasets. Support vector machine (SVM) is a supervised computer learning technique used for data classification. It performs classification by constructing an optimal hyper plane which separates the data into two classes.

Many researchers have developed and demonstrated different classification techniques for cancer classification based on micro array gene expression data. Feature selection techniques [1],[2] have been suggested before classification, which finds the top features that discriminate various classes. Kernel based techniques [3],[4] like SVM have already been used for binary disease classification problems. Gene selection[5] and neural networks[6] based classifications were also reported in microarray data analysis.

In this paper SVM is used for the cancer classification based on microarray gene expression data. SVM is trained using different kernels like Poly Kernel, Normalized Poly Kernel and RBF and found that SVM performs better or equal classification than Neural Network.

2. MATERIALS AND METHODS

2.1 DNA Microarray

DNA microarrays can be used to measure changes in expression levels of genes in different biological conditions. The principle behind microarrays is hybridization between two DNA strands, the property of complementary nucleic acid sequences to specifically pair with each other by forming hydrogen bonds between complementary nucleotide base pairs. A high number of complementary base pairs in a nucleotide sequence mean tighter non-covalent bonding between the two strands. After washing off of non-specific bonding sequences, only

strongly paired strands will remain hybridized. So fluorescently labeled target sequences that bind to a probe sequence generate a signal that depends on the strength of the hybridization. Microarrays use relative quantization in which the intensity of a feature is compared to the intensity of the same feature under a different condition.

2.2 Support Vector Machine

Support vector machine[7] is a powerful data mining technique for classifying data. The support vector machine is a training algorithm for learning classification and regression rules from data. SVM was developed from statistical learning theory and was first suggested by Vapnik[8] in the 1960s for data classification. SVM classifies data in large data sets by identifying a linear or non-linear separating surface in the input space of a data set. The separating surface depends only on a subset of the original data known as a set of support vectors. A support vector machine constructs a hyper plane or set of hyper planes in a high dimensional space, which can be used for classification. A good separation is achieved by the hyper plane that has the largest distance to the nearest training data points of any class, called functional margin. If this functional margin is large, then the generalization error of the classifier will be small. SVM models are built around a kernel function [9],[10] that transforms the input data into an n-dimensional space where a hyper plane can be constructed to partition the data.

2.3 Dataset Used

In this paper, the acute leukemia benchmark dataset described by Golub et al [1] is used for classification and it is downloaded from Broad Institute's website[11]. The leukemia data set includes expression profiles of 7,129 human DNA probes spotted on Affymetrix Hu6800 microarrays of 72 patients with either acute myeloid leukemia (AML) or acute lymphocytic leukemia (ALL). Tissue samples were collected at time of diagnosis before treatment, taken either from bone marrow (62 cases), or peripheral blood (10 cases) and reflect both childhood and adult leukemia. The gene expression profiles of the original data set are represented as log10 normalized expression values. This data set was used as a benchmark for various machine learning techniques. The data set is divided into training set containing 38 samples and a validation set containing 34 samples.

2.4. Feature Selection

The proposed SVM based classification method, uses a feature selection algorithm to find the top features, which classifies the data sets effectively. The F(x) score[2] helps to find features that discriminate between the two classes. In this application genes are the features. The feature selection algorithm described below identifies the genes whose expression shows great change in both the classes.

1. Obtain the mean of the expression values for each gene of ALL samples and mean of the expression values for each gene of AML samples.
2. Obtain absolute difference between the mean of ALL samples and the mean of AML samples.
3. Arrange the genes based on absolute difference in decreasing order.
4. Select Top 250 genes.
5. Apply the following formula on selected 250 genes.

$$F(x_i) = (\mu(ALL) - \mu(AML)) / (\sigma(ALL) + \sigma(AML))$$
 where μ is the mean and σ is the standard deviation.
6. Select 200 genes with highest absolute F(x_i) scores as our top features.

2.5 SMO (Sequential Minimal Optimization) Algorithm

The learning task in SVM can be formulated as a convex optimization problem, which can be solved by using Lagrange Multiplier method. Sequential Minimal Optimization (SMO) [12] is a simple algorithm that can quickly solve the SVM QP problem without any extra matrix storage and without using numerical QP optimization. The advantage of SMO is its ability to solve the Lagrange multipliers analytically.

4. RESULTS AND DISCUSSION

We have used the WEKA version 3.6.4[13] software for performing the classification. WEKA contains an implementation of SMO algorithm which supports SVM. Feature selection algorithm is implemented in C#.

First SVM is trained by using the bench mark training set. After training, the classification accuracy is validated using the training set as well as testing set. The training dataset contains 38 training samples and all the samples were classified without error using poly kernel, Normalized poly kernel and RBF kernel during training as shown in Figure 1. On 10 fold cross validation of training dataset all the 38 samples were classified without error using poly kernel, Normalized poly kernel and one AML sample was misclassified using RBF kernel as shown in Figure 2. Then we applied 34 test data samples to the trained SVM, 2 AML samples were misclassified using RBF kernel and 3 AML were misclassified using Poly kernel and Normalized poly kernel. All other samples were classified correctly. Figure 3 shows this result.

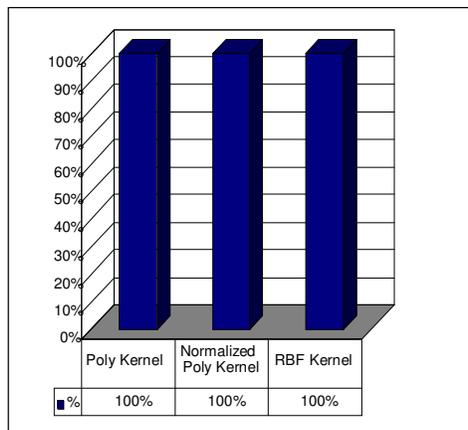


FIGURE 1: Classification accuracy on training set for different kernels

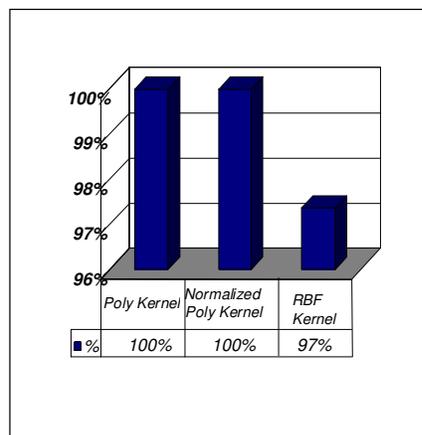


FIGURE 2: Classification accuracy on 10-fold cross validation of training set for different kernels

