Functional Anatomy of Stereopsis: Effective Connectivity identified using NARMAX and fMRI data

Hector-Gabriel Acosta-Mesa[1], John Mayhew[2], John Frisby[2], David Buckley[2], Ying Zheng[3], and Iain Wilkinson[4]

[1] Department of Artificial Intelligence, University of Veracruz,
Sebastián Camacho 5, CP 91090, Xalapa, Ver., México.
heacosta@uv.mx

[2] Artificial Intelligence Vision Research Unit, University of Sheffield,
Western Bank, S10 2TP, Sheffield, United Kingdom.
{j.e.mayhew, j.p.frisby, d.buckley}@sheffield.ac.uk

[3] Academic Department of Psychiatry, SCANLab, University of Sheffield,
The Longley Centre, S5 7JT, Sheffield, United Kingdom.
ying.zheng@sheffield.ac.uk

[4] Section of Academic Radiology, University of Sheffield,
Glossop Road, S10 2JF, Sheffield, United Kingdom.
i.d.wilkinson@sheffiel.ac.uk

Abstract. Functional magnetic resonance imaging was used to investigate the relationship between stereo and motion visual processing. Red/green random dot anaglyph stereograms with radial motion were used as visual stimuli. Three main areas of cortical activations were identified. One was sensitive to motion corresponding to V5, one sensitive to sterosis (V5a) and one more responsive to both stimuli (V5+). Time series from the activated regions were extracted from the raw data. Non linear system identification techniques were used to identify a model of the interregional connectivity. The statistical validity of the functional relationship between the different regions was assessed using Structural Equation Modelling.

1 Introduction

Although many psychophysical studies have investigated how the human brain computes stereoscopic information [14],[3], it is quite uncertain which cortical areas are involved in its implementation. Some electrophysiological studies in monkeys report the sensitivity of V1 to absolute disparities, suggesting that this area could be a preliminary stage of processing for stereo information [7]. MT/V5 in monkeys shows a columnar organization tuned for disparity [8]. MT/V5 in human brains has been widely reported as a motion sensitive area [13],[4],[17],[12].

Rudomín, J. Vázquez-Salceda, J. Díaz de León. (Eds.). e-Health: Application of Computing Science in medicine and Health Care. © IPN, México 2003.

Given the similarity between the visual system of the monkey and the human [10], it is not unreasonable to think that V5 in human brains is involved in the processing of stereo information as well. However some studies of patients with lesion in parieto –temporal / occipito-temporal (V5) region reported no loss of performance in a stereoscopic depth task [18],[16],[19]. This could be attributed either V5 is not related to stereo disparities processing or there is another region (beyond V5) more sharply tuned to stereoscopic information. The goal of the present study is not only to investigate the cortical areas involved in the processing of stereo information but also how these areas interact with the V5 region.

Functional magnetic resonance imaging (fMRI) is a non-invasive technique that provides the opportunity to monitor activity in time from many regions in the brain. Connectionist approach permits us to use these time series to determine modulatory interactions between connected regions. McIntosh *et all* [11] have demonstrated the use of path analysis to interpret fMRI data in order to assess the validity of models of effective connectivity [9].

Generally speaking these analyses have been done considering only linear relationships among regions however many neuronal interactions are non linear and is important to identify them. In the present study we used a non-linear system identification algorithm (NARMAX) which combines structure selection and parameter estimation. The NARMAX algorithm has been successfully used in many engineering applications due to its facility to capture in few terms the non-linear dynamic of the systems [1].

Structural Equation Modelling (SEM) is a mathematical technique used to assess models that define relations among variables [2]. Although SEM theory is not new, the application of this tool to estimate effective connectivity in the brain is relatively novel [11]. The basic idea of SEM is to test how well a proposed model fits the data observed, this evaluation is given in terms of covariance analysis i.e. the degree to which the activities of two regions are related. The goodness of the model is given by the chi-square statistic (χ^2). SEM was used in this study to evaluate how well the proposed model fits the observed data. A short description of the basic ideas underlying path analysis used in this work was included in an appendix.

