

# **A HUMAN BINOCULAR VISION MODEL BASED ON INFORMATION THEORY FOR LUMINANCE AND CHROMATICITY AT ISOLUMINANCE SUPRATHRESHOLD CHANGES**

J. M. Medina<sup>1</sup>, J. R. Jiménez<sup>2</sup> and L. Jiménez del Barco<sup>2</sup>.

<sup>1</sup> University Miguel Hernandez, Applied Physics. Avda. del Ferrocarril s/n, E-03202, Elche, Spain (e-mail: jmanuel@umh.es).

<sup>2</sup> University of Granada. Department of Optics. Facultad de Ciencias, E-18071, Granada, Spain,

## **1. Introduction.**

The binocular-vision system is a fundamental characteristic throughout mammalian species. In humans, this system provides a unifying single perception or cyclopean vision. Multiple visual phenomenon reveal the importance and complexity of this system, providing important neurobiological and psychophysical data such as those found in the phenomenon of stereopsis, binocular summation or binocular rivalry (Julesz, 1971; Blake, Sloane & Fox, 1981; Reading, 1983; Blake & Logothetis, 2001). From these areas, a variety of growing disciplines seek to understand how the binocular system combines monocular inputs, notably through research on binocular cognitive models, the stereoscopic-correspondence problem, and their dynamics (Marr 1980; Blake et al., 1981; Westendorf & Blake, 1988; Hughes & Townsend, 1998; Grossberg & Grunewald, 2002). Although a variety of physiological studies have explained how simple and complex binocular cells make the monocular combination in the striate cortex, it remains unclear how binocular signals are taken into account in higher processing areas and their internal and external dynamics with other non-visual areas such as the motor cortex. In the present work, we study these topics at the temporal-processing level. We manually measured simple visual reaction times (VRT) for both binocular and monocular vision, for different luminance and isoluminance stimuli under suprathreshold conditions. Since Helmholtz proposed reaction times as a supplement to other psychophysiological indicators, this non-invasive temporal parameter has provided key information on the neural conduction-time dynamics at the large scale of integration (Luce, 1986; Meyer, Osman, Irwin & Yantis, 1998; Deiber, Ibañez, Sadato, & Hallet, 1998; Miniussi, Girelli & Marzi, 1998). In a simple reaction-time procedure, this involved the simplest process (there is only one possible stimulus and one possible response). Thus in visual perception, there is a noteworthy possibility to scale the magnitude of VRT with the intensity of the signal following Piéron's law (Piéron, 1914; Luce, 1986; Díaz, Jiménez del Barco, Jiménez & Hita, 2002). On the other hand, since binocular VRT are shorter than monocular (Blake et al., 1981; Westendorf & Blake, 1988; Jiménez, Medina, Jiménez del Barco & Díaz, 2002), binocular interactions are captured within this paradigm (a binocular summation effect). Starting from these results, we have developed a new binocular model based on the existing information theory of neural coding adapted to the study of reaction times (Norwich, Seburn & Axelrad, 1989; Norwich, 1993).

## 2. Method.

The stimuli were presented on a CRT colour monitor connected to a microcomputer. The chromaticity and luminance of the stimuli were controlled by periodic calibrations using a SpectraScan PR-704 Photoresearch spectroradiometer. To produce stimuli with specific CIE-1931 chromatic and luminance coordinates, we used a calibration method proposed by Jiménez del Barco, Díaz, Jiménez and Rubiño (1995). Observers were seated 70 cm from the monitor in a dark room, and head was stabilized with a chin rest. The experimental stimuli were circular 1.5° patches of uniform colour on a dark background and were observed foveally by two normal observers with the natural pupil under three observational conditions: right and left monocular (in this case one eye was covered with a black patch) and binocular vision.

On the one hand, we used 12 achromatic stimuli with chromaticity coordinates equal to those of the equal-energy illuminant and with luminance values between 3 and 27 candela per square meter ( $\text{cd/m}^2$ ) in increments of  $2\text{cd/m}^2$ . On the other hand, we generate chromatic variations at isoluminance along L&M and S cone axis according to the Boynton's (1986) two-stage vision-vision model. This approach is more informative since it takes into account the first stages in vision processing presented in most colour-vision models. In each case, 12 stimuli were chosen according to red-green and yellow-blue confusion lines (Boynton, 1986). For each block of conditions (achromatic, red-green or yellow-blue signals), an achromatic stimulus of  $15\text{cd/m}^2$  was chosen to produce suprathreshold variations. The chromaticity coordinates were also identical to the equal-energy illuminant. Stimuli were distributed symmetrically around this and all could be clearly discriminated. This procedure scales the magnitude of VRT according to the polarity and intensity of the signals, enabling the adjustment of the mean value of VRT by Piéron functions. It should be noted that at isoluminance conditions, a minimum-flicker photometric procedure at a frequency of 12.3Hz was used to match the luminance of each stimulus with the reference stimulus (Díaz et al., 2001). Under these particular conditions, the VRT was determined following the standard procedure (Nissen & Pokorny, 1977; Díaz et al., 2001; Jiménez et al., 2002).

