# State of the art on human visual system modeling

Fatima Zohra Allan, Hicham Bousbia-Salah and Latifa Hamami

Signal and Communication Laboratory, Department of Electronics, National Polytechnic School, 10, Avenue des Frères Oudek, El Harrach, 16200 Alger, Algeria

E-mail: fatima_zohra.allam@g.enp.edu.dz, hicham.bousbia-salah@g.enp.edu.dz, latifa.hamami@g.enp.edu.dz

**Abstract**

Perceptual processes are becoming more and more complex as new observations and research in the field of Digital Image Processing are published. The discoveries of the last decades have notably modified the image that we traditionally had of the Human Visual System and visual perception. The main objective of this work is to draw up an inventory of the main functions of the Human Visual System and its components. We will highlight the interest of using models of the Human Visual System to develop image processing tools, more particularly, in the field of individual recognition. Indeed, it is recognized that the HVS can overcome a certain number of difficulties commonly encountered in computer vision. We will also develop aspects relating to functions and processes associated more immediately with the Human Visual System. We will highlight the most important properties, namely, Local Band-Limited Contrast, Contrast Sensitivity Function and masking effect.

metadata91F2003D10, 68T45visual perception, HVS modeling, limited band local contrast, contrast sensitivity function, masking effect

## 1 Introduction

Human vision remains a great mystery to researchers in the field of neuroscience. Indeed, of our five senses, it is vision that uses the most neurons in our brain. Human vision remains a great mystery to researchers in the field of neuroscience. Indeed, of our five senses, it is vision that uses the most neurons in our brain.

In recent years, a new research has been undertaken to study the behavior of the Human Visual System. Knowledge of this is particularly interesting in image processing insofar as most applications are intended to provide images viewed and used by humans. This line of research remains very broad and diverse. It opens so many doors to explore. It has been widely developed in the literature. This has allowed the recent appearance of new information and communication technologies. This discipline has different categories, the most inclined are those of compression, coding [1] [2], quality evaluation [3] [4] [5], pattern recognition, cryptography and many more.

The objectives are multiple and aim, firstly, to explain how the brain produces elaborate visual reproductions of the world, based on scarce visual information and, secondly, to create mathematical models to meet the requirements of certain very specific applications.

The visual system of the brain receives very little information despite the multitude and quantity of information contained in the images perceived by humans.

Modeling any biological system requires knowledge of the biological stages of information processing [6]. We are interested, more particularly, in the modeling of the HVS. The HVS appears as a multi-sensor system, gradually integrating spatial information, color, depth, spatial frequencies and movement of our environment.

**Advanced Studies: Euro-Tbilisi Mathematical Journal** 16, supplement issue 2 (2023), pp. 119–133.
DOI: 10.32513/asetmj/1932200823210

Tbilisi Centre for Mathematical Sciences.

*Received by the editors:* 15 November 2022.
*Accepted for publication:* 15 January 2023.

The mechanisms and functioning of the HVS have already been the subject of detailed studies in numerous theses, notably [3] [6] [7], ...

In this article, we briefly present some basic notions of visual perception. Subsequently, we will present the main characteristics and functionalities of the HVS with an emphasis on the calculation of the local contrast, the CSF, the multi-channel decomposition and the masking effect. Some of these properties have caught our attention in the context of our modeling work on HVS applied to the biometric recognition of individuals [8].

## 2 Visual perception

Multitudes of questions arise: What is visual perception? How does it work? What is the process of perception? What are the mechanisms of perception?

From a more general perspective, the role of visual perception often goes beyond the simple function of gathering information. Visual perception contributes significantly to conceptualization [9] [10].

Visual sensitivity is considered to be one of the most developed sensory modalities in humans. The human eye is complex and is made up of many organs involved in human visual perception, from the retina to the sensorimotor system [11].

Among the various ways of studying visual perception, researchers have adopted, for the past forty years, a concept known as "information processing". According to this approach, perceptual mechanisms are a set of operations performed by the brain on the signals that our sensory receptors pick up in the environment.

The information formed by these biological signals will be used in two different ways. On the one hand, they are used more or less automatically in the regulation of motor behavior. However, some of these automatic behaviors also involve knowledge. The information received can, on the other hand, be interpreted in terms of objects and events of the outside world, based on the knowledge, representations that we previously have in memory [10].

## 3 Modeling of the HVS

The modeling of the Human Visual System stems mainly from its biological and functional structure, as well as from psychophysical experiments [6]. We are going to present the properties of the human visual system that we think are the most important. We can find in [12], an exhaustive description of the anatomical aspects of the HVS and the visual models of the literature.

In this section, we will present respectively the perception of luminance, the computation of the contrast, the function of sensitivity to the contrast.

### 3.1 Perception of luminance

Understanding the process of visual perception is an element to consider when modeling the HVS. The perception of an area of the image generates three types of sensations. The sensations of hue and saturation are linked to the chromaticity of the observed area while the sensation of luminosity reflects the perceived luminance [6].

The Human Visual System, being naturally confronted with a significant dynamic of light intensity, adaptation mechanisms have been put in place allowing it to maintain its sensitivity both in conditions of significant illumination (photopic conditions), as well as in low illumination conditions (scotopic conditions) [7].