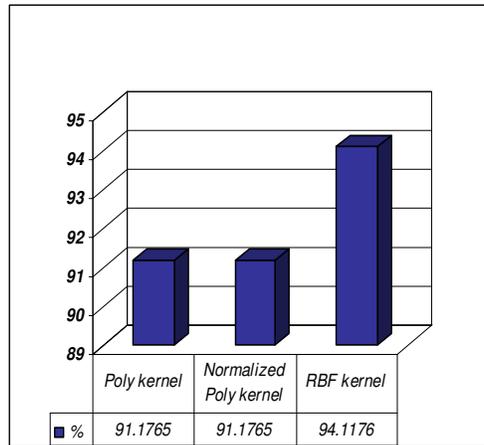


FIGURE 3: Classification accuracy on test dataset for different kernels

In order to evaluate the performance of SVM, we have applied the same dataset to the neural network learning algorithm available in WEKA. We found that both SVM and neural network classifies the data with same accuracy. But SVM is taking less time than neural network. Figure 4 and figure 5 show the comparison on time.

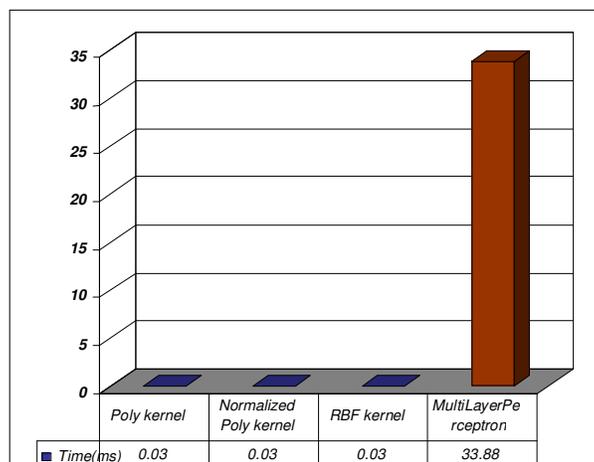


FIGURE 4: SVM Vs NEURAL NETWORK(Training)

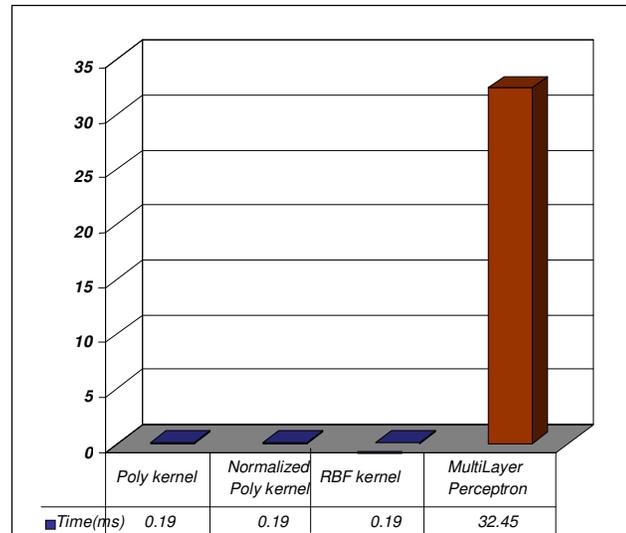


FIGURE 5: SVM Vs NEURAL NETWORK(Testing)

5. CONCLUSION

We have proposed an efficient and powerful method for microarray gene expression data classification and prediction using support vector machine. We applied SVM on ALL/AML dataset. In order to evaluate the performance of SVM, we have applied the same dataset to the neural network learning algorithm available in WEKA. We found that both SVM and neural network classifies the data with same accuracy. But SVM is taking less learning time than neural network.

6. REFERENCES

1. Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gassenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., Lander, E.S., "Molecular classification of cancer: class discovery and class prediction by gene expression monitoring", *Science*, 286(15):531–537, 1999.
2. Terrence S. Furey, Nello Cristianini, Nigel Duffy, David W. Bednarski, Michèl Schummer, and David Haussler, "Support vector machine classification and validation of cancer tissue samples using microarray expression data ", *Bioinformatics*6(10): 906-914 , 2000
3. Zhang, X. and Ke, H., " ALL/AML cancer classification by gene expression data using SVM and CSVM approach", *Genome Informatics*, Universal Academy Press, pp. 237-239, 2000
4. Xin Zhao, Leo Wang-Kit Cheung, "Kernel-imbedded Gaussian processes for disease classification using microarray gene expression data", *BMC Bioinformatics*.,8:67,2007.
5. Wenlong Xu, Minghui Wang, Xianghua Zhang, Lirong Wang, Huanqing Feng," SDED: A novel filter method for cancer-related gene selection", *Bioinformation* 2(7): 301-303,2008.
6. D.P. Berrar, C.S. Downes, W. Dubitzky, "Multiclass Cancer Classification Using Gene Expression Profiling and Probabilistic Neural Networks", *Pacific Symposium on Biocomputing* 8:5-16, 2003.
7. Pang-Ning Tan, Michal Steinbach, Vipin Kumar, "Introduction to Data Mining.", Pearson Education Inc., pp. 256-276, 2009
8. Vapnik V , *The nature of statistical learning theory*. 2nd edition. Springer,1999

9. Joachims, T., "Making large-scale SVM learning practical", Advances in Kernel Methods – Support Vector Learning, B. Schölkopf et al. (ed.), MIT Press, 1999.
10. Ben-Hur A, Ong CS, Sonnenburg S, Schölkopf B, Rätsch G, "Support Vector Machines and Kernels for Computational Biology.", PLoS Comput Biol 4(10), 2008.
11. ALL/AML Bench Mark Dataset:
12. www.broadinstitute.org/cgi-bin/cancer/publications/pub_paper.cgi?mode=view&paper_id=43
13. Platt, J. C., "Fast training of support vector machines using sequential minimal optimization". Advances in kernel methods: Support vector machines, B. Schölkopf et al. (ed.), MIT Press, 1999.
14. WEKA software: www.cs.waikato.ac.nz/~ml/WEKA