Visual-reaction-time data were taken independently for each block of conditions. At the beginning of each experimental session, while observational conditions (binocular vision, right eye or left eye) were randomly varied and fixed for the duration of that session. Afterwards, the subject was allowed 3 min to adapt to darkness and 3 additional min to adapt to the achromatic reference stimulus. At this point, a tone followed by a 7-sec pause signalled the start of a trial. After a random delay (3-7 sec, uniform sampling distribution) to avoid anticipation, the reference stimulus was changed to the test stimulus with the change synchronized with the beginning of the video refresh cycle. The test stimulus that replaced the reference remained on until the subject responded by pressing the button on the mouse connected to the microcomputer to indicate that a stimulus change had been perceived. Immediately following a response, the test stimulus was replaced by the reference stimulus. Observers did not know which stimulus was the next in the sequence, and therefore their task consisted only of responding as soon as possible to a luminance variation. Each test stimulus was randomly presented a total of 8 times during a session. The computer clock was programmed to provide 1 ms timing.

The first four reaction times from each session were excluded from the analyses. In accordance with a "kill the twin" procedure, reaction times shorter than 110 ms were discarded as false alarms while those exceeding 1500 ms were also excluded as misses. These

cut-off values were similar to those used in other VRT studies (Blake et al., 1981; Westendorff y Blake, 1988; Díaz et al., 2001). We ran a number of sessions until no fewer than 80 VRTs had been collected for each test stimulus under each experimental condition. The VRT value for each change in stimulus was taken as the arithmetic mean of the distribution recorded for a given stimulus.

### 3. Results

#### 3.1 A reaction time model with memory based on information theory.

Norwich et al., (1989) presented a reaction time model based on a receptor sampling model without memory founded on the information theory adapted to perception (commonly called The Entropic Theory of Perception), to predict the empirical Piéron equations (Norwich et al., 1989; Norwich 1993). Later, Norwich y Wong, (1994) generalized this receptor model considering the effects of sampling when a stimulus intensity dependent local memory is activated. Nevertheless, depicted the equations proposed by Norwich et al. (1989) are correct, they was written in a special form that obstruct an examination of the underlying neural-coding dynamics. At the same time, the case that takes into account local memory effects have not been developed, so it would be important to investigate these aspects. Regarding the particular case of visual perception, if the environment could be represented by a visual stimulus with a fixed mean signal intensity  $I$ , and duration of time  $t$ , the model proposed by Norwich and Wong (1994) asserts that photoreceptors detect stimuli according to the following entropy function:

$$H(I,t) = \frac{I}{2} \cdot \ln \left( 1 + \frac{\beta \cdot I^p}{I - e^{-at}} \right); \beta > 0, p > 0, a > 0 \quad (1)$$

where  $H(I,t)$  represents the entropy function (the drop in uncertainty, or, the equivalent, the gain in information in natural units), where as  $\beta$ ,  $p$  and  $a$  are constants, the latter representing the retinal sampling frequency (Norwich & Wong, 1994). Expression (1) represents a generalization of the entropy function proposed by Norwich et al. (1989), which can predict the empirical well-established Piéron functions. For this, introducing the existence of a threshold condition, the subject can only react when  $\Delta H$  units of information are received (Norwich et al., 1989):

$$\Delta H = H(I, t_0) - H(I, t_{VRT}) \quad (2)$$

where  $t_0$  is the value of  $t$  for which  $H$  is maximum (it is assumed that the stimulus is generated at  $t=0$ ) and  $t_{VRT}$  signifies the time to transmit  $\Delta H$  natural units of information. The informational principle of reaction time reflects the traditional assumption concerning the existence of serial processing stages in a VRT procedure (Luce, 1986), and thus it is involved in Piéron functions. Therefore, it is desirable to take it into account to express the functions calculated in a suitable manner. From (1), it is possible to solve (2) for reaction time ( $t_{VRT}$ ). At the same time, if we take limits when  $t_{VRT}$  and  $I$  tends to infinity, we get the intensity threshold ( $I_0$ ) and the asymptotic or irreducible part of the VRT ( $VRT_0$ ), respectively:

$$\langle VRT \rangle = t_{VRT} \approx VRT_0 + \frac{I}{a} \cdot \ln \left( I + \left( \frac{I_0}{I} \right)^P \right), \forall (I \gg I_0) \quad (3)$$