2 Methods

2.1 Subjects

Seven healthy right-handed volunteers (4 female,3 male) aged from 20 to 30 years participated in the present preliminary study. One of the male subjects was scanned twice. All subjects gave informed written consent.

2.2 Experiment Design

All subjects did 4 sequential scans each lasting 6 min. 12 sec. (17 epoch) with a 5 min interscan interval to permit subjects to rest. One hundred and twenty four image volumes were acquired in each run. Each condition lasted 21 sec., giving 7 multislice volumes per condition (TR=3 sec.). A dummy condition of a blank screen was presented during the first 15 sec (5 scans) of each run to eliminate magnetic saturation. To avoid habituation the conditions were counterbalanced using a Latin Squares design. The motion stimulus was radial to facilitate fixation following by Buchel's experiments [5],[6]. The subjects were no trained and during all conditions they were instructed to fixate a point (0.3 °) in the middle of the screen (circular field of view 13 °) and foveate among one of the following visual stimuli:

Fixation : In this condition only the fixation point is displayed in the centre of the active area, this condition was taken as a base line. *Stationary*: two hundred and fifty dots (with radio 0.1 °) were randomly positioned within the circular field of view (mean dot density 8 dot deg^{-2} at the centre and 1 dot deg^{-2} in the edges) , the aim of this condition was to activate the visual cortex areas sensitive to the luminance produced by the dots. *Motion*: the same set of dots moving (constant speed 6.8 deg s^{-1}) radially, changing from expansion to contraction every 3 seconds. The dot density was maintained constant by replacing each dot moving outside the visual field with one appearing at the centre. With this stimulus we expected to activate the motion sensitive regions. *Stereo*: the same number of dots positioned in depth (red/green anaglyph stereogram) forming a 3D cone structure (maximum disparity ±0.08 deg.). A 3D cone shape was used to provide a wide range of disparities to stimulate binocular neurons. *StereoMotion*: the previously *Stereo* and *Motion* stimuli were combined.

2.3 Stimulus Presentation

Subjects lay on their back in the magnet. They wore red/green anaglyph glasses and looked via a mirror angled at ~45° from their visual axes at a back illuminated screen located at the extremity of the magnet. The viewing distance was 2.4 m. Stimuli were projected on to the screen using an EPSON (EMP-7300) projector driven by a 3G Mac running Psychophysics Tool Box ver. 2.44 [5],[15] under MATLAB ver. 5.3. The mean luminance of the image was 2.15 cd/m^2. Although the stimuli were displayed at a video frame rate of 60 Hz, the image was only updated on alternate frames, producing an effective frame rate of 6 Hz.

2.4 Data Acquisition

Subjects were scanned in a 1.5 T whole-body MRI scanner (Eclips Marconi Systems) with BOLD contrast echo planar imaging (TR= 3s, TE= 40 ms, 128 x128 voxel, voxel size 1.875 x 1.875 x 4 mm.). Eighteen slices covering the whole visual cortex were acquired.

2.5 Data Analysis

The data was pre-processed and analysed using SPM99 (Welcome Department of Cognitive Neurology). The first five scans of each run were discarded to exclude magnet saturation artefacts. All volumes were slice timed, motion corrected and normalised in the Talairach stereotactic space. The data were smoothed using a 6 mm FWHM (full width at half maximum) isotropic Gaussian kernel . Data analysis was performed using a boxcar design matrix of the different conditions (fixation, stationary, motion, stereo and stereomotion) convolved with the haemodynamic response function. Specific effects were tested by applying the corresponding linear contrast to the parameters obtained applying General Linear Model (GLM) using the design matrix previously described. The statistical parametric maps (SPMs) were then interpreted by referring to the probabilistic behaviour of Gaussian random fields. The threshold adopted was $P < 0.05$ (corrected).

3 Statistical Analysis

Five analyses were performed in order to identify the areas involved in each particular stimulus. In contrast with the consistent and significant activation (V5 region) among subjects following the motion condition, only three subjects showed a consistent activation during stereo. We attribute this to the fact that some people have difficulty persiving stereograms and the activation produced by the stimuli is not strong enough to be significant in the analysis. For reasons of space we show only the regions of activation found in the best subject (fig. 1).

