

Modeling of Biometric Recognition Based on Human Visual System



Fatima Zohra Allam^{1,2}, Hicham Bousbia-Salah¹ and Latifa Hamami¹

¹ Laboratoire Signal et Communication, Ecole Nationale Polytechnique d'Alger, 10, Avenue des Frères Oudek El Harrach 16200 Alger Algérie. <http://www.enp.edu.dz>.

² Laboratoire de Recherche en TIC, Institut National de la Poste et des TIC, Route de Larbaâ Eucalyptus 16220 Alger Algérie. <http://www.inptic.edu.dz> {fatima_zohra.allam, hicham.bousbia-salah, latifa.hamami}@g.enp.edu.dz

ABSTRACT

Research on biometric techniques has a wide range of applications. The methods developed have shown their robustness, especially in the field of biometric recognition. These methods are generally based on the processing of information in the space domain. In this paper, another alternative is presented for realizing a biometric system for biometric recognition based on the processing of spectral information. It is a more elaborate model on the perception of images based on the behavior of the Human Visual System, or HVS. The Human Visual System consistently models the key functions of perception such as Perceptual Channel Decomposition, contrast sensitivity functions, and so on. We are interested in developing a bi-modal biometric system. We merge two biometric signatures, in this case, the face and the iris to improve the accuracy of biometric recognition of individuals.

Key words: Spectral information, Human Visual System, Perceptual Channel Decomposition, Fusion of biometric signatures.

1. INTRODUCTION

The design of multimodal biometric systems is of great importance in the field of recognition. Biometric techniques, widely developed in the literature, have grown considerably, mainly in the field of people recognition.

Unimodal biometric systems make it possible to recognize a person using a single biometric modality, but cannot guarantee with certainty a good identification. In addition, these systems are sensitive, among other things, to the noise introduced by the single sensor, to the non-universality and the lack of individuality of the biometric modality chosen, as well as to intrusion attempts. These constraints can be reduced by setting up multimodal biometric systems using several biometric signatures of the same person.

This paper aims to approach the concept of bimodal Biometric Techniques. The methods used

in our approach are the face and the iris. We briefly explain the principle of facial recognition and iris recognition. Next, we discuss multimodal fusion. We will focus on the score level merger that was chosen for the merger of our two modalities.

Facial recognition, as one of the basic biometric technologies, has become increasingly important in the field of research. It is relatively very old and widely used. It is certainly the biometric characteristic that humans most naturally use to identify themselves. This explains why it is very well accepted by users and considered non-intrusive and its acquisition, via a sensor, is relatively easy and inexpensive.

Like any biometric system, facial recognition has certain limitations and drawbacks, including, mainly, the difficulty of identifying individuals in movement, the impossibility of differentiating between twins, sensitivity to the variation of lighting and to the change of the position of the face [1].

The second modality retained in our work is the iris. Iris recognition is a relatively newer technology. It is one of the most precise and discriminating methods. Hardly falsifiable, it is considered very reliable. The irises are characterized by their uniqueness; the two irises of the same individual are different. One of the main advantages of iris is the large amount of information it contains. The iris is considered intrusive because of the discomfort of individuals towards the laser applied to the eye during the acquisition. It is mainly reserved for high security applications. Some problems may arise during acquisition such as reflection, or variation in the size of the pupil [2].

One of the alternatives to improve security and enhance the performance of biometric systems is to merge these two modalities[3]-[7].

We plan to develop a new recognition technique based on the concept of the Human Visual System by trying to simulate or mimic human perception.

In our approach, it is interesting to see how to exploit the behavior of Human Visual System in biometric systems for improving self-recognition.

Several questions arise, among others: is it possible to develop algorithms capable of providing a better performance compared to other methods used and already widely developed in the literature? It should be noted that these methods have shown their performance and robustness in various fields of application such as compression or image quality evaluation[8]-[11], Watermarking of images [12]-[14],tattoo of images [15], etc.