Efficient Small Template Iris Recognition System Using Wavelet Transform

Mohammed A. M. Abdullah

*Computer Engineering Department,
University of Mosul,
Mosul, 41002, Iraq*

m.am_86@yahoo.com

F. H. A. Al-Dulaimi

*Computer Engineering Department,
University of Mosul,
Mosul, 41002, Iraq*

fhali310@yahoo.com

Waleed Al-Nuaimy

*Department of Electrical Engineering and Electronics,
University of Liverpool,
Liverpool, L69 3GJ, UK*

wax@liv.ac.uk

Ali Al-Aataby

*Department of Electrical Engineering and Electronics,
University of Liverpool,
Liverpool, L69 3GJ, UK*

ali.al-aataby@liv.ac.uk

Abstract

Iris recognition is known as an inherently reliable biometric technique for human identification. Feature extraction is a crucial step in iris recognition, and the trend nowadays is to reduce the size of the extracted features. Special efforts have been applied in order to obtain low templates size and fast verification algorithms. These efforts are intended to enable a human authentication in small embedded systems, such as an Integrated Circuit smart card. In this paper, an effective eyelids removing method, based on masking the iris, has been applied. Moreover, an efficient iris recognition encoding algorithm has been employed. Different combination of wavelet coefficients which quantized with multiple quantization levels are used and the best wavelet coefficients and quantization levels are determined. The system is based on an empirical analysis of CASIA iris database images. Experimental results show that this algorithm is efficient and gives promising results of False Accept Ratio (FAR) = 0% and False Reject Ratio (FRR) = 1% with a template size of only 364 bits.

Keywords: Biometrics, Iris Recognition, Wavelets Transform, Feature Extraction, Pattern Recognition.

1. INTRODUCTION

The term "Biometrics" refers to a science involving the statistical analysis of biological characteristics. This measurable characteristic, biometric, can be physical, such as eye, face, retina vessel, fingerprint, hand and voice or behavioral, like signature and typing rhythm. Biometrics, as a form of unique person identification, is one of the subjects of research that is growing rapidly [1].

The advantages of unique identification using biometric features are numerous, such as fraud prevention and secure access control. Biometrics systems offer great benefits with respect to other authentication techniques. In particular, they are often more user friendly and can guarantee the physical presence of the user [1].

Iris recognition is one of the most reliable biometric technologies in terms of identification and verification performance. The iris is the colored portion of the eye that surrounds the pupil as depicted in Figure 1. It controls light levels inside the eye similar to the aperture on a camera. The round opening in the center of the iris is called the pupil. The iris is embedded with tiny muscles that dilate and constrict the pupil size. It is full of richly textured patterns that offer numerous individual attributes which are distinct even between the identical twins and between the left and right eyes of a person. Compared with other biometric features such as face and fingerprints, iris patterns are highly stable with time and unique, as the probability for the existence of two irises that are same is estimated to be as low as, one in 10^{72} [1,2].

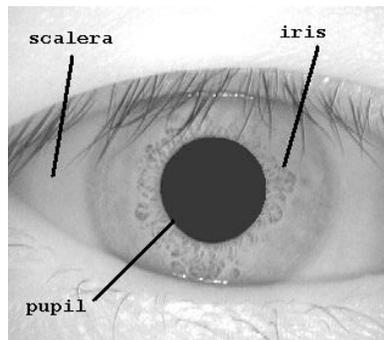


FIGURE 1: Image of the eye.

In this paper, the iris is efficiently normalized such that only useful data are encoded. Image enhancement techniques are applied. Moreover, the best combination of wavelet coefficients is found and used for successful identification and the best number of bits used for encoding the feature vector have been deduced while maintaining low template size.

The paper is organized as follows. Section 2 presents the main related works. Section 3 explains the typical stages of iris recognition and the proposed eyelid removing method. Section 4 presents the proposed feature extraction method. Experimental results are given in section 5, and Section 6 concludes this paper.

2. RELATED WORK

Iris identification using analysis of the iris texture has attracted a lot of attention and researchers have presented a variety of approaches in the literature.

Daughman [3] proposed the first successful implementation of an iris recognition system based on 2-D Gabor filter to extract texture phase structure information of the iris to generate a 2048 bits iris code. A group of researchers have used the 1-D wavelet transform as the core of the feature extraction module [4,5,6,7]. For instance, Boles and Boashash [4] extracted the features of the iris pattern by using the zero-crossings of 1-D wavelet transform of the concentric circles on the iris.

On the other hand, another group of researcher utilized 2-D wavelet transform to extract iris texture information [8,9,10,11,12,13,14]. For instance, Narote *et al* [11] proposed an algorithm for iris recognition based on dual tree complex wavelet transform and explored the speed and accuracy of the proposed algorithm. Hariprasath and Mohan [13] described iris recognition based on Gabor and Morlet wavelets such that the iris is encoded into a compact sequence of 2-D wavelet coefficient, which generate an iris code of 4096 bits. Kumar and Passi [14] presented a comparative study of the performance from the iris identification using different feature extraction methods with different templates size. Even though the previous systems have good recognition ratios, the template size remains rather large.

3. IRIS RECOGNITION SYSTEM

Generally, an iris recognition system is composed of many stages as shown in Figure 2. Firstly, an image of the person's eye is captured by the system and preprocessed. Secondly, the image is localized to determine the iris boundaries. Thirdly, the iris boundary coordinates are converted to the stretched polar coordinates to normalize the scale of the iris in the image. Fourthly, features representing the iris patterns are extracted based on texture analysis. Finally, the person is identified by comparing their features with an iris feature database.

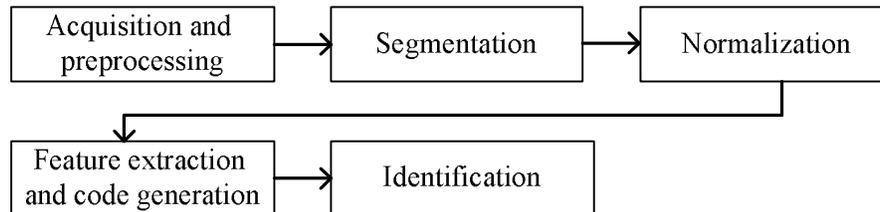


FIGURE 2: Block diagram of an iris recognition system.

3.1 Segmentation

For the purpose of identification, the part of the eye image carrying useful information is only the iris that lies between the sclera and the pupil [2]. Therefore, prior to performing iris matching, it is very important to localize the iris in the acquired image. The iris region, shown in Figure 3, is bounded by two circles, one for the boundary with the sclera and the other, interior to the first, with the pupil.