Motion against Stationary: In accordance with previous research the V5 complex area showed the highest activation to the directional motion presented in the stimulus. *Stereo against Stationary:* Much of the cortical tissue sensitive to motion was also sensitive to stereo stimulation, and it is important to note that this contrast reveals some areas that were not activated by the motion stimulus. *Motion against Stereo:* This contrast shows the region (V5 region, Talairach space (45, ±65, -3)), which had higher activation to motion than to stereo. *Stereo against Motion:* The cortical area which was more sensitive to stereo than motion was very small but consistent among three subjects. We could not find in the literature any report of an area with the characteristics shown on this location. We called this area V5a (Talairach space (35, ±83, -3)).

StereoMotion against Motion and Stereo: Given that some regions are sensitive to both stimuli, we used this contrast to assess which areas become more active when both stimuli are presented at the same time instead of when they were presented separately.

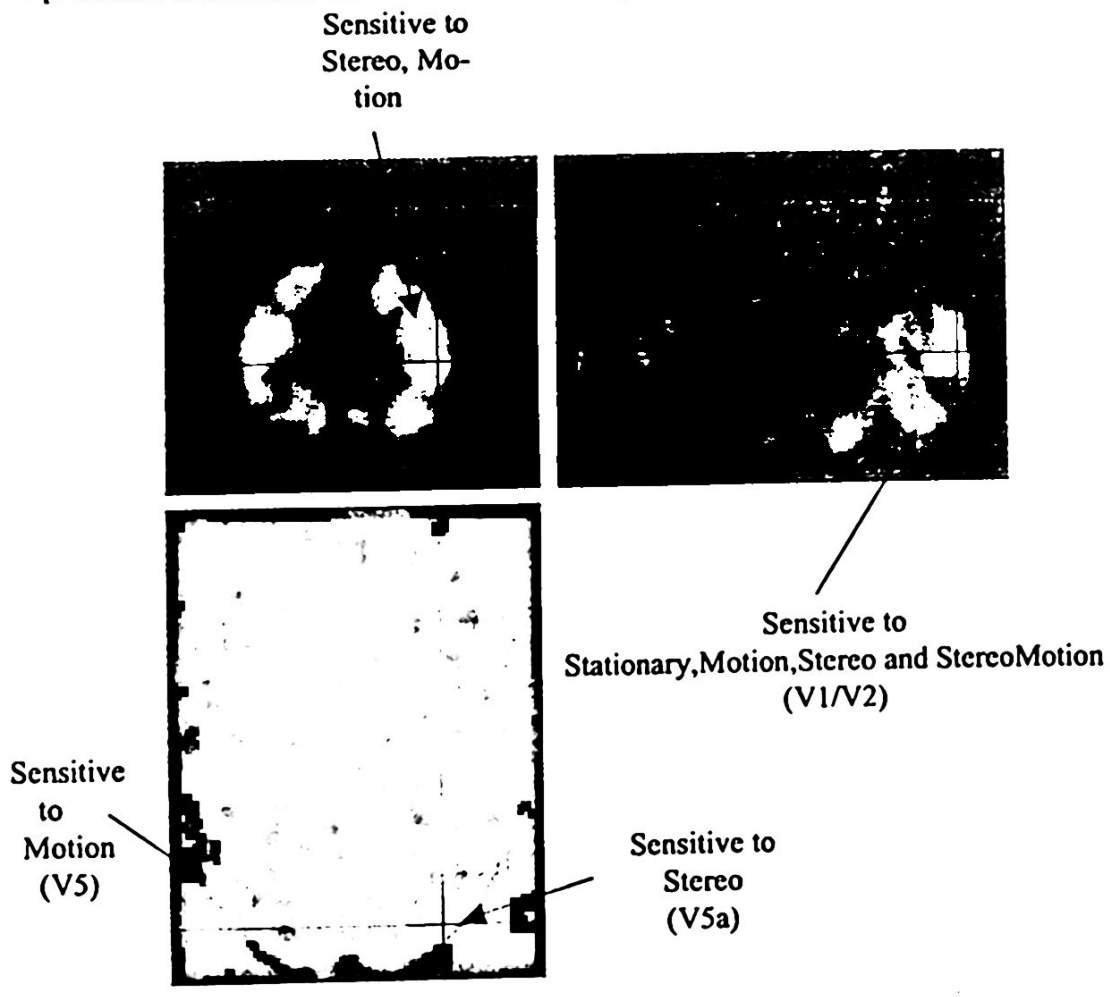

Figure 1. Axial, sagittal and coronal views of slice z=-3 according Talairach space. As we can appreciate V5 and V5a are specifically responsive to motion and stereo respectively. Otherwise V5+ is sensitive to both stimuli.

