"Functional Anatomy of Stereoscopic Visual Process Assessed Using Functional Magnetic Resonance Imaging and Structural Equation Modelling"

Thesis submitted in September 2002 for the Degree of
Doctor in Philosophy

By

Héctor Gabriel Acosta Mesa

Artificial Intelligence Vision Research Unit.
Department of Psychology. The University of Sheffield.

This thesis is dedicated to my wife, thanks for your love, patience and understanding. This success belongs equally to us both.

Thanks to my mother, my grandmother and all my family for their love and confidence on me.

Acknowledgements

I gratefully acknowledge all those people who have contributed and encouraged me to complete this thesis.

Abstract

The purpose of this thesis is to study the functional anatomy of stereoscopic vision. Although many studies have investigated the physiological mechanisms by which the brain transforms the retinal disparities into three-dimensional representations, the invasive nature of the techniques available have restricted them to studies in non-human primates, whilst the research on humans has been limited to psychophysical studies.

Modern non-invasive neuroimaging techniques now allow the investigation of the functional organisation of the human brain. Although PET and fMRI studies have been widely used, few researchers have explored the functional anatomy of stereoscopic vision. Most of these studies appear to be pilot work, showing inconsistency, not only in the areas sensitive to stereo disparities, but also in the specific role that each of these possesses in the perception of depth.

In order to investigate the cortical regions involved in stereoscopic vision, four fMRI studies were performed using anaglyph random dot stereograms. Our results suggest that the stereo disparity processing is widespread over a network of cortical regions which include V1, V3A, V3B and B7. Functionally, the V3A region seems to be the main processing centre of pure stereo disparities and the V3B region to be engaged in motion defined purely by spatio-temporal changes of local horizontal disparities (stereoscopic -cyclopean- motion).

Interregional connectivity was investigated with two approaches. Structural Equation Modelling (SEM), as the classical technique for the analysis of effective connectivity, was used to assess one connectivity model proposed to explain the cortical interaction observed in the first experiment. The implementation and application of this technique permitted us to identify some of its weaknesses in representing fMRI data. An extension of the SEM technique was introduced as a Non-linear Auto-Regressive Moving Average with eXogenous variables (NARMAX) approach. This can be thought of as an attempt to bring SEM towards a non-linear dynamic system modelling technique which permits a more appropriate representation of effective connectivity models using fMRI time series.

Chapter 1: Background and Motivation.

Chapter 2: Neural Structural Equation Modelling.

Chapter 3: NARMAX.

Chapter 4: Experiment # 1: Stereo Motion Interactions.

Chapter 5: Experiment # 2: Stereo Motion in Depth.

Chapter 6: Experiment # 3: Global Stereo Tracking.

Chapter 7: Experiment # 5: Shape Discrimination from Stereopsis.

Chapter 8: Discussion and Future Work.

Appendix: Code.

Bibliography

Chapter 1

Background and motivation

Abstract.

The current chapter shows a general framework of the work described in this thesis. The operational principles of stereoscopic vision are explained and the basic anatomical organization of the visual cortex is presented. A review of the literature concerning studies of stereovision made in monkeys and humans is discussed. The concepts and techniques to assess neural connectivity are introduced. Lastly a summary of objectives and relevance of the work is given.

1.1 Stereoscopic Vision.

Visual perception theory establishes four main sources of retinal image information to recover the shape of the world. Texture, luminance, optic flow and disparity are the intrinsic sources of information from which the visual system gets clues about the three dimensional properties of objects (Marr 1982). Based on psychophysical observations, computational implementations of visual functions have been developed to explore the operational principles to recover 3D structure. Although most of the algorithms proposed to process these sources of intrinsic visual information work under many assumptions, they have helped to explore the mechanism that implements three-dimensional vision. Shape from texture algorithms use the local deformation of a "pattern of texture" to estimate the slope in the surface of the object (fig. 1.1). This procedure assumes the existence of repetitive patterns from which gradients of local deformation can be estimated (Aloimonos 1988; Malik and Rosenholtz 1994).
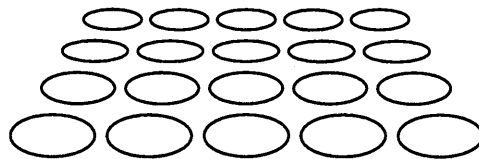
Figure 1.1 Local deformation of a pattern of texture. A circle in the fronto parallel plane deforms to an ellipse as the slope of the plane change.

Shape from shading uses the reflective property of surfaces to estimate slope (Ramachandran 1988; Seyama and Sato 1998). Given that the quantity of light reflected by the surface is dependent on the slant on each particular point, the luminance gradients are used to estimate surface slant relative to the viewer (fig.1.2).

Shape from motion uses the gradient of speed and direction (optic flow) involved in one point in space when it changes its position. Optic flow is estimated by calculating the gradient of change in position of one point in the surface.
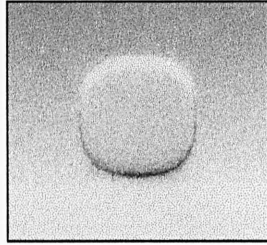
Figure 1.2 Shape from shading. Three dimensional form can be observed using gradients of luminance. The quantity of light reflected in one point depends on the angle of reflection created by the surface slant.

For example, let us imagine a transparent sphere with some kind of dot pattern. If the sphere is still, it is not possible to recognise the form, but if the sphere is permitted to rotate around one of its axes, it is easy to identify the structure (Marr 1982; Poggio and Koch 1987; Hubel 1995), (fig. 1.3).
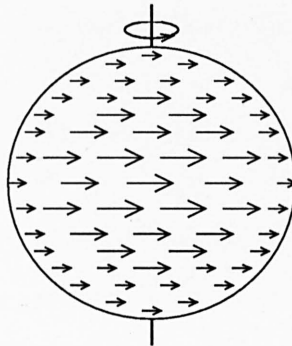
Figure 1.3 Optic flow. In a rotating sphere, the spatiotemporal derivatives (with respect to the viewer) in every position of the sphere vary over the surface. The gradient of optic flow is bigger on the periphery and decreases toward the centre.

Because one of the topics of this work is how disparity is used by the brain to get shape from stereopsis, a fuller explanation about principles of stereoscopic vision will be given in the next section.

It is important to notice that the previous analyses of intrinsic visual information are complementary in the sense that they estimate slant using different physical properties of the world.

It is believed that all of these sources of information are processed and integrated by the brain to estimate surface orientation relative to the viewer (Tittle, Norman et al. 1998).

The present work is only concerned with the interaction between brain regions involved in the process of stereoscopic vision. However, a natural extension of this work is the study of the relationships between the cortical regions which recover shape from different intrinsic signals.

1.1.1   Geometry of binocular vision.

Due to the separation of the eyes the visual field is projected to spatially different parts of each retina. Let us imagine each retina as a bidimensional array of cells, each point in the visual field is projected to one cell of each array, if the point projects to the same corresponding cell in each retina it is considered to have zero horizontal disparity. Helped by the extraocular muscles, the eyes converge to fixate items of interest in the world, these points have zero disparity (fig.1.4).

The set of points whose projections have zero disparities form the horopter. Points in front of the horopter have crossed or negative disparities, those beyond the horopter have uncrossed or positive disparities (fig.1.5).

F

Left eye.   Right eye.
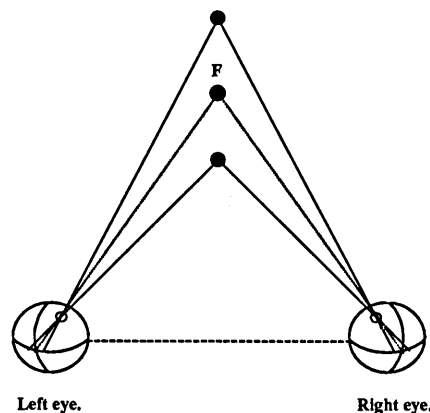
Figure 1.4 Horizontal disparity for points lying in the horizontal plane. Points at different depth projects in different position in the x axis defined by the horizontal eccentricity in the retina.
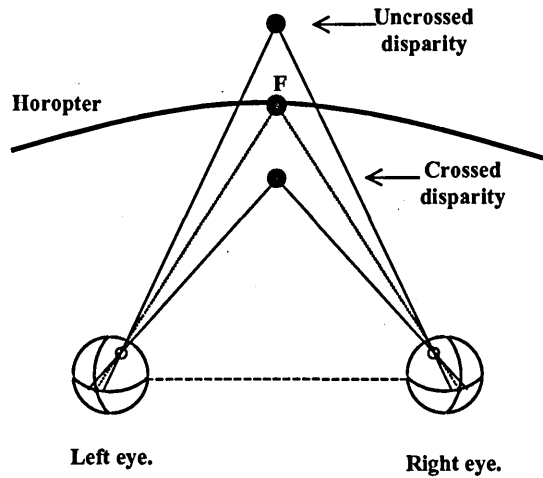
Figure 1.5 Horopter. The horopter is defined by the family of points which projects onto the retina with zero horizontal disparity.

The space around the fixation point that can be fused is limited, points outside certain values of disparities are perceived as two images (diplopia). Panum's fusional area is defined as the space of disparities that can be fused around the horopter (fig.1.6). Some points outside the Panum's fusional area can be seen in depth despite diplopia, however the precision of the depth recovered by this fusion is low.
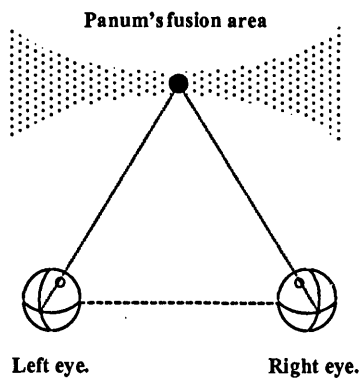
Figure 1.6 Panum's fusion area.

So far, only horizontal disparities have been considered, however, points that in the visual field are closer to one eye project with different angles on each retina, and this fact generates image differences that are called vertical disparities (fig.1.7).
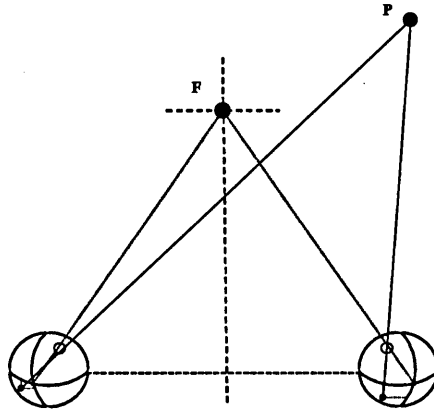
Figure1.7 Vertical disparity. Points physically located near to one eye project in different position in the *y* axis defined by the vertical eccentricity in the retina.

Ogle's experiments proved that the brain is able to compute vertical disparities (Ogle 1962). He showed that the binocular perception of depth is altered by distorting one of the retinal images in the vertical dimension, this visual phenomenon is called the induced effect (fig.1.8).
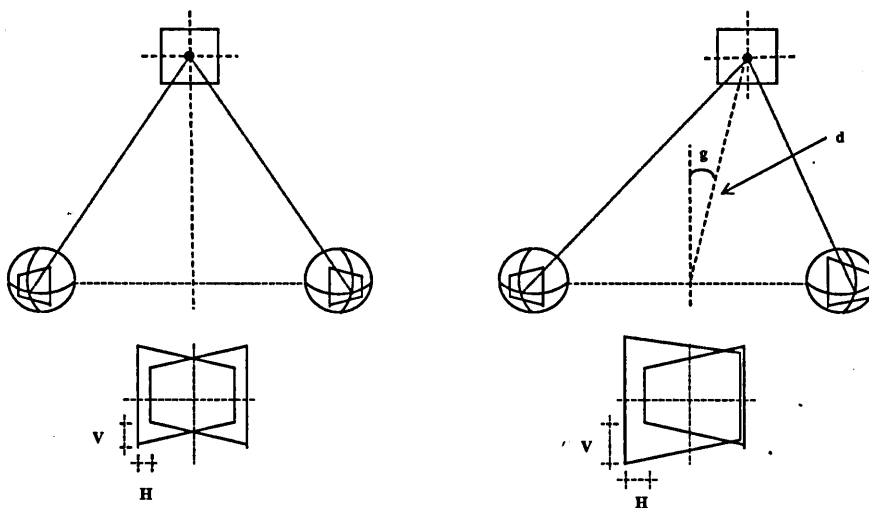
Figure 1.8 Induced effect. Symmetrically converged eyes observing a square produce a symmetrical trapezoidal shape in the retinas (left). Non-symmetrically convergence observing the same square produces a non-symmetrically trapezoidal shape (right). Vertical (V) and Horizontal (H) disparities are defined as the *x* and *y* coordinates of matching points.

The stereoscopic perception of depth can be summarised in two processes: the solution of the correspondence problem and the interpretation of the disparities. The correspondence problem consists of mapping the points in the left image to points of the right image. This is a difficult task in which ambiguities could arise, for example, from some perspectives there are more than one possible match to correspond. Many computational approaches have studied this problem, but given the complexity of it, most of the proposed algorithms work under restrictive assumptions (Marr and Poggio 1979).

The interpretation of the disparities consists of using the horizontal and vertical disparities, calculated by solving the correspondence problem, to recover the absolute location of one point in three dimensional space (Mayhew 1982). Mayhew *et al* (Mayhew and Longuet-Higgins 1982) proposed a computational model that uses both vertical and horizontal disparity. One consequence of this approach was that it is able to explain the induced effect. The algorithm works on the principle of continuity, i.e., it assumes the reconstruction of a planar surface defined by a plane z=PX+QY+K (fig 1.9).
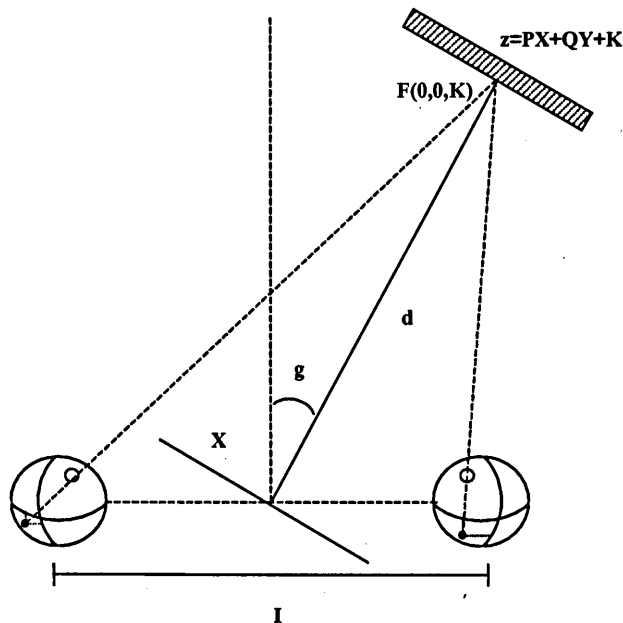
Figure 1.9 Stereo geometry. The origin of the coordinate system that defines the plane *z* has its origin at the middle point of the axis that separates the eyes. X axis can be thought of as the rotation (g) of this axis. Z axis is defined by the projection from the origin to the position of fixation point. Y axis is orthogonal to X and Z (no represented in the figure).

After solving the correspondence problem, vertical and horizontal disparities can be used to estimate the position and orientation of the plane Z. The geometry of the system is defined by 1.1 and 1.2, where V is an approximation.

$$V = Ixy/d + Iyg/d \qquad (1.1)$$

$$H = Ix^2/d + Ixg/d + Iz/d^2 \qquad (1.2)$$

Where:

H .- Horizontal disparity.

V .- Vertical disparity measured at the retinal location.

I .- Inter-ocular separation.

d .- Viewing distance.

g .- Angle of gaze.

x .- Horizontal component of the eccentricity measured at the retinal location .

y .- Vertical component of the eccentricity measured at the retinal location.

z = PX+QY+K .- Planarity assumption.

P, Q .- Coefficients related to the surface slant.

K .- Translation along the line of gaze.

Using equation 1.1 it is possible to estimate g and d. Then using these values in the equation 1.2, one can estimate value of *z* with respect to the fixation point. Contrary to the theory that convergence angles of the eyes are used to estimate absolute depth, the previous approach revealed that no extra-retinal information is needed. For a fuller explanation of the derivation of the previous equations consult (Mayhew and Longuet-Higgins 1982).

1.2 Visual cortex.

Most of the experiments made to investigate the functional neuroanatomy of visual systems have been done on monkeys and cats, due to the invasive nature of the available techniques.

Electro-recording, electro-stimulation and cytocrome oxidase staining have been the most frequent techniques applied to monitor neural activity. Given the physiological similarities between the human and monkey visual systems, it is reasonable to expect a similar functional architecture, however, this assumption should be taken very cautiously and many investigations have to be developed to assess the scope of its validity (Van Essen, J.W. et al. 2001).

The visual cortex is one of the most studied parts of the brain. It is believed that understanding the functional organisation of the visual cortex could help understanding the structural organisation in the rest of the brain. Cortical areas sensitive to colour or motion are well identified, but it is not clear which parts of the brain are involved in stereopsis.

1.2.1   Visual pathways.

After the light is captured by the retinas information is transferred to the brain via the optic nerve. One important characteristic of the information transmitted through the optic nerves is that it preserves the spatial relationship with respect to the part of retina where the information was taken (retinotopic organisation). Both optic nerves join in the optic chiasm, at this point each hemiretina combines to form the optic tracts. Each optic tract contains information about one epsilateral temporal hemiretina and one contralateral nasal hemiretina.

Optic tracts propagate information into the Lateral Geniculate Nucleus (LGN). This seems to be a thalamic relay for the visual information coming from the retina. The LGN is divided into six layers which make up the parvocellular and magnocellular divisions. The four most external layers form the parvocellular division, which receives input from the fovea, and consequently, contains a high resolution representation and also it is sensitive to colour.

The two most internal layers of the LGN form the magnocellular division, which contains information from the periphery of the visual field and thus it has less spatial resolution than the parvocellular division. The LGN is connected to the striate cortex (primary visual cortex), (fig. 1.10).
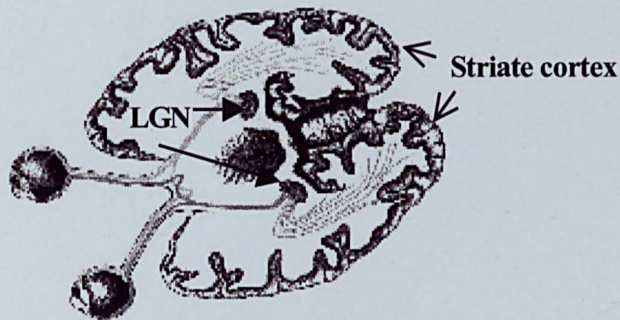
Figure 1.10 Visual pathway. Visual information flows from hemiretinas, via the optic nerve, optic chiasm, optic tracts, and the lateral geniculate nucleus to the visual cortex.

1.2.2. Striate and extrastriate visual areas.

Visual cortex is spread over the parietal, occipital and temporal lobes. It is divided into different cytoarchitectonically and functionally defined areas. Cytoarchitectonically defined areas are commonly reported using Brodmann's map, in which the extriate visual cortex corresponds to area 17 (Amunts, Malikovic et al. 2000), (fig. 1.11).
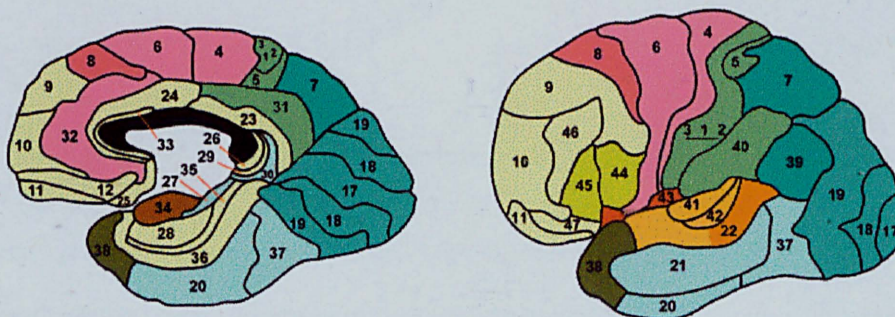
Figure 1.11 Brodmann areas. Visual cortex occupies areas 17, 18, 19, 37 and 7. Left.-medial view. Right.- Lateral view.

One problem of three dimensional visualization of the cerebral cortex is that neocortex is folded in the sulcus, so it is not possible to visualise brain regions under the cortical surface (fig.1.12).
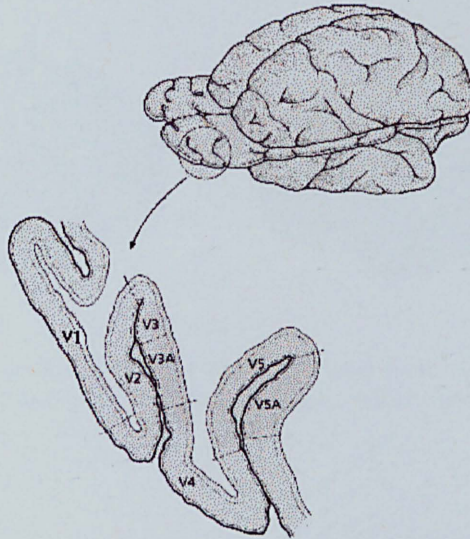
Figure 1.12 Cerebral cortex. Some cortical regions in the visual system are not visible in three dimensional representations due to their location in highly folded structures in which the cerebral cortex is organised. Modified from (Zeki 1993).

One alternative is to visualise the cerebral cortex as a cortical flat map, it allows the entire surface of the brain to be visualised in a single view (Carman, Drury et al. 1995),(fig. 1.13).
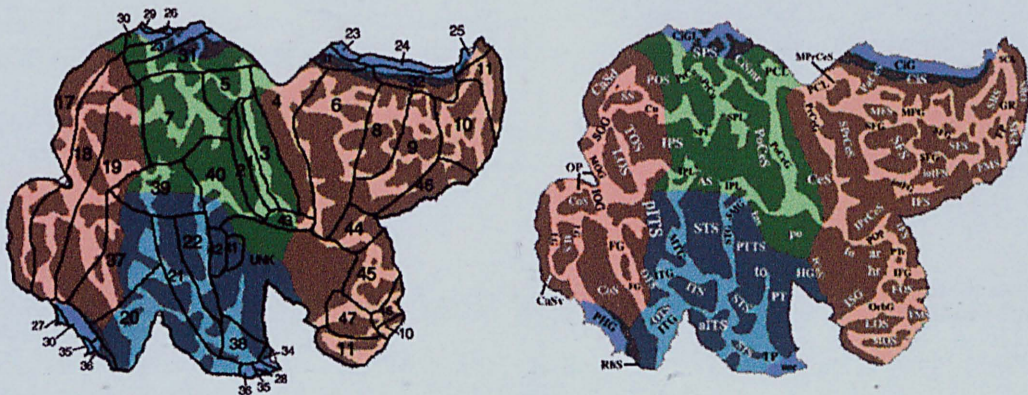
Figure 1.13 Surface representation of the right hemisphere. Left.- Brodmann's areas distribution. Right.- Geographical atlas based on the atlas of the cerebral sulci (Ono, Kubick et al. 1990).

Studies in monkeys and humans have revealed the existence of functional defined areas related with specific aspects of visual function, like edge detection (V1), motion (V5), colour (V4), space, form, etc. (DeYoe, Carman et al. 1996; Van Essen and Drury 1997), (fig. 1.14 and 1.15).

Figure 1.14 Visual cortex. Estimated extent of visual areas. Left.- Lateral view. Right.- Medial view. Blue area represents visual cortex for which functional properties have not been identified.

Figure 1.15 Functionally defined visual areas. Partial hemisphere display show the estimated extent of visual areas as reported by Van Essen *et all* (Van Essen and Drury 1997). Blue area represents visual cortex without assigned name.

The functional divisions shown above were defined using as a reference the activations reported by neuroimaging studies (Van Essen and Drury 1997; Hasnain, Fox et al. 1998). Commonly, activations related to visual function are reported in stereotaxic coordinates in Talairach space (Talairach and P. 1988).

One problem in creating accurately functional defined areas is the inter-subject variability in the cerebral cortex (Hasnain, Fox et al. 1998; Amunts, Malikovic et al. 2000). For a detailed description of the functional properties and stereotaxic locations of visual areas consult the fold out at the end of this chapter (fig.1.19).