Which model of Human Visual System will be most appropriate? How will the performance of the biometric system be evaluated?

This article is organized as follows: In section 2, we present multimodal biometric systems and their principle of operation. We also present multimodal fusion, its objectives and the levels of data fusion. In section 3, we present the CASIA database which provides us with images relating to the face and iris. The Human Visual System model, explained in section 4, is the most relevant part of our work. In section 5, we explain the concept of contrast by emphasizing the calculation of Local Contrast with Limited Band. It is the calculation of this contrast that simulates the behavior of the Human Visual System. Finally, section 6 explains the general structure of the system we have developed.

2. BIOMETRIC SYSTEMS: PRINCIPLE AND OPERATION

2.1 Principle

Generally, biometric systems consist of five main modules, namely:

a. Acquisition module

It is responsible for acquiring an individual's biometric data.

b. Characteristic extraction module

It takes the biometric data acquired by the sensor as input and extracts only the relevant information in order to form a new representation of the data.

c. Correspondence module

It compares all the characteristics extracted with the model saved in the system database and determines the degree of similarity (or divergence) between the two.

d. Matching module

It is used to increase the similarity rate between two objects. A matching score calculation is performed between the model of the claimed identity and the vector of the recorded biometric parameters. The proclaimed identity will be granted if and only if the matching score exceeds a certain threshold which is considered sufficient.

e. Decision module

This module makes it possible to decide whether the identity of the user corresponds or not to the identity proclaimed or sought.

2.2 Operation of a biometric system

Biometric systems have three main functions [16]:

a. Enrollment

It corresponds to the biometric registration of individuals in the system database.



Figure 1:Functional diagram of a biometric system in enrollment mode

b. Authentication

It consists in verifying the concordance between the biometric data entered and that of the same person, assembled in a database. This is an individual comparison.

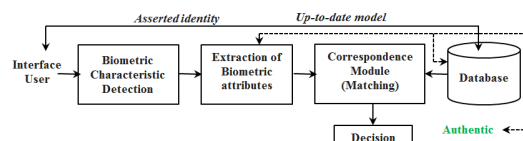


Figure 2: Functional diagram of a biometric system in authentication mode

c. Identification

It allows you to search for an individual's identity by comparing their biometric data with other people in a database. It is a collective comparison.

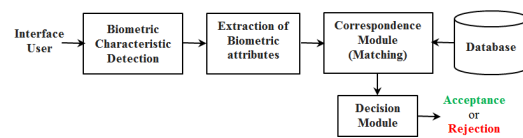


Figure 3:Functional diagram of a biometric system in identification mode

2.3 Multimodal fusion

Data fusion occupies an important place in multimodal biometric systems. Various ways of merging these data have been developed in the literature [17]-[19]. This diversity depends mainly on:

- Level where the merger takes place
- Objective of the merger
- Type of sensors used

There are four levels of merge that can be classified into two families, the merge before the matching and the merge after the matching.

In our approach, we are interested in fusion after correspondence, known in the literature as Multiple Classifier Systems. This merging mode can be applied either at the level of scores or at the level of decisions. Score level merger is the most used type of merger because it can be applied to all types of systems, in a space of limited size and with relatively simple and effective methods.

Given that the two biometric modalities are characterized by a certain homogeneity and / or compatibility, this prompts us to merge at the level of scores, therefore, after the matching phase [16].

For the acquisition of both modalities, we use a single capture device with a very high resolution. This will allow us to simultaneously analyze the texture of the iris and the face of the same individual[20] and [21].

3. CASIA IRIS IMAGE DATABASE

The database used is CASIA-Iris-Distance, a subset of CASIA-IrisV4. It contains iris images, in JPEG format and 8-bit grayscale, captured using a long-range, high-resolution system, collected under near infrared lighting [22] and [23]. The different characteristics shown in the iris images make it possible to study research questions specific to iris recognition, such as the robustness of iris recognition against changes in lighting, the recognition of iris of twins, etc.