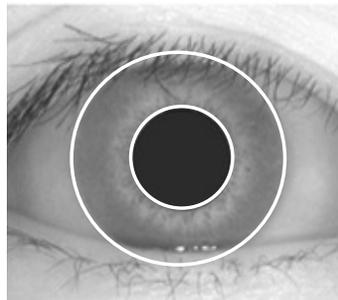


FIGURE 3: Segmented eye image.

To detect these two circles the Circular Hough transform (CHT) has been used. The Hough transform is a standard computer vision algorithm that can be used to determine the geometrical parameters for a simple shape, present in an image, and this has been adopted here for circle detection [15]. The main advantage of the Hough transform technique is its tolerance for gaps in feature boundary descriptions and its robustness to noise [16].

Basically, the first derivatives of intensity values in an eye image are calculated and the result is used to generate an edge map. From the edge map, votes are cast in Hough space for the parameters of circles passing through each edge point. These parameters are the center coordinates x_c and y_c , and the radius r , which are able to define any circle according to the following equation:

$$x_c^2 + y_c^2 - r^2 = 0 \quad \dots (1)$$

A maximum point in the Hough space will correspond to the radius and center coordinates of the best circle defined by the edge points [15].

3.2 Normalization

The size of the iris varies from person to person, and even for the same person, due to variation in illumination, pupil size and distance of the eye from the camera. These factors can severely affect iris matching results. In order to get accurate results, it is necessary to eliminate these factors. To achieve this, the localized iris is transformed into polar coordinates by remapping each point within the iris region to a pair of polar coordinates (r, θ) where r is in the interval $[0, 1]$ with 1 corresponding to the outermost boundary and θ is the angle in the interval $[0, 2\pi]$ as shown in Figure 4 [17,18].

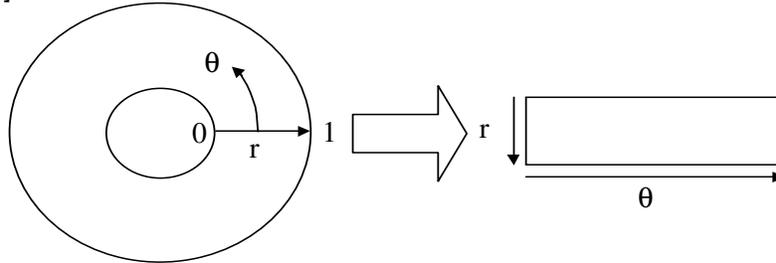


FIGURE 4: Rubber sheet model [17].

With reference to Figure 5, the remapping of the iris region from (x,y) Cartesian coordinates to the normalized non-concentric polar representation is modeled by the following equations:

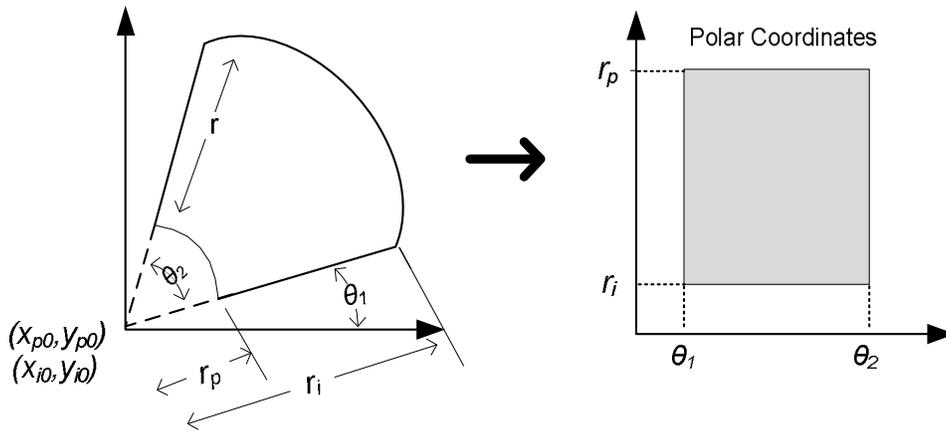


FIGURE 5: Image mapping from Cartesian coordinates to dimensionless polar coordinates.

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta) \quad \dots (2)$$

$$x(r, \theta) = (1 - r)x_p(\theta) + rx_i(\theta) \quad \dots (3)$$

$$y(r, \theta) = (1 - r)y_p(\theta) + ry_i(\theta) \quad \dots (4)$$

with

$$x_p(r, \theta) = x_{p0}(\theta) + r_p \cos \theta \quad \dots (5)$$

$$y_p(r, \theta) = y_{p0}(\theta) + r_p \sin \theta \quad \dots (6)$$

$$x_i(r, \theta) = x_{i0}(\theta) + r_i \cos \theta \quad \dots (7)$$

$$y_i(r, \theta) = y_{i0}(\theta) + r_i \sin \theta \quad \dots (8)$$

Where I is the iris picture, r_p and r_i are respectively the radius of pupil and the iris, while $x_p(\theta), y_p(\theta)$ and $x_i(\theta), y_i(\theta)$ are the coordinates of the papillary and iris boundaries in the direction θ . (x_{p0}, y_{p0}) and (x_{i0}, y_{i0}) are the centers of pupil and iris respectively.

For a typical eye image of dimension 320×280 pixel, the previous normalization method is performed to produce 50 pixels along r and 600 pixels along θ which result in 600×50 unwrapped strip.

On account of asymmetry of pupil (*not being a circle perfectly*) and probability of overlapping outer boundaries with sclera, we select 45 pixels from 50 pixels along r in the unwrapped iris. Therefore, the unwrapped iris becomes of dimensions 600×45 . The normalized iris image is shown in Figure 6.

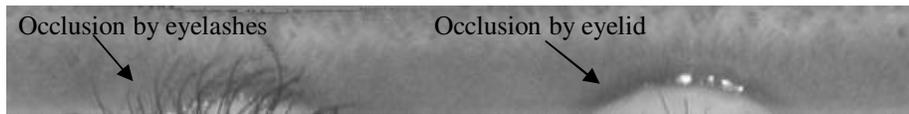


FIGURE 6: Normalized iris image.

3.3 Proposed Eyelash and Eyelid Removing Method

Since in most cases the upper and lower parts of the iris area are occluded by eyelids, it was decided to use only the left and right parts of the iris with a partial area of the upper and lower region for the iris recognition. Therefore, the whole iris $[0, 360^\circ]$ is not transformed in the proposed system. Experiments were conducted by masking the iris from $[148, 212^\circ]$ and $[328, 32^\circ]$ for the right and left parts while for the upper and lower parts, a semi circle with a radius equals to the half of the iris radius is used to mask the iris as depicted in Figure 7. Hence, the regions that contain the eyelids and eyelashes have been omitted while the remaining eyelashes are treated by thresholding, since analysis reveals that eyelashes are quite dark when compared with the rest of the eye image [15]. The corresponding rectangular block is show in Figure 8. Afterward, the block is concatenated together as shown in Figure 9.