Figure 1 illustrates the spectral sensitivity of the human eye:

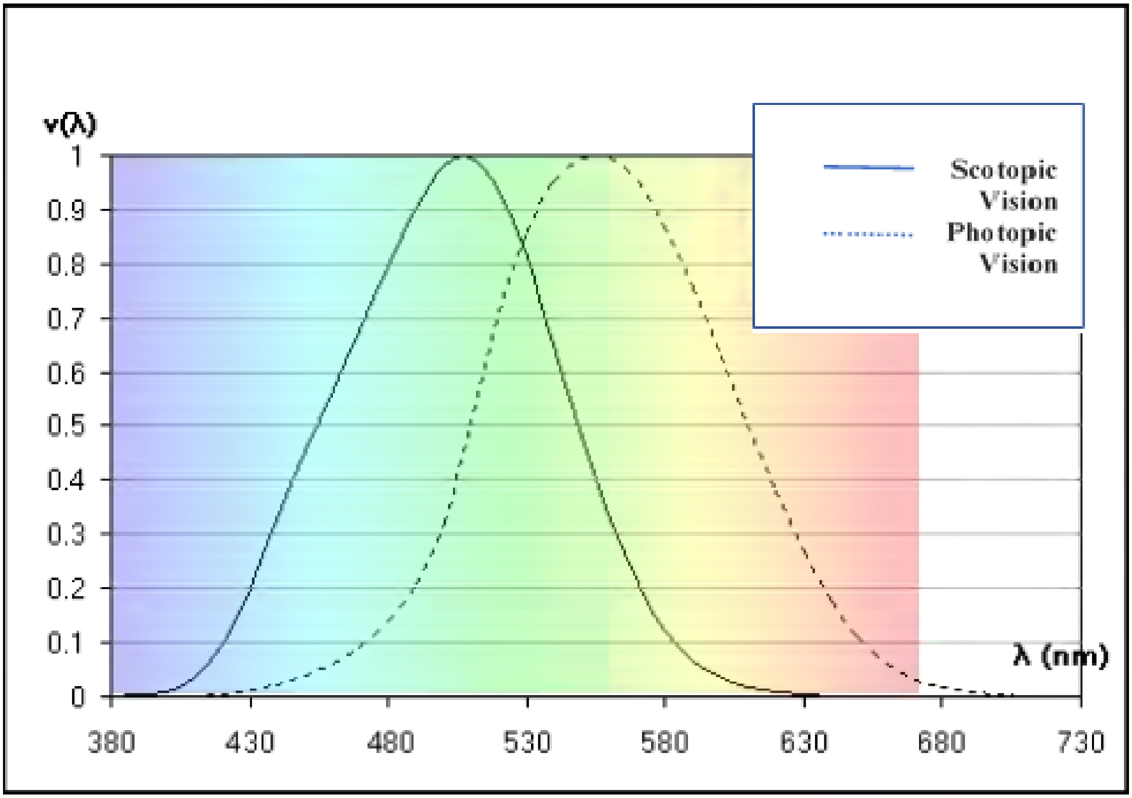


FIGURE 1. Relative Spectral Sensitivity of the human eye in photopic and scotopic vision [13]

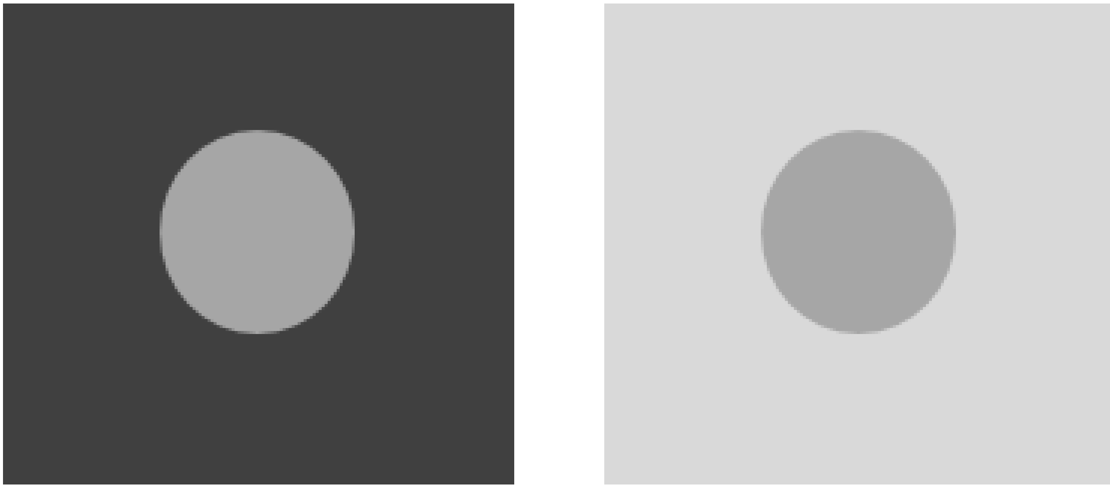

FIGURE 2. Brightness contrast

### 3.2 Contrast sensitivity

The response of the Human Visual System depends more on local variations in luminance ($\Delta L$) with respect to the surrounding luminance ($L$), than on absolute values of luminance. This property is known under the Weber-Fechner law. This is one of the first quantitative models of the perception of luminance.