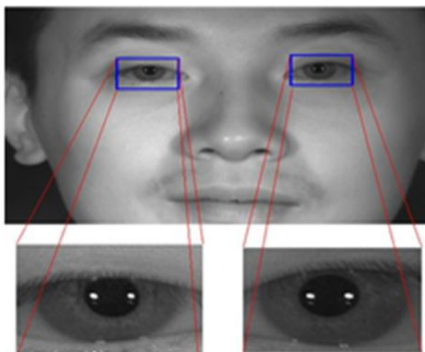


Figure 4: Image of the CASIA-Iris-Distance database.

4. MODELING THE HUMAN VISUAL SYSTEM

Of our five senses, vision takes the most important place in the brain. It is the one that uses the most neurons in our brain [24]. Many researchers have been interested in the functioning of the Human Visual System. The knowledge of it is of crucial interest in Image Processing. Indeed, most applications are intended to provide images evaluated and exploited by humans.

The Human Visual System is a multi-sensor system with a high-performance interpretation and recognition capability.

In this section, we briefly discuss the modeling of Human Visual System frequently used in Digital Image Processing. It derives, mainly, from its biological and functional structure as well as from psychophysical experiments [25] and [26].

A comprehensive description of both the anatomical aspects of Human Visual System and the visual models of the literature is given in[15].

The modeling of the Human Visual System in perceptual channels considered here, widely described, *inter alia*, in [15], [24], [25]and [27], is given in Figure 5[15]. It consists of four radial frequency domains, called rings, indexed from I to IV.

The ring I corresponds to the spatial frequencies between 0 and 1.5 cy/d° (cycles per degree). Domain II corresponds to frequencies between 1.5 and 5.7 cy/d° , domain III at frequencies between 5.7 and 14.2 cy/d° and domain IV at frequencies between 14.2 and 28.2 cy/d° [9].

The angular selectivity depends on the frequency domain considered. There is no selectivity highlighted for low frequencies (ring I). For domain II, an angular selectivity of 45° was measured in which four channels oriented for this domain, indexed from 1 to 4, were defined. For domains III and IV, the angular selectivity measured was 30° . For each of these two domains, six oriented channels, indexed from 1 to 6, are defined[25]. It should be noted that the band widths of crowns II, III and IV are respectively 1.9 octaves, 1.3 octaves and 1 octave[27].

We use Human Visual System behavior modeling to extract the relevant information from face and iris images. The images of the face and iris come from the contours detection phase and from the segmentation phase respectively.

The representation space is therefore decomposed into seventeen visual channels distributed as follows [15]:

- A mono-directional BF channel without angular selectivity.
- Three radial frequency bands decomposed into angular channels whose number depends on the radial band considered:
 - One band (zone II) 1.5 cy/d° containing four angular channels (45°).
 - Two bands (zone III) 5.7 cy/d° containing six angular channels (30°).
 - Three bands (zone IV) 14.2 cy/d° containing six angular channels (30°).

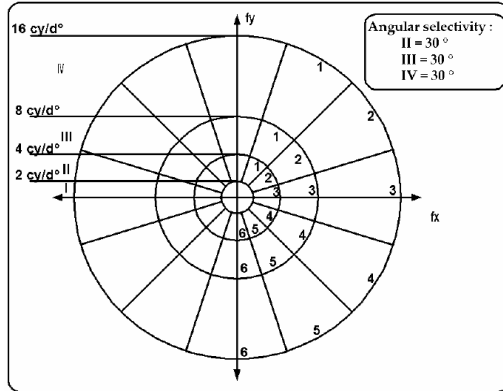


Figure 5: Decomposition of the representation Space into Perceptual Channels [15].

5. CALCULATION OF CONTRAST

Understanding the process of visual perception is an important element to consider when modeling the Human Visual System.