Although the visual cortex has been one of the most studied areas in the brain, many functional areas involved in different processes of visual information remain unknown. That is the case of stereoscopic vision. In the next section the most significant findings in the anatomy of stereo perception are reviewed.

1.3 Functional anatomy of stereopsis.

Psychophysical and computational models of stereopsis have been widely studied (Marr and Poggio 1979; Mayhew and Longuet-Higgins 1982), however, how the brain implements this process at a neural level is still unknown. The intuitive idea behind binocular fusion implies convergence of information from both eyes into unique neurones, in other words, to compute depth from disparities, it is necessary to solve the correspondence problem.

1.3.1 Studies in monkeys.

The first stage in the visual pathway that shares information from both eyes is the LGN (Schroeder, Tenke et al. 1990). Studies in monkeys reveal disparity sensitivity properties of magnocellular cells (Hubel and Livingstone 1987), nevertheless the conclusions among studies are not consistent, for example DeValois et al (DeValois, Smith et al. 1958) found no evidence of stereo sensitivity in these cells. Following the visual pathway, the second structure is the striate cortex, which contains information from both eyes (Cumming and DeAngelis 2001). Neurons dominated by each eye are segregated into the system of alternating strips known as the ocular dominance columns (ODC) (Blasdel 1992).

Some studies suggest that the information about the spatial relationships of disparities is encoded in the ODC (Chklovskii 2000) . Something important to note is that stereovision and ODC mature at almost the same time in young monkeys (Chklovskii 2000).

There is general agreement about the disparity sensitivity of V1 (Hubel and Wiesel 1970; Poggio and Fischer 1977; Gonzalez, Krause et al. 1993a; Gonzalez, Relova et al. 1993b). Cumming and Parker (Cumming and Parker 1999) used random dot stereograms (RDS) (Julesz 1960) to manipulate the absolute and relative components of the stereo stimuli, proving that neurones in V1 are sensitive to absolute but not to relative disparity. In the same study Cumming and Parker showed that the activity in V1 to process stereo disparities is not enough to create the sensation of depth, so they concluded that other cortical areas are necessary to create depth sensation and V1 represents only the first stage in the stereo information process.

V1 projects forward to V2, Poggio and Fisher (Poggio and Fischer 1977) investigated the sensitivity of cells in V1 and V2 using bright and dark bars moving at different depths. Hubel and Wiesel (Hubel and Wiesel 1970) found that disparity sensitive neurones in V2 do not respond to monocular stimulation. The main inputs to V1/V2 neurones come from the foveal region, these have sharply tuned disparity (0.5 arcmin). V1/V2 are also sensitive to vertical disparities (Gonzalez, Krause et al. 1993a). Contrary to the results cited above, Cowey and Wilkinson (Cowey and Wilkinson 1991) report reduced but not total impairment on stereoacuity after lesion in V1/V2.

V3 and V3A are narrow strips anterior to V2, Poggio *et al* (Poggio, Gonzalez et al. 1988) found that the sensitivity to horizontal disparities is higher in this regions than in V1/V2. Felleman and Van Essen (Felleman and Van Essen 1987) study reports sensitivity of V4 neurones to stereo disparities.

Contrary to these results Schiller (Schiller 1993) reports no loss of stereo acuity in monkeys after lesions in V4. Its important to note that V4 is well characterised by its sensitivity to colour (Zeki 1978; Hubel and Livingstone 1987).

Deangelis and Newsome (DeAngelis and Newsome 1999) have proved the existence of columnar organisation of disparity-selective neurones in MT. Preferred disparity changes smoothly across the MT surface whereas neurones with similar disparity preference are allocated in the same column.

This discovery is consistent with the well known columnar organisation of neurones sensitive to directional motion in MT (Zeki 1974; Salzman, Murasugi et al. 1992). These finding were confirmed in the same study using cortical electro stimulation in monkeys, the induced activity in binocular neurones created illusory sensation of depth.

It is believed that neurones in MT could combine motion and stereo information to create representation of dynamic environments (Saito, Yukie et al. 1986; Roy and Wurtz 1990; Cumming and Parker 1994). Because the tuning disparity preferred by MT binocular neurones is coarse it is believed that other areas beyond MT could be involved in the processing of fine disparities (Cumming and DeAngelis 2001). This thought is reinforced by the fact that some studies in monkeys report no major deficit in stereo perception after MT lesions (Schiller 1993).

Far from producing a general theory about how the stereoscopic information process is implemented at neural level, no general agreement exists about the areas involved in this process. As we saw above, many cortical areas are sensitive to stereo disparities but no general conclusions can be made about the specific role that each region of the brain takes in the complete process. Studying the interpretation of the results, many different characteristics are implied in the design of almost the same stereo stimulus, so the results of each experiment have to be interpreted carefully in order to avoid contradictory conclusions.

On the other hand, psychophysical assumptions about the physical perception of the stimulus and the correctness of the task development in monkeys make it difficult to assess, due to the nature and physical conditions of the subjects in the previous experiments.

Although similar neural configurations between monkey and human visual systems are assumed, new non-invasive brain imaging techniques have been used to explore and to investigate stereo vision in human beings. Probably the most promising of these techniques is functional Magnetic Resonance Imaging (fMRI) which applications in human stereo vision will be introduced in the next section.

1.3.2 Studies in humans.

As we saw previously, the invasive nature of the functional brain mapping techniques available in the past limited the study of the human brain. Recently, the availability of new non invasive brain imaging technique has permitted the exploration of the functional organisation of the human brain. Binocular vision research projects using Positron Emission Tomography (PET) started less than twenty years ago, this technique allows one to monitor virtually the whole brain, with a spatial resolution from 5 to 10 mm. It's disadvantages are the poor temporal resolution in the signal and the necessity of using radioactive tracers (Frackowiak and Friston 1994).

Revolutionary applications of physics to medical sciences have permitted the creation of functional Magnetic Resonance Imaging (fMRI). This technique permits one to record neural activity in the whole brain at spatial resolutions of few millimetres and with temporal resolutions of less than one second. No tracers or special subject preparation are needed and it is completely non-invasive. Although the physical principles used in fMRI comes from Felix Bloch's theories established in 1946, the complete implementations available to medical applications started only 10 years ago.

1.3.2.1 PET studies.

Gulyas et al (Gulyas and Roland 1993) used polaroid random dot stereograms to create depth perception. The stereoscopic stimulus was constructed by forming a geometrical shape in the centre of the screen (range from 0.06 – 1.13 degrees of horizontal disparity). The control task was made by a random noise pattern forming a geometrical shape in the centre. This was recognisable by twice its spatial frequency with respect to the background. The areas of activation with respect to stereo stimulus were located bilaterally in the occipital pole around the caudal end of the calcarine sulcus, bilateral in the occipital medial gyri, bilateral in the superior occipital gyri, bilateral in the precuneus, and bilateral along the banks of the intraparietal sulcus. This study did not support the right hemisphere dominance hypothesis believed to be found in stereoscopic vision. The authors concluded that stereoscopic vision is implemented through a segregated network in which both occipital and parietal areas are involved.

Ptito (Ptito, Zatorre et al. 1993) used random dot stereograms to create a rectangular cyclopean shape (disparity information not reported). The stereo task was defined as describing the orientation of the rectangle. The control task consisted of a two dimensional shape (black outline) without stereo information. The activations raised from the contrast were located in Brodmann areas BA17 (V1) and BA18 (V2) in the right hemisphere. Deactivation during stereo tasks occurred in the right inferotemporal cortex. This is one of the few studies which report sensitivity in V1/V2 to stereo stimulus.

A closely related topic in the study of stereo vision is the performance of vergence eye movements. Hasebe (Hasebe, Oyamada et al. 1999) studied this phenomenon using two LCD panels to create an approaching vertical bar (1.9 deg.). Subjects were instructed to obey one of these instructions: *follow* the approaching bar, *ignore* the bar and fix on a stationary cross, or *fix* on the cross while the vertical bar is not moving. Comparison between ignore and fix revealed activation in the left temporo-occipital junction (MNI coordinates –52 –60 -16), including V5.

It is not surprising to observe activation in V5 during this contrast given that the vertical bar is moving closer and V5 is sensitive to different kinds of motion (Howard, Brammer et al. 1996). Activations obtained between the contrast of follow and ignore tasks were found in the temporo-occipital junction and inferior parietal lobe.

Fortin (Fortin, Ptito et al. 2000) used red-green random dot stereograms to create a stereo condition in which a central square produced a vivid sensation of depth (disparity information not reported). The control condition was made of static red, green and yellow dots with no stereo information at all. The contrast between these conditions revealed a bilateral activation on V2, bilateral activation in the middle temporal lobe (MT), right side activation on V3 and B7 (precuneus). This work supported the theory of right cerebral dominance in stereo vision.

1.3.2.2 fMRI studies.

Kwee's (Kwee, Fujii et al. 1999) study used two sets of images of magnetic resonance angiography (MRA) in slightly different planes (disparity information not reported) from which it was possible to recover three dimensional representation. The control task was the use of the pair of MRA images projected in an identical plane, so only two dimensional information was available. This contrast revealed bilateral activation along the intraparietal sulcus.

Mendola's (Mendola, Dale et al. 1999) experiment used a static red-green random dot stereogram to create a depth defined square (0.56 degrees). The control task defined square contours in a flat plane. Activations found were located in the areas V3A and V7. However, Hanazawa (Hanazawa, Kawashima et al. 2000) applied dynamic red-green random dot stereograms to create two vertical bars defined by depth (disparity information not reported). Dynamic random noise was used to remove motion clue accompanied with change in disparity.

In the control task the vertical bars were defined by luminance (50 % more than the background). Activation was found in the posterior parietal and in the right occipito-temporal regions.

Instead of using anaglyph, Nishida's study (Nishida, Hayashi et al. 2001) used an array of two monitors to show the stereoscopic pair of images. This study utilised random dot stereograms to create a three dimensional cone in the centre of the screen (disparity gradually changed from 0 at the bottom to 1.6 degrees at the top). The control task was made by creating a flat circle shape within high density dots. Activations were found in the parieto-occipital region (BA18, BA19, Talairach coordinate 28 −86 22) and in the right superior parietal lobule (BA7, Talairach coordinate 14 −70 62). This study supports the right hemisphere dominance in stereo information processing. The author concludes that the superior parietal lobule is a higher level processing centre for stereopsis.

The most resent published stereo vision study used a random dot stereogram created by two independent displays (Backus, Fleet et al. 2001). Subjects viewed the stimuli through a pair of angled mirrors. The stereo stimulus consisted in two planes at different depth. The base line condition was set by a flat plane (zero disparity), whereas stereo condition varies the interplane disparity from 0 to ± 4 degrees. Backus study reported high activation in V3A and less higher but consistent activation in V1, V2, V3 and V3B.

1.3.2.3 Lesion studies.

It is always difficult to assess the magnitude of damage to the areas involved in a brain injury, however some interesting observations can be made from the behaviour of patients with brain abnormalities (Cowey and Porter 1979). One of the first reports of stereo acuity deficits was presented by Vaina (Vaina 1989). Two groups of patients with right occipito-parietal and occipito-temporal lesions respectively, were requested to identify the three dimensional form created using random dot stereograms.

Although both groups of patients had normal visual acuity, they were impaired in doing the task. In a second study made by Vaina (Vaina, Lemay et al. 1990), one patient with bilateral lesion in the posterior visual pathways affecting the lateral parieto-temporal-occipital cortex was tested for different visual capabilities. Visual acuity, form discrimination, colour, and contrast-sensitivity discrimination were normal whereas spatial localisation, line bisection, and binocular vision were severely impaired.

Finally, although it is not a lesion study, the following work is discussed because it applies a methodology which permits the assessment of neuronal performance during electromagnetic stimulation. Takayama *et al* (Takayama and Sugishita 1994) applied Transcranial Magnetic stimulation (TMS) to a set of subjects whilst they were viewing a red-green random dot stereogram. The coil was positioned over the midline of the bilateral occipital lobes. The subjects reported loss of stereoscopic perception during stimulation.

Although the spatial resolution of TMS is very low, and it is not possible to ensure the scope of the regions affected, the present results shows the contribution of the superior occipital cortex in the processing of stereo disparities.

The stereoscopic studies developed so far using brain-imaging show inconsistencies between them. Most of these studies are reported in brief communications and no detailed description of methodologies used by them are presented (Table 1.1).

Reference	Task	Base line	Activation
Gulyas,1993 (PET study) Proc. Natl. Acad. Sci.	Geometrical shape using Polaroid RDS (0.06-1.13 degrees of horizontal disparity).	Geometrical shape within higher density dots.	Bilateral on occipital gyri. Bilateral on superior occipital gyri. Bilateral in precuneus. Bilateral on intraparietal sulcus.
Ptito,1993 (PET study) Neuroreport	Rectangular shape using Red-Green RDS.	Rectangular contours in a flat plane.	Right BA17. Right BA18.
Hasebe,1999 (PET study) Neuroimage	Approaching vertical bar.	Static vertical bar.	Left temporo-occipital junction. $[-52 -60 -16]_{Tal}$
Mendola, 1999 (fMRI study) Neuroscience	Square shape using Red-Green RDS (0.56 degrees).	Square defined by luminance.	Bilateral on occipital superior gyrus (V3A). $[-21 -89 16]_{Tal}$ $[23 -88 16]_{Tal}$ V7
Kwee,1999 (fMRI study) Neurology brief communications	Stereoscopic pair of medical images.	Flat medical image.	Right intraparietal cortex. Bilateral on the intraparietal sulcus.
Fortin,2000 (PET study) HBM poster	Square shape using Red-Green RDS.	Square shape within higher density dots.	Bilateral on V2. Bilateral on MT. Right on V3. Right on precuneus.
Hanazawa,2000 (fMRI study) HBM poster	Two vertical bars defined by depth.	Two vertical bars defined by luminance.	Posterior parietal. Right accipito-temporal region.
Nishida,2001 (fMRI study) Neuroreport	3D cone using Red-Green RDS (0 – 1.6 degrees)	Flat circle shape within higher density dots.	Parieto-occipital region (BA18 and BA19).$[28 -86 22]_{Tal}$ Right superior parietal lobule (BA7). $[14 -70 62]_{Tal}$
Backus, 2001 (fMRI study) Neurophysiology	Two planes in different depths (0- ±4 degrees), two panel display.	Flat plane.	V1, V2, V3, V3B and the highest sensitivity was found in V3A.

Table 1.1 Literature review of PET and fMRI studies in stereoscopic vision.

1.4 Neural connectivity.

The traditional theory of the existence of functional segregated areas, highly specialised in specific cognitive tasks, has been replaced by the concept that brain functionality is the result of the interaction between several regions. In this new approach the functional profile of a cerebral region is determined by its relationship with other regions (McIntosh 2000). Hence, to evaluate a functional hypothesis under this integrated approach, it is necessary: i) to identify the regions involved in the task, ii) to identify the pathways that connect these regions and iii) to assess the interaction between the regions.

Using fMRI makes it relatively easy to identify the brain regions activated during a given cognitive task. The anatomical connections could be approximated using as a reference the inter-regional connections across the brain, which have been widely researched in monkeys (Gerstein and Perkel 1969), (fig.1.16). Another method is the new technique called Diffusion Tensor Imaging (DTI) which allows us to determine anatomical connectivity in the human brain using the anisotropy of diffusion existing in the white matter (Le Bihan, Mangin et al. 2001).

To explain how these procedures can be used to assess the interactions between regions we must to introduce the concept of effective connectivity in the next section (Friston, Frith et al. 1993).

1.4.1 Functional and effective connectivity.

The key point in the study of brain activations is the analysis of the hemodynamic behaviour during neuronal stimulation, i.e. the analysis of time dependent changes (time series) in the neural activity (Friston, Holmes et al. 1994). *Functional connectivity* is defined as "the temporal correlation between remote neurophysiological events" (Friston, Frith et al. 1993) . This idea comes from the analysis of spikes recorded by multiunit micro-electrodes (Gerstein and Perkel 1969).
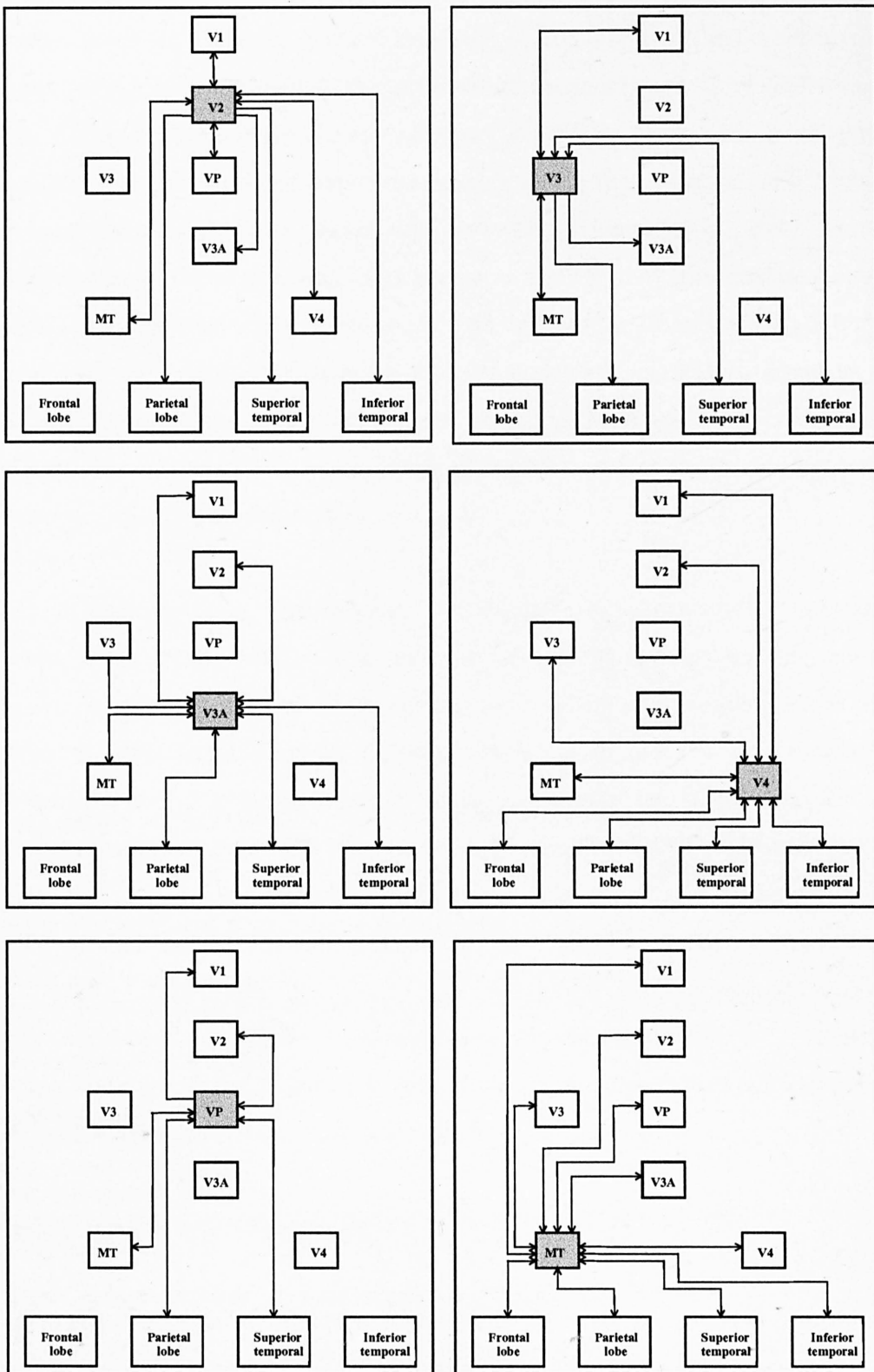
Figure 1.16 Connectivity maps. Intra-hemispheric connections of functional defined visual areas in macaque monkeys, modified from (Rockland and Kaas 1997).

As the main objective of functional connectivity is to identify correlation between the time series of different cortical regions, it is possible to apply Principal Component Analysis to identify the different (orthogonal) patters of activations within the set of time series (Friston, Frith et al. 1993). These orthogonal time series (principal components) represent families of cortical regions with highly correlated temporal response, thus spatial maps of *functional connectivity* can be constructed using these principal components as a model under a General Linear Model (GLM) approach. It is important to note that *functional connectivity* is only related with the temporal correlation between neuronal activations, however it does not say anything about how these temporal patterns of activation are mediated. The matter of how cortical regions interact with each other is the main objective of *effective connectivity*.

Effective connectivity is "the influence one neural system exerts over another" (Friston, Frith et al. 1993). This concept can be thought of as the neural connectivity at synaptic level. "*Effective connectivity* should be understood as the simplest possible circuit-diagram that would replicate the observed time relationship between the recorded neurones"(Aertsen and Preissl 1991). For any set of neural activations recorded, there are a huge number of possible models to explain the temporal correlation. For instance, with 3 areas of activation, it is possible to construct 25 different models, therefore the consequence of this is that an effective connectivity model is very dependant on the *apriori* known anatomical connections. *Effective connectivity* models are divided into a neuro-anatomical model (which areas are structurally connected) and a mathematical model (how areas are connected functionally).

1.4.1.1 Linear neural relationships.

Mathematical models of effective connectivity can be expressed as a multiple linear regression problem, where the variance in the post-synaptic neurone is explained as a weighted sum of variance in the pre-synaptic neurones (Buchel and Friston 2000).

Let us imagine one neuronal region y which receives input from 5 different neural regions x_n.

$$y = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4 + \theta_5 x_5 + \psi \qquad (1.3)$$

$$y = \sum \theta_i x_i + \psi \qquad (1.4)$$

In matrix notation:

$$y = \theta x + \psi \qquad (1.5)$$

y represents a column vector containing the time course of the post-synaptic neurone, x is a matrix which contain the time courses of the pre-synaptic neuronal input, θ is a vector of weights, which represents the strength of connections between the post-synaptic neurone and each of the pre-synaptic inputs. Finally, ψ represents the error on y, i.e. the part of y that can not be explained by the system. Using least-squares it is possible to solve for the values of the weights θ.

$$\theta = (x^T \cdot y)(x^T \cdot x)^{-1} \qquad (1.6)$$

1.4.1.2 Non-linear neural relationships.

Although many neural systems can be modelled using linear representations (Friston, Frith et al. 1993), some neural interactions behave in a non-linear fashion (Friston, Josephs et al. 1998). In these cases it is necessary to extend the previous representation to capture the non-linearity of the system. A good example of a non-linear neural system is the modulatory interaction between V1 and V2 regions in the visual cortex.

Reversible cooling experiments in areas V1 and V2 of the visual cortex of the monkey have proved that neural activity in V2 depends on inputs from V1 (Sandell and Schiller 1982).

On the other hand neural activity in V1 is modulated by re-entrant connections from V2. The experiment consisted of cooling V1 while recording from V2 and vice versa. In the first case, it was demonstrated that activation in V2 depends on the activity of V1, however, in the second case the activity in V1 was affected (modulated) by the lack of activity in V2. This means that the activity in V1 depends not only on the main input of V1, but also in the modulatory effect that V2 exerts over V1. On the contrary, the strength of this modulatory interaction depends also on the input from V1 to V2.

This non-linear behaviour in V1 was studied by Friston *et al* (Friston, Ungerleider et al. 1995) using fMRI in humans. They proposed an extension to the linear model by including a non-linear interaction term.

$$y = \sum \theta_i \, x_i \; + \sum \beta_i \, x_i \, y \; + \psi \qquad (1.7)$$

$$y = \theta \, x + \beta \, x \, y + \psi \qquad (1.8)$$

In this case θ represents the obligatory connection because any change in x_i produces a change of magnitude $\theta_i \, x_i$ in y. β is more a modulatory connection in the sense that the rate of change produced by this term not only depends on the value of x_i but also on the value of y. Although this equation is linear in the parameters, it captures the non-linearity of the system through to the interaction term included in the second term of the equation (Kenny and Judd 1984). The structure of interactive terms could be as complex as necessary, i.e. it could include terms of higher order or even include the influences of past activities on the present.

$$y(t) = \sum \theta_i \, x_i(t) + \; \sum \beta_i \, y(t\text{-}1) \; + \psi(t) \qquad (1.9)$$

However, it is important to have in mind that every model is a simplification of the real system and in practice it is necessary to compromise between accuracy, complexity, and interpretability.

In order to identify the non-linear relationships among regions we propose the use of a Non-linear AutoRegressive Moving Average with eXogenous inputs (NARMAX) algorithm (Billings, Chen et al. 1989). The underlying idea behind this algorithm is to represent the model as a linear-in-the-parameters non-linear difference equation system and estimate the parameters doing an orthogonal decomposition. This non-linear system identification approach is described in chapter 3.