In other words, the contrast is defined as a relative measure of variation in the luminance of a region (target luminance) relative to the luminance of surrounding regions (background luminance).

Indeed, a medallion, of average luminance, presented on a dark background appears lighter than this same medallion presented on a light background (Figure 2).

The relative change in luminance is measured through what is called contrast. Mathematically, Weber's contrast is modeled by the relation:

$$\frac{\Delta L}{L} = C_W \tag{3.1}$$

This law indicates that if on a uniform background of luminance $L$, called adaptation, a medallion-type stimulus of luminance $\Delta L+L$ is superimposed (Figure 2), the ratio $\Delta L/L$ is practically constant in a wide range of luminosity.

Figure 3 illustrates Weber's experiment

In the range of medium and high luminance, Sakmann and Creutzfeld, in [14], confirmed and adopted Weber's formulation. On the other hand, in low luminances, Moon and Spencer, in [2] and [15], detected a slight inconsistency, and therefore they made a correction giving rise to the following expression:

$$\frac{\Delta L}{L} = (\frac{C_W}{L})(0.456 + \sqrt{L})^2 \tag{3.2}$$

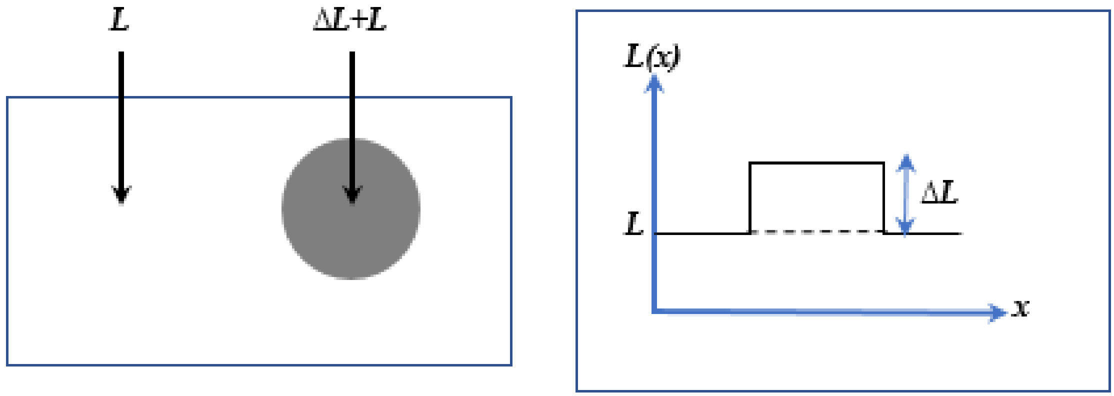


Figure 3. Weber's experiment

At the detection threshold, and when the luminance increases, the ratio $\Delta L \;/\; L$ tends towards the Weber constant $C_W$. This constant depends on the geometry and size of the stimulus [6]. The threshold contrast, also called the detection threshold or visibility threshold, is the minimum value of contrast necessary for a change in light intensity to be detected.

Often, models are built from experimental results on the detection of sinusoidal signals that use Michelson's definition of contrast.

Contrast is therefore defined as the ratio between the local intensity of an image and its average intensity:

$$C = \frac{L_{\max} - L_{\min}}{2L_{moy}} \tag{3.3}$$

Where: $L_{\max}$ and $L_{\min}$ correspond to the maximum and minimum luminance values respectively.

It should be noted that there is a large number of definitions of contrast, adapted to more or less complex stimuli. In the case of sinusoidal stimuli, Michelson's formula (3) is generally adopted.

The Local Band-Limited Contrast proposed by E. Peli [16] is much more complex and reflects the very important fact that the perception of a detail of the image also depends on its local environment.

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 given sub-band and the local mean luminance relating to this channel, i.e. for the considered site *(m, n)*, the sum of the luminance of all the radial sub-bands lower than the sub-band:

$$C_i(m,n) = \frac{L_i(m,n)}{\sum_{k=0}^{i-1} L_k(m,n)} \tag{3.4}$$

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.

### 3.3 Contrast sensitivity functions (CSF)

In order to determine the contrast sensitivity of the HVS, several studies have been carried out. One of the first experiments on frequency behavior is that of Campbell and Robson [17].

One of the most important features of creating visual models is the decrease in sensitivity for high spatial frequencies. This phenomenon is modeled by the contrast sensitivity function (CSF).

Several achromatic CSF are detailed by Barten in [18]. The CSF curve (Figure 4) is usually assimilated to the transfer function of a linear spatial filter at low and high frequencies. Visual sensitivity is maximum for medium spatial frequencies; it decreases for higher frequencies. This filtering comes from the low-pass behavior of the lens and the pupil, as well as the limited number of retinal cells per unit length.

The CSF therefore describes the evolution of visual sensitivity, i.e. the inverse of the contrast of a signal at its visibility threshold, most often as a function of spatial frequencies and the orientation of the signal to detect.

Most of the time, the models proposed in the literature are developed, most of the time, from experimental results on the detection of sinusoidal signals using Michelson's definition for contrast. The fact that the real signals are composed of multiple frequencies implies variations of the visibility thresholds, these variations reflect what is called the masking effect. This phenomenon will be explained later.