In this section, we reveal the different definitions of contrast. One of the best known is the Michelson contrast, which was introduced to give a measure of the visibility of the interference fringes on test patterns whose luminance varied in a sinusoidal way from L_{max} to L_{min} [28]:

$$C^M = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \tag{1}$$

Where:

L_{max} and L_{min} are the maximum and minimum luminance values in the image.

The joint use of sinusoidal test patterns and this definition of contrast has been very successful in psychophysics. This allowed in particular to study the sharpness of the human eye by building the Functions of Contrast Sensitivity (FCS).

For his part, Weber defines contrast as a relative luminance variation ΔL on a uniform background L [29]and [30]. This tool has been widely used, among other things, to measure the visibility of targets.

$$C^W = \frac{\Delta L}{L} \tag{2}$$

This is mainly because the perception of contrast is local.

Meanwhile, Gordon defines the notion of local contrast. This calculates the Michelson contrast between the mean gray levels of two regions [31].

Beghdadi and Negrate proposed another concept taking into account the strong sensitivity of the visual system to the contours of objects.

They integrate in the measure of the local contrast the average level of contours estimated in the analysis window. This method is inspired by that of Gordon [32].

Peli noted the inadequacy and inadequacy of classic contrasts to real images. He proposed another definition which takes into account the sensitivity of the Human Visual System to spatial frequencies [33]. He therefore introduced the concept of local contrast with limited band. This contrast reflects the more important fact that the perception of a detail of the image also depends on its local environment[34].

The calculation of the local contrast supposes a decomposition of the image into visual sub-bands and is defined by the ratio between the local luminance of a sub-band and the local average luminance relative to this channel, which, for a pixel considered (m, n), refers to the sum of the luminance of all the radial sub-bands lower than the sub-band.

For a breakdown such as that illustrated in Figure 1, this contrast is written [34] and [35]:

$$C_i(m, n) = \frac{L_i(m, n)}{\sum_{k=0}^{i-1} L_k(m, n)} \tag{3}$$

Where: i represents i^{th} the radial channel

The denominator represents the local average luminance, which corresponds to all the channels of spectral support lower than that of the i^{th} channel.

From these different definitions, we can, a priori, opt for local contrast with limited band, because it seems to us the most suitable for our application.

6. PROPOSED SYSTEM ARCHITECTURE

In this section, we will explain the methodology we adopted in our work. The figure6 shows the overall architecture of our system.

The enrollment, pretreatment and feature extraction phases are applied to each of the two modalities, face and iris, taken separately in order to retain only the representative discriminative attributes of the face and the iris.

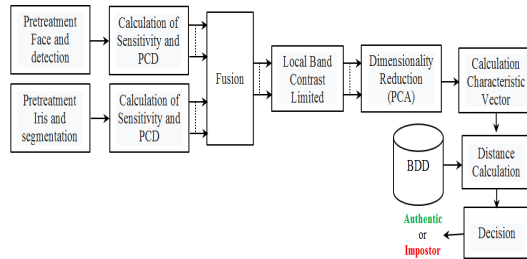


Figure 6: Overall architecture of proposed recognition system based on Human Visual System

The two images, obtained at the end of this phase, undergo a Perceptual Channels Decomposition (PCD). This decomposition, modeling the behavior of the Human Visual System, exploits a set of seventeen frequency channels distributed in a Low Frequencies channel and four radial rings. Each ring is broken down into angular sectors.

The advantage of performing a Perceptual Channel Decomposition is to generate images, confined to a certain frequency range, perfectly uncorrelated. The goal of this phase is to calculate the luminance of each pixel. Luminance is a function of frequency. It represents the contrast sensitivity.

As mentioned in section 2, we merge the characteristics of the two modalities before matching. At the end of this phase, we obtain seventeen images containing the merged characteristics of the two modalities: Face and Iris. We calculate the contrast for each of the seventeen luminance sub-bands from the Perceptual Channel Decomposition. This calculation is necessary taking into account the sensitivity of the Human Visual System to this quantity. We chose the Peli model to calculate the Limited Band Local Contrast given by the equation (3) in section 5.