One of the most important applications of *effective connectivity* is assessing changes in effective connectivity connections as a function of task requirements, for example, Buchel (Buchel, Josephs et al. 1998) shows how the coefficients between neural connections in the visual cortex change as a function of attentional requirement. In the next section a general framework to define and assess models of effective connectivity is introduced.

1.4.2 Structural equation modelling of neural networks.

Structural Equation Modelling (SEM) is a statistical technique to assess causal models that define relations amongst variables in terms of the analysis of covariance (Bollen 1989). It is also called path analysis or analysis of variance-covariance structures. SEM permits analysis of complex models of *effective connectivity* in which many regions could be included. The advantage of this approach over others like ANCOVA is that SEM evaluates the model as a complete network (where the output of one unit can be the input to another), whereas ANCOVA only considers individual multi-input single-output relationships (as shown above).

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 structural equations, for example consider the following model that assumes the connectivity between three variables A, B and C:

$$B = \theta_1 A + \psi_2 \qquad (1.10)$$

$$C = \theta_2 A + \theta_3 B + \psi_3 \qquad (1.11)$$

Matrix notation:

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

Vector notation:

$$v = K\,v + \psi \qquad (1.12)$$

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 (1.12) we have:

$$v = (I - K)^{-1}\psi \qquad (1.13)$$

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

$$S = v \bullet v^T \qquad (1.14)$$

Then substituting:

$$\Sigma(\theta) = (I - K)^{-1}\psi\,((I - K)^{-1}\psi)^T \qquad (1.15)$$

The estimation of the parameters in θ that minimise the difference between S and $\Sigma(\theta)$ is usually done using the Maximum Likelihood fit 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. A fuller explanation of the mathematical properties of SEM will be given in the next chapter.

Although the concept of SEM is not new (Wright 1918; Wright 1921; Wright 1934) having been applied widely in econometrics (Chow 1983), the idea of using it to assess models of *effective connectivity* using functional brain imaging is relatively recent (McIntosh and Gonzalez-Lima 1994b). McIntosh and Gonzalez-Lima were the first investigators who applied SEM to asses neural interactions (McIntosh and Gonzalez-Lima 1992). They measured 2-deoxyglucose uptake in the visual system of rats during either patterned light or darkness. With this study they demonstrated the ability of SEM to assess changes in functional strengths between anatomical connections within the visual cortex.

Since then, many studies have been made using SEM with PET or fMRI data (McIntosh, Grady et al. 1994c; Nyberg, McIntosh et al. 1996; Buchel and Friston 1997; Jennings, McIntosh et al. 1998; McIntosh, Cabeza et al. 1998; Bavelier, Tomann et al. 2000; Bullmore, Horwitz et al. 2000; Friston and Buchel 2000). The most common examination followed by the application of SEM in neuro-imaging is the staked model analysis. This process permits one to evaluate changes in *effective connectivity* under different experimental conditions (Friston, Frith et al. 1993).

Classical application of SEM has been performed in the Social Sciences where the connections among variables are usually hypothetical with models that contain variables that cannot be measured directly. In neuroimaging, the connection between variables are anatomically constrained and the measurement of each variable can be made directly with functional imaging. These features make the application of SEM to functional imaging data a useful tool to explore models of *effective connectivity* on the human brain.

In a preliminary study seven subjects were scanned in a 1.5 T whole-body MRI scanner (chapter 4). The subjects wore red/green anaglyph glasses, and were instructed to fixate on a point in the middle of the screen 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. *Stationary*: two hundred and fifty dots were randomly positioned within the circular field of view. *Motion*: the same set of dots moving radially, changing from expansion to contraction every 3 seconds. *Stereo*: the same number of dots positioned in depth (red/green anaglyph stereogram) forming a 3D cone structure. *StereoMotion*: the previously *Stereo* and *Motion* stimuli were combined. The data were pre-processed and analysed using SPM99. The statistical parametric map of one subject with $P < 0.05$ (corrected) is shown projected on the anatomical images below (fig. 1.17).
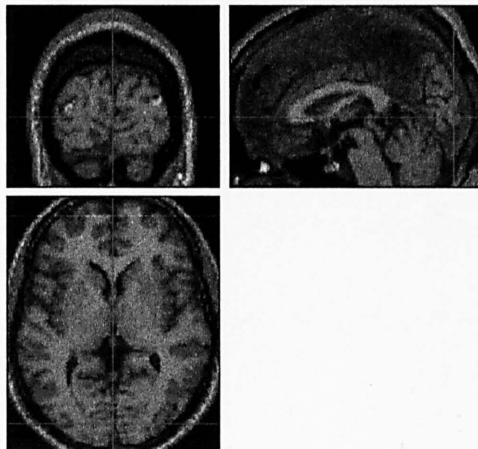
Figure 1.17 Axial, sagittal and coronal views of activations at point $[0 - 90\ 2]_{MNI}$. The anatomical images are used to show the location of the regions sensitive to: Motion V5 (blue), Stereo V3B (red) and Stereo and Motion V3A (yellow).

We tested the hypothesis that there is effective connectivity (Friston, Frith et al. 1993; Friston 1994) between these areas (V5: motion, V3B :stereo and V3A:stereo and motion) using the model in figure 1.18. Path analysis was done using the first principal component time series from the regions involved. First the NARMAX algorithm (Billings, Chen et al. 1989; Chen, Billings et al. 1989) was used to identify the path coefficients and the non linear term from the times series of the regions.
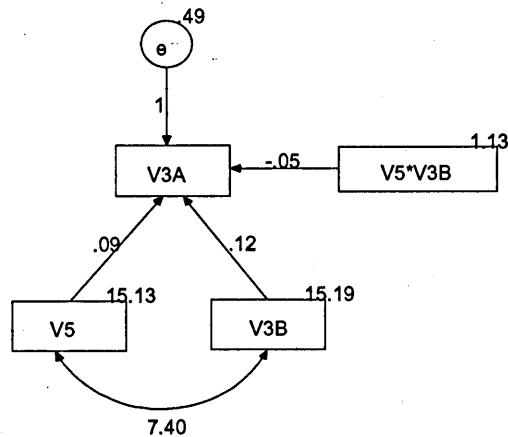
Figure 1.18 Effective connectivity model. The model identified using the NARMAX algorithm was evaluated using SEM (AMOS 4.01). The model fitted the data ($\chi^2 = 0.2$, dof=2, p= 0.99).

Next the statistical fit of the model to the data was evaluated using structural equation modelling and the AMOS software. According to the statistical test, the proposed model is able to account for the interactions among regions, particularly the non-linear relationship occurred when the inputs of V3A are given at the same time. For a fuller explanation of this pilot study consult chapter 4.

1.5 Conclusions.

Many psychophysical studies in humans have been made to understand the principles that govern stereoscopic vision. However, not much work has been done to understand the binocular perception at a physiological level. Electrophysiological studies in monkeys suggest that the MT/V5 region, commonly related to process directional motion, is involved in processing binocular disparities.

Modern non-invasive neuroimaging tools, like fMRI or PET, have opened new pathways to explore the functional anatomy of the human brain. Despite the novelty of these techniques, few studies have researched stereo vision and the results reported in them show inconsistencies.

Although some advances have been made in the search to understand the functional anatomy of stereopsis, much work still has to be done before obtaining firm evidence regarding the cortical regions and its functional interactions which produce stereoscopic vision.

Furthermore, although most of the studies are in agreement with the fact that a network of cortical regions is involved in performing this task, no one has reported a connectivity analysis to assess the interaction between the stereo-sensitive regions. Therefore, the interregional connectivity of stereo sensitive regions requires further investigation.

The purposes of this work can be summarised in four objectives: i) Using functional magnetic resonance imaging to study the cortical regions involved in human stereovision. ii) Comparing the results of our experiments (testing the sensitivity of humans' V5 to stereo stimulus) with those reported on similar studies with monkeys. iii) Assessing the interactions between stereo sensitive regions by applying an effective connectivity approach. iv) Developing a tool to assess dynamic non-linear models of effective connectivity.

V1.-
Edges (Tootell, Hadjikhani et al. 1998),
Texture (Kastner, De Weerd et al. 2000),
Motion (Sunaert, Van Hecke et al. 1999),
Random motion (Paradis, Cornilleau-Peres et al. 2000),
Absolute disparities (DeAngelis and Newsome 1999).

V2.-
Texture (Kastner, De Weerd et al. 2000),
Random motion (Paradis, Cornilleau-Peres et al. 2000),
3D surfaces (Bakin, Nakayama et al. 2000).

V3.-
Optic flow (DeJong, Shipp et al. 1994),
Shape from motion (Paradis, Cornilleau-Peres et al. 2000),
Second order motion (Smith, Greenlee et al. 1998a),
Motion and direction (Felleman, Burkhalter et al. 1997),
Direction selective (Burkhalter, Felleman et al. 1986).

V3A.-
Texture (Kastner, De Weerd et al. 2000),
Shape from motion (Paradis, Cornilleau-Peres et al. 2000),
Motion (Tootell, Mendola et al. 1997),
Motion (Sunaert, Van Hecke et al. 1999),
Coherent motion (Braddick, O'Brien et al. 2001).

V4.-
Colour (Zeki, Watson et al. 1991; Zeki, Watson et al. 1993),
Colour (Lueck, Zeki et al. 1989),
Form and colour (Burkhalter and Van Essen 1986),
Form (Ghose and Ts'o 1997),
Form (Tanaka, Saito et al. 1991),
Texture (Kastner, De Weerd et al. 2000),
Optic flow (DeJong, Shipp et al. 1994).

VP.-
Form and colour (Burkhalter and Van Essen 1986),
Second order motion (Smith, Greenlee et al. 1998a).

V5.-
Coherent motion (Uusitalo, Virsu et al. 1997),
Speed and direction (Zeki 1974),
Coherent motion (Braddick, O'Brien et al. 2001),
Motion (Watson, Myers et al. 1993),
Depth from motion (Orban, Sunaert et al. 1999),
Depth from motion (Xiao, Marcar et al. 1997),
Motion (Sunaert, Van Hecke et al. 1999),
Coherent motion (Zeki, Watson et al. 1991),
Stereo disparities (DeAngelis and Newsome 1999).
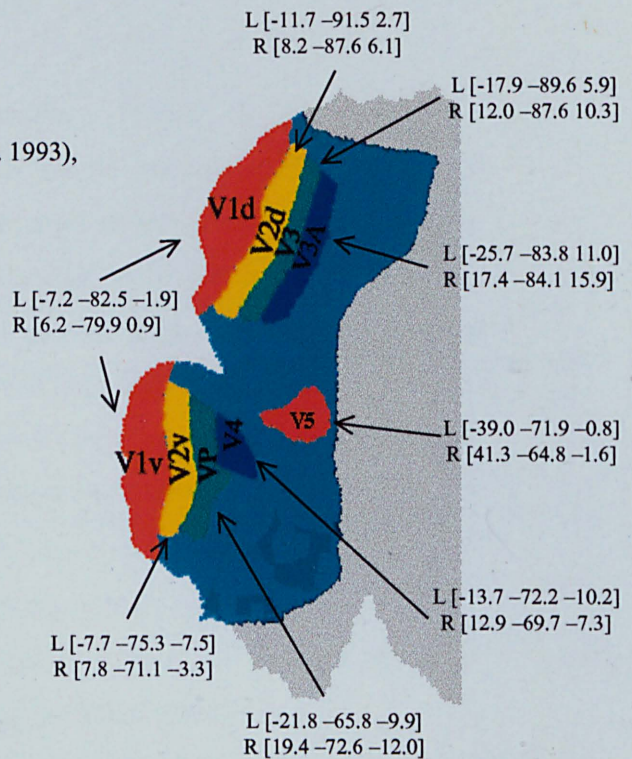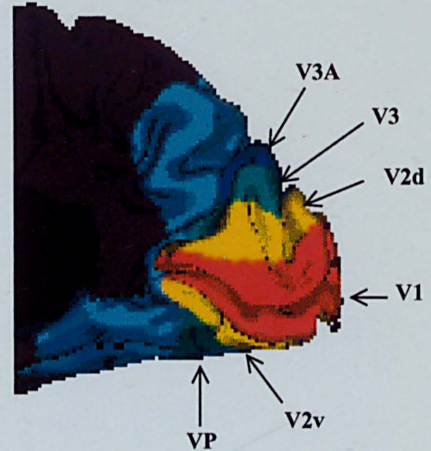
Figure 1.19 Functional defined visual regions. Left column summarises the functiona properties reported in each visual region. Flat representation of the right hemisphere shows the Talairach space coordinates of visual regions in both hemisphere.

Chapter 2

Neural Structural Equation Modelling

Abstract.

Structural Equation Modelling (SEM) is a mathematical technique to assess models that define relations among variables in terms of the analysis of covariance. Although SEM theory is not new, the application of this technique to quantify functional relationships among neural regions in the brain is a relatively new area of research. The methodology to apply this mathematical technique to represent models of effective connectivity using fMRI data is introduced. Finally, the performance of this technique is assessed using artificial fMRI time series.

2.1 Introduction.

Structural equation modelling (SEM) is a statistical technique to assess interactions within causal models (Bollen 1989). The basic idea of SEM comes from Wright's studies in path analysis (Wright 1921; Wright 1934). Path analysis technique consists in drawing a graphical model which specifies causal relationships between variables, then, the graphical model is transformed into a simultaneous equation system.

In the context of neuroimaging, variables represent cortical regions and causal relationships represent the strength of influence of one region has into another through their anatomical connection (McIntosh and Gonzalez-Lima 1994b).

It is important to clarify that the specification of the path diagram implies an *a priori* knowledge of the relationships between variables. SEM does not discover the structure of the model; it only assesses the goodness of the proposed model to fit the observed data. Moreover, "anatomical foundation is the key feature for neural structural equation modelling" (McIntosh and Gonzalez-Lima 1994b).

One important characteristic of SEM is that it evaluates the complete model as a whole unit, that is to say, it estimates all the relationships (covariance matrix) implied by the model to assess its goodness of fit. This is a huge difference with respect to other methods, like regression analysis or ANOVA, in which only individual dependent variables are explained from the set of predictors.

The working hypothesis of SEM is that if the proposed model is the correct one, the observed covariance matrix (Σ) and the covariance matrix implied by the model ($\Sigma(\theta)$) are equal. Finding the values of parameters (θ) which minimise the difference $\Sigma - \Sigma(\theta)$, provides an estimate of how well the model can represent the data.

Although SEM is not new, the application of this technique to quantify functional relationships among neural regions from fMRI or PET neuroimaging data is a new area of research. The major application of SEM has been in the Social Sciences where the connections among variables are usually causal-hypothetical with models that contain variables that cannot be measured directly. In neuroimaging, the connections among the variables representing brain regions are anatomically constrained and the measurement of each variable can be made directly with functional imaging. These features makes the application of SEM to functional imaging data an useful tool to explore the functional organisation in the human brain.

2.2 Model notation.

The two main parts that make any structural equation model are the path diagram and the structural equations implied by the model. SEM is able to work with *latent* and *observed* variables. Most of the applications of SEM have been made using latent variables, i.e., hypothetical variables that cannot be measured directly, such as loyalty, motivation, affection, etc. in contrast to variables such as income, educational qualification, age, etc. Figure 2.1 shows an example of a complete SEM model (*measurement model*) which includes latent and observed variables.

Square boxes represent observed variables which are indicators of latent variables (ellipses). The grey square represents the structural model that defines the causal relationships between the latent variables. Circles represent two kinds of different noises: *system noise* (e.g. inputs coming from regions not included in the model) and *measurement noise* (e.g. variation in the signal due to the limitations of the equipment). One of the advantages of the use of SEM in neuroimaging is that the variables representing specialised brain regions involved in the model are observed, i.e. are physically recorded by the fMRI scan using the BOLD response as a indicator of neural activity. For that reason the present chapter explains only the basis of structural equation modelling with observed variables.
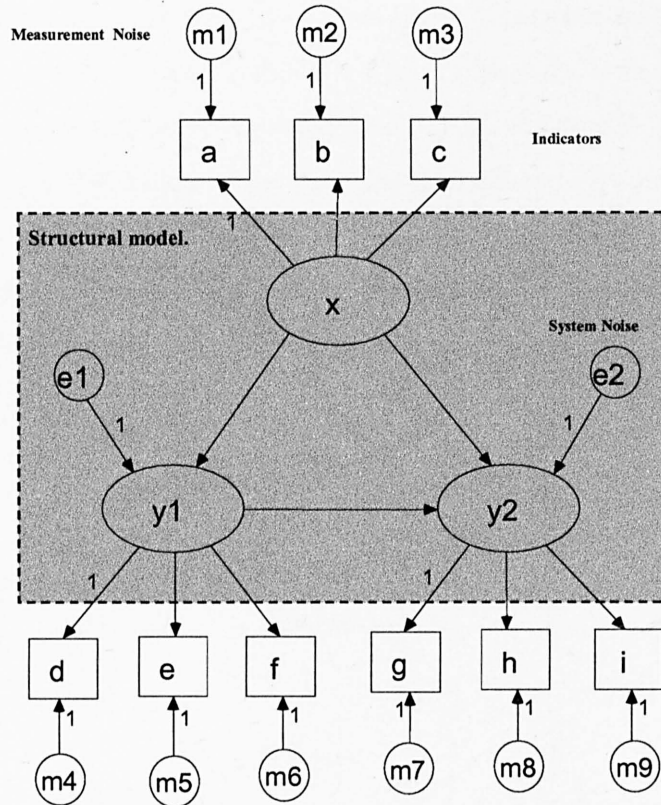
Figure 2.1 Structural equation measurement model.

2.2.1 Terminology and notation.

Observed variables can be *exogenous* or *endogenous*. Exogenous variables are independent of any other variable. For example in figure 2.3, x_1 is an exogenous variable because it does not receive any influence from another variable, i.e. its variance is not explained by the system. Endogenous variables always receive influence from other variables, their value is determined by inputs within the model.

In figure 2.3, y_1 and y_2 are endogenous variables. Because the value of any endogenous variable is explained by the behaviour of one or more variables, an associated error has to be introduced to account for the amount of variance that is not explained by its predictors. These error terms can be thought of as a mixture of biological and measurement noise.

It is assumed that this error term has a zero mean (Gaussian noise) and it is not correlated with any other variable in the model. Perhaps the strongest assumption of SEM is that the system noise is propagated through the model. This assumption is often violated by fMRI data where the measurement error frequently dominates the error term. It will be shown in subsequent MonteCarlo simulations, the violation of this assumption leads to under-determined models. A summary of the symbols used in path diagrams is presented in the figure 2.2.

⬭	Latent variable.
▭	Observed variable.
◯	Error term.
⟶	Causal relationship.
↻	Correlation between variables.
⇄	Reciprocal causation.
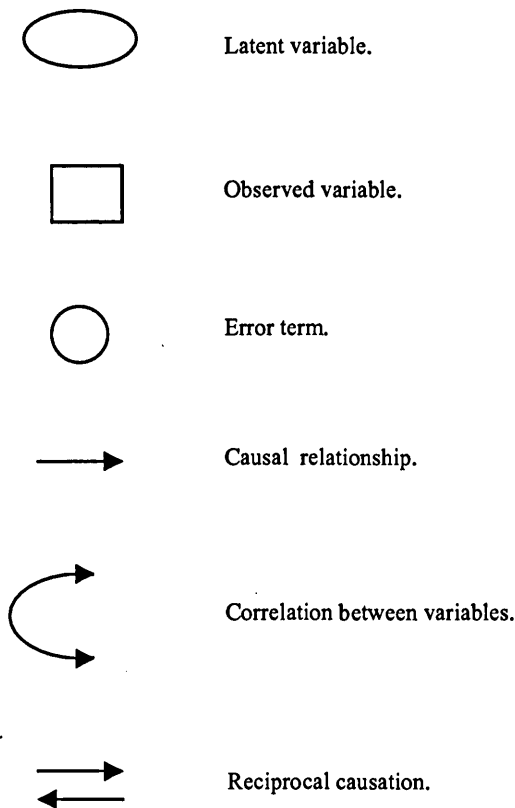
Figure 2.2 Path analysis symbols.

2.2.2 Structural equation models with observed variables.

The general representation of a SEM separates the model into endogenous-endogenous relationships, exogenous-endogenous relationships and error terms:

$$y = B\,y + \Gamma\,x + \zeta$$

where:

B Matrix of endogenous - endogenous coefficients.

Γ Matrix of exogenous - endogenous coefficients.

y Vector of endogenous variables.

x Vector of exogenous variables.

ζ Vector of error terms.

Let us imaging that we have a model defined by the following path diagram.
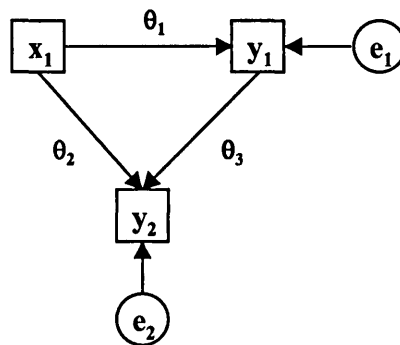
Figure 2.3 Path diagram.

The structural equations related to this model are:

$$y_1 = \theta_1 \, x_1 + e_1$$

$$y_2 = \theta_2 \, x_1 + \theta_3 \, y_1 + e_2$$

Where x_1 is an exogenous variable and y_1 and y_2 are endogenous variables.

The matrix notation of the general representation of SEM is:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \theta_3 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} [x_1] + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$y \quad = \quad B \quad y \quad + \quad \Gamma \quad x \quad + \quad \zeta$$

In this case **B** contain the parameter which defines the relation between the two endogenous variables. Γ contains the parameter of the relation between exogenous and endogenous variables. The positions in the matrixes which contain zeros represent the fact that no causal relationship exists between the variables involved. The covariance matrix of **x** is called Φ and the covariance matrix of ζ is called Ψ. For this example they are equal to:

$$\Phi = [\phi_1] \qquad\qquad \Psi = \begin{bmatrix} \psi_1 & 0 \\ 0 & \psi_2 \end{bmatrix}$$

2.2.3 Recursive and non-recursive models.

There are two types of structural equation models, *recursive* and *non-recursive*. Contrary to the intuition and common usage, recursive models differ from non-recursive models in the fact that they *do not* contain reciprocal causation or feedback loops. The previous example corresponds to a recursive model. A recursive model implies a lower triangular **B** matrix, that is to say, there are only unidirectional relationships and no reciprocal relationships are considered. Recursive models assume a diagonal covariance matrix of error terms (Ψ), i.e. the errors are orthogonal. The next path diagram shows a non-recursive model.
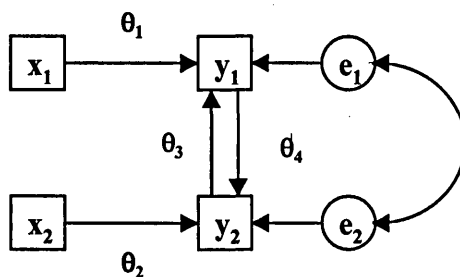
The matrix notation on the general representation is:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & \theta_3 \\ \theta_4 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{bmatrix} [x_1 \, x_2] + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$y \;=\; B \qquad y \;+\; \Gamma \qquad x \;+\; \zeta$$

In this case the covariance matrix of exogenous (Φ) and the covariance matrix of error terms (Ψ) are:

$$\Phi = \begin{bmatrix} \phi_1 & \\ \phi_2 & \phi_3 \end{bmatrix} \qquad \Psi = \begin{bmatrix} \psi_1 & \\ \psi_2 & \psi_3 \end{bmatrix}$$

2.3 Implied covariance matrix.

The main hypothesis of SEM is that the observed covariance matrix (Σ) is equal to the implied covariance matrix ($\Sigma(\theta)$).

$$\Sigma = \Sigma(\theta)$$

Ordering the observed variables in endogenous (y) and exogenous (x), the observed covariance matrix is constructed as:

$$\Sigma = \begin{bmatrix} yy' & yx' \\ xy' & xx' \end{bmatrix} = \begin{bmatrix} \Sigma_{yy}(\theta) & \Sigma_{yx}(\theta) \\ \Sigma_{xy}(\theta) & \Sigma_{xx}(\theta) \end{bmatrix}$$

Where x' stands for x transpose. The implied covariance matrix is constructed using the general representation of SEM, which as explained above, contains the structural relationship of the model being assessed. Let us construct the section $\Sigma_{yy}(\theta)$ of the implied covariance matrix. From the general representation:

$$y = B\,y + \Gamma\,x + \zeta$$

$$y = (I - B)^{-1}\,(\Gamma\,x + \zeta)$$

$$\Sigma_{yy}(\theta) = E(yy')$$

$$= E[(I - B)^{-1}\,(\Gamma\,x + \zeta)((I - B)^{-1}\,(\Gamma\,x + \zeta))']$$

$$= E[(I - B)^{-1} (\Gamma x + \zeta)(x'\Gamma' + \zeta') (I - B)^{-1'}]$$

$$= (I - B)^{-1} (E(\Gamma xx'\Gamma') + E(\Gamma x\zeta') + E(\zeta x'\Gamma') + E(\zeta\zeta')) (I - B)^{-1'}$$

$$= (I - B)^{-1} (\Gamma\Phi\Gamma' + \Psi) (I - B)^{-1'}$$

Then $\Sigma_{xx}(\theta)$ is equal to:

$$\Sigma_{xx}(\theta) = E(xx') = \Phi$$

Finally $\Sigma_{xy}(\theta)$ is:

$$\Sigma_{xy}(\theta) = E(xy')$$

$$= E[x((I - B)^{-1} (\Gamma x + \zeta))']$$

$$= \Phi\Gamma' (I - B)^{-1'}$$

So, the complete implied covariance matrix is:

$$\Sigma(\theta) = \begin{bmatrix} (I - B)^{-1} (\Gamma\Phi\Gamma' + \Psi) (I - B)^{-1} & \Gamma(I - B)^{-1}\Gamma\Phi \\ \Phi\Gamma' (I - B)^{-1'} & \Phi \end{bmatrix}$$

The equality hypothesis $\Sigma = \Sigma(\theta)$ implies that for each element in the observed covariance matrix there exists one element in the implied covariance matrix whose values are the same. The solution of the simultaneous equation system is given by the estimation of the values in θ which minimise the difference between the matrixes. The most common method to identify the values of parameters θ is maximum likelihood (ML) or least squares (LS).