4 Effective Connectivity Analysis

Our results showed one area in the visual cortex which was both activated by the motion stimulus as well as by the stereo stimulus. We refer to this area as a V5+ (Talairach space (33, -88, 17)).

The mean time series of this region was highly correlated with the boxcar time series of stereo and motion conditions. It is not unreasonable to hypothesise that this area is an early stage of processing of motion and stereo information and that some more specialised areas are necessary to make a further treatment of the respective stimuli. Considering the correlation of the region with the individual models of motion and stereo, we believed that V5 could be a good candidate to realise the specialised processing to motion information and V5a could be the primary place in the processing of stereo information. In order to asses this hypothesis of effective connectivity [Friston, 1994], path analysis (see appendix) was done using the time series from the regions involved.

First, we applied the NARMAX algorithm to identify the interaction terms (i.e. non linear relationships) among regions. Some apriori constraints were introduced as an anatomical connection between regions. An interactive term (V5xV5a) was suggested by NARMAX to explain the response activation of the joint activity of V5 and V5a. The statistical significance of the model was evaluated applying SEM. The Analysis of Moment Structures (AMOS ver. 4) software package was used in this analysis. As an indicator of correctness, the probability level achieved was near to one (fig. 2).

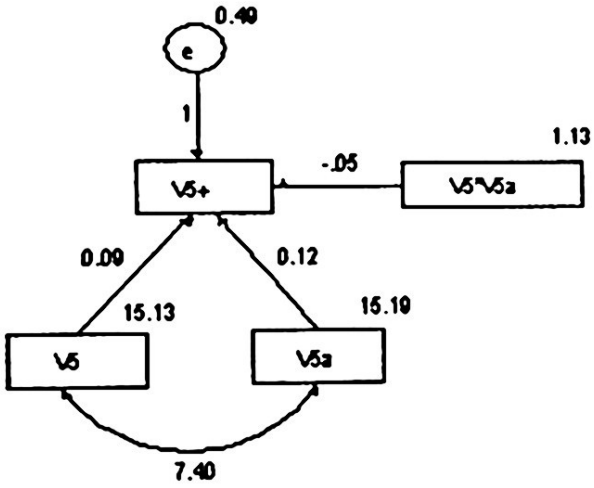

Figure 2. The model following the use of NARMAX algorithm was assessed using AMOS 4.01 SEM software. The covariance analysis with 2 degrees of freedom achieved a chi-Square = 0 and a probability level of 0.99, implying that the model is a very good fit of the data.

5 Conclusions

A conclusion of our research is that area V5 in humans appears to be functionally similar to MT in monkeys, not only in the sense that both of them process motion information, but also because both are involved in the processing of stereo disparities.

Our results suggest that V5 complex has a specific set of neurons tuned exclusively to compute stereo disparities. We called V5a this new functional defined area which is specialised in the treatment of stereoscopic information.

A model of effective connectivity that accounts for the functional architecture of stereoperception was presented. The validity of the proposed model was tested statistically using structural equation modelling. Finally, we believe that the inconsistency in V5a activation among subjects is caused by the large range of people ability to solve stereograms. Owing to all subjects were not trained future work has to be done to explore the effects of training to ensure a good level of stereo acuity between subjects.

6 Appendix: Path Analysis

Path analysis is a mathematical technique which is used to find a model that best explains the causal relationship among the data. Path analysis can be split in two different but complementary procedures, the first one is called *exploratory analysis*, in this stage the objective is to find the structure of the model (system identification problem). The second one, called *confirmatory analysis*, consists of assessing how well the proposed structure fits our data.