## 4 Multi-channel organization

Several physiological evidences reveal that the cells of the visual system are tuned to certain types of visual information such as color, orientation or frequency.

The results of psychophysical experiments suggest the existence of groupings of information prior to its processing [20], which consolidates the hypothesis of the spatio-frequency decomposition of HVS into visual or perceptual channels.

Sakrison remarks that if we present a stimulus containing several frequency components, only the fundamental component will set the visibility threshold [21]. This decomposition leads to the definition of several filters which characterize the sub-bands or the perceptual channels [22]. The characteristics of the decomposition are generally described in terms of radial and angular selectivities.

Watson and Daly have detailed the necessary conditions and properties of linear transformations used to model the natural spatio-frequency selectivity of the Human Visual System [23] [24].

In the literature, there are many ways to implement the decomposition into spatial frequency channels aimed at approaching the behavior of the Human Visual System [25].

Regarding the design of our biometric recognition system [8], we opted for the use of perceptual channel decomposition of the Human Visual System. This modeling, widely described, among others, in [3] [6] [26], is given in figure 5 [7]. It consists of four radial frequency domains, called crowns, 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° [20].

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

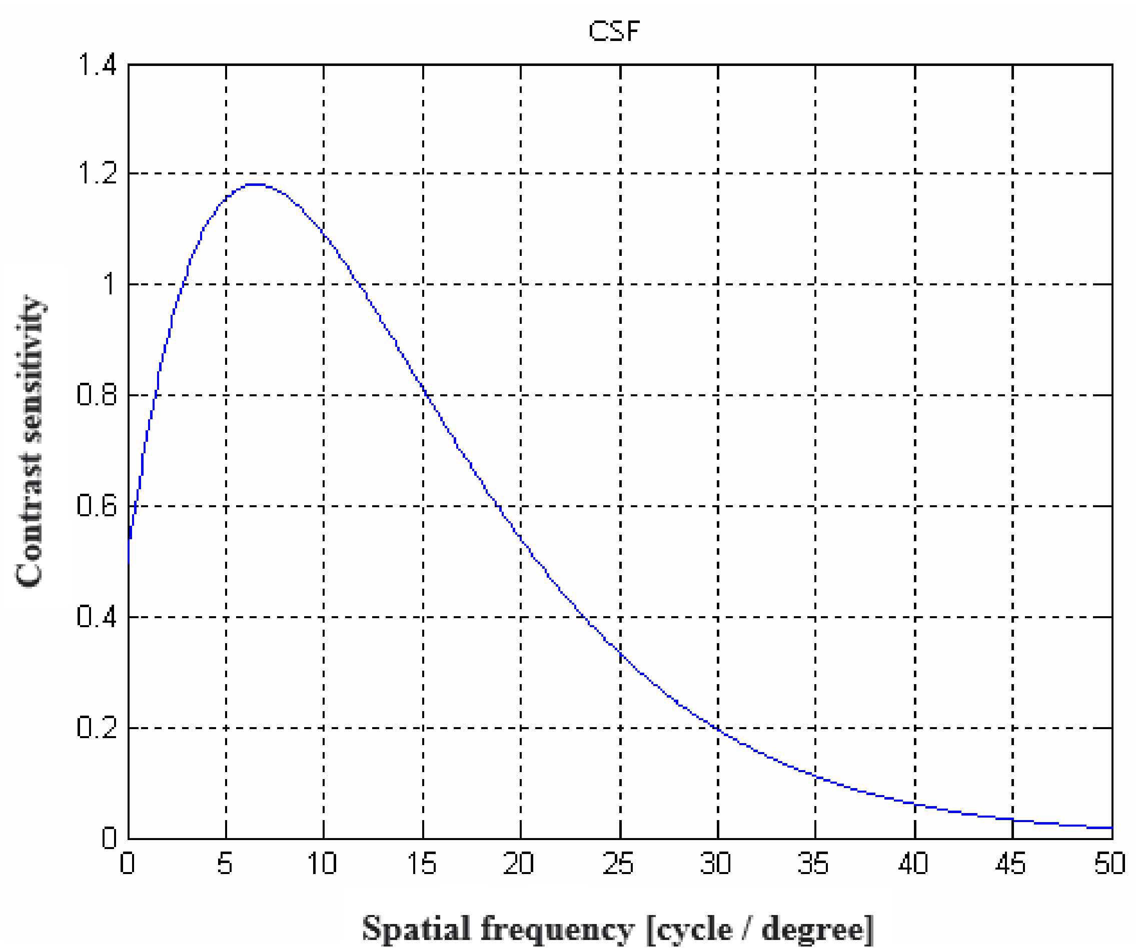


FIGURE 4. Curve of the Contrast Sensitivity Function according to the Mannos model [19]

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 [6]. 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 [26].

The representation space is therefore decomposed into seventeen visual channels distributed as follows [7]:

- A mono-directional BF channel (zone I) 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 6 illustrates part of the architecture of our biometric recognition system [8].