As in all numerical methods to solve equation systems, there are some criteria that have to be reached to successfully identify the structural parameters. Most of the time these constraints have to be introduced in the system in order to obtain unique solutions. The implementation of the SEM routines was made in MATLAB 5.3 for Windows NT 4.0. For a detailed description of the code consult the appendix at the end of the thesis.

2.4 Identification.

According to the general hypothesis of SEM, each element of the observed covariance matrix is a function of the implied covariance matrix. Each function represents an equation that contains parameters that need to be estimated. The identification issue consists in finding unique values for the parameters of the equation system. If it is possible to find one and only one value for each parameter of the model, the model is *identified*. However, if it is not possible to find a value or a unique value, then these parameters are *unidentified* and the model becomes *under-determined*.

2.4.1 The *t* rule (Bollen 1989).

If the number of equations is smaller than the number of unknowns, the system does not have unique solution. For example, in an observed covariance matrix with p endogenous and q exogenous variables there are $t = \frac{1}{2}(p+q)(p+q+1)$ non-redundant elements. If t is smaller than the number of unknown parameters in the model, the model becomes under-determined (the *t* rule).

The *t* rule is a necessary but not sufficient condition to get identification, that is to say, although the *t* rule is satisfied, there are other reasons for which a model can be under-identified. "Model identification in structural equations with observed variables is not possible without placing restrictions on model parameters" (Bollen 1989).

2.4.2 Fixing error variances.

Applications of SEM to network analysis in functional brain imaging frequently have problems of model under-determination due to the greater number of unknown parameters. One common practice is to fix the values of Ψ with arbitrary constants. This not only permits identification of a model but also improves its goodness of fit by increasing its degrees of freedom.

For example, McIntosh *et al* (McIntosh and Gonzalez-Lima 1992; McIntosh and Gonzalez-Lima 1994a; McIntosh and Gonzalez-Lima 1994b; McIntosh, Grady et al. 1994c), fix the residual values at 35 to 50 % of the variance of the endogenous variable associated with it. Similarly, they pre-compute the variance of the exogenous variables to fix their value and so reduce the number of parameters to estimate in the matrix Φ.

Bullmore *et al* (Bullmore, Horwitz et al. 2000) proposed an empirical estimation of Ψ using principal component analysis (PCA). They applied PCA to the set of times series from each region (*i*) recorded in a group of subjects. Then used the following equation:

$$\psi_i = 1 - \frac{\lambda_1^2}{\sum_{j=1}^{m} \lambda_j^2}$$

Where the variance magnitude of each *j* component (eigenvector) is represented by their eigenvalue (λ_j), *m* is the number of components specified in the PCA analysis (Lai and Fang 1999). The idea is that λ_1 represents the amount of 'neural variance' contained in the region *i* , ψ_i is 1 minus the normalised ratio of the noise variance in region *i*. Finally, it is important to remark that different conclusions can be reached from the same group of data, depending on the assumptions used to estimate the error term (Bullmore, Horwitz et al. 2000).

2.5 Estimation.

Estimation of a model is the process of finding the values of the unknown parameters in the system equations. Thus the estimation of the unknown parameters in **B, Γ, Φ, Ψ** is made by finding values that solve the hypothesis $\Sigma = \Sigma(\theta)$. Because the population observed covariance matrix is not know, the best estimation of it is the sampled covariance matrix (**S**). The solution of the system is obtained by finding the parameters which minimise the difference between the sampled covariance matrix and the implied covariance matrix (S-$\Sigma(\theta)$).

2.5.1 Fitting function.

The criteria to assess how big is the difference S-$\Sigma(\theta)$, is given by a fitting function. The most common used fitting function is Maximum Likelihood.

$$F_{ML}(S, \Sigma(\theta)) = \log|\Sigma(\theta)| + \text{tr}(S\Sigma^{-1}(\theta)) - \log|S| - (p+q)$$

In which, $F_{ML}(S, \Sigma(\theta))$ is a real number greater than or equal to zero, $F_{ML}(S, \Sigma(\theta)) = 0$ means that the proposed model represents a perfect match of the observed covariance matrix, and thus, the data supports the proposed model. For example, if S=$\Sigma(\theta)$, substituting in $F_{ML}(S, \Sigma(\theta))$, then $F_{ML} = 0$. Generally speaking, iterative numerical techniques are used to find the parameters which minimise the Maximum Likelihood fitting function. For a complete derivation of the ML fitting function (F_{ML}) consult (Bollen, 1989, p131:135).

There are other fitting functions, for example Unweighted Least Squares (F_{ULS}) which minimise the sum of the squares of the observed and implied matrix (Σ-$\Sigma(\theta)$).

$$F_{ULS} = \tfrac{1}{2}\text{tr}[(S - \Sigma(\theta))^2]$$

The disadvantage of this fitting function is that it does not lead to the asymptotically best parameter estimate. The advantage is that it does not assume normal distribution on the variables (Bollen).

2.5.2 Goodness of fit.

It is important to remember that SEM does not discover the model which best fits the observed data, it only says how well the proposed model fits the data. It could be the case that more than one model shows a good fit with the data. In which case, the researcher has to use his expertise to select the correct model.

The statistical validation of the goodness of fit of the model is made using the fact that the value of the minimised maximum likelihood fitting function multiplied by the length of the sampling minus one $((N-1)*F_{ML})$ is asymptotically distributed as chi-square (χ^2), with $\frac{1}{2}(p+q)(p+q+1)-k$ degrees of freedom, where k is the number of free parameters in $\mathbf{B}, \mathbf{\Gamma}, \mathbf{\Phi}, \mathbf{\Psi}$.

The statistical test is given under the null hypothesis: $\mathbf{H_o}: \Sigma=\Sigma(\theta)$. It is assumed that the observed variables are normally distributed and N is large (> 200) (Bollen 1989). If the observed discrepancy $\Sigma-\Sigma(\theta)$ has a small probability under this distribution then $\mathbf{H_0}$ is formally rejected. It is important to notice that a small number of observations (N) overestimates the goodness of fit, for example, the same model with the same data can be rejected or accepted depending on the length of the time series used in the analysis (Bullmore, Horwitz et al. 2000).

There is no general agreement about the criteria to define the value of the threshold (*p-value*) for refutation. This is because the *p-value* assigned to a given model can not be considered as an absolute estimator, due to its dependence on the length of the sampling (*N*) and on the number of *a priori* estimations of residual term variances (Bullmore, Horwitz et al. 2000).

Furthermore, there is a lack of standardised ways of reporting SEM analysis in neuroimaging, and most of the cases, publications do not contain enough information to replicate or even interpret the results.

Sometimes it is important to assess the goodness of fit not only in terms of how well it is able to represent the observed data, but also in terms of how easy it is to interpret it. By definition, every model is a simplification of a system and it is necessary to find a balance between accuracy and complexity. In these cases, a more appropriate approach is to use Akaike Information Criterion (AIC) which favours the simple over the complex model (Bollen, 1989).

$$AIC = x^2 + 2t$$

Where t is the number of parameters in the model and x^2 is the value of the chi-square estimator.

2.6 Stacked models.

Recent studies of brain functionality suggest the idea that cognitive process are not confined to specific cortical regions, instead, networks of cortical regions are responsible to perform specific tasks (McIntosh 2000). The functionality for certain cortical regions is dependent on the interactions with other regions. Thus one cortical region can be involved in more than one cognitive function. In order to investigate the changes in connectivity in the cognitive activity, it is necessary to assess the behaviour of a given neural system under different task requirements.

The most common application of neural structural equation modelling is to assess a model of effective connectivity using the same group of subjects under different experimental conditions or by using the same experimental condition with two groups of subjects. The hypothesis behind this approach, called *stacked model analysis*, is that the same set of parameters within the structural model, are adequate for both sets of experimental data.

The procedure consists in defining a *null model,* in which parameters are constrained to be equal between conditions and an *alternative model,* in which parameters are allowed to differ.

The statistical inference is made by subtracting the χ^2 value from the models and assessing its *p-value* with degrees of freedom equal to the difference of degrees of freedom between null and alternative models. If the *p-value* is not smaller than a threshold (e.g. 0.05), then the difference in functional connectivity between experimental conditions is not statistically significant. In other words, the same model (with same parameter values) can explain the data under the different experimental conditions.

There is no general agreement about the need to assess the validity of the model to account for the different data sets. For most researchers it is enough to prove that the parameters of the model are different under different experimental conditions, without worrying if the individual data sets achieve a statistical valid goodness of fit.

2.7 Non-linear models.

Although most of the neural interactions can be represented by linear approximations, some neural systems have non-linear relationships between components (Friston, Josephs et al. 1998). Modulatory interactions under hierarchical organisations of cortical regions are a common example of it.

2.7.1 Interaction terms.

Let us imagine a neural system in which one cortical region (y_1) receives inputs from two cortical regions (x_1, x_2).
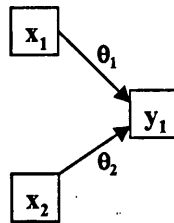
Suppose in an experiment when the inputs are received individually (x_1 or x_2), the activity in the dependent region (y_1) has a linear relation response with each of its predictors ($y_1 = \theta_1 x_1 + \theta_2 x_2$).

However, in another experiment when both inputs are present, it is found that the response is non-linear ($y_1 = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2$). This kind of non-linear relationship can be represented by creating new variables (interaction terms) to account for the non linearity (Kenny and Judd 1984; Schumacker and Marcoulides 1998).

For the previous example it is possible to create a new variable from x_1 and x_2 called $x_1 * x_2$. Because the product of x_1 and x_2 is going to be highly correlated with them, it is necessary to orthogonalise it, i.e. to delete the part of x_1 and x_2 contained in $x_1 * x_2$ (Buchel and Friston 1997). In this representation, the model is still linear in the parameters but non-linear in the variables (Bollen 1989).
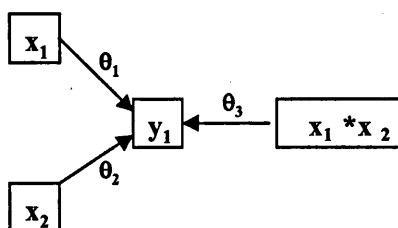
The main issue in this approach is the identification of the interaction terms. There are different techniques to solve this problem, for example, polynomial expansions or Volterra kernels. The next chapter explains in detail a non-linear identification algorithm based on orthogonal decomposition.

2.7.2 Feedback loops.

Another common source of non linear relationships in neural models is found in systems which contain reciprocal causation or feedback loops (Berry 1984). Let us consider the following model.
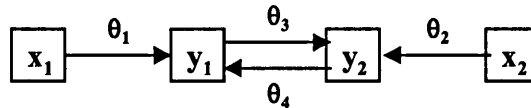
This modality of a non-linear system can be solved using partial correlation analysis or two stages least squares (2SLS). Although most of the anatomical connection between regions are reciprocal, it is not possible to specify all the bi-directional paths because the model easily becomes under-determined, and a compromise between adequacy and complexity is necessary (McIntosh and Gonzalez-Lima 1994b).

Feedback loops can be easily modelled in SEM as non-recursive models. As was shown above, these models imply that **B** is not a lower triangular because the relationships $y_1 \rightarrow y_2$ and $y_2 \rightarrow y_1$ are not symmetrical.

2.8 Structural equation modelling software.

There are many commercial software packages for structural equation modelling analysis. The general operation of these SEM tools require that the user specifies the model to be assessed, including the fixed parameters. The estimation of the model is given by the minimisation function preferred by the user.

The most common packages used to assess SEM are: **AMOS** (Arbuckle and Wothke 1995), **LISREL** (Jöreskog and Sörbom 1989), **EQS** (Bentler 1985), **TETRAD II** (Glymour, Spirtes et al. 1994). All of these above packages perform almost the same procedures as multivariate regression, confirmatory factor analysis, bootstrapping, and multiple population comparisons.

As a difference with the other packages AMOS 4.0 has an excellent graphical interface. TETRAD II has the capability to develop a Bayesian network analysis when the input data is discrete. For further references about the individual characteristics consult (Tabachnik and Fidell 2001).

In the next section examples of structural equation models will be given to illustrate the application of the SEM's theory explained above. All the results shown in these examples were obtained using AMOS 4.0 software.

2.9 MonteCarlo Simulations.

The current section presents examples of data generation, model specification, and model assessment of different types of structural equation models. The data for each example was generated simulating fMRI data. Model assessment was made via AMOS 4.0 software using the Maximum Likelihood estimate.

In the first part of this simulations section, theoretical (ideal) data is used to illustrate the SEM principles explained above. In the second part, the features of fMRI experimental data are replicated as closely as possible and one example model is used to show the behaviour of SEM analysis under non-ideal conditions. For both parts of this section, a stimulus paradigm ABABAB was considered. The boxcar stimulus was 200 data points long with 10 data points per epoch.

2.9.1 Theoretical examples.

The estimated hemodynamic response time series were computed by convolving a boxcar stimulation paradigm with a canonical hemodynamic response function (hrf). The hrf was generated using the statistical parametric map (SPM) software (spm_hrf function) with TR=3 sec (WDCN 1999). Lastly, random noise was added to the signal at 50 % level (figure 2.4). MonteCarlo simulation techniques were used to generate data sets with the statistical properties implied by each structural model, i.e., for given specified values of $\mathbf{B}, \mathbf{\Gamma}, \mathbf{\Phi}, \mathbf{\Psi}$.

2.9.1.1 Linear recursive model.

Refering to the first model presented at the beginning of this chapter.
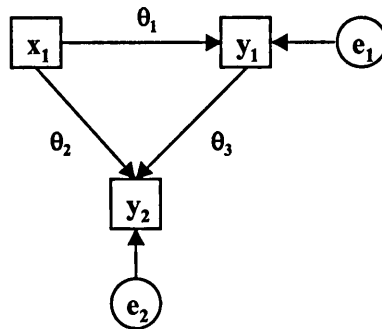
Let us suppose that we want to generate data that fits the following features, modified from Sharma (Sharma 1996).
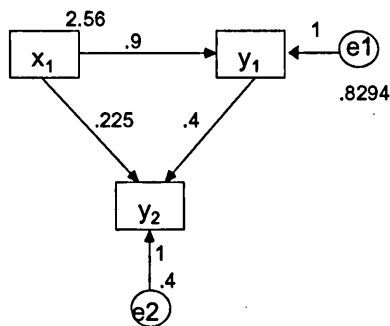
The structural equations related to this model are:

$$y_1 = .9 \, x_1 + e_1$$

$$y_2 = .225 \; x_1 + .4 \, y_1 + e_1$$

Figure 2.5 shows the time series generated by these structural equations.
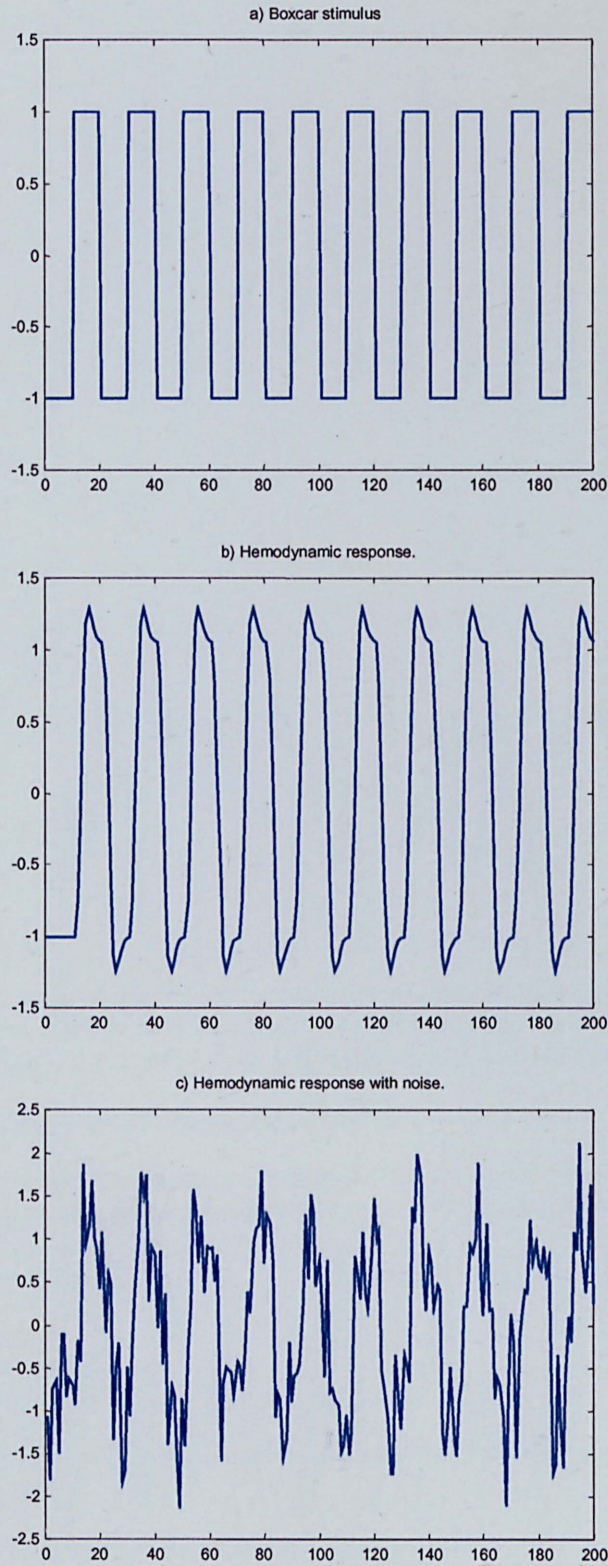
Figure 2.4 MonteCarlo simulation. a) Shows the boxcar paradigm used in the experimental stimulation. b) Represents the box car stimulus convolved with the hemodynamic response function (hfr). c) The hrf with added noise.
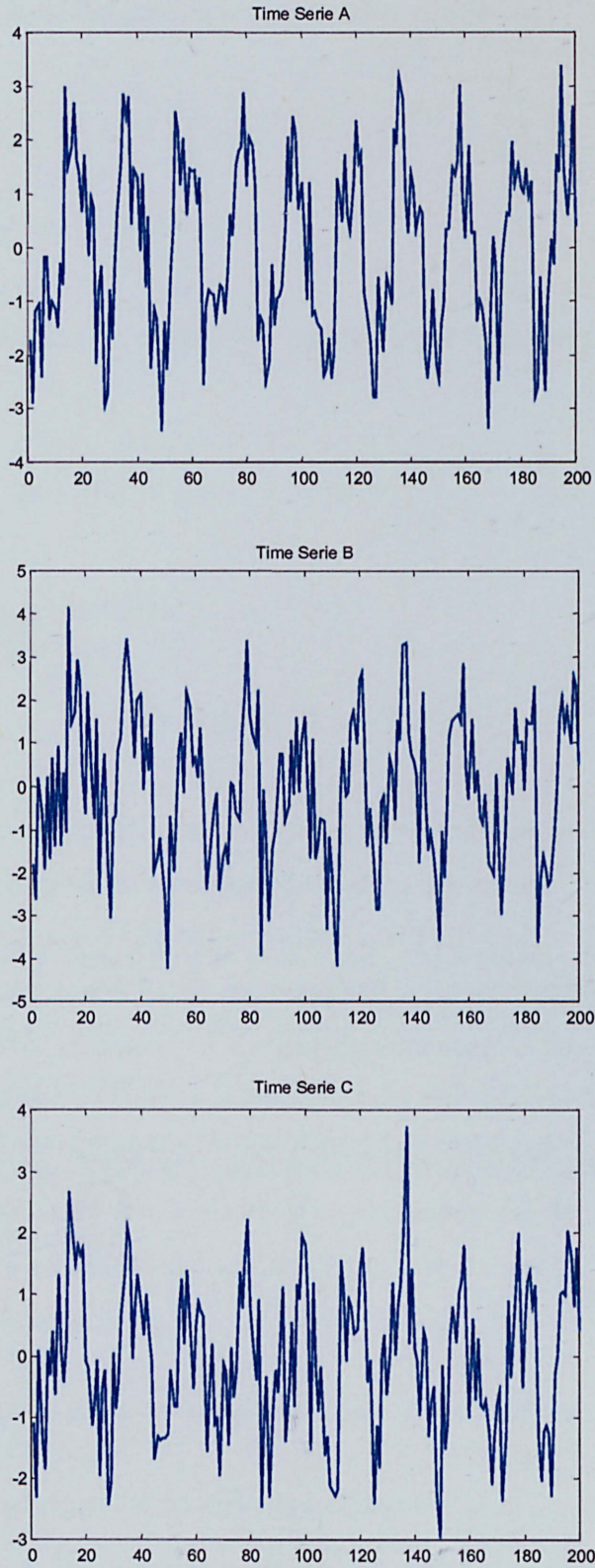
Figure 2.5 Mean time series. The hemodynamic response of each simulated region was obtained using the structural equations specified by the model.

The matrix notation of the general representation of SEM is:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ .4 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} .9 \\ .225 \end{bmatrix} [x_1] + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\mathbf{y} \ = \ \mathbf{B} \quad \mathbf{y} \ + \ \mathbf{\Gamma} \quad \mathbf{x} \ + \ \zeta$$

The exogenous covariance matrix (Φ) and the error term covariance matrix (Ψ) are:

$$\Phi = [2.56] \qquad \Psi = \begin{bmatrix} .8294 & 0 \\ 0 & .4 \end{bmatrix}$$

So the data can be generated as:

$$\mathbf{y} = (\mathbf{I} - \mathbf{B})^{-1}(\mathbf{\Gamma} \, \mathbf{x} + \zeta)$$

Given that the present model contains 1 exogenous variable (p) and 2 endogenous variables (q) the number of non-redundant elements in the observed covariance matrix is computed as $t = \frac{1}{2}(p+q)(p+q+1) = \frac{1}{2}(2+1)(2+1+1) = 6$, then the degrees of freedom of the model is df$=\frac{1}{2}(p+q)(p+q+1)-k = 6-6=0$. Because $k=3$ path coefficients + 2 error variances + 1 exogenous variances. In this case, although the model is just identified, is it not possible to assess the goodness of fit of the model because it has zero degrees of freedom. However, because the variance of the exogenous variable is known *a priori*, it is possible to fix Φ and so reduce the number of unknowns and increase the degrees of freedom. The parameters identified by AMOS 4.0 were:
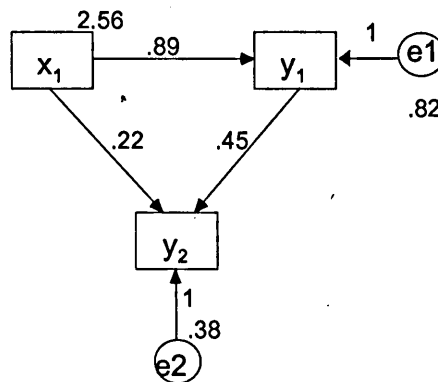
The statistical analysis revealed a chi-square equal to 0.002 with 1 degree of freedom. This corresponds to a probability level equal to 0.96, which means that the model is a good representation of the data.