6.1 Exploratory Analysis

In order to solve the system identification problem we propose the use of Multi input Multi output Non-linear AutoRegressive Moving Average with eXogenous inputs (MIMO-NARMAX) algorithm [1]. The underlying idea behind this algorithm is to represent the model as linear-in-the-parameters non-linear difference equation system and estimate the parameters doing an orthogonal decomposition. For example, let the dependent variable Y be represented as the following discrete-time multivariable non-linear stochastic system:

$$Y(t)=f(Y(t-1),\ldots,Y(t-n_y),\ U(t-1),\ldots,U(t-n_u),\ E(t-1),\ldots,E(t-n_e))+E(t) \qquad (1)$$

Where Y, U and E represents endogenous, exogenous and error terms respectively with n_y, n_u and n_e maximum lags (memory of the system). Due to each term in (1) represents a family of subterms it can be decomposed in:

$$\begin{aligned} y_i(t) = f_i(&y_1(t-1),\ldots, y_1(t-n_y),\ldots, y_m(t-1),\ldots, y_m(t-n_y),\\ &u_1(t-1),\ldots, u_1(t-n_u),\ldots, u_r(t-1),\ldots, u_r(t-n_u),\\ &e_1(t-1),\ldots, e_1(t-n_e),\ldots, e_m(t-1),\ldots, e_m(t-n_y)) + e_i(t) \end{aligned} \quad (2)$$

The non-linear form of f_i can be approximated by the polynomial expansion of order l ,where multiplicative terms between monomials are calculated and included as new terms of the system:

$$y_i(t) = \theta_0^{(i)} + \sum_{i1=1}^{n} \theta_{i1}^{(i)} x_{i1}(t) + \sum_{i1=1}^{n}\sum_{i2=i1}^{n} \theta_{i1i2}^{(i)} x_{i1}(t)x_{i2}(t) + \ldots + \sum_{i1=1}^{n}\cdots\sum_{il=1}^{n} \theta_{i1\ldots il}^{(i)} x_{i1}(t)\ldots x_{il}(t) + e_i(t) \quad (3)$$

This representation capture the non linearity of the system and is linear in the parameters. The expanded system can be expresed as:

$$y(t) = \sum_{i=1}^{M} p_i(t)\theta_i + e(t) \quad (4)$$

$$Y = P\theta + e \quad (5)$$

Where P represents the full model set (design matrix). The problem of structure selection and parameter estimation can be sorted out finding the subterms P_i of P which minimise *e*. This minimisation can be formulated as a standard least squares problem, however, due to some terms in P are very similar and P^TP easily becomes singular, we used a forward regression orthogonal decomposition (Gram-Schmidt orthogonalization) to estimate the θ_i parameters. The complexity of the system (how many terms include as a predictors) can be controlled using a parsimony criteria like Akaike's information or Bollen's parsimonious fit index.

6.2 Confirmatory Analysis

The use of Structural Equation Modelling as a confirmatory tool implies the existence of a theoretical model proposed to explain the observed data. The hypothesis to assess is that the observed covariance (S) taken from the data is equivalent to the implied covariance ($\Sigma(\theta)$) of the model. The relationships between the variables included in the model are expressed as a structural equations, for example consider the following model that establish the connectivity between three variables A, B and C:

$$B = \theta_1 A + \Psi_2 \quad (6)$$

$$C = \theta_2 A + \theta_3 B + \psi_3 \quad (7)$$

Matrix notation:

Vector notation:

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \theta_1 & 0 & 0 \\ \theta_2 & \theta_3 & 0 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} + \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \end{bmatrix}$$

$$v = K v + \psi \quad (8)$$

Where v is a vector that represents the observed variables, K is the matrix of coefficients and ψ is a vector of residuals. The positions in K that contain 0 denote the lack of connection between these variables. Factorising (8) we have:

$$v = (I - K)^{-1}\psi \quad (9)$$

Then the implied covariance matrix is constructed following the hypothesis that $S = \Sigma(\theta)$.

$$S = v \bullet v' \quad (10)$$

Then substituting (9) in (10)

$$\Sigma(\theta) = (I - K)^{-1}\psi \, ((I - K)^{-1}\psi)' \quad (11)$$

The estimation of the parameters in θ that minimise the difference between S and $\Sigma(\theta)$ is usually done using the maximum likelihood description function (ML) that is asymptotically distributed as chi-square statistic. The goodness of fit of the model can be estimated using χ^2 distribution with degrees of freedom equal to the number of non repeated terms in the observed covariance matrix minus the number of parameters to be estimated in the model.