### 4.1 Spatial decomposition of information

The spatial decomposition of the HVS information into different channels occurs according to radial selectivity (from 1 to 2 octaves), and angular selectivity (from 20 to 60 degrees) [6].

There are several decompositions. We can cite as examples:

- Watson's Cortex Transform [23].
- The Two-Dimensional Transformations of Gabor [27].
- Pyramidal Transformations [28].
- Transforms into Classic 2D Wavelets [29].

## 5 Time decomposition of information

In the literature, the temporal decomposition of information is not as clear as the spatial decomposition. However, what follows is the existence of two mechanisms, one low pass and the other band pass [23] [30]. Sustained and transient channels are mentioned respectively.

## 6 Masking function

The main problem for modeling masking effects lies in the complexity of the phenomena studied which results in a multitude of experimental conditions. Classical experiments measure the probability of detection of a sinusoidal signal in the presence of a signal also masking sinusoidal [20].

These experiments relate to masking:

- between signals of different orientations [31].
- between signals of different spatial frequencies [32].
- between chromatic and achromatic signals [33].

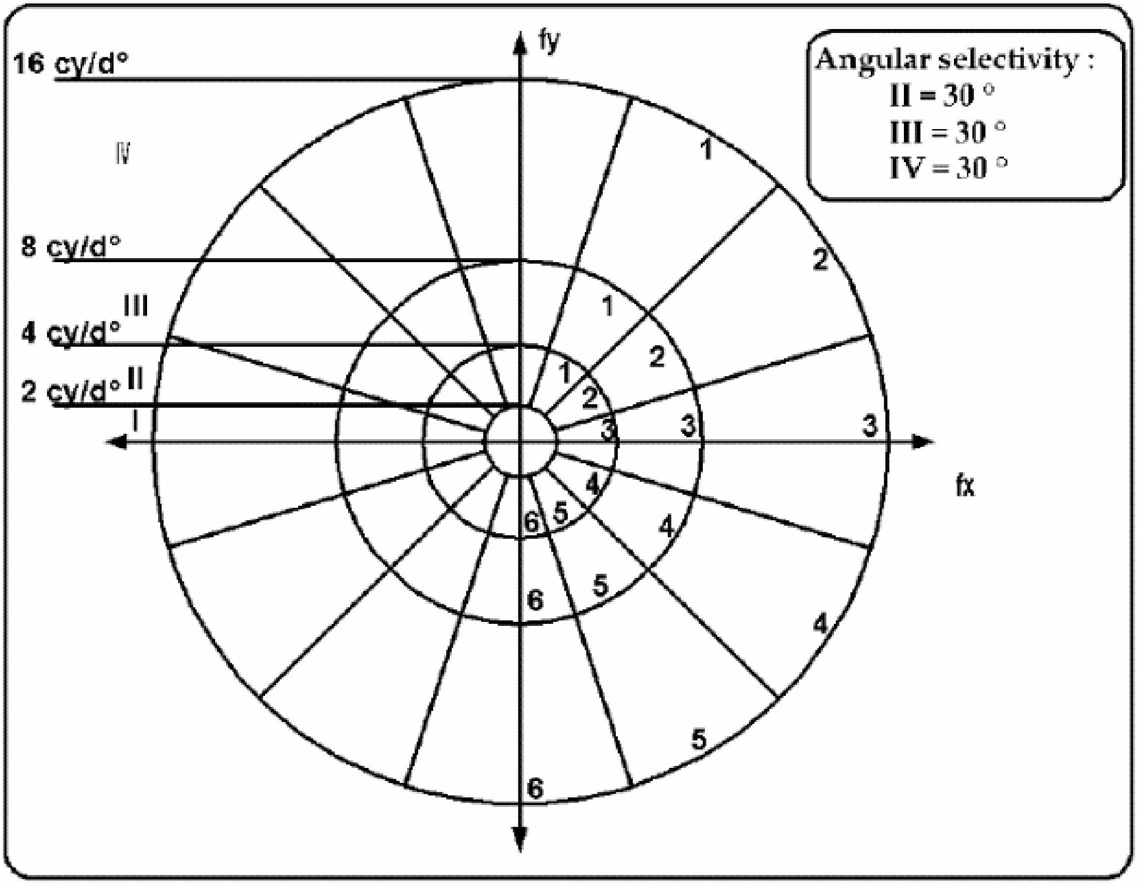


Figure 5. Decomposition of the representation Space into Perceptual Channels [7]