2.9.1.2 Non-linear recursive model.

Let us imagine that in our previous example the response in y_2 was non-linear to the joint inputs x_1 and x_2, i.e. the response in y_2 to individual inputs from x_1 or y_1 is linear, however, when y_2 receives a joint input from x_1 and y_1, it is not equal to the sum of their individual inputs. As mentioned before these sort of non-linear relationship can be modelled creating a new variable which represents the gain of the interaction between variables. For example.
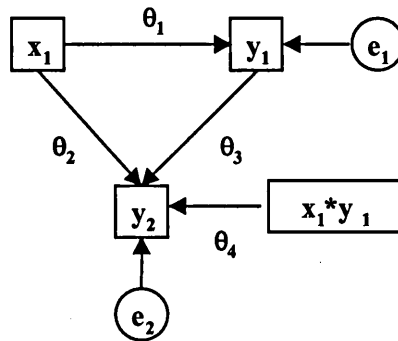
In a similar way that data was generated for the previous model, we generated data for the current model with the following features.

- 54 -
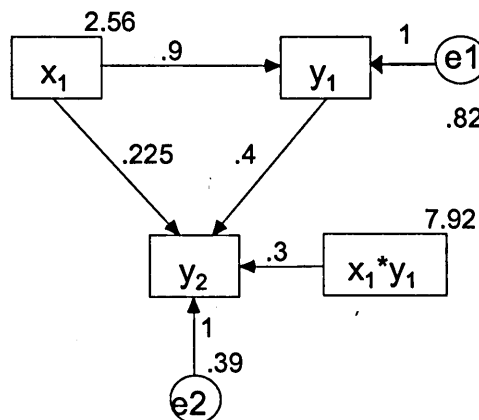
It is important to emphasise that the product of two variables is going to be highly correlated with the source variables. So, it is necessary to orthogonalise it (Buchel and Friston 1997). There is no rule to select the source variables to make the interaction term. In practice, the expertise of the researcher has to be used to construct the interaction term based on his *a priori* knowledge. After analysis, the parameters identified by AMOS 4.0 were:
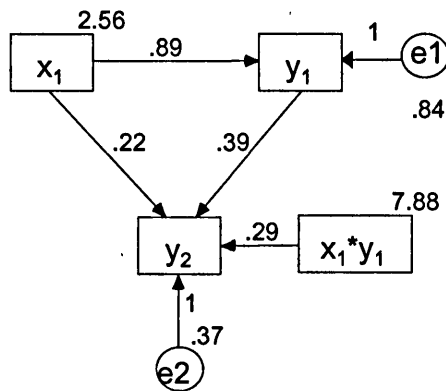
The chi-square of the model is equal to 0.002 with 3 degrees of freedom, then the probability level is equal to 0.99.

2.9.1.3 Non-recursive model.

Non-recursive models are those which contain feed back loops or reciprocal causation. They provide another form to represent non-linear relationship. For example, let us imagine a region y_1 which receives a main input from x_1. Then the sensitivity of y_1 is modulated by another region y_2, i.e. the responsiveness of y_1 depends on the modulatory influence received from y_2 but at the same time y_2 is dependent on the input received from y_1. This modulatory loop can be modelled by a bi-directional causation. Let us generate data for the following model.
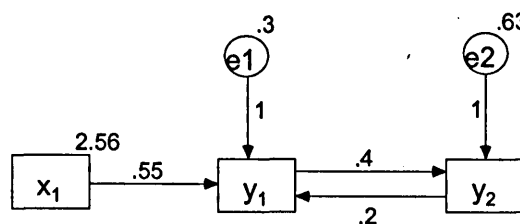
The influence of y_2 over y_1 can be seen as an addition of information, that is to say, given that the main input of y_1 is x_1, then the remaining part of y_1, which is not explained by x_1, is the part that is added by y_2. Models containing reciprocal causation are generally solved using partial regression or two stages least squares (2SLS) methods (Schumacker and Marcoulides 1998). The estimated parameters identified by AMOS 4.0 are:
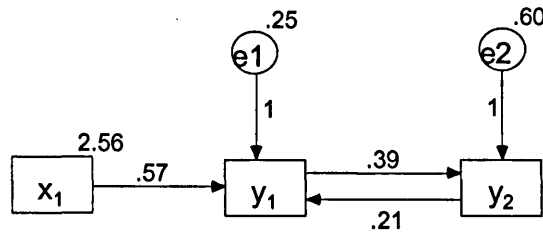
Which have a chi-square of 0.002 with 1 degree of freedom. The probability level of the model is 0.96.

2.9.1.4 Stacked models.

The principal way to test the use of stacked models is to assess one model under two or more data sets. Lets imagine that we want to assess the differences of a theoretical model of effective connectivity using two groups of subjects (healthy and unhealthy). The null hypothesis is that the connectivity in a model using the different groups is the same (H_o: $\theta_1=\theta_2$), i.e. the parameters for the model under the two conditions are the same. The first step is to assess the model with the data of one of the groups (data1). Let us take the model assessed in the first example, where $\theta_1=0.9$, $\theta_2=0.225$, $\theta_3=0.4$.
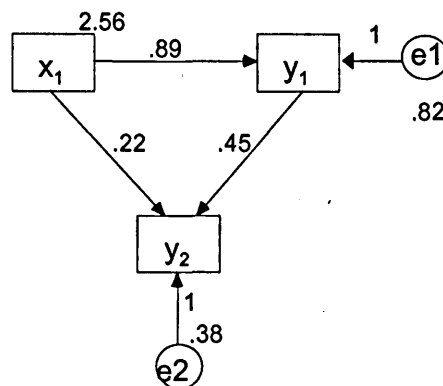
Then, generate a different set of data (data2) for the second group, we use $\theta_1=0.5$, $\theta_2=0.4$, $\theta_3=0.1$. The estimated parameters are shown below.
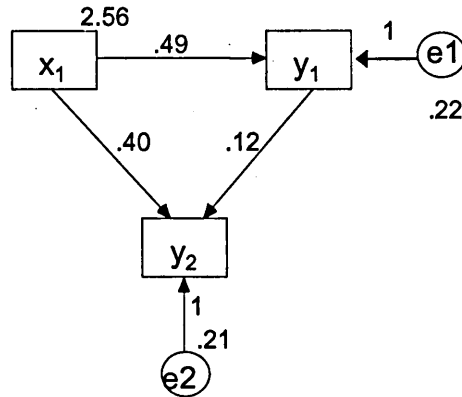
Under the assumption that there is no difference between both data sets, the stacked analysis consists in defining a null model (n), in which the base model is assessed using the second data set (data2) while the parameters are *constrained* to be equal to those identified using the data set (data1). For the current example, the analysis of the null model reports a chi-square (χ_n^2) equal to 230.247 with 4 degrees of freedom (df_n).

The alternative model (a) is assessed in a similar form, however, in this case the parameters are not constrained to take any specific values, i.e., they are *free*. The results of this second analysis are: chi-square (χ_a^2) equal to 0.002 with 1 degree of freedom (df_a).

The statistical significance (p-value) of the parameter difference for the two data sets is computed assessing the chi-square difference ($\chi_n^2 - \chi_a^2 = 230.247 - 0.002 = 230.245$) with degrees of freedom equal to $df_n - df_a = 4-1=3$. Where the p-value < 0.05. Thus the null hypothesis (H_o) is formally rejected.

2.9.2 fMRI models.

In the previous section, some examples of different structural equation model were presented to illustrate the potential application of this technique to assess connectivity models. However, ideal conditions were assumed. In empirical applications of SEM to investigate neural systems with fMRI data, some of these assumptions are not attained and this considerably affect the conclusions that can be made about the observations.

The first and most important assumption that is violated is the fact that the error terms used in models with observed data represent system error (Pearl 2000), i.e. unobserved variances that affect the activity in one node and are *propagated* through its anatomical connection to other nodes. This assumption makes a huge difference with respect to fMRI data in which this error system is almost null and the error term represents measurement error (which is not propagated). The following figure represents a more realistic model of fMRI data.
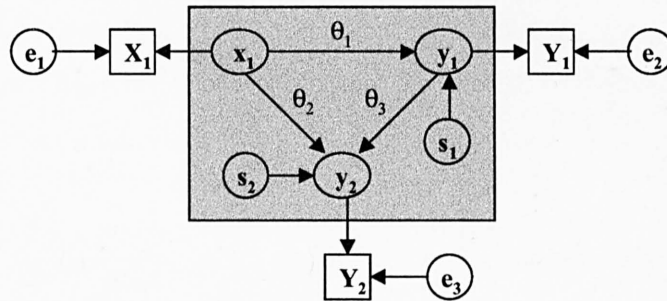
Where the grey square represents the structural model which contain the relationships between the latent variables. It also contain the system noise (s_n) which is almost null because fMRI data is dominated by measurement error (e_n) (Friston, personal communication). Outside the structural model there are the observations, that it to say, the measurements of the activity in different regions (fMRI time series). These observations are affected by measurement errors, which as can be seen in the above figure are not propagated through the system.

In models in which two or more collinear variables are predictors of the same variable, the lack of system noise is a cause of under-determined models. For example, considering the structure of the previous structural model.

$$y_1 = \theta_1 x_1 + s_1$$
$$y_2 = \theta_2 x_1 + \theta_3 y_1 + s_2$$

If $var(s_1)=0$ then, there are infinite number of values for θ_2 and θ_3 which satisfy the equation, in other words, the variance of y_2 can be explained using different combination of values in θ_2 and θ_3 (McIntosh, personal communication). Thus, it is not possible to know the true values which generated the data in the physical model (Bollen 1989). The model is *theoretically identified but empirically under-determined*. The second invalid assumption is to expect linear relationships between variables. For example, one evidence of these non-linearity in neural responses is the fact that different cortical regions have different hemodynamic response (Aguirre, Zarahn et al. 1998; Gossl, Fahrmeir et al. 2001).

Although neural interactions can be approximated by linear relationships, these approximation can bias the parameter estimation. For the next example, a data set was generated for the model with the following parameters.
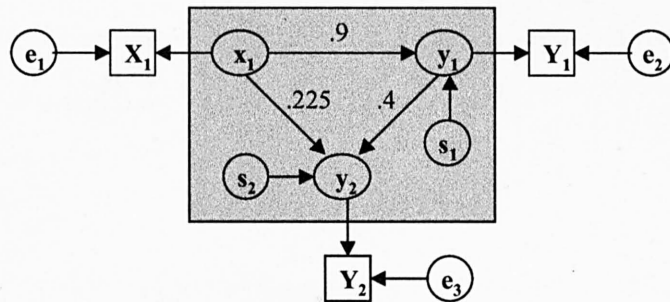
For this simulation, the structural model data was generated previously convolving the experimental stimulation times with the hemodynamic response function. It is equivalent to generating the interactions implied by the model at 'neural level'. For simplicity, it was assumed that the neural response between regions is linear.

After generating the structural model, each time series was convolved with a different hemodynamic response, see figure 2.5.
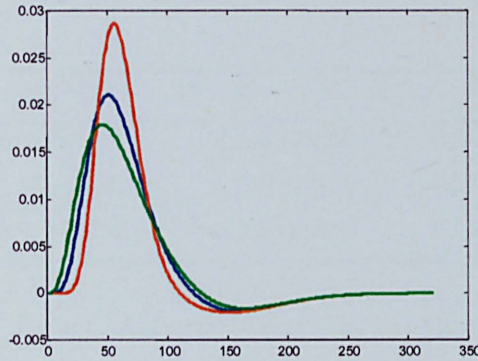
Figure 2.5 Hemodynamic responses.

Three set of data were generated under the following assumptions: *i*) Only measurement noise was including in the model (35 percent of the variance), no system noise was considered. *ii*) System noise and measurement noise were included in the model, both of them at 35 percent of the variance. *iii*) Only system noise was including in the model (35 percent of the variance), no measurement noise was considered. Twenty data sets were generated for each of the categories previously explained (*i,ii,iii*). Figure 2.6 shows the parameters estimated in each simulation.
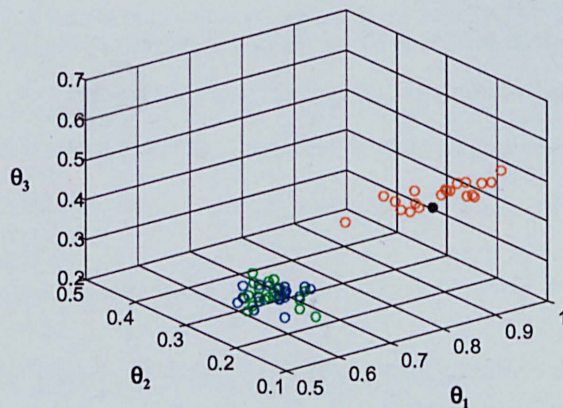
Figure 2.6 Model parameter estimation under different assumptions. Blue circles represent the estimated vectors obtained under case *i*. Green and red circles represent estimations for cases ii and iii respectively. The black dot represents the solution vector which was used to generate the model.

The means and standard deviation of the parameters estimation in each category (i,ii,iii) are shown in table 2.1. Standard deviations are displayed in brackets. It is important to say that all the estimations were significant at p-value > 0.05.

	$\theta_1 = 0.9$	$\theta_2 = 0.225$	$\theta_3 = 0.4$
i	0.616 (0.032)	0.246 (0.029)	0.266 (0.032)
ii	0.601 (0.033)	0.240 (0.023)	0.274 (0.033)
iii	0.907 (0.042)	0.220 (0.068)	0.431 (0.062)

Table 2.1 Summary of parameter estimation.

The previous simulations show how parameter estimation can be biased by the properties of the system being modelled. More over, confidence in the correctness of the model is not only given by the goodness of fit , but also, dependent on the assumptions made about the system. For example, in the previous model the parameters are theoretically identified, however if the assumption is that no system noise is present, then the model becomes empirically under-determined because there is not a unique solution. Otherwise if system noise is assumed, the model has a unique solution and the parameters can be used to make conclusion about the relationships in the model.

fMRI studies applying SEM do not make an explicit distinction between system noise and measurement noise. Actually, the system noise is, in fact, seldom 'noise' because the system is responding to designed inputs that have a large deterministic component (Friston, personal communication). So, system noise has to be correlated with its associated variable, i.e. given $y = \theta x + e$, then $cov(x,e) \neq 0$.

In the next chapter, an alternative technique is explored to complement the limitation of SEM to represent models of effective connectivity using fMRI data. This technique uses the dynamics of the system to account for non-linear relationships.

2.9.3 Comments on stacked analysis.

The invalid assumptions refered to above and the noise dominance in the fMRI time series can lead to a wrong conclusion when stacked model analysis is carried out. For example, let us use the previous model to generate various sets of data under different proportion of noise, i.e. noise at 10, 30 and 50 percent of the signal. Three data sets (D1,D2,D3) were generated for each category of noise proportion. The parameters estimated (using AMOS 4.0) per each data set were:

	Theta 1	Theta 2	Theta 3
D1_10	0.81	0.25	0.35
D2_10	0.84	0.28	0.29
D3_10	0.84	0.31	0.30
D1_30	0.57	0.22	0.34
D2_30	0.69	0.22	0.33
D3_30	0.58	0.27	0.24
D1_50	0.53	0.09	0.29
D2_50	0.49	0.25	0.16
D3_50	0.48	0.20	0.25

According with stacked analysis technique, the goodness of fit in terms of chi-square was estimated creating a null and an alternative model per each combination of data sets. The chi-squares are shown in the following table.

	D1_10	D2_10	D3_10	D1_30	D2_30	D3_30	D1_50	D2_50	D3_50
P1_10	0.0	4.3	3.4	36.3	8.6	35.8	70.1	60.0	51.2
P2_10	4.3	0.0	3.7	45.2	11.3	37.4	75.0	56.5	52.9
P3_10	5.3	5.2	0.0	51.3	15.1	43.8	87.6	64.9	61.2
P1_30	79.1	87.4	114.2	0.0	5.4	5.1	25.5	23.3	11.8
P2_30	28.7	34.5	51.4	9.0	0.0	10.4	34.7	32.1	22.6
P3_30	92.5	94.0	125.0	5.9	8.7	0.0	25.3	9.9	8.0
P1_50	225.8	227.6	265.3	42.2	37.3	34.0	0.0	18.6	9.6
P2_50	209.7	209.2	249.4	30.0	37.9	14.6	16.2	0.0	4.2
P3_50	187.1	186.6	229.5	17.9	26.0	11.8	9.4	4.9	0.0

The p-values for the model comparison are show in the table below.

	D1_10	D2_10	D3_10	D1_30	D2_30	D3_30	D1_50	D2_50	D3_50
P1_10	**0.996**	0.115	0.182	0.000	0.013	0.000	0.000	0.000	0.000
P2_10	0.112	**0.997**	0.160	0.000	0.003	0.000	0.000	0.000	0.000
P3_10	0.067	0.071	**0.999**	0.000	0.000	0.000	0.000	0.000	0.000
P1_30	0.000	0.000	0.000	**0.997**	0.065	0.076	0.000	0.000	0.002
P2_30	0.000	0.000	0.000	0.010	**0.992**	0.005	0.000	0.000	0.000
P3_30	0.000	0.000	0.000	0.051	0.012	**0.993**	0.000	0.006	0.018
P1_50	0.000	0.000	0.000	0.000	0.000	0.000	**0.993**	0.000	0.008
P2_50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.997**	0.121
P3_50	0.000	0.000	0.000	0.000	0.000	0.002	0.009	0.084	**0.992**

Although all the data sets were generated by the same model, the stacked model analysis showed differences in most of the data sets. It is important to notice that all of the right conclusions (displayed in grey rectangles) where made in data sets with the same percentage of noise, whereas data sets with different noise amounts where always considered unequal. It reflects the high instability of the method under different noise conditions.

2.10 Conclusions.

Structural equation modelling is a potentially useful tool to assess models of effective connectivity using functional magnetic resonance imaging (fMRI). Although SEM has been widely applied in Social Sciences, the advantage of its application in neuroimaging is that the variables are observed and the connections between variables are anatomically defined, whereas in the traditional application of SEM in social sciences the variables are latent and the relationships between them are hypothetical.

However, it is important to remark that the goodness of fit of a given model can not be treated as an absolute value. There are three main reasons for it: i) Although the model can be theoretical identified, it could be empirically under-determined.

ii) A small number of observations (N) over-estimates the goodness of fit because the chi-square of the model is a function of the ML estimator times N. iii) The *a priori* estimation of the error terms can bias the parameter estimation at the same time that it biases the goodness of fit due to increases in the degrees of freedom.

Moreover, MonteCarlo simulation showed the dependency between the assumptions and the conclusion that can be made about the estimated parameters. The collinearity between pre-cursors of the same node, produces empirically under-determined models, because the propagation of the error system (generally assumed by SEM modellers) does not exist in fMRI time series or is relatively much smaller than measurement noise.

One important feature of SEM is its ability to represent non-linear relationships through feedback loops or interaction terms. However the identification of non-linear relationships is not a trivial task and part of this work is to integrate lagged covariance structures to capture the dynamic nature of fMRI time series. This brings SEM toward non-linear dynamical modelling. The next chapter presents a technique to incorporate these aspects to SEM through the use of a Non-linear AutoRegresive Moving Average with eXogenous approach.

Chapter 3

NARMAX

Abstract.

In the present chapter a Non-linear Auto-Regressive Moving Average with eXogenous variables (NARMAX) algorithm is introduced as an alternative approach to represent models of effective connectivity using fMRI time series. This approach can be thought of as an attempt to bring Structural Equation Modelling to a non-linear dynamic system modelling technique which permits a more appropriate representation of effective connectivity models using fMRI time series. The theoretical basis and implementation details of the algorithm are explained. The performance of the NARMAX algorithm is assessed using simulated data.

3.1 Introduction.

Classical applications of structural equation modelling (SEM) in social sciences create models of hypothetical connections to represent relationships among events that cannot be measured directly. Contrary to this, in neuroimaging, the constraints imposed by anatomical pathways and the use of BOLD signals, as a neural activity indicator, bring significant advantages for the application of this technique to assess relationships among cortical areas using fMRI time series (Friston 1994; Buchel and Friston 1997).

Effective connectivity analysis has been focused on finding changes in connectivity, through the evaluation of an anatomically defined model under different experimental conditions. Using structural equation modelling, these changes can be detected by assessing individually each set of data and then by comparing the estimated parameters in each data set (stacked models). It is important to clarify, that these changes in connectivity refer to changes in the values of the path coefficients and not to changes in the model structure, which is assumed to be correct (Goncalves, Hall et al. 2001).

In this approach the data is partitioned into as many data sets as experimental conditions, so only unique linear parameters are estimated per each data set. This means that the estimated parameters explain the relationships under each particular experimental condition only, but nothing is learnt about the cortical interactions which modulate the changes in effective connectivity under different experimental conditions. For example, in our first experiment (chapter 4) we are going to assess the non-linear behaviour of the activation in V3A when it is stimulated by Stereo and Motion stimuli at the same time rather than when each stimulus is presented independently.

A more robust approach is to model the system to account for changes in connectivity over time, i.e., to let the model explain all the data sets of the individual experimental conditions.

As this approach implies the identification of hidden variables, which explain the non-linear relationships of the cortical interactions, the effective connectivity analysis becomes not only a method for parameters estimation but also for model identification. Another huge difference between the classical applications of SEM in social sciences and neuroimaging is that social sciences work with static linear systems whereas in neuroimaging the nature of the relationships between cortical regions is inherently a non-linear dynamic system.

It is well known that non-linearity can be present in fMRI data due to the facts that hemodynamic response varies among cortical regions (Aguirre, Zarahn et al. 1998), or because hemodynamic responses are non-linear functions of experimental conditions. For example, passive listening to words produces different responses in periauditory regions depending on the rate at which the words are spoken (Friston, Josephs et al. 1998).

Moreover, different hemodynamic response shape is going to be obtained by applying the same stimulus for different periods of time, i.e., hemodynamic responses are time dependant in the sense that actual neural response is affected by previous activity (autoregressive). This suggests that 'time' is another important aspect to consider in this relatively new application.

On the other hand, as different voxels are acquired at different times (multislice acquisition), artificial lags are introduced into the time series, and although fMRI analysis tools include utilities for slice time correction, the consideration of lagged inputs in effective connectivity analysis can optimise the parameter identification, maximising the correlation between variables. In the present chapter a Non-linear Auto-Regressive Moving Average with eXogenous variables (NARMAX) algorithm is introduced as an alternative to obtain a more accurate representations of effective connectivity models. Its theoretical basis and implementations details are also introduced. The capabilities of this approach to represent lagged inputs and non-linear relationships can be thought of as an attempt to bring SEM towards a non-linear dynamic system modelling technique.

3.2 Non-linear dynamic system identification.

System identification refers to the problem of constructing mathematical models which represents a mapping between input and output space of a particular system, i.e. to predict an output ($y(t)$), given a set of inputs ($u(t)$) or even also to give the lagged input itself. When modelling time-changing relationships, the problem becomes one of non-linear dynamic system identification. System identification is one of the most active fields in engineering. In the present section, two of the most common approaches to represent non-linear systems are going to be introduced and its application to fMRI time series analysis is going to be explored. The first approach, Volterra series, has been used by Friston *et al* (Friston, Josephs et al. 1998) to characterise the hemodynamic response of a single region (Multiple Input Single Output, MISO) in a event-related fMRI study. The second approach, polynomial expansions, is introduced as an attempt to use the basic ideas of Volterra series in a multiple region effective connectivity analysis (Multiple Input Multiple Output, MIMO).