Acknowledgement

The first author thanks the Mexican National Council for Science and Technology (CONACyT) for the economic support of this project under the scholarship number 70355.

References

1. Billings S A, Chen S and Korenberg M J, 1989 "Identification of MIMO non-linear systems using a forward-regression orthogonal estimatior". Int. J. Control 49 (6) 2157 – 2189.
2. Bollen K A, 1989 "Structural equation with latent variables". Wiley interscience.
3. Burt P, Julesz B, 1980 "A disparity gradient limit for binocular fusion". Science 208 : 615 – 617.
4. Braddick O J, O'Brien J M D, Wattam-Bell J, Atkinson J, Hartley T and Turner R, 2001 "Brain areas sensitive to coherent visual motion". Perception 30 : 61 – 72.
5. Brainard, D H, 1997 "The Psychophysics Toolbox". Spatial Vision 10 : 433 - 436. Buchel C, Josephs O Rees G, Turner R, Frith C D and Friston K, 1998 "The functional anatomy of attention to visual motion". Brain 121 : 1281 – 1294.
6. Buchel C and Friston K, 1997 "Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI". Cerebral cortex 7 : 768 – 778.
7. Cumming B G, Parker A J, 1999 "Binocular neurons in V1 of awake monkey are selective for absolute, not relative disparity". The journal of neuroscience 19 (3) : 5602 – 5618.
8. DeAngelis G C and Newsome W T, 1999 "Organization of disparity-selective neurons in macaque area MT". The journal of neuroscience 19 (4) : 1398 – 1415.
9. Friston K J, 1994 "Functional and effective connectivity in neuroimaging: A synthesis". Human brain mapping 2 : 56 – 78.
10. Gonzalez F, Perez R, 1998 "Neural mechanisms underlying stereoscopic vision". Progress in neurobiology 55 : 191 – 224.
11. McIntosh A R and Gonzalez-Lima F, 1994 "Structural equation modelling and its application to network analysis in fuctional brain imaging". Human brain mapping 2 : 2 – 22.
12. McKeefry D J, Watson J D G, Frackowiak R S J, Fong K and Zeki S, 1997 "The activity in human areas V1/V2, V3 and V5 during the perception of coherent and incoherent motion". Neutoimage 5 : 1-12.
13. Totell R B H, Mendola J D, Hadjikhani N K, Ledeen P J, Liu A K, Reppas J B, Sereno M I and Dale A M, 1997 "Functional analysis of V3A and related areas in human visual cortex". The Journal of neuroscience 17 (18) : 7060 – 7078.
14. Tyler C W, 1973 "Stereoscopic vision: cortical limitations and disparity scaling effect". Science 181 : 276 – 278.
15. Pelli, D G, 1997 "The VideoToolbox software for visual psychophysics: Transforming numbers into movies". Spatial Vision 10 : 437 - 442.
16. Schiller P H, 1993 "The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. Visual neuroscience 10 : 717 – 746.
17. Smith A T, Greenlee M W, Singh K D, Kraemer F M and Henning J, 1998 "The processing of first an second order motion in human visual cortex assessed by functional magneticresonance imaging (fMRI)". The journal of neuroscience 18 (10) : 3816 – 3830.
18. Vaina L M, 1989 "Selective impairment of visual motion interpretation following lesions of the right occipito-parietal areas in human". Biol. Cybern. 61 : 347 – 359.
19. Vaina L M, Lemay M, Bienfang D C, Choi A Y and Nakayama K, 1990 "Intact biological motion and structure from motion perception in a patient with impaired motion mechanism: a case of study". Visual neuroscience 5 : 353 – 369.