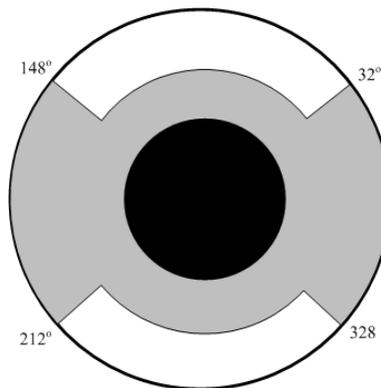


FIGURE 7: Masking the iris.



FIGURE 8: The normalized masked iris image.

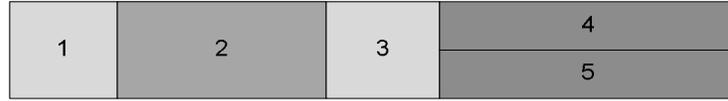


FIGURE 9: The concatenated block after removing the ignored parts.

The size of the rectangular block is reduced accordingly. By applying this approach, detection time of upper and lower eyelids and some cost of the polar transformation are saved. Saving ratio can be calculated from this equation:

$$\text{Saving ratio} = (\text{ignored parts of the iris} / \text{whole iris region}) * 100\% \quad \dots (9)$$

where

$$\begin{aligned} \text{ignored parts} &= ((148-32) + (328-212))/2 = 116 \\ \text{Saving rati}on &= 116/360 * 100\% = 32.22\% \end{aligned}$$

Figure 10 illustrates applying the proposed masking method on a normalized iris.

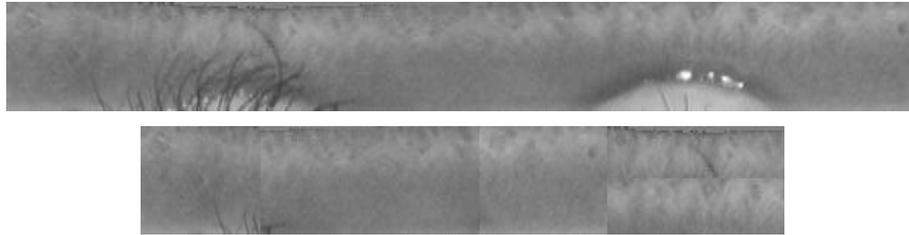


FIGURE 10: Applying the proposed masking method on a normalized iris.

Although the homogenous rubber sheet model accounts for pupil dilation and imaging distance it does not compensate for rotational inconsistencies. Rotational inconsistencies are treated in the matching stage (section 3.6).

3.4 Image Enhancement

Due to the effect of imaging conditions and situations of light sources, the normalized iris image does not have an appropriate quality. These disturbances may affect the performance of feature extraction and matching processes [12].

Hence for getting a uniform distributed illumination and better contrast in iris image, the polar transformed image is enhanced through adjusting image intensity values by mapping the intensity values in the input grayscale image to new values such that 1% of the pixel data is saturated at low and high intensities of the original image. This increases the contrast in a low-contrast grayscale image by remapping the data values to fill the entire intensity range [0, 255]. Then, histogram equalization has been used. Results of images before and after enhancement are shown in Figure 11.

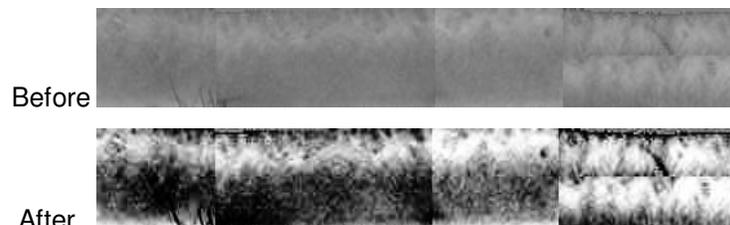


FIGURE 11: Image enhancement of the normalized iris.

3.5 Proposed Feature Extraction Method

In order to provide accurate recognition of individuals, the most discriminating information present in an iris pattern must be extracted. Only the significant features of the iris must be encoded so that comparisons between templates can be made. Most iris recognition systems make use of a band pass decomposition of the iris image to create a biometric template. For the encoding process the outputs of any used filter should be independent, so that there are no correlations in the encoded template, otherwise the filters would be redundant [19].

The Wavelet transform is used to extract features from the enhanced iris images. Haar wavelet is used as the mother wavelet. The Wavelet transform breaks an image down into four sub-sampled images. The results consist of one image that has been high-pass filtered in the horizontal and vertical directions (*HH* or Diagonal coefficients), one that has been low-pass filtered in the vertical and high-pass filtered in the horizontal (*LH* or Horizontal coefficients), one that has been low-pass filtered in the horizontal and high-pass filtered in the vertical (*HL* or Vertical coefficients), and one that has been low-pass filtered in both directions (*LL* or details coefficient) [8].

In Figure 12, a conceptual figure of basic decomposition steps for images is depicted. The approximation coefficients matrix *cA* and details coefficients matrices *cH*, *cV*, and *cD* (horizontal, vertical, and diagonal, respectively) obtained by wavelet decomposition of the input iris image. The definitions used in the chart are as follows [12].

- $C \downarrow$ denote downsample columns.
- $D \downarrow$ denote downsample rows.
- Lowpass_D denotes the decomposition low pass filter.
- Highpass_D denotes the decomposition high pass filter.
- I_i* denotes the input image.

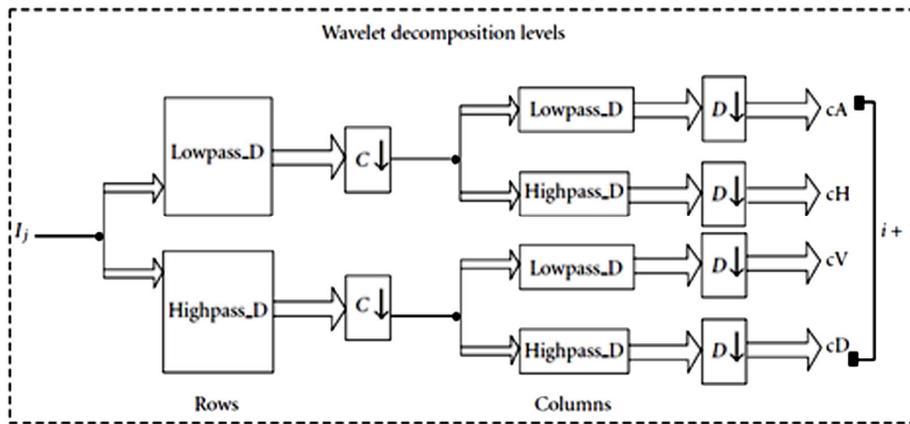


FIGURE 12: Wavelet decomposition steps diagram.