3.2.1 Volterra series.

Volterra series are extensions of the Tailor series representations to model non-linear dynamic systems. Volterra showed that these series are capable of representing any non-linear dynamic system (Volterra 1959) . Volterra series can be thought of as a high order extension of linear convolutions of the inputs to produce the output. The general form of the Volterra representation is:

$$y(t) = h_0 + \int_{-\infty}^{\infty} h_1(\tau_1)u(t-\tau_1)d\tau_1 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2)u(t-\tau_1)u(t-\tau_2)d\tau_1 d\tau_2 + ...$$

$$+ \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} h_n(\tau_1,...,\tau_n)u(t-\tau_1)...u(t-\tau_n)d\tau_1 ... d\tau_n + ...$$

Where $u(t)$ and $y(t)$ are the input and output respectively at time t, and the $h_n(\tau_1,... \tau_n)$ is the nth order Volterra kernel. The system identification requires one to estimate the kernels.

In the context of fMRI BOLD response characterisation, Friston *et al* (Friston, Josephs et al. 1998) used Volterra series to represent the relationship between evoked neuronal activity and the resulting hemodynamic response. In this case, the Volterra kernels were a predetermined set of basis functions (gamma density functions) and the corresponding first derivatives (to accommodate differences in time), see figure 3.1. The choice of the gamma density functions used was motivated by prior knowledge of the hemodynamic response function to represent.
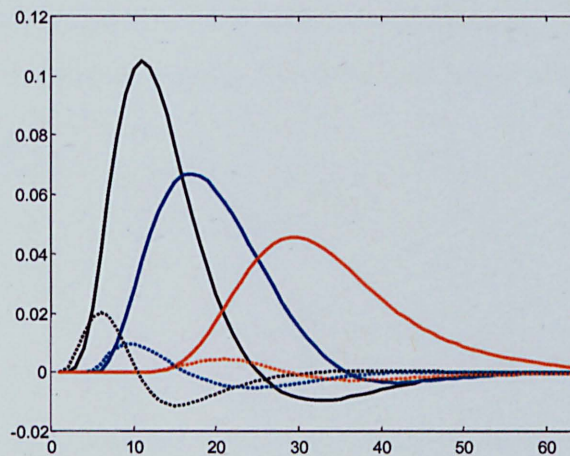
Figure 3.1 Basis functions. Three gamma density functions were used by Friston to approximate the hemodynamic response observed in an event-related study. Dashed lines represent the fist derivative of each gamma function. These gamma functions are used just as an example and they are not the same as those used in Friston's study.

Considering a linear system, the response in any region *y(t)* can be represented as a linear convolution of the basis functions $(h_1(\tau_1))$ with the experimental stimulus *u(t)*. Given that fMRI BOLD response is inherently non-linear, the estimation of the Volterra kernels has to be extended to a high order $(h_n(\tau_1, ... \ \tau_n))$. For example, Friston *et al* (Friston, Josephs et al. 1998) used Volterra representations to model the non-linear behaviour of the periauditory regions under different rates of word presentation. In order to linearise the problem to identify the kernels, they used three gamma density functions (fig. 3.1) as a basis to construct the hemodynamic response of the periauditory region.

The second order representation used by Friston was defined as follows:

$$y(t) = h_0 + \int\limits_0^T h_1(\tau_1)u(t-\tau_1)d\tau_1 + \int\limits_0^T\int\limits_0^T h_2(\tau_1,\tau_2)u(t-\tau_1)u(t-\tau_2)d\tau_1 d\tau_2$$

Where the scope of the integral from 0 to T, means that the system is causal (responses happen after the input) and the system has a finite memory T. The zero order kernels (h_0) represent the base line. The first order kernels ($h_1(\tau_1)$) represent the linear part of the system and the second order kernel ($h_2(\tau_1, \tau_2)$) represent the non-linear part of the system. Each kernel was defined as a linear combination of a number (P) of gamma density functions and its corresponding derivatives ($b_i(\tau_1)$).

$$h^0 = g^0$$

$$h^1(\tau_1) = \sum_{i=1}^P g^1 b_i(\tau_1)$$

$$h^2(\tau_1,\tau_2) = \sum_{i=1}^P \sum_{j=1}^P g_{ij}^2 b_i(\tau_1) b_j(\tau_2)$$

Then, the estimation of the coefficients (g) was made using least squares methods. It is important to notice that Volterra series are useful to characterise the behaviour of a single region. The non-linear analysis is restricted to the time series of the region in isolation, without considering its interactions with other regions.

However, keeping the idea of the use of time delayed inputs and basis function to create multiplicative terms to represent the non-linear part of the system, it is possible to use a polynomial expansion approach to represent non-linear multiple region effective connectivity models. It is important to clarify that the intention of the present section was only to introduce the antecedents of the ideas to be used in the present chapter, for a meticulous explanation of Volterra series consult (Bendat 1990).

3.2.2 Polynomial expansions.

The polynomial expansion approach is one of the most common techniques used to represent non-linear systems. One of its main advantages is that few polynomial terms (less than 10) are usually enough to capture the non-linear relationships (Chen, Billings et al. 1989). Under this approach it is possible to integrate multiple regions on the non-linear modelling. As the basic idea of the algorithm proposed in this work is based on this approach, an in depth explanation of its basis is given in the next section.

3.3 NARMAX model.

The Non-linear Autoregressive Moving average with eXogenous variables (NARMAX) model was proposed by Billings *et al* (Leontaritis and Billings 1985; Korenberg, Billings et al. 1988; Billings, Chen et al. 1989; Chen, Billings et al. 1989), in the department of Automatic Control and System Engineering of the University of Sheffield. This approach has been widely tested under simulated data and used in different engineering applications.

3.3.1 Representation.

A discrete-time multi-input multi-output (MIMO) non-linear dynamic system with m outputs and r inputs can be described as a Non-linear Autoregressive Moving average with eXogenous variables model (Leontaritis and Billings 1985).

$$y_i(t) = f_i\,(y_1(t-1), ..., y_1(t-n_y), ..., y_m(t-1), ..., y_m(t-n_y),$$
$$u_1(t-1), ..., u_1(t-n_u), ..., u_r(t-1), ..., u_r(t-n_u), \qquad (3.1)$$
$$e_1(t-1), ..., e_1(t-n_y), ..., e_m(t-1), ..., e_m(t-n_y)) + e_i(t)$$

Where,

Output	Input	Noise
$$y(t) = \begin{bmatrix} y_1(t) \\ \cdot \\ \cdot \\ \cdot \\ y_m(t) \end{bmatrix}$$	$$u(t) = \begin{bmatrix} u_1(t) \\ \cdot \\ \cdot \\ \cdot \\ u_r(t) \end{bmatrix}$$	$$e(t) = \begin{bmatrix} e_1(t) \\ \cdot \\ \cdot \\ \cdot \\ e_m(t) \end{bmatrix}$$

And

f_i is an unknown non-linear function.

t is the time index with values from 1 to N.

i is the sub-system index with values from 1 to m.

y is a vector of outputs.

U is a vector of inputs.

e is a vector of random noises.

n_y, n_u, n_e are the maximum lags.

The non-linear form of f_i can be approximated by a polynomial expansion of degree l under y, u and e terms (Chen, Billings et al. 1989).

$$(3.2)$$

$$y_i(t) = \theta_0^{(i)} + \sum_{i_1=1}^{n_i} \theta_{i_1}^{(i)} x_{i_1}(t) + \sum_{i_1=1}^{n_i} \sum_{i_2=i_1}^{n_i} \theta_{i_1 i_2}^{(i)} x_{i_1}(t) x_{i_2}(t) + ... \sum_{i_1=1}^{n_i} ... \sum_{i_2=l-1}^{n_i} \theta_{i_1...i_l}^{(i)} x_{i_1}(t)...x_{il}(t) + e_i(t)$$

Where

$\theta^{(i)}$ is a vector of parameter regressors (coefficients) for the sub-system i.

$\theta_0^{(i)}$ is equal to dc.

$n_i = m*n_y + m*n_e + r*n_u$

$x_{i...l}$ represents each term in the function f_i. Each term can be thought of as an independent component (time series) in each column in the design matrix.

For example,

$x_1(t) = y_1(t-1)$,

$x_2(t) = y_1(t-2)$,...

$x_{m*ny}(t) = y_m(t- n_y)$,

$x_{m*ny+1}(t) = e_1(t-1)$, ...

$x_{m*ny+m*ne}(t) = e_m(t- n_e)$,

$x_{m*ny+m*ne+1}(t) = u_1(t-1)$, ...

$x_{m*ny+m*ne+ r*nu}(t) = u_r(t- n_u)$

Then, the non-linear function can be represented by a difference equation system which is linear in the parameters and non linear in the terms, lets us now represent $y_i(t)$ as $z_i(t)$.

$$z_i(t) = \sum_{k=1}^{M_{il}} p_k^{(i)}(t)\theta_k^{(i)} + \varepsilon^{(i)}(t) \tag{3.3}$$

Where M_{il} can be computed recursively as

$M_{il}=0.5 \, M_{il-1} \, (M_{il-1} + 1) \quad \therefore \quad M_{i1}=n_i$

$p_k^{(i)}$ terms are monomials of degree from zero to l.

For example:

$p_0=1,$	$p_1= y_1(t-1),$
$p_2= y_1(t-2),$	$p_3= y_2(t-1),$
$p_4= y_2(t-2),$	$p_5= y_1(t-1)\,y_2(t-1),$
$p_6= y_1(t-1)\,y_2(t-2),$	$p_7= y_2(t-1)\,y_2(t-1), \ldots$

Then, for the subsystem i the solution can be formulated as a least squares regression of the form:

$$Z=P\theta + E \tag{3.4}$$

Where

$$Z = \begin{bmatrix} z(t) \\ \cdot \\ \cdot \\ z(N) \end{bmatrix} \quad P = \begin{bmatrix} p_1(t) & \cdots & p_M(t) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ p_1(N) & \cdots & p_M(N) \end{bmatrix} \quad \theta = \begin{bmatrix} \theta_1 \\ \cdot \\ \cdot \\ \theta_M \end{bmatrix} \quad E = \begin{bmatrix} e(t) \\ \cdot \\ \cdot \\ e(N) \end{bmatrix}$$

Because the design matrix (P) is generated by a polynomial expansion in which each column of P is a product of a set of basic terms in y, u and e, many of the columns of P are highly correlated, then, the covariance matrix ($P'P$) is singular (ill-conditioned) and the conventional least square method is not applicable to solve the equation system (Korenberg, Billings et al. 1988).

On the other hand, it is clear that only some of the terms of P are needed to explain Z. The question which arises is, which terms have to be selected from P to explain Z, avoiding the ill-conditioning problem? The answer is given by the Gram-Schmidt orthogonalisation algorithm.

3.3.2 Gram-Schmidt orthogonalisation.

Any inner product space V is defined by an orthogonal basis $S=\{v_1, v_2, \dots, v_n\}$ in which any vector $w \in V$ can be represented as a linear combination of the basis vectors in S (Lay 2000).

$$w = \frac{\langle w, v_1 \rangle}{\|v_1\|} v_1 + \frac{\langle w, v_2 \rangle}{\|v_2\|} v_2 + \dots + \frac{\langle w, v_n \rangle}{\|v_n\|} v_n$$

Where $\dfrac{\langle w, v_i \rangle}{\|v_i\|} v_i$ is the orthogonal projection of w on v_i. For example, let us

define the vector space $S=\{v_1, v_2\}$ and represent in it the vectors w_1 and w_2.
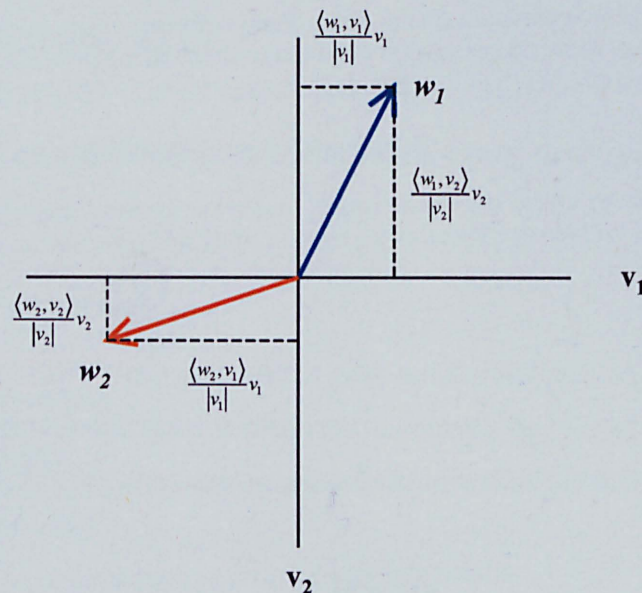
The idea behind the Gram-Schmidt algorithm is that given an arbitrary basis $\{ w_1, w_2, \dots, w_n \}$ for an n-dimensional space V, it constructs an orthogonal basis $S=\{ v_1, v_2, \dots, v_n \}$ for V.

For example, given the vectors w_1, w_2 in our previous example Gram-Schmidt orthogonalisation would create the following orthogonal basis $S=\{ v_1, v_2\}$.
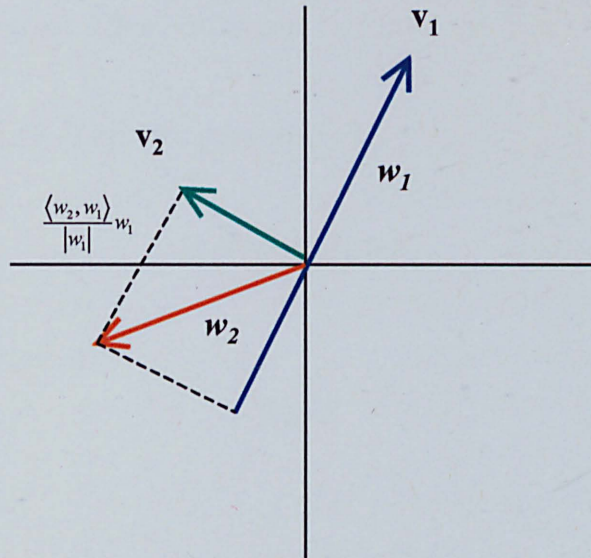
Using the Gram-Schmidt approach the problem of solving (3.4), given that the $(P'P)$ covariance matrix is ill-conditioned, can be sorted out using orthogonal least squares.

3.3.3 Orthogonal least squares.

Any symmetric positive definite square matrix L can be decomposed in

$$L=A'DA \qquad (3.5)$$

Where A is an upper triangular matrix with unitary diagonal elements and D is a diagonal matrix with all positive elements. Assuming that the column vectors of P are orthogonal, $P'P$ is symmetric positive definite and it can be represented as:

$$P'P=A'DA \qquad (3.6)$$

Because the design matrix P is generated by a polynomial expansion, it does not form a positive definite covariance matrix $(P'P)$.

Consequently, it is necessary to orthogonalise the columns of P to find the orthogonal components (W) which explains the vector Z. These components can be found following an orthogonal least squares approach inspired by the Gram-Schmidt orthogonalisation (Chen, Billings et al. 1989).

Because $A^{-1}A=I$, Z (in 3.4) can be represented as

$$Z=P(A^{-1}A)\theta+E \tag{3.7}$$

It is possible to factorise as

$$Z=Wg+E \tag{3.8}$$

Where

$$W=PA^{-1} \tag{3.9}$$

$$g=A\theta \tag{3.10}$$

Where W represents the orthogonal basis of P, and g represents the orthogonal projection coefficients of W on Z. Therefore,

$$D=W'W \tag{3.11}$$

Where D (as defined above) is a diagonal matrix with all positive elements.

$$D = \begin{bmatrix} ||w_0|| & & & \\ & ||w_1|| & & 0 \\ & & ||w_2|| & \\ & & & \ddots \\ 0 & & & ||w_M|| \end{bmatrix}$$

The components of W can be obtained recursively from

$$W=P-W(A-I) \tag{3.12}$$

It represents the orthogonalisation (subtraction) of the projection (common part) between the components of P.

Multiplying (3.9) by W' and by A:

$$W = PA^{-1}$$
$$AW'W = W'PA^{-1}A$$
$$AW'W = W'P$$
$$A = (W'W)^{-1}W'P \qquad (3.13)$$

So, the matrix A can be computed as

$$A = D^{-1}W'P \qquad (3.14)$$

With the form

$$A = \begin{bmatrix} 1 & \alpha_{01} & \alpha_{02} & \cdots & \alpha_{0M} \\ & 1 & \alpha_{12} & \cdots & \alpha_{1M} \\ & & 1 & & \vdots \\ & & & \ddots & \vdots \\ 0 & & & & 1 \end{bmatrix}$$

From (3.8) the estimation of g is given by

$$Z = Wg + E$$
$$W'Z = W'Wg + W'E$$
$$g = (W'W)^{-1}W'Z$$
$$g = D^{-1}W'Z \qquad (3.15)$$

Finally, from (3.10) the parameters θ can be computed as

$$\theta = A^{-1}g \qquad (3.16)$$

3.4 NARMAX algorithm.

The Non-linear Autoregressive Moving average with eXogenous variables (NARMAX) algorithm is an orthogonal forward regression least square estimator which implements a modification of the classical Gram-Schmidt algorithm (Billings, Chen et al. 1989).

NARMAX algorithm computes the matrix A one row at a time, orthogonalising from the kth column (at the kth iteration), all the remaining columns from k+1 to M. Considering the subsystem (3.4), the algorithm can be structured as follows.

3.4.1 Algorithm.

Step 1:

Given the input and output vectors generate the polynomial expansion as showed in (3.2) to construct (3.4).

Step 2:

With (3.15) consider *all* the elements of P as candidates to be w_1 (the first orthogonal basis). Select the element of P which gives more information (inf$_k$) about Z. In this context, information means the amount of variance on the dependent variable (Z) that is explained by the candidate (P_i). Let us consider an auxiliary vector $W=P$. For k=1 to M.

$$g_k = \frac{\langle w_k, z \rangle}{\|w_k\|}$$

$$\inf_k = \frac{g_k^2 \|w_k\|}{\|z\|}$$

Select the term of W which gives most information about Z. Make it equal to w_l and make $g_1 = g_k$. Permute the vector w_l to the first position of the matrix W and move the vector which used to be in the first column to the position in which the more informative element was found. Save the quantity of information given by the term as $\exp_1 = \inf_k$ because it can be used to define the convergence criteria.

Step 3:

Orthogonalizes the terms of W (from 2 to M) with respect to w_l, i.e. using (3.12 and 3.13) subtract the orthogonal projection of the first component from the rest of the components. For k=2 to M.

$$\alpha_k = \frac{\langle w_1, w_k \rangle}{\|w_1\|}$$

Update the vector W and construct the matrix A as:

$$A_{1,k} = \alpha_k$$

$$w_k = w_k - \alpha_k w_1$$

Iterate steps 2 and 3 until a *convergence criterion* is reached. It is important to note that in each iteration: i) In step 3, the search for candidates to form new orthogonal components has to be restricted to those terms that have not been selected to be part of the orthogonalised bit of W. ii) In step 3, the orthogonalisation has to be done with respect to the component selected in step 2.

The convergence criteria can be defined in different ways, two of the most common used are:

i) As the error term in the difference:

$$\varepsilon(t) = z(t) - \sum_{i=1}^{Ms} w_i g_i , \|\varepsilon\| < \text{threshold}$$

ii) As a percentage of the unexplained variance:

$$\left(1-\sum_{i=1}^{Ms}\exp_i\right)*100 < \text{threshold}$$

Where exp_i represents the percentage of variance (information, $\inf_k$) given by each orthogonal component selected to explain Z. Finally, compute the parameters θ using (3.16) as

$$\theta = A^{-1}g$$

Because the error term is only known when the system has been identified with input and output terms, the delayed noise terms have to be identified in a second run of the algorithm in which the error term is included in the polynomial expansion (Billings, Chen et al. 1989; Chen, Billings et al. 1989).

3.4.2 Implementation comments.

The ill-conditioning characteristic of the covariance matrix of P can be sorted out in step 3, by checking the variance of the orthogonalised terms of W. If the variance of any of the orthogonalised terms ($\|w_k\|$) is less that a threshold (near to zero), it means that this term is a linear combination of a previous selected terms and then it can be deleted. They can be also identified before orthogonalisation by checking if the angle between the term to orthogonalise and the term to substract ($Arc\cos(\langle w_k, w_n \rangle)$) is smaller than a desired threshold.

The NARMAX algorithm explained above includes a general framework to represent non-linear auto-regressive moving average systems with endogenous inputs. Linear autoregression, non-linear autoregression, non-linear moving average, etc. are particular instances of the NARMAX model. So, for the special characteristics of particular models, it could be the case that specific instances of the NARMAX model could be sufficient to capture the dynamics of the system.

For example, if the assumption of the unknown system is that there are not lagged correlated errors, it is possible to create a Non-linear AutoRegressive with exogenous inputs model (NARX). On the other hand, if the assumption is that the system is not autoregressive, the explained variable can be left out of the polynomial expansion which generates the design matrix (P). It is always advisable to create a parsimonious representation model of the system.

Because the polynomial expansion of order l generates monomials which are multiplicative terms of the basis terms. They are going to be highly correlated with the basis terms. In order to give priority in the regression to the basis terms, a residualisation procedure is suggested to delete (residualise) the projections of these terms (Buchel, Josephs et al. 1998).

Lastly, it is necessary to remember that every model is a simplification of the real system and it is convenient to obtain a balance between accuracy and complexity. The convergence criteria shown above do not take account of the complexity of the terms included in the model. However some measures like Akaike (AIC) information criteria or Minimum Description Length (MDL) can be used to address this problem (Bollen 1989; Pearl 2000).

The implementation of the NARMAX algorithm was made in MATLAB 5.3 for Windows NT 4.0. For a detailed description of the code consult the appendix at the end of the thesis.

3.4.3 Model validation.

The goodness of fit of the model can be assessed in the same form as in Structural Equation Modelling (SEM), that its to say, evaluating (in terms of χ^2) the difference between the observed covariance matrix and the implied covariance of the model proposed by the forward regression orthogonal algorithm (see section 2.5.2).

3.5 MonteCarlo simulations.

In the present section the efficiency of the forward regression orthogonal algorithm is going to be tested using artificial data. In the first two examples, classical exemplification used by the developers of NARMAX are shown. In the second and third examples, potential applications of the algorithm to model non-linear models of effective connectivity are going to be introduced using fMRI simulated data. As made in effective connectivity analysis within SEM, effective connectivity analysis within NARMAX approach assumes that the anatomical pathways are known before hand. This approach can be thought as a non-linear dynamic confirmatory analysis in which the hidden non-linear terms are identified. The convergence criteria in all the analyses was defined as the minimum explained variance per predictor does not go below one percent.

In each case the model validation was made assessing the output model under AMOS 4.0 software. It is important to remember that in SEM the goal is to know if the model fits the data, i.e. the parameter estimates of the model produce an implied matrix that is close to the observed covariance matrix. The null hypothesis is that the implied covariance matrix and the observed covariance matrix are equal ($\Sigma=\Sigma(\theta)$). This hypothesis is assesses in terms of chi-square (χ^2) test statistics.

If the difference between the implied covariance matrix and the observed covariance matrix is "small" (for example, p-value > 0.05) the χ^2 is going to be small and the null hypothesis has to be retained. Otherwise, if the implied covariance matrix and the observed covariance matrix are sufficiently different, the null hypothesis has to be rejected, i.e. the model is not a good representation of the data.

Finally, it is important to notice that, in order to be consistent with fMRI data characteristics, the error term of the simulated data is added as a measurement noise, i.e. it is not propagated.

3.5.1 Linear autoregressive with exogenous variables model.

The present example defines a linear system with lagged inputs and autoregressive behaviour. Let us consider the following model (modified from (Billings, Chen et al. 1989)).

$$y_1(t) = -0.4\, y_1(t-1) + 0.5\, y_1(t-3) - 0.2\, u_1(t-1) + 0.5\, u_1(t-3) + 0.7\, u_2(t-2)$$

$$y_2(t) = 0.6\, y_1(t-2) - 0.8\, y_2(t-2) + 0.3\, y_2(t-3) + 0.4\, u_1(t-1) + 0.3\, u_2(t-1)$$

$$y_1(t) = y_1(t) + e_1(t)$$
$$y_2(t) = y_2(t) + e_2(t)$$

Where, u_1 and u_2 are independent sequences of uniform distribution of variance 1 (see figure 3.2). e_1 and e_2 are uncorrelated gaussian noise of variance 0.04. The time series generated were 200 time points length.