For the reduction of dimensionality, we proceed to the Principal Component Analysis. It is one of the best known and most used algorithms in facial recognition[36] and [37]. This step makes it possible to determine the characteristic vector.

Subsequently, we calculate the distance to compare the characteristic vector generated with that from

the database. This allows us to decide if we have a best possible match according to the entry presented to it. To reach the final decision, we use the “Majority voting” [38] and [39] algorithm which gives conclusive results.

7. CONCLUSION

Multimodal biometric systems have grown considerably in the areas of identification and surveillance. Several research works, relating to image processing methods applied to individual recognition, have been developed and have demonstrated their robustness, their effectiveness and their performance. Our approach is different from that proposed in the literature.

In this paper, we have previously introduced the general concepts of biometric systems. We have briefly explained the main steps in the operation of any biometric system, in this case, enrollment, authentication and identification. Next, we discussed the modules that make up a biometric system: the capture module, the module for extracting relevant attributes, the correspondence module and the decision module. The pretreatment and feature extraction step is followed by the most important phase, namely, the recognition phase. We also approached multimodal fusion, specifying the objectives as well as the levels of multimodal fusion.

A second part, of not least importance, was treated in this paper. We have developed a new recognition technique, based on the concept of the Human Visual System by trying to simulate or mimic human perception. In this context, the visual approach seems to be an innovative and interesting line of research. This approach has been successfully applied in several areas, such as image compression, degraded quality assessment, water marking...

Our approach is focused on the Perceptual Channels Decomposition (PCD) in order to generate images, confined around a certain frequency range, perfectly decorrelated. For the extraction of the characteristic vectors, we opted for the Principal Component Analysis (PCA)

The principle of the visual approach consists in exploiting one or more characteristics of the peripheral parts of the Human Visual System. These characteristics can integrate the sensitivity of the HVS to spatial frequencies, its sensitivity to local contrast, its spacio-frequency selectivity, etc.

During the extraction of relevant attributes, the image of each of the two modalities was broken down into perceptual channels. This decomposition modeled the spacio-frequency selectivity of the Human Visual System. Generally speaking, this decomposition uses a set of seventeen channels, each adjusted to a given frequency and radial frequency band. The seventeen images output from this Perceptual Channels Decomposition (PCD) contain the same spatial information but are perfectly decorrelated from a spectral point of view.

We calculated the Local Contrast with Limited Band for each image. Local Bandwidth Contrast, which models the sensitivity of the Human Visual System more to contrast than to luminance, has shown its value in image processing. It is the calculation of this magnitude that simulates the behavior of the Human Visual System.