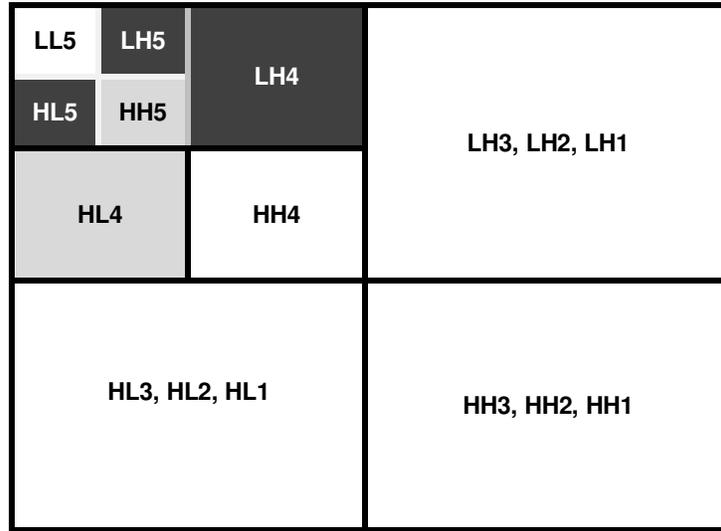


FIGURE 13: Five-level decomposition process with Haar wavelet. (Black indicates 4 levels quantization, Grey indicates two levels quantization).

Experiments were performed using different combinations of Haar wavelet coefficients and the results obtained from different combinations were compared to find the best. Since unwrapped image after masking has a dimension of 407×45 pixels, after 5 times decompositions, the size of the 5th level decomposition is 2×13 while for the 4th level is 3×26. Based on empirical experiments, the feature vector is arranged by combining features from *HL* and *LH* of level-4 (vertical and horizontal coefficients [*HL4 LH4*]) with *HL*, *LH* and *HH* of level-5 (vertical, horizontal and diagonal coefficients [*HL5 LH5 HH5*]). Figure 13 shows a five level decomposition with Haar wavelet.

In order to generate the binary data, features of *HL4* and *HH5* are encoded using two-level quantization while features of *LH4*, *HL5* and *LH5* are encoded using four-level quantization. After that these features are concatenated together as shown in Figure 14 which illustrates the process used for obtaining the final feature vector.

LH4 156 bits 2×[3×26]	HL4 78 bits [3×26]	LH5 52 bits 2×[2×13]	HL5 52 bits 2×[2×13]	HH5 26 bits [2×13]
------------------------------------	---------------------------------	-----------------------------------	-----------------------------------	---------------------------------

FIGURE 14: Organization of the feature vector which consists of 364 bits.

3.6 Matching

The last module of an iris recognition system is used for matching two iris templates. Its purpose is to measure how similar or different templates are and to decide whether or not they belong to the same individual or not. An appropriate match metric can be based on direct point-wise comparisons between the phase codes [18]. The test of matching is implemented by the Boolean XOR operator applied to the encode feature vector of any two iris patterns, as it detects disagreement between any corresponding pair of bits. This system quantifies this matter by computing the percentage of mismatched bits between a pair of iris representations, *i.e.*, the normalized Hamming distance.

Let *X* and *Y* be two iris representations to be compared and *N* be the total number

$$HD = \frac{1}{N} \sum_{j=1}^N X_j \oplus Y_j \quad \dots (10)$$

In order to avoid rotation inconsistencies which occur due to head tilts, the iris template is shifted right and left by 6 bits. It may be easily shown that scrolling the template in Cartesian coordinates is equivalent to an iris rotation in polar coordinates. This algorithm performs matching of two templates several times while shifting one of them to different locations. The smallest *HD* value amongst all these values is selected, which gives the matching decision [18, 19].

4. EXPERIMENTAL RESULTS AND COMPARISON

The images are obtained from the Chinese Academy of Sciences Institute of Automation (CASIA) [20] which is available in public domain. The database consists of 756 iris images from 108 classes. Images from each class are taken from two sessions with one month interval between the sessions. For each iris class, we choose three samples taken at the first session for training and all samples captured at the second session serve as test samples. This is also consistent with the widely accepted standard for biometrics algorithm testing [21, 22].

Experiments were performed using different combinations of wavelet coefficients and the results obtained from different combinations are compared to find the best as shown in Table 1. The selected combination gives the best Correct Recognition Rate (CRR) for a minimum feature vector length of 364 bits only.

Combinations	Quantization	CRR	Vector Size
CH4 (D&V)	2 bits	69%	156 bits
CH4 (V&H)	2 bits	73%	156 bits
CH4 (D&H)	2 bits	70%	156 bits
CH4 (D&V) + CH5 (V)	2 bits	76%	182 bits
CH4 (D&V) + CH5 (H)	2 bits	82%	182 bits
CH4 (D&V) + CH5 (D)	2 bits	77.8%	182 bits
CH4 (D&V) + CH5 (D&V)	2 bits	83%	208 bits
CH4 (D&V&H)	2 bits	85%	162 bits
CH4 (H) + CH5 (H)	4 bits	92%	208 bits
CH4 (H) + CH5 (V)	4 bits	89%	208 bits
CH4 (H) + CH5 (V&H)	4 bits	95%	260 bits
CH4 (D) + CH5 (V&H)	4 bits	72%	260 bits
CH4 (V) + CH5 (V&H)	4 bits	68.5%	260 bits
CH4 (D&H)	4 bits	92%	312 bits
CH4 (D&V)	4 bits	62%	312 bits
CH4 (V&H)	4 bits	88%	312 bits
CH5 (V&H)	4 bits	54%	312 bits
CH5 (V&D)	4 bits	49%	312 bits
CH4 (V&H) + CH5 (V)	4 bits	90%	368 bits
CH4 (V&H) + CH5 (H)	4 bits	93%	368 bits
CH4 (D&V) + CH5 (D&V)	4 bits	71%	416 bits
CH4 (V&H) + CH5 (V&D)	4 bits	90.5%	416 bits
CH4 (V&H) + CH5 (V&H)	4 bits	96%	416 bits
CH4 (V&D&H)	4 bits	91%	468 bits
CH4 (H) ₄ + CH4 (V) ₂ CH5 (V) ₄ + CH5 (H) ₄ + CH5 (D) ₂	2 bits and 4 bits	99%	364 bit

TABLE 1: Comparison among multiple wavelet coefficients
(D: Diagonal coefficients, H: Horizontal coefficients, and V: Vertical coefficients).