Using an autoregressive linear model of dynamic order=3 and polynomial order=1, (where the dynamic and polynomial orders are defined as the maximum lag and polynomial degree respectively permitted in the expansion), the forward regression orthogonal algorithm identified the following model.

$$y_1(t) = \quad -0.394\, y_1(t-1) + 0.486\, y_1(t-3) +$$
$$-0.225\, u_1(t-1) + 0.502\, u_1(t-3) +$$
$$0.702\, u_2(t-2) \qquad \textit{where } var(e_1) = 0.062$$

$$y_2(t) = \quad 0.590\, y_1(t-2) - 0.786\, y_2(t-2) + 0.293\, y_2(t-3) +$$
$$0.397\, u_1(t-1) + 0.292\, u_2(t-1)$$
$$\textit{where } var(e_2) = 0.077$$

The statistical validity of the model reported by AMOS 4.0 was $p > 0.05$ for each subsystem. The time series estimated by the algorithm are shown in figure 3.3.

3.5.2 Non-linear autoregressive with exogenous variables model.

The non-linear system presented below, represents a non-linear relationship in terms of multiplicative lagged inputs. Let us consider the following model structure, modified from (Billings and Voon 1986b; Billings, Chen et al. 1989).

$$y_1(t) = \quad 0.5\ u_1(t\text{-}3) + 0.7\ u_2(t\text{-}2) +$$
$$0.35\ y_1(t\text{-}3) - 0.2\ u_1(t\text{-}1)\ u_2(t\text{-}3) +$$
$$-0.4\ y_1(t\text{-}1)\ u_1(t\text{-}2) + 0.5\ y_1(t\text{-}3)\ u_2(t\text{-}1)$$

$$y_2(t) = \quad 0.3\ u_2(t\text{-}1) - 0.7\ y_1(t\text{-}2) +$$
$$0.3\ y_2(t\text{-}3) + 0.4\ u_1(t\text{-}1)\ u_1(t\text{-}2) +$$
$$-0.8\ y_2(t\text{-}2)\ u_2(t\text{-}3) + 0.1\ y_1(t\text{-}3)\ y_2(t\text{-}1)$$

$$y_1(t) = \quad y_1(t) + e_1(t)$$

$$y_2(t) = \quad y_2(t) + e_2(t)$$

Figure 3.2 Exogenous inputs of the linear autoregressive model.

y_1

y_2
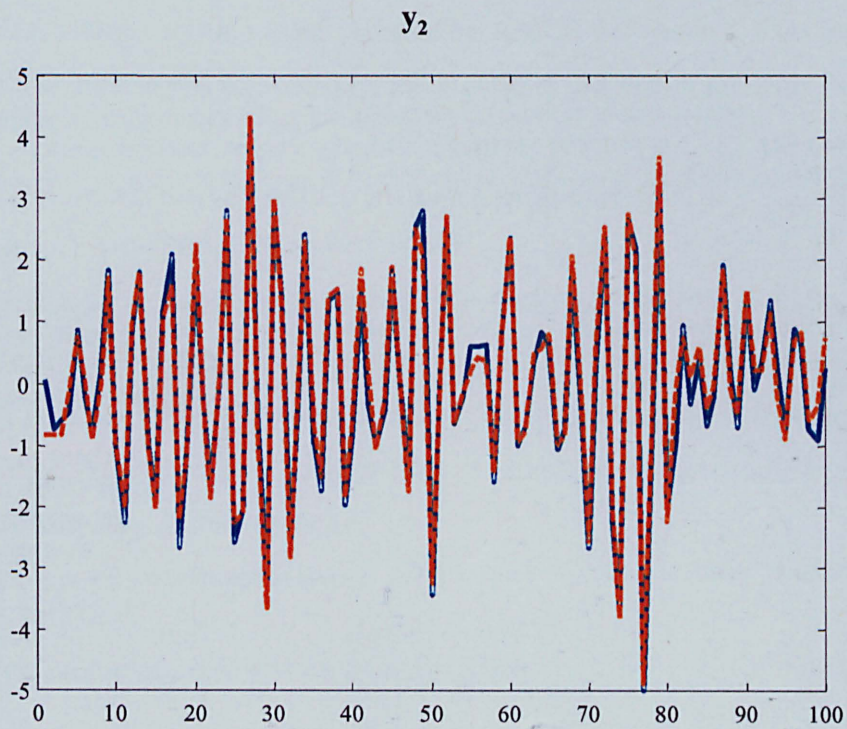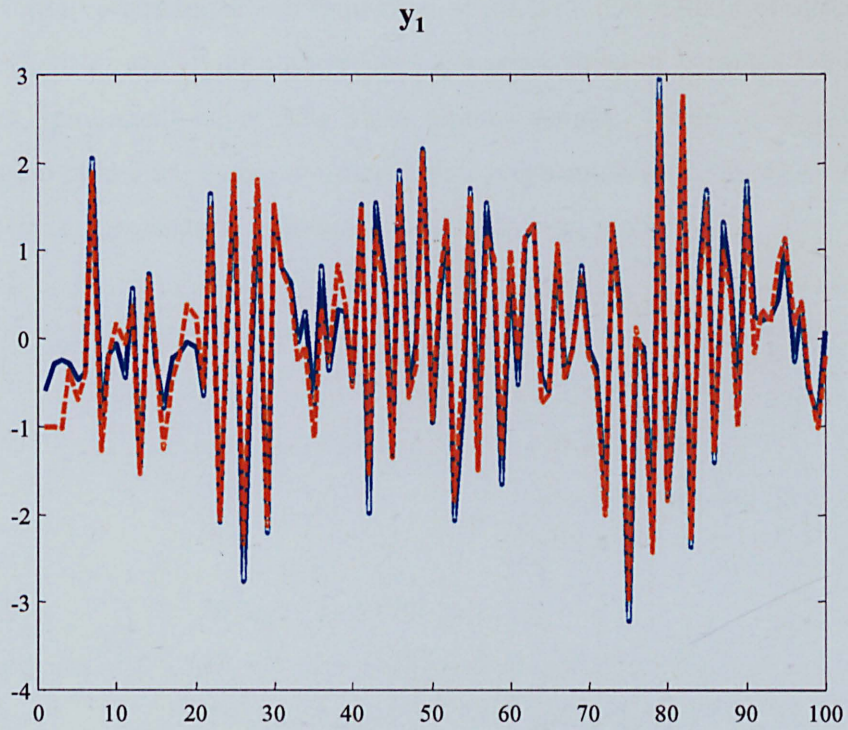
Figure 3.3. System output. For display convenience, only the first 100 time points of the time series are shown. The real output is shown in blue, the predicted model is shown in red dashed lines.

Where, u_1 and u_2 are independent sequences of uniform distribution of variance 1 (see figure 3.4). e_1 and e_2 are uncorrelated gaussian noise of variance 0.04. The time series generated were 200 time points length. Using a Non-linear autoregressive model of dynamic order=3 and polynomial order=2. The forward regression orthogonal algorithm identified the following model.

$y_1(t) =$ $\quad$ 0.517 $u_1(t-3)$+ 0.718 $u_2(t-2)$+

$\quad$ 0.341 $y_1(t-3)$ -0.224 $u_1(t-1)$ $u_2(t-3)$ +

$\quad$ -0.376 $y_1(t-1)$ $u_1(t-2)$ +0.483 $y_1(t-3)$ $u_2(t-1)$

$\qquad$ where var(e_1) =0.083

$y_2(t) =$ $\quad$ 0.236 $u_2(t-1)$ - 0.680 $y_1(t-2)$+

$\quad$ 0.309 $y_2(t-3)$ + 0.421 $u_1(t-1)$ $u_1(t-2)$ +

$\quad$ -0.779 $y_2(t-2)$ $u_2(t-3)$ +0.077 $y_1(t-3)$ $y_2(t-1)$

$\qquad$ where var(e_2)=0.122

The statistical validity of the model reported by AMOS 4.0 was $p > 0.05$ for each subsystem. The time series estimated by the algorithm are shown in figure 3.5.

3.5.3 Linear autoregressive interactions using simulated BOLD.

Let us consider a cortical network in which two regions become active under a certain experimental stimulation. The behaviour to model is the decrease in the response of the regions y_1 and y_2 over time, i.e., the magnitude of their response at the time t depends on their past activity. The dynamics of the cortical interactions can be represented by following model.

$y_1(t) =$ $\quad$ $u_1(t)$- 0.7 $y_1(t-2)$

$y_2(t) =$ $\quad$ 0.6 $y_1(t-1)$- 0.5 $y_2(t-2)$

$y_1(t) =$ $\quad$ $y_1(t)$ + $e_1(t)$

$y_2(t) =$ $\quad$ $y_2(t)$ + $e_2(t)$

$u_1(t) =$ $\quad$ $u_1(t)$ + $e_3(t)$
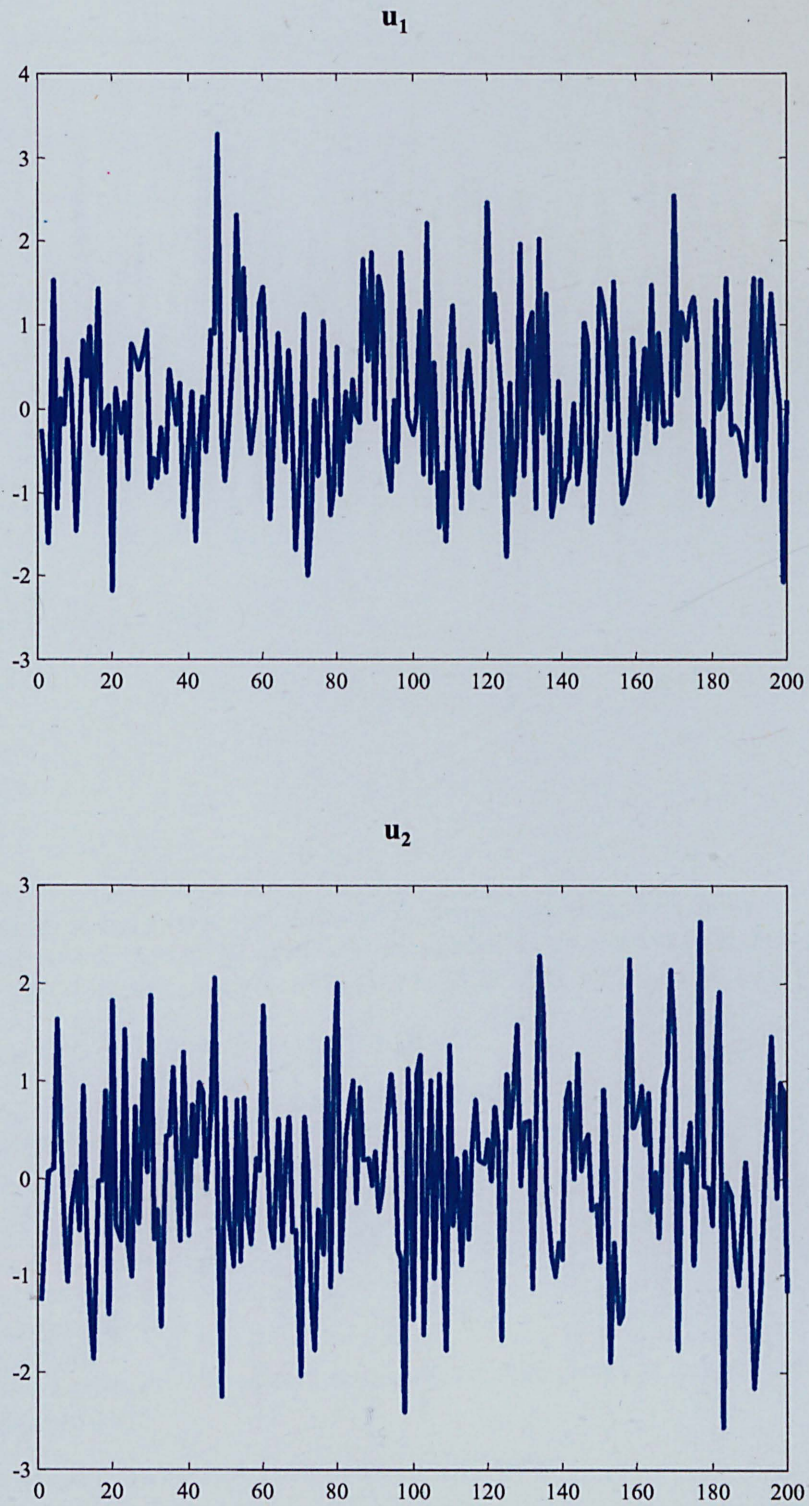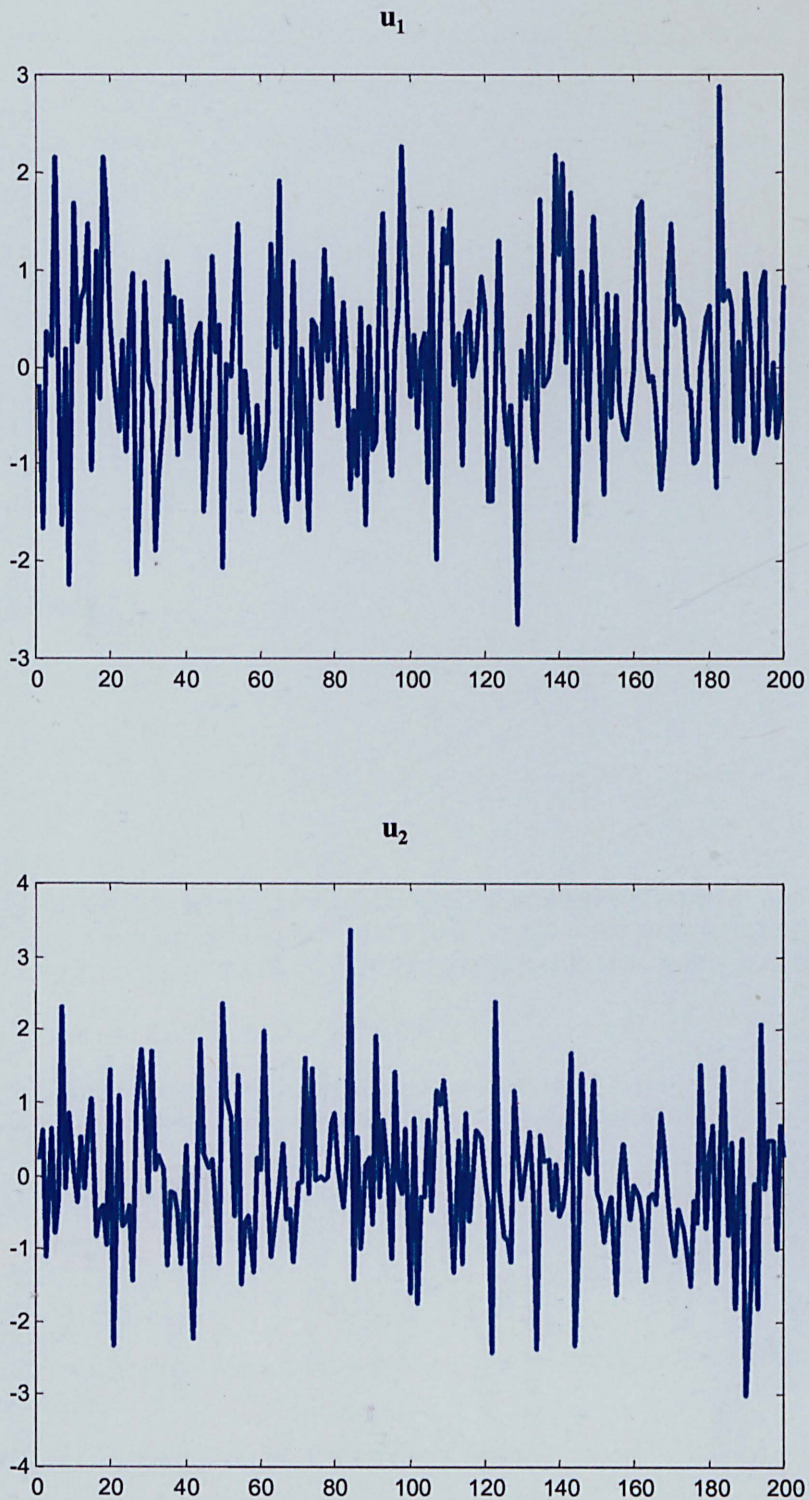
Figure 3.4. Exogenous inputs of the non-linear autoregressive model.

$$y_1$$

$$y_2$$
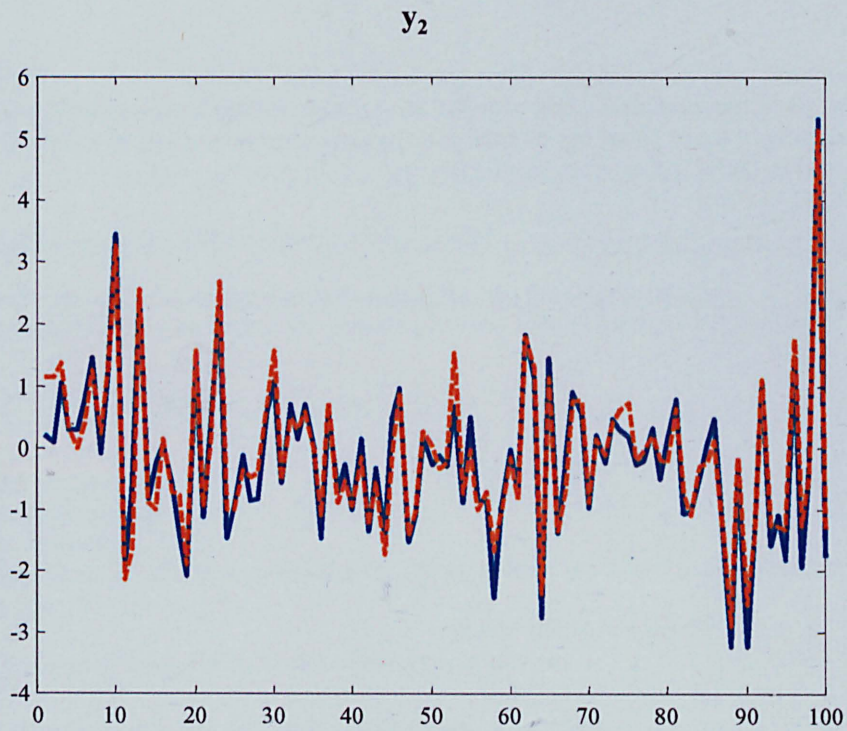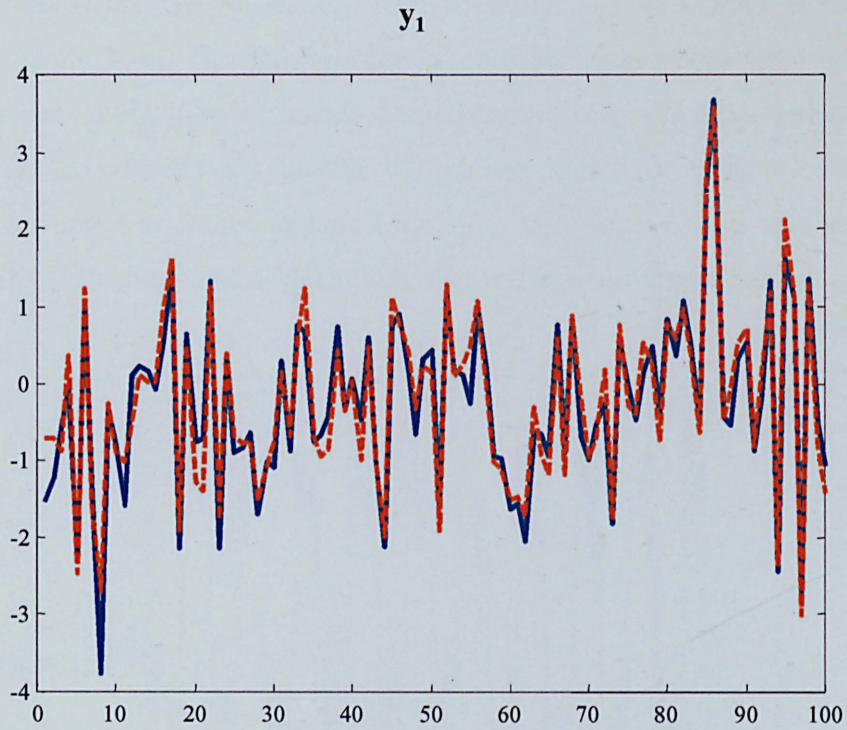
Figure 3.5. System output. For display convenience only the first 100 time points of the time series are shown. The real output is shown in blue, the predicted model is shown in red dashed lines.

Where u_1 represents the measured hemodynamic response after a boxcar (ABAB...) stimulation. The boxcar stimulus was 200 data points long with 10 data points per epoch. The estimated hemodynamic response time series were computed by convolving the boxcar stimulation paradigm with a canonical hemodynamic response function (consult 2.9.1). e_1, e_2 and e_3 are uncorrelated gaussian noise of variance 0.04, 0.02, 0.08, respectively (see figure 3.6).

$$u_1$$
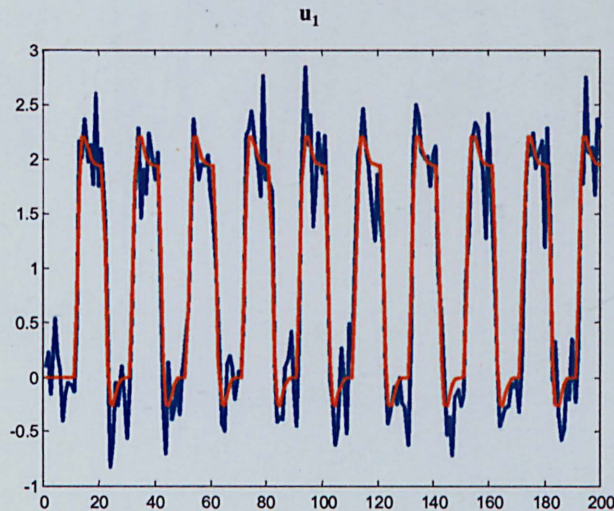
Figure 3.6 Exogenous input. Hemodynamic response of the boxcar experimental stimulation is shown in red (model). The blue time series represents the hemodynamic response plus random noise (e_3).

Using a linear autoregressive model of dynamic order=3 and polynomial order=1. The forward regression orthogonal algorithm identified the following model.

$$y_1(t) = \qquad 0.844\ u_1(t) - 0.556\ y_1(t-2)$$

$$where\ var(e_1)\ = 0.127$$

$$y_2(t) = \qquad 0.56\ y_1(t-1) - 0.432\ y_2(t-2)$$

$$where\ var(e_2) = 0.036$$

The statistical validity of the model reported by AMOS 4.0 was $p > 0.05$ for each subsystem. The time series generated by the proposed model are shown in figure 3.7.
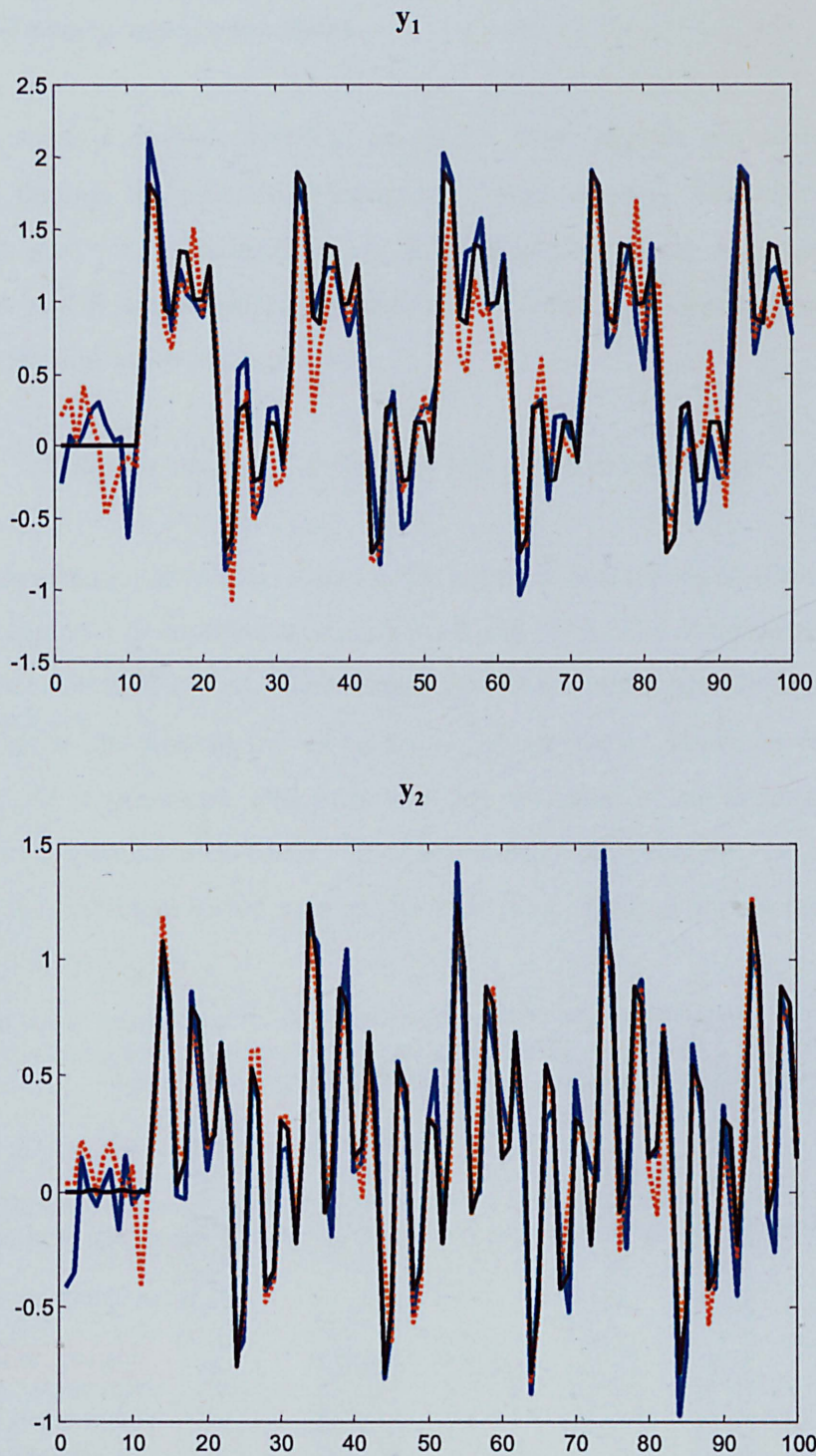
Figure 3.7. System output. For display convenience only the first 100 time points of the time series are shown. The model output (as it was generated by the model without noise) is shown in black, the real output plus noise is shown in blue, and the predicted model is shown in red dashed lines. As defined by the model of system, the output responses (y1 and y2) show a decrease in sensitivity over time, i.e. the activity on the past modulates the sensitivity of the system to new inputs (autoregressive system).