where  $\langle VRT \rangle$  expresses the expected value associated with the reaction-time process. Equations (3) express a generalized version of the Piéron functions. Traditionally, the asymptotic term  $VRT_0$  has been associated with the residual process contained in the VRT process such as motor execution (Luce, 1986; Díaz et al., 2001). Nevertheless, this hypothesis is not admissible. If we consider the case when the intensity of the stimulus  $I$ , tends to infinity, transient temporal glare effects should appear, so that the visual system cannot operate as under normal conditions, and it is not possible to make a highly accurate motor response such as those involved in VRT experiments. Therefore, this term must include another unknown process. If we express  $VRT_0$  as a sum of terms we get to the following expression:

$$VRT_0 = t_0 + \frac{I}{a} \cdot \ln \left( I + \left( \beta \cdot I_0^P \right) \right) \quad (4)$$

Now, Equations (3) and (4) enable us to describe the empirical Piéron functions for simple-reaction-time experiments as a three-serial-stage coding communication system with memory. If we accept the principle of efficiency of information representation (Ziemer and Tranter, 1981; Atick, 1992), it follows that the term  $t_0$  represents the coding time associated with the photoreceptor-signals encoder to reduce redundancy. At the same time, if we admit the generic diagram associated with a standard communications system (Ziemer and Tranter, 1981; Carlson, 1986), the second term in (4) comes behind  $t_0$  and represents the time associated with the transmission of the minimal quantity of information processed (a signal noise ratio) within a continuous communication channel that follows the Shannon-Hartley law, i.e., the second term represents the threshold characteristic processing time or  $t_{CPU}$ , which it is controlled by the retinal sampling time or  $1/a$ . Similar arguments establish the third term in (3) as the time associated with the transmission of a fixed quantity of information processed that come from the second stage, i.e., the third term represents the motor-processing time or  $t_{PM}$  (in our case, the act of pressing a button with the index finger).

### 3.2. A binocular vision model.

The visual-motor model presented implies a break with traditional models when the motor component is considered constant or with a positive correlation respect to the stimulus intensity (Ulrich & Stapf, 1984; Luce, 1986). At the same time, our model takes into account the glare effects involved when the intensity of the stimulus becomes infinity (there is no motor processing) and thus in consonance with the experiment. On the other hand, according to the magnitude of VRT and the Piéron's law, it establishes multiple distributed neural networks that depend on the type of signal (achromatic and chromatic at isoluminance) and its polarity (light, dark, red, green, yellow or blue). Therefore, to discern the combination rules in binocular vision, we adjusted the VRT data to these functions, comparing the parameters obtained in each observational condition. For each type of signal-polarity induced, the result was described by a similar Piéron's law in both monocular and binocular vision. Comparing

the coding time and the signal to noise ratios found at each stage ( $\beta(I_0)^P$ ,  $(I_0/I)^P$ ), our results show the existence of a binocular distributed neural network in both visual and motor-processing levels:

$$t_{0B} = \frac{t_{0L}^{LR} \cdot t_{0R}^{LR}}{t_{0L}^{LR} + t_{0R}^{LR}}, \text{ (binocular combination rule)} \quad (5)$$

$$\left(\frac{S}{N}\right)_{B_{CPU}} = g_{CPU} \sqrt{\left(\frac{S}{N}\right)_{R_{CPU}} \left(\frac{S}{N}\right)_{L_{CPU}}}, \text{ (visual-processing stage)} \quad (6)$$

$$\left(\frac{S}{N}\right)_{B_{PM}} = \frac{g_{PM} \cdot g_{CPU}}{\sqrt[4]{\left(\frac{S}{N}\right)_{R_{CPU}} \left(\frac{S}{N}\right)_{L_{CPU}}}}, \text{ (motor-processing stage)} \quad (7)$$

where  $t_{0i}$ , represents the photoreceptor-coding time for the  $i$ th observational condition ( $i \in B, R, L$ ; superscripts  $R$  and  $L$  denote the right and left monocular coding time when both monocular channels are activated in binocular vision),  $(S/N)_{ij}$  the signal-to-noise ratio for the  $i$ th observational condition at the  $j$ th stage and  $g_{CPU}$ ,  $g_{PM}$ , the visual and motor gain factors. The visual coding phase generate a multiplicative combination rule normalized to the sum of the monocular inputs (Expression (5)). This form of time-division multiplexing is maintained in the successive stages, which, according to the communication systems theory, could produce intermodulation distortion or crosstalk effects. The signal will be demultiplexed, and attenuated (the roots in (6) and (7)) being processing (compressed  $g_i < 1$  or expanded  $g_i > 1$ , depending on the type of signal-polarity induced), to become be multiplexed again in order to transmit to the next stage. This form of processing justified the decision centres subsequent to  $t_0$  involved in  $t_{CPU}$  and  $t_{PM}$  and is compatible with the temporal asynchronies found in binocular summation tasks (Blake et al., 1981; Reading, 1983; Cogan, Clarke, Chan and Rossi (1990). On the other hand, our results show that these companding systems are coupled, in the sense that a compressed signal in the visual stage implies an expanded factor in the motor-processing stage and vice versa. Finally, the form of the signal-to-noise ratios found at the motor-processing level suggest a form of quasi-synchronous multiplexing system, when monocular inputs are accommodated in the motor memory to produce a motor execution response with little variability.