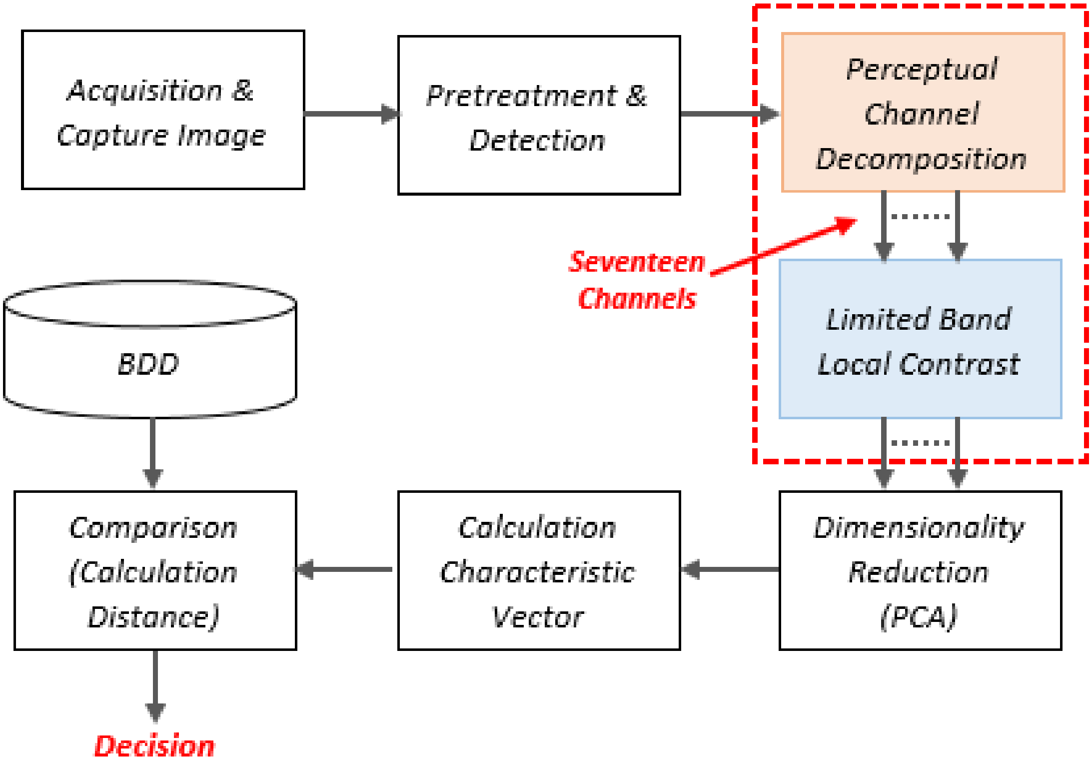


FIGURE 6. Block of the architecture of biometric recognition

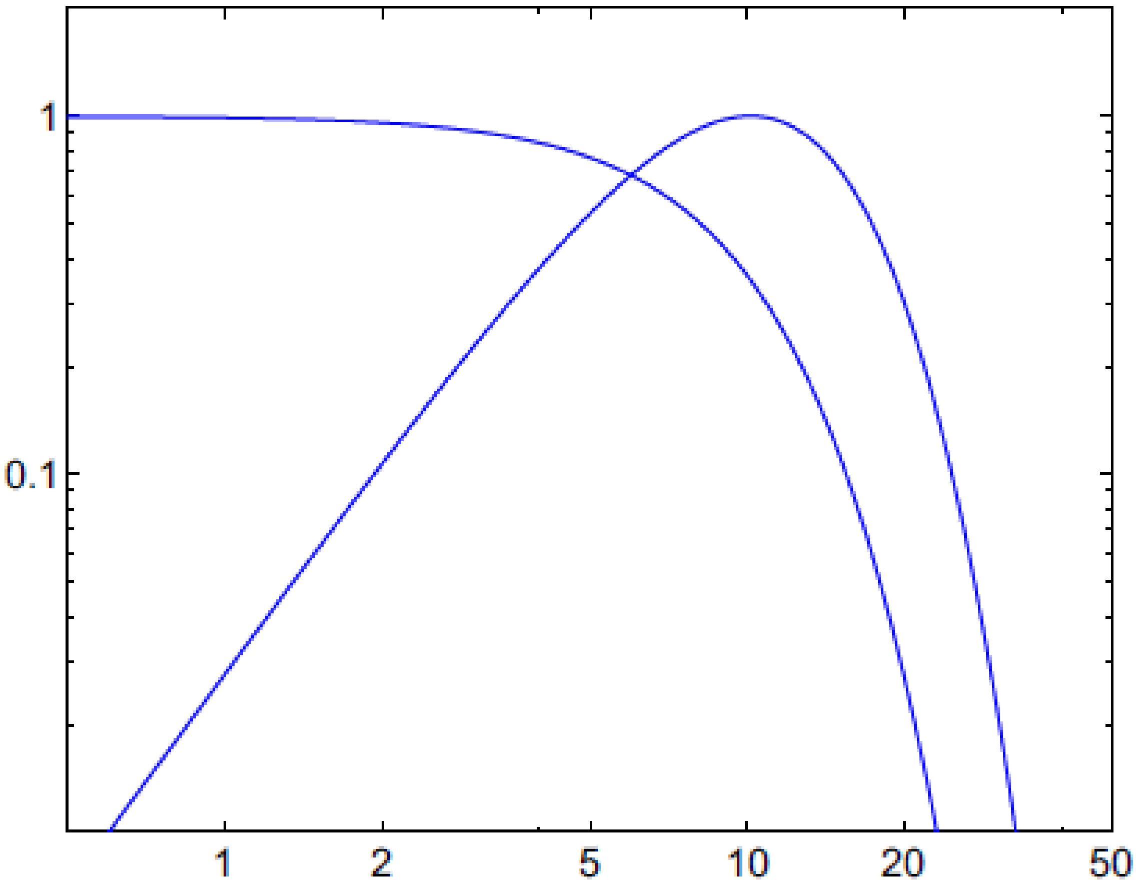


FIGURE 7. Frequency response of the sustained and transient visual channels of the time decomposition of information [30]

Recently, contrast gain control models have seen notable success because they also predict visibility thresholds of single signals well. Initially developed by Teo and Heeger [34], these models are constantly refined in order to explain the multiple interactions between visual channels or the components of the color space [35] [36].