With a pre-determined separation Hamming distance, a decision can be made as to whether two templates were created from the same iris (*a match*), or whether they were created from different irises. However, the intra-class and inter-class distributions may have some overlap, which would result in a number of incorrect matches or false accepts, and a number of mismatches or false rejects. Table 2 shows the FAR and FRR associated with different separation points.

Threshold	FAR (%)	FRR (%)
0.20	0.00	59.34
0.24	0.00	28.80
0.26	0.00	10.30
0.28	0.00	3.87
0.29	0	1.00
0.30	1.51	0.86
0.32	5.43	0.00
0.36	26.47	0.00
0.38	48.68	0.00

TABLE 2: False accept and false reject rates for CASIA database with different separation points.

Figure 15 shows the distribution of inter-class and intra-class distribution of the system with a Hamming distance separation point of 0.29. With this separation point, false accept rate and false reject rate of 0% and 1% respectively are achieved. Such FRR are appeared, due to the overlap between the classes but it still allows for accurate recognition.

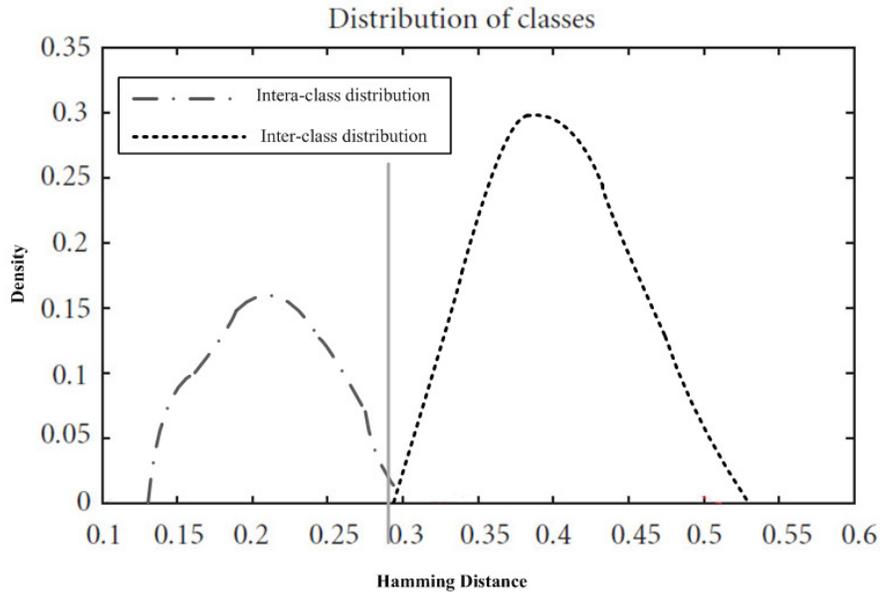


FIGURE 15: The distribution of intra-class and inter-class distances with a separation point of 0.29.

This system scored a perfect 0% FAR and 1% FRR. Table 3 shows the classification rate compared with a well known methods.

Method	Feature Lengths (bits)	CRR (%)	Used Database
Narote <i>et al</i> [11]	1088	99.2	CASIA [20]
Poursaberi [12]	544	97.22	CASIA [20]
Hariprasath [13]	4096	99.0	UBIRIS [24]
Xiaofu [23]	1536	98.15	CASIA [20]
Proposed	364	99.0	CASIA [20]

TABLE 3: Comparison of feature vector length and the Correct Recognition rate (CRR).

In the two methods; [11] and [13], the CRR is equal or a little bit better than ours. In fact, the dimensionality of the feature vector in both methods is much higher than ours. The feature vector consists of 1088 bits in [11] and of 4096 in [13], while it consists only of 364 bits in the proposed method. In addition, neither [11] nor [12] have suggested a method for removing eyelids or eyelashes.

Furthermore, [12] proposed a method to produce 544 bits of feature vector by applying four-level wavelet transform on the lower part of the iris assuming that only the upper part is occluded by the eyelashes and eyelids while the lower part is not. On the other hand, [23] employed two dimensional complex wavelet transform to produce 1536 bits of feature vector, however no method for noise removing has been applied also.

5. CONCLUSION

In this paper, we proposed an iris recognition algorithm using wavelet texture features based on a novel masking approach for eyelid removing. A masked area around the iris is used in the iris detection method. This area contains a complex and abundant texture information which are useful for feature extraction. The feature vector is quantized to a binary one, reducing the processing time and space, while maintaining the recognition rate.

Experimental results using CASIA database illustrate that relying on a smaller but more reliable region of the iris, although reduced the net amount of information, improve the recognition performance.

The experimental results clearly demonstrate that the feature vector consisting of concatenating *LH4*, *HL4*, *LH5*, *HL5*, and *HH5* gives the best results. On the other hand, Haar wavelet is particularly suitable for implementing high-accuracy iris verification/identification systems, as feature vector is at the least with respect to other wavelets. In identification mode, the CRR of the proposed algorithm was 99% with template size of 364 bits. Such vector size can be easily stored on smart cards and participate to reduce the matching and encoding time tremendously.

The proposed algorithm is characterized by having less computational complexity compared to other methods. Based on the comparison results shown in Table 3, it can be concluded that the proposed method is promising in terms of execution time and performance of the subsequent operations due to template size reduction.