#### 4. References.

- Atick, J. J. (1992). Could information theory provide an ecological theory of sensory precessing?. *Network*, 3, 213-251
- Blake, R. y Logothetis, N. (2001). Visual competition. *Nature Reviews of Neuroscience*, 3, 1-11.
- Blake, R., Sloane, M. y Fox, R. (1981). Further developments in binocular summation. *Perception & Psychophysics*, 30, 266-276.
- Boynton, R. M. (1986). A system of photometry and colorimetry based on cone excitations. *Color Research and Application*, 11, 244-252.

- Carlson, A. B. (1986). *Communication Systems*. McGraw-Hill Book Company, New York.
- Cogan, A. I., Clarke, M. Chan, H. y Rossi, A. (1990). Two-pulse monocular and binocular interactions at the differential luminance threshold. *Vision Research*, 30, 1617-1630.
- Deiber, M. P., Ibañez, V., Sadato, N. y Hallet, M. (1998). Cerebral structures participating in motor preparation in humans: a positron emission tomography study. *Journal of Neurophysiology*, 75, 233-247.
- Díaz, J. A., Jiménez del Barco, L., Jiménez, J. R. e Hita, E. (2001). Simple reaction time to chromatic changes along L&M-constant and S-constant cone axes. *Color Research and Application*, 26, 223-233.
- Grossberg, S., Grunewald, A. (2002). Temporal dynamics of binocular disparity processing with corticogeniculate interactions. *Neural Networks*, 15, 181-200.
- Hughes, H. C. y Townsend, J. T. (1998). Varieties of binocular interaction in human vision. *Psychological Science*, 9, 53-60.
- Jiménez, J. R. Medina, J. M. Jiménez del Barco, L. & Díaz, J. A. (2002). Binocular summation of chromatic changes as measured by visual reaction time. *Perception & Psychophysics*, 64, 140-147.
- Jiménez del Barco, L., Díaz, J.A., Jiménez, J. R. y Rubiño, M. (1995). Considerations on the calibration of color displays assuming constant-channel chromaticity. *Color Research and Application*, 20, 377-387.
- Julesz, B. (1971). *Foundations of cyclopean perception*. University of Chicago Press, Chicago.
- Luce, R. D. (1986). *Response times*. Oxford University Press, New York.
- Marr, D. (1980). *Vision*. W. H. Freeman and Company, San Francisco.
- Meyer, D. E., Osman, A. M. & Irwin, D. E. (1988). Modern mental chronometry. *Biological Psychology*, 26, 3-67.
- Miniussi, C. Girelli, M. y Marzi, C. A. (1998). Neural site of the redundant target effect: electrophysiological evidence. *Journal of Cognitive Neuroscience*, 10, 216-230.
- Nissen, M. J. & Pokorny, J. (1977). Wavelength effects on simple reaction time. *Perception & Psychophysics*, 457-462.
- Norwich, K. H. (1993). *Information, Sensation and Perception*. Academic Press, San Diego.
- Norwich, K. H., Seburn, C. N. L. y Axelrad, E. (1989). An informational approach to reaction times. *Bulletin of Mathematical Biology*, 51, 347-358.
- Norwich, K. H. y Wong, W. (1994). A universal model of single-unit sensory receptor action. *Mathematical biosciences*, 125, 83-108.
- Piéron (1914). Recherche sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices. *L'Année Psychologique*, 20, 17-96.
- Reading, R. W. (1983). *Binocular vision: foundations and applications*. Butterworth Publishers, London.
- Ulrich, R. y Stapf, K., H. (1984). A double response paradigm to study stimulus intensity effects upon the motor system in simple reaction time experiments. *Perception & Psychophysics*, 36, 545-558.
- Westendorf, D. y Blake, R. (1988). Binocular reaction time data to contrast increments. *Vision Research*, 28, 355-359.
- Ziemer, R. E. y Tranter, W. H. (1981). *Principles of communications: Systems, Modulation and Noise*. Houghton Mifflin, Boston.