### 6.1 Spatial masking

Signals with similar characteristics are processed by the same visual channels and therefore follow the same path from the eye to the cortex. There are interactions with nonlinear effects between such neighboring signals.

Masking, or masking effect, is one such effect. It reflects the variation of the differential visibility threshold of a stimulus due to the presence of another signal in its vicinity, qualified as a masking signal, having a higher level [6].

The masking effect is all the more important as the two signals have similar characteristics. This is referred to as a masking effect both in the case of the increase in the differential visibility threshold, and in the case of the decrease in the threshold value. In the first case, it is about masking in the proper sense (masking effect), while in the second case it is about what is called facilitation (pedestal effect) where a signal will increase the visibility of another [6].

The masking effect has been the subject of numerous studies because of its importance in the various axes of image processing. Different models have been proposed in the literature [24] [31] [32] [34] [37] [38].

In a space of visual representation of information with several channels and components, it is possible to dissociate the different origins of masking effects [6]:

- The most important is intra-channel masking which explains the relationships between signals processed by the same channel. In other words, it results in an interaction between stimuli and masking signal having similar characteristics (frequency, orientation, component).
- Inter-channel masking which characterizes the influences between stimuli and signal masking different characteristics, that is to say not belonging to the same channel, but belonging to the same component.
- Inter-component masking which reflects the interactions between signals carried by different components.

### 6.2 Temporal masking

As in the case of spatial masking, temporal masking reflects a change in the visibility threshold of a signal due to the presence of another signal. This change in the visibility threshold is due to the interaction of temporally adjacent stimuli. These masking effects are less well known than those encountered in the space field.

Temporal masking is concerned with how two temporally close stimuli interact. The answer is complex and depends on many factors such as:

- the spatial structure of the masking signal,
- the similarity between the spatial characteristics between the target and the masking signal,
- the time interval between the target and the masking signal,

- the difference in luminance between the target and the masking signal,
- etc.

In studies on temporal masking, we are interested in the masking effects due to strong temporal discontinuities, such as changes of plan or rapid transitions (dark-light, light-dark) [12] [39].

## 7 Conclusion

Recognition methods have been widely studied for several years with the aim of achieving performance close to those observed by humans. The fields of application are very diverse. They are found in digital image compression, pattern recognition, ...

Studies on psycho-visual modeling are likely to bring innovative elements to the field of recognition.

The purpose of this work was to present the HVS and more particularly the existing models of some of its properties. We have covered some important concepts on visual perception. We have also made a summary of the main characteristics of the HVS.

The properties of the HVS that interest us are those of interest in a recognition context. We have chosen to focus on the Local Band-Limited Contrast and the Contrast Sensitivity Function because of their preponderant role in perception.

# References


[1] L. Bedat, *Aspects psychovisuels de la perception des couleurs. Application au codage d'images couleurs fixes avec compression de l'information*, PhD Thesis, University of Nantes, France, 1998.

[2] H. Sénane, *Représentation d'images en sous-bandes visuelles. Application au codage d'images de télévision sans défauts visibles*, PhD Thesis, University of Nantes, France, 1996.

[3] N. Bekkat, *Objective criterion of subjective quality of monochrome images. Model design and experimental validation*, Specialty: Signal Processing and Applied Informatics, PhD Thesis, Polytechnic School of the University of Nantes, France, 1999.

[4] P. Le Callet, *Critères objectifs avec référence de qualité visuelle des images couleur*, PhD Thesis, Polytechnic School of the University of Nantes, France, 2001.

[5] M.S. Kankanhalli, Rajmohan, and K.R. Ramakrishnan, *Content based watermarking of images*, ACM Multimedia - Electronic Proceedings (1998).

[6] 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*, Specialty: Signal Processing and Applied Informatics, PhD Thesis, Polytechnic School of the University of Nantes, France, March 2009.

[7] F. Autrusseau, *Image tattoo based on the modeling of the human visual system and the mojette transformation*, Specialty: Image Processing, PhD Thesis, Polytechnic School of the University of Nantes, France, November 2002.

[8] F.Z. Allam, H. Bousbia-Salah, and L. Hamami, *Modeling of Biometric Recognition Based on Human Visual System*, International Journal of Advanced Trends in Computer Science and Engineering **9** (2020), no. 1-2, 198-204.

[9] J. Lajoie and A. Delorme, *Les Fonctions et les Processus Visuels – Perception et Réalité – Une introduction à la Psychologie des Perceptions. Chapitre 4*, Neurosciences & Cognition, De Boeck Supérieur, 2003.

[10] A. Benoît, *Le Système Visuel Humain au secours de la vision par ordinateur*, Doctoral dissertation, Grenoble Images, Parole, Signal Automatique Laboratoire (GIPSA-Lab), Institut National Polytechnique de Grenoble, France, 2007.