REFERENCES

1. S. Gunasekar, J. Ghosh, and A.C. Bovik. **Face Detection on Distorted Images Augmented by Perceptual Quality-Aware Features.** *IEEE Transactions on Information Forensics and Security*, Vol. 9, pp. 2119-2131, December 2014.
<https://doi.org/10.1109/TIFS.2014.2360579>
2. J. Zuo, and N.A. Schmid. **Adaptive quality-based performance prediction and boosting for iris authentication: methodology and its illustration.** *IEEE Transactions on Information Forensics and Security*, Vol. 8, no. 6, pp. 1051-1060, 2013
3. S. Bharadwaj, M. Vatsa, and R. Singh. **Biometric quality: A review of fingerprint, iris and face.** *EURASIP Journal on Image Video Processing*. Vol. 34, pp. 1-28, 2014.
<https://doi.org/10.1186/1687-5281-2014-34>
4. A. Dogra, B. Goyal, and S. Agrawal. From Multi-Scale Decomposition to Non-Multi-Scale Decomposition Methods: A Comprehensive Survey of Image Fusion Techniques and Its Applications. *IEEE Access*, Vol. 5, pp. 16040-16067, August 2017.
<https://doi.org/10.1109/ACCESS.2017.2735865>
5. N. Khiari. **Multimodal biometry based on the iris and the face.** *Doctoral thesis in Signal and image processing. Paris-Saclay University, jointly with the National School of Engineers, Tunis, Tunisia*, May 2016.
6. T. Ma, J. Ma, B. Fang, F. Hu, S. Quan, and H. Du. **Multi-scale decomposition based fusion of infrared and visible image via total variation and saliency analysis.** *Infrared Physics & Technology*, Vol. 92, pp. 154-162, August 2018.
7. O. Sharifi, and M. Eskandari. **Optimal Face-Iris Multimodal Fusion Scheme.** *Symmetry in Complex Networks II*, Vol. 8 no. 6, June 2016.
<https://doi.org/10.3390/sym8060048>
8. Z. Ahmed Seghir. **Evaluation of the quality of images,** *PhD thesis, specialty: Computer Science, Mentouri University Constantine. Computer Science Department*, 2012.
9. P. Callet, and D. Barba. **Color perception model. Application to image quality assessment.** *Signal Processing, Color Digital Image*, Vol. 21, pp. 61-77, 2004.
10. C. Charrier, C. Larabi, and H. Saadane. Evaluation of the quality of images. *Winter School on Digital Color Image*, 2005.
11. D. Cherifi, A. Beghdadi, and A. Belbachir. **A Multiscale and multi-orientation contrast enhancement method for digital mammography.** *IEEE Proceeding. 3rd International Symposium on Communications, Control and Signal Processing. St Julians, Malta*, pp. 312-315, March 2008.
<https://doi.org/10.1109/ISCCSP.2008.4537241>
12. M.J. Nadenau, S. Winkler, D. Alleysson, and M. Kunt. **Human vision models for perceptually optimized image processing.** *A Review. Submitted to Proceedings of the IEEE*, 2000.
13. W. Abdul, P. Carre, H. Saadane, and P. Gaborit. **Watermarking using multiple visual channels for perceptual color spaces,** *2010 IEEE International Conference on Image Processing. Hong Kong, China*, September 2010.
<https://doi.org/10.1109/ICIP.2010.5651627>
14. E. Simonetto, and H. Saadane. **Watermarking of color images fixed by decomposition in perceptual channels.** *Conference CORESA. At Lille, France*, 2004.
15. F. Autrusseau. **Image tattoo based on the modeling of the human visual system and the mojette transformation.** *PhD thesis, specialty: Image Processing. Polytechnic School of the University of Nantes, France*, November 2002.
16. N. Morizet. **Biometric recognition by multimodal fusion of Face and Iris.** *PhD thesis, specialty: Signal and Images. National School of Telecommunications. Paris, France*, March 2009.
17. C. Sanderson and K.K. Paliwal. **Information fusion and person verification using speech and face information.** *IDIAP Research Report 02-33*, September 2002.
18. L. Longbin, Z. Xinman, X. Xuebin. **Fusion of face and visual speech information for identity verification.** *2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, November 2017
19. C. Sanderson and K.K. Paliwal. **Identity verification using speech and face information.** *Digital Signal Processing, Vol.14, no. 5*, pp. 449-480, 2004
<https://doi.org/10.1016/j.dsp.2004.05.001>