6. REFERENCES

- [1] D. Woodward, M. Orlans and T. Higgins. *"Biometrics"*. McGraw-Hill, Berkeley, California, pp. 15-21 (2002)
- [2] H. Proença, A. Alexandre. *"Towards noncooperative iris recognition: A classification approach using multiple signatures"*. IEEE Transaction on Pattern Analysis, 29(4): 607-612, 2007
- [3] J. Daugman, *"High confidence visual recognition of persons by a test of statistical independence"*. IEEE Transaction on Pattern Analysis, 15(11): 1148-1161, 1993
- [4] W. Boles, B. Boashash. *"A Human Identification Technique Using Images of the Iris and Wavelet Transform"*. IEEE Transactions on Signal Processing, 46(4): 1085–1088, 1998
- [5] C. Chena, C. Chub. *"High Performance Iris Recognition based on 1-D Circular Feature Extraction and PSO–PNN Classifier"*. Expert Systems with Applications journal, 36(7): 10351-10356, 2009
- [6] H. Huang, G. Hu. *"Iris Recognition Based on Adjustable Scale Wavelet Transform"*. 27th Annual International Conference of the Engineering in Medicine and Biology Society, Shanghai, 2005
- [7] H. Huang, P.S. Chiang, J. Liang. *"Iris Recognition Using Fourier-Wavelet Features"*. 5th International Conference Audio- and Video-Based Biometric Person Authentication, Hilton Rye Town, New York, 2005
- [8] J. Kim, S. Cho, R. J. Marks. *"Iris Recognition Using Wavelet Features"*. The Journal of VLSI Signal Processing, 38(2): 147-156, 2004
- [9] S. Cho, J. Kim. *"Iris Recognition Using LVQ Neural Network"*. International conference on signals and electronic systems, Porzan, 2005

- [10] O. A. Alim, M. Sharkas. "*Iris Recognition Using Discrete Wavelet Transform and Artificial Neural Networks*". IEEE International Symposium on Micro-Nano Mechatronics and Human Science, Alexandria, 2005
- [11] S.P. Narote, A.S. Narote, L.M. Waghmare, M.B. Kokare, A.N. Gaikwad. "*An Iris Recognition Based on Dual Tree Complex Wavelet Transform*". TENCN IEEE 10th conference, Pune, India, 2007
- [12] A. Poursaberi, B.N. Araabi. "*Iris Recognition for Partially Occluded Images: Methodology and Sensitivity Analysis*". EURASIP Journal on Advances in Signal Processing, 2007(1): 12-14, 2007
- [13] S. Hariprasath, V. Mohan. "*Biometric Personal Identification Based On Iris Recognition Using Complex Wavelet Transforms*". Proceedings of the 2008 International Conference on Computing, Communication and Networking (ICCCN) IEEE, 2008
- [14] A. Kumar, A. Passi. "*Comparison and Combination of Iris Matchers for Reliable Personal Identification*". Computer Vision and Pattern Recognition Workshops, IEEE, 2008
- [15] R. Wildes, J. Asmuth, G. Green, S. Hsu, and S. McBride. "*A System for Automated Iris Recognition*", Proceedings IEEE Workshop on Applications of Computer Vision, Sarasota, FL, USA, 1994
- [16] T. Moravčík. "*An Approach to Iris and Pupil Detection in Eye Image*". XII International PhD Workshop OWD, University of Žilina, Slovakia, 2010
- [17] K. Dmitry "*Iris Recognition: Unwrapping the Iris*". The Connexions Project and Licensed Under the Creative Commons Attribution License, Version 1.3. (2004)
- [18] R. Schalkoff. "*Pattern Recognition: Statistical, Structural and Neural Approaches*". John Wiley and Sons Inc., pp. 55-63 (2003)
- [19] J. Daugman. "*Statistical Richness of Visual Phase Information: Update on Recognizing Persons by Iris Patterns*". International Journal of Computer Vision, 45(1): 25-38, 2001
- [20] Chinese Academy of Sciences, Center of Biometrics and Security Research. Database of 756 Grayscale Eye Images. <http://www.cbsr.ia.ac.cn/IrisDatabase.htm>
- [21] A. Mansfield and J. Wayman, "*Best practice standards for testing and reporting on biometric device performance*". National Physical Laboratory of UK, 2002
- [22] L. Ma. "*Personal identification based on iris recognition*". Ph.D. dissertation, Institute of Automation, Chinese Academy of Sciences, Beijing, China, 2003
- [23] H. Xiaofu, S. Pengfei. "*Extraction of Complex Wavelet Features for Iris Recognition*". Pattern Recognition, 19th International Conference on Digital Object Identifier, Shanghai, 2008
- [24] Department of Computer Science, University of Beira Interior, Database of eye images. Version 1.0, 2004. <http://iris.di.ubi.pt/>

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Biometric and Bioinformatics (IJBB)* brings together both of these aspects of biology and creates a platform for exploration and progress of these, relatively new disciplines by facilitating the exchange of information in the fields of computational molecular biology and post-genome bioinformatics and the role of statistics and mathematics in the biological sciences. Bioinformatics and Biometrics are expected to have a substantial impact on the scientific, engineering and economic development of the world. Together they are a comprehensive application of mathematics, statistics, science and computer science with an aim to understand living systems.

We invite specialists, researchers and scientists from the fields of biology, computer science, mathematics, statistics, physics and such related sciences to share their understanding and contributions towards scientific applications that set scientific or policy objectives, motivate method development and demonstrate the operation of new methods in the fields of Biometrics and Bioinformatics.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJBB.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 5, 2011, IJBB appears in more focused issues. Besides normal publications, IJBB intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

LIST OF TOPICS

The realm of International Journal of Biometrics and Bioinformatics (IJBB) extends, but not limited, to the following:

- Bio-grid
- Bioinformatic databases
- Biomedical image processing (registration)
- Biomedical modelling and computer simulation
- Computational intelligence
- Computational structural biology
- DNA assembly, clustering, and mapping
- Fuzzy logic
- Gene identification and annotation
- Hidden Markov models
- Molecular evolution and phylogeny
- Molecular sequence analysis
- Bio-ontology and data mining
- Biomedical image processing (fusion)
- Biomedical image processing (segmentation)
- Computational genomics
- Computational proteomics
- Data visualisation
- E-health
- Gene expression and microarrays
- Genetic algorithms
- High performance computing
- Molecular modelling and simulation
- Neural networks

CALL FOR PAPERS

Volume: 5 - Issue: 3 - May 2011

i. Paper Submission: May 31, 2011

ii. Author Notification: July 01, 2011

iii. Issue Publication: July /August 2011

CONTACT INFORMATION

Computer Science Journals Sdn Bhd

M-3-19, Plaza Damas Sri Hartamas
50480, Kuala Lumpur MALAYSIA

Phone: 006 03 6207 1607
006 03 2782 6991

Fax: 006 03 6207 1697

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2011
COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6207 1607
006 03 2782 6991

FAX: 006 03 6207 1697
EMAIL: cscpress@cscjournals.org