[11] C. Massot, *Perception 3D dans le Cortex Visual*, Doctoral dissertation, Université Joseph Fourier, Grenoble, France, 2006.

[12] Marcus J. Nadenau, Stefan Winkler, David Alleysson, and Murat Kunt, *Human vision models for Perceptually Optimized Image Processing - A Review*, Submitted to Proceedings of the IEEE (2000).

[13] D. Lagadic-Gossmann, *The Human Visual System*, Available from: `http://people.rennes.inria.fr/Eric.Marchand/ESIR/ESIR2/BINP/svh.pdf`, 2017. University of Rennes I.

[14] B. Sakmann and O.D. Creutzfeldt, *Scotopic and Mesopic light adaptation in the cat's retina*, Pügers archives **313** (1969), 168-185.

[15] P. Moon and D.E. Spencer, *Visual data applied to lighting design*, Journal of the Optical Society of America **34** (1944, October), 605-617.

[16] E. Peli, *Contrast in complex images*, Journal of the Optical Society of America **7** (1990), no. 10, 2032-2040.

[17] F.W. Campbell and J.G. Robson, *Application of Fourier analysis to the visibility of gratings*, Journal of Physiology **197** (1968), 551-566.

[18] P.G.J. Barten, *Contrast sensitivity of the human eye and its effects on image quality*, SPIE (1999).

[19] J. Mannos and D. Sakrison, *The effects of a visual fidelity criterion on the encoding of images*, IEEE Transactions on Information Theory **IT-20** (1974), 525-536.

[20] P. Le Callet and D. Barba, *Color perception model. Application to image quality assessment*, Signal Processing: Image Communication **21** (2004), 61-77.

[21] D.J. Sakrison, *On the role of the observer and a distortion measure in image transmission*, IEEE Transactions on Communications **COM-25** (1977), no. 11, 1251-1267.

[22] J.G. Daugman, *Spatial visual channels in the Fourier plane*, Vision Research **24** (1984), no. 9, 891-910.

[23] A. B. Watson, *The Cortex Transform: Rapid Computation of Simulated Neural Images*, Computer Vision, Graphics, and Image Processing **39** (1987), 311–327.

[24] S. Daly, *A Visual Model for Optimizing the Design of Image Processing Algorithms*, IEEE Int. Conf. on Image Processing, 1994, pp. 16–20.

[25] A. Sénane, A. Saadane, and D. Barba, *The Computation of Visual Components in a Sub-band Decomposition of Images*, International Conference on Signal Processing Applications and Technology, October 1993.

[26] H. Saadane, D. Barba, and H. S'enane, *The Estimation of Visual Bandwidth and Their Impact in Image Decomposition and Coding*, Proceedings of Visual Communications and Image Processing '93, 1993.

[27] J. Jones and L. Palmer, *An Evaluation of the Two-Dimensional Gabor Filter Model of Simple Receptive Fields in Cat Striate Cortex*, Journal of Neurophysiology **58** (1987), no. 6, 1233–1258.

[28] P. J. Burt and E. H. Adelson, *The Laplacian Pyramid as a Compact Image Code*, Transactions on Communications **31** (1983), 532–540.

[29] H. S. Li, P. Fan, H. Y. Xia, S. Song, and X. He, *The Multi-level and Multi-dimensional Quantum Wavelet Packet Transforms*, Scientific Reports **8** (September 2018), no. Article number: 13884.

[30] R. E. Fredericksen and R. F. Hess, *Estimating Multiple Temporal Mechanisms in Human Vision*, Vision Research **38** (1998), no. 7, 1023–1040.

[31] J. M. Foley, *Human luminance pattern-vision mechanisms: Masking experiments require a new model*, Journal of the Optical Society of America **11** (1994), no. 6, 1710–1719.

[32] G. E. Legge and J. M. Foley, *Contrast masking in human vision*, Journal of the Optical Society of America **70** (1980), no. 12, 1458–1471.

[33] M. A. Losada and K. T. Mullen, *The spatial tuning of chromatic mechanisms identified by simultaneous masking*, Vision Research **34** (February 1994), no. 3, 331-341.

[34] P. C. Teo and D. J. Heeger, *Perceptual image distortion*, Proceedings of 1st International Conference on Image Processing, August 2002, pp. 127-141.

[35] A. B. Watson and J. A. Solomon, *Model of visual contrast gain control and pattern masking*, Journal of the Optical Society of America **14** (1997), no. 9, 2379–2391.

[36] N. Graham and A. Sutter, *Normalization: Contrast-gain control in simple (Fourier) and complex (non-Fourier) pathways of pattern vision*, Vision Research **40** (2000), no. 20, 2737–2761.

[37] D. J. Heeger, *Normalization of cells responses in cat striates cortex*, Visual Neuroscience **9** (1992), 181–198.

[38] P. Le Callet and D. Barba, *A robust quality metric for color image quality assessment*, Proceedings 2003 International Conference on Image Processing, 2003.

[39] W.J. Tam, L.B. Stelmach, L. Wang, D. Lauzon, and P. Gray, *Visual masking at video scene cuts*, Proceedings of SPIE, Human Vision, Visual Processing, and Digital Display VI, 1995, pp. 111-119.