20. L. Zoubida, and R. Adjoudj. **Integrating face and the both irises for personal authentication.** *International Journal Intelligent Systems and Applications. Modern Education and Computer Science*, Vol. 3, pp. 8-17, 2017.
21. Q. Wang, B. Zhu, Y. Liu, L. Xie, and Y. Zheng. **Iris-Face fusion and security analysis based on fisher discriminant.** *International Journal on Smart Sensing and Intelligent Systems*, Vol. 8 no. 1, pp. 387-407, 2015.
<https://doi.org/10.21307/ijssis-2017-764>
22. T. Tieniu, H. Zhaofeng, S. Zhenan. **Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition.** *Image and Vision Computing*, Vol.28, no. 2, pp. 223-230, 2010.
23. S. Zhenan, T. Tieniu. **Ordinal measures for iris recognition.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, no. 12, pp. 2211-2226, 2009.
<https://doi.org/10.1109/TPAMI.2008.240>
24. N. Bekkat. **Objective criterion of subjective quality of monochrome images. Model design and experimental validation.** *PhD thesis, specialty: Signal Processing and Applied Informatics. Polytechnic School of the University of Nantes, France*, 1999.
25. A.Ninassi. **From the local perception of coding distortions to the global appreciation of the visual quality of images and videos. Contribution of the visual attention in the judgment of quality.** *PhD thesis, specialty: Signal Processing and Applied Informatics. Polytechnic School of the University of Nantes, France*, March 2009.
26. P.Callet, and D. Barba. **A robust quality metric for color image quality assessment.** *Proceedings 2003 International Conference on Image Processing*, 2003.
<https://doi.org/10.1117/12.503106>
27. H. Saadane, D. Barba, and H. Senane. **The estimation of visual bandwidth and their impact in image decomposition and coding.** *Proceedings of Visual Communications and Image Processing '93*, 1993.
<https://doi.org/10.1117/12.157910>
28. A.A. Michelson. **Studies in optics. Department of Physic. The University of Chicago. USA**, 1927.
29. T.N. Cornsweet. **Visual Perception. Academic Press. Introduction to Applications and Current Approaches to Visual Perception**, 1970.
30. T. Caelli. **Introduction to applications and current approaches to visual perception.** *Visual Perception*, 1981.
31. R. Gordon, and R.M. Rangayyan. **Feature enhancement of film mammograms using fixed and adaptive neighborhoods.** *Applied Optics* Vol. 23 no. 4, pp. 560-564, 1984.
<https://doi.org/10.1364/AO.23.000560>
32. A. Beghadadi, and A. Negrata. **Contrast enhancement technique based on local detection of edges.** *Computer Vision, Graphics and Image Processing*, Vol. 46, pp. 162-174, May 1989.
33. E. Peli. **Contrast in complex images.** *Journal of the Optical Society of America*, Vol. 7 no. 10, pp. 2032-2040, 1990.
34. N. Hautiere, D. Aubert, and M. Jourlin. **Measurement of local contrast in images, application to the measurement of distance visibility through use of an onboard camera.** *Signal Processing*, Vol. 23 no. 2, pp. 145-158, 2006.
35. N. Hautiere, and D. Aubert. **Contrast restoration of foggy images through use of an onboard camera.** *Proceedings. 2005 IEEE Intelligent Transportation Systems*, 2005.
36. N.H. Barnouti, S.S.M. Al-Dabbagh, W.E. Matti, and M.A.S. Naser. **Face detection and recognition using Viola-Jones with PCA-LDA and square euclidean distance.** *International Journal of Advanced Computer Science and Applications*, Vol. 7 no. 5, pp. 371-377, 2016.
<https://doi.org/10.14569/IJACSA.2016.070550>
37. N.T. Deshpande, and S.Ravishankar. **Face detection and recognition using Viola-Jones algorithm and fusion of PCA and ANN.** *Advances in Computational Sciences and Technology*, Vol. 10 no. 5, pp. 1173-1189, 2017.
38. A.Hajdu, L.Hajdu, A. Jónás, L.Kovács, and H. Tomán, H. **Generalizing the Majority Voting Scheme to Spatially Constrained Voting.** *IEEE Transactions on Image Processing*, Vol. 22 no. 11, pp. 4182-4194, June 2013.
<https://doi.org/10.1109/TIP.2013.2271116>
39. L. Lam, and C. Suen. **Application of majority voting to pattern recognition: on analysis of its behavior and performance.** *IEEE Transactions on Systems, Man and Cybernetics, Part A*, Vol. 27 no. 5, pp.553-568, 1997.
<https://doi.org/10.1109/3468.618255>