3.5.4 Non-linear neural interactions.

Let us consider a cortical network in which three regions are anatomically connected through the following structure: $u_1 \rightarrow y_1$, $u_2 \rightarrow y_1$. The experimental stimulation involves four conditions (A, B, A+B and N), where A+B means that stimulus A and B are presented together, and N is the base line condition. The complete stimulus sequence is given by:

$$N\,A\,N\,(A+B)\,N\,B\,N\,A\,N\,(A+B)\,N\,B\,N\,A\,N\,(A+B)\,N\,B\,N$$

Region u_1 is sensitive to stimuli A and A+B, region u_2 is sensitive to stimuli B and A+B, and region y_1 is sensitive to stimuli A, B and A+B. The non-linearity of the systems relies on the fact that, when stimuli A or B are presented individually, the relationships in the system ($u_1 \rightarrow y_1$, $u_2 \rightarrow y_1$) are linear. However when the stimulus A+B is presented, and u_1 and u_2 are activated at the same time, the relationship becomes non-linear. In other words, the response to the A+B condition it is different to the sum of the individual stimulus inputs (see figure 3.8).
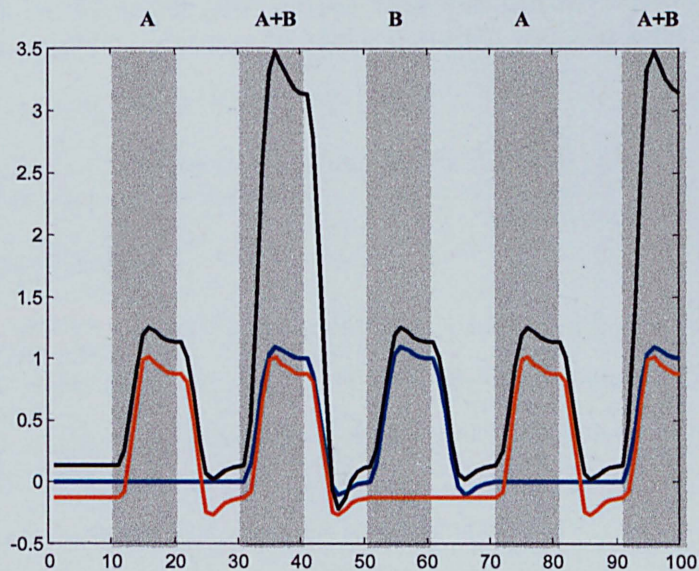
Figure 3.8 . Non linear response. The first 100 time points of the stimulus paradigm are shown. For display convenience the dc of each time series is moved. Red (u_1), blue(u_2), and black (y_1) time series represent the hemodynamic response of the regions.

The behaviour of the system can be represented by the following model.

$$y_1(t) = u_1(t) + u_2(t) + u_1(t)u_2(t)$$

$$y_1(t) = y_1(t) + e_1(t)$$
$$u_1(t) = u_1(t) + e_2(t)$$
$$u_2(t) = u_2(t) + e_3(t)$$

Where u_1, u_2 and y_1 represent the measured hemodynamic response under stimulation. Each condition was 10 time length (see figure 3.9), so each time series was 190 time point length. In order to simulate more realistic data, the data was firstly generated using a unitary response and then the time series of each region was convolved with a different hemodynamic response (see section 2.9.2 of chapter 2).

e_1, e_2 and e_3 are uncorrelated gaussian noise of variance 0.6, 0.15, 0.15, respectively. These variances were selected to introduce 45 % signal of measurement noise. Using an autoregressive linear model of dynamic order=0 and polynomial order=2. The forward regression orthogonal algorithm identified the following model.

$$y_1(t) = 0.908\ u_1(t) + 0.923\ u_2(t) + 0.667\ u_1(t)u_2(t)$$

$$where\ var\ (e_1) = 0.951$$

The statistical validity of the model reported by AMOS 4.0 was $p > 0.05$. The time series generated by the proposed model are shown in figure 3.10.
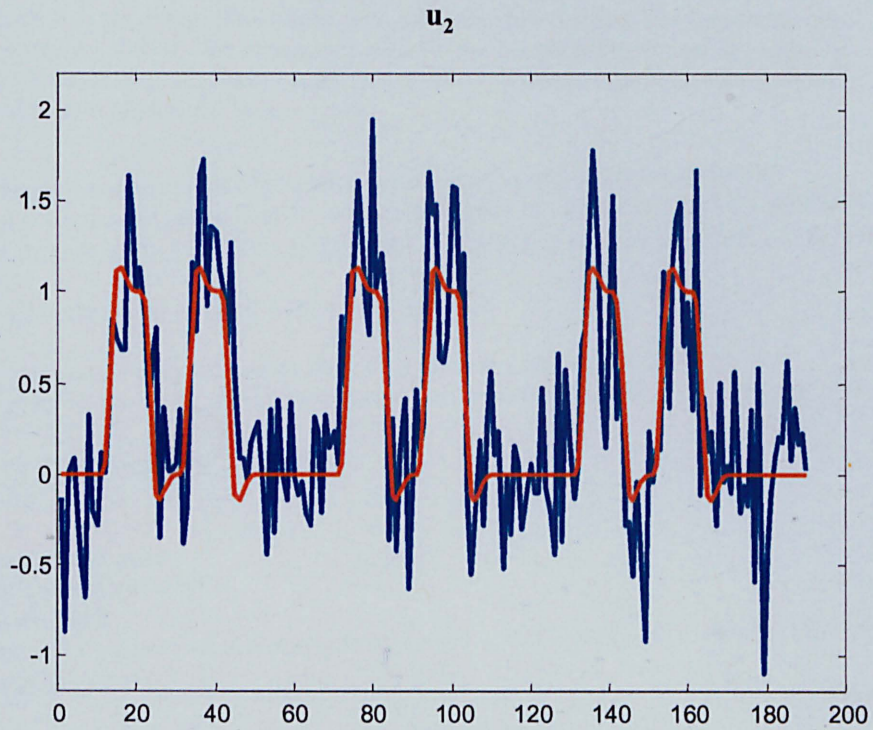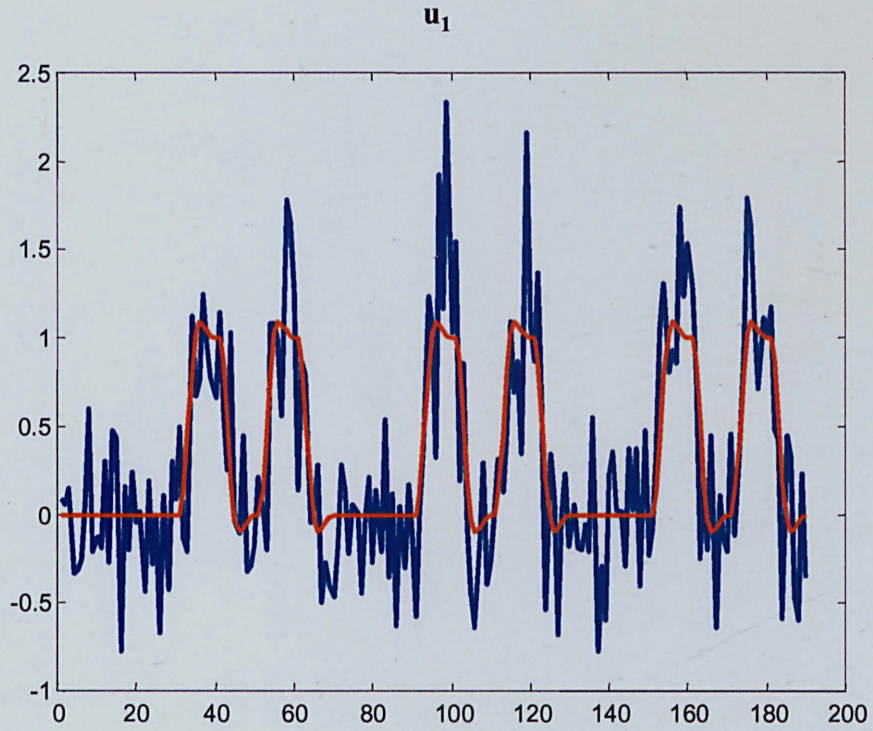
Figure 3.9. System inputs. For display convenience only the first 100 time points of the time series are shown. The model signal (without noise) is shown in red, the measured input is shown in blue.

y_1
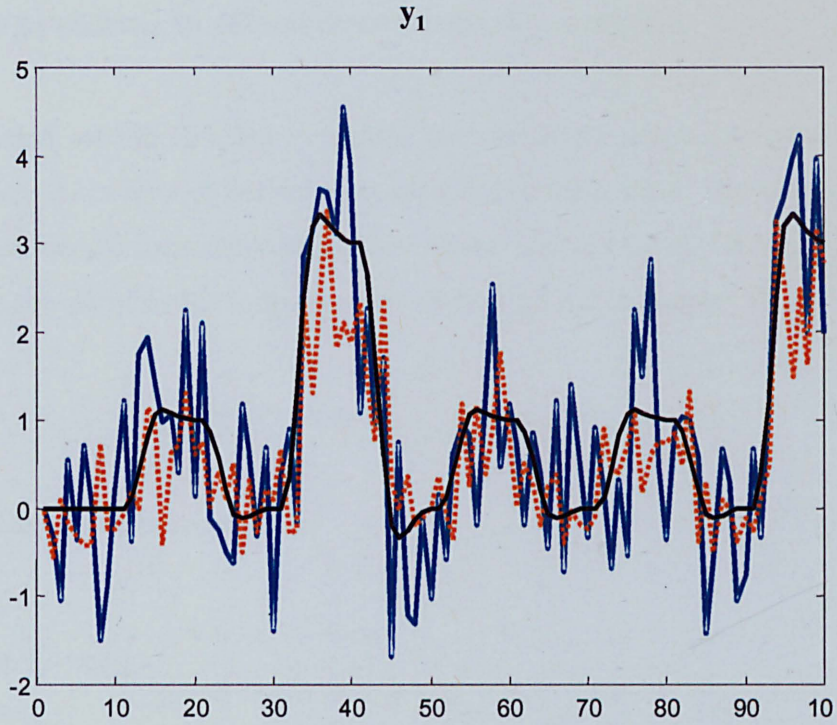
Figure 3.10. System output. For display convenience only the first 100 time points of the time series are shown. The model output (as it was generated by the model without noise) is shown in black, the real output plus noise is shown in blue, and the predicted model is shown in red dashed lines.

3.6 Open problems in effective connectivity analysis.

Ill-conditioned models are those models for which no unique solution exists because there is not enough information specified in the system. The solution of a linear system can be formulated using the General Linear Model (GLM), in which each linear equation has the form (see also section 1.4.1.1 in chapter 1):

$$y = \theta_1 x_1 + \theta_2 x_2 + \ldots + \theta_n x_n + e \qquad (3.17)$$

And the system is represented as:

$$y_1 = \theta_1 x_{1,1} + \theta_2 x_{1,2} + \ldots + \theta_n x_{1,n} + e_1$$
$$y_2 = \theta_1 x_{2,1} + \theta_2 x_{2,2} + \ldots + \theta_n x_{2,n} + e_2$$

$$\vdots$$

$$y_m = \theta_1 x_{m,1} + \theta_2 x_{m,2} + \ldots + \theta_n x_{m,n} + e_m$$

Where m represents the number of equations in the system and n represents the number of variables in the system. In matrix notation:

$$Y = X \theta + E \qquad (3.18)$$

The solution of the system is given by:

$$\theta = (X^T X)^{-1} (X^T Y) \qquad (3.19)$$

Where ^{T} denotes for matrix transposition and $^{-1}$ denotes for matrix inversion. For the sake of clarity, let us to illustrate the ill-conditioned problem using a simple linear system which is a particular case of the General Linear Model (the noise term is not represented). Consider the following system of equations in which each equation represents a line.

$$-1 = \theta_1 - 2\theta_2$$

$$3 = -\theta_1 + 3\theta_2$$

Where

$$\mathbf{Y} = \mathbf{X} \quad \theta$$

$$\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

Graphically this system of equations looks like:
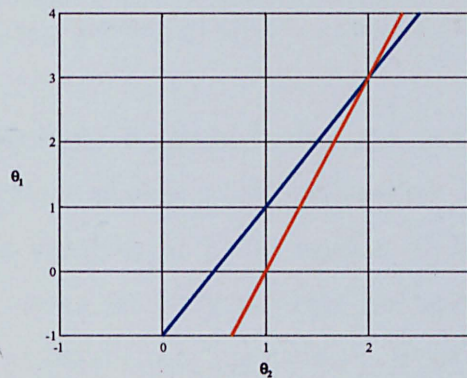
Where the blue and red lines represent the first and second equations of the system. The solution of the system is given by the point in which the lines intersect each other (θ_1=3 and θ_2=2). This point can be identified using formula 3.19. However, two lines can not necessarily intersect in one point, i.e. they could be parallel (no solution) or they could coincide, in which case they intersect in every point on the line (not unique solution), see figure 3.11.

For those cases in which the system has infinite number of solutions it is said that the model is ill-conditioned or bad conditioned. In these kind of systems the inverse of the matrix X^TX in equation 3.19 is not computable because the determinant of X^TX is equal to zero and the matrix is call singular or non-positive definite. This problem arises because some variables in X are highly correlated or they are linear combinations of each other.

$-1 = \theta_1 - 2\theta_2$

$-1 = \theta_1 + 2\theta_2$

(a)

$-1 = \theta_1 - 2\theta_2$

$1 = -\theta_1 + 2\theta_2$

(b)
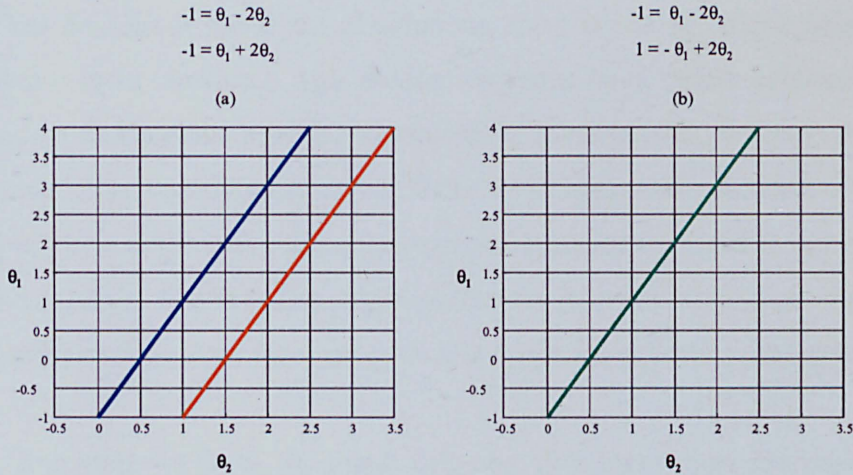
Figure 3.11 Linear systems. A linear system can also have no solution (a) parallel system or infinitely number of solutions (b) collinear system (modified from (Lay 2000)).

Strictly speaking, singularity is reached when the variables (terms) in X are perfect linear combinations of each other (collinearity), however, although for a less highly correlated variables in X the inverse of $X^T X$ is computable, the solutions of these systems are very unstable because there are many local solutions instead of a global solution. For example, let us imagine an hypothetical space of solutions represented as a three-dimensional surface (fig. 3.12), in which the axis θ_1 and θ_2 represent the parameters to be identified in the model and the axis $f(\theta_1, \theta_2)$ represents the function to be maximised (consult section 2.5.1).
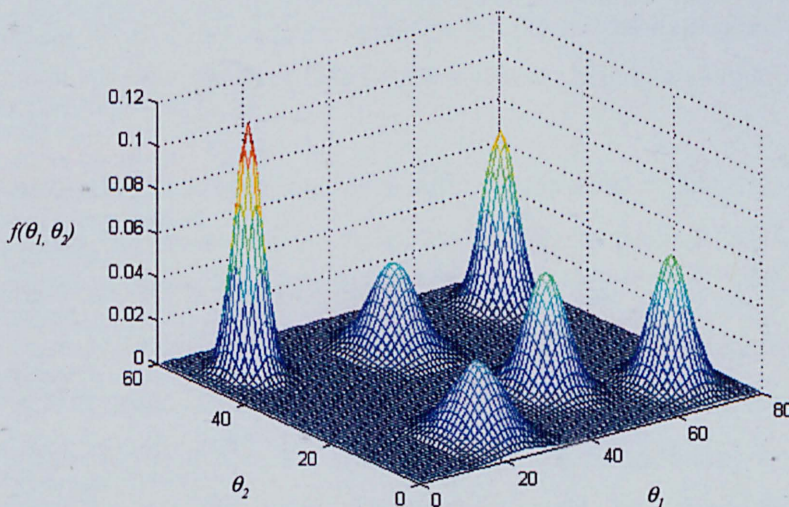
Figure 3.12 Space of solutions.

As it can be seen in the space of solutions, there is not an unique solution (there are many local maximal) and the convergence to a given solution is highly dependent on the place in which the searching is started. So, although the model is mathematically consistent (because there is at least one solution), for practical purposes it is empirically under-determined because it is not possible to find a unique solution. For the sake of simplicity the previous examples engaged only two variables, however, the same principle apply for n dimensional systems.

It is important to have in mind that the difference between model under-determination due to the insufficient number of equations with respect to the number of parameters to being estimated (see *the t rule* in section 2.4.1) and model under-determination due to ill-conditioned conditions. As stated in chapter 3, *the t rule* is a necessary but not sufficient condition to get identification, i.e., although exists more equations than unknown parameters the model could still have many possible solutions.

3.6.1 Effective connectivity analysis of fMRI data.

Effective connectivity analysis can be thought of as the simplest causal model that would replicate the observed time relationships between cortical regions. The structure of the model is defined using *a priori* knowledge about the anatomical connections. Because the cortical activations are determined by the controlled experimental stimulation, their fMRI time series are highly correlated.

Let us exemplify the problem using a very simple model. Suppose that fMRI is used to record the neural activity raised under a given boxcar (ABABABABAB) stimulation (10 data points per condition). After data analysis, three cortical regions (x_1, x_2, x_3) are identified and their time series are extracted (fig. 3.13). Assuming that based on anatomical maps, the structural model shown in figure 3.14 is proposed. As it is easily seen in the system of equations that defines the structure of the model, the variance in x_3 can be explained either by x_1 or by x_2.
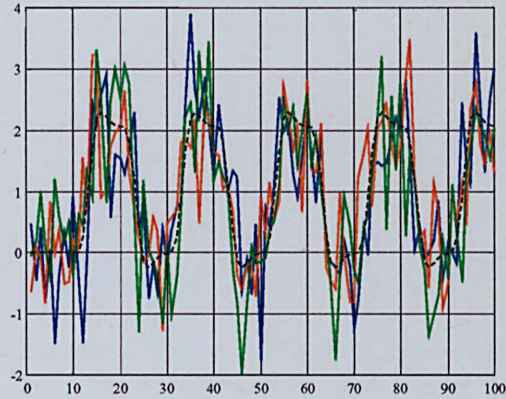
Figure 3.13 Time series. Three artificial time series were generated convolving the boxcar stimulation with a canonical hemodynamic response. Measurement noise was added as gaussian noise at 50% of the variance (terms e_1 and e_2 in figure 3.14). Blue, red and green time series represent the BOLD response of regions x_1, x_2, x_3 respectively. Dashed black line represents the hemodynamic response model. X axis represents time (volumes) and Y axis represents an arbitrary scale of signal change.

Therefore the model is ill-conditioned since there are different sets of parameters (θ_2, θ_3) that satisfy the system. In other words, there is redundant information in the system, since the variance of x_3 can be entirely explained by one (x_1) or another term (x_2).
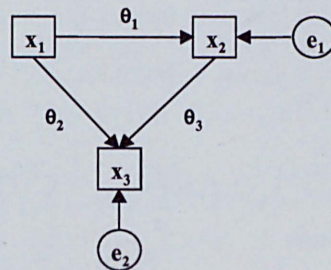
Structural equations

$$x_2 = \theta_1 x_1 + e_1$$

$$x_3 = \theta_2 x_1 + \theta_3 x_2 + e_2$$

Figure 3.14 Structural model.

It is important to realise that, contradictorily, the less noisy the data the worst the bad conditioning problem, that is to say, if the measurement noise is almost null the matrix becomes singular and the inverse of $X^T X$ does not exist (fig. 3.15). On the other hand, although the big amount of noise characteristic in fMRI data let us to estimate the inverse of $X^T X$, the solution is very unstable. This is a typical problem on fMRI models in which predictors of an independent variable are correlated as a result of the functional profile of the regions involved.
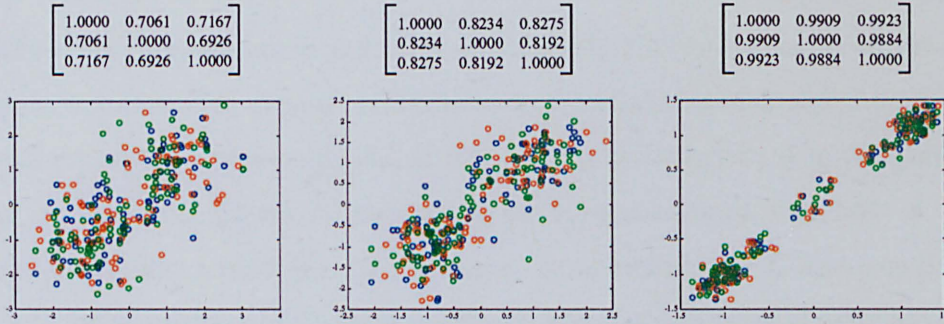
Figure 1.15 Singularity problem. Three sets of time series were generated for the model showed in figure 3.14. using different amounts of noise. The graphs show the correlation matrix for x_1, x_2, x_3 and the plots: x_1, x_3 (blue), x_2, x_3 (red), x_1, x_2 (green). From let to right 50, 20 and 1 percent of noise were used.

Even using a very simple model like that showed in figure 3.14 it is not possible to make any reliable conclusion about the meaning of the parameters estimated, because the solution obtained represents only one of many. That is to say, by definition, parameters of effective connectivity represent the strength of the influence that one cortical region exerts over another, however, under ill-conditioned models it is only possible to made conclusion about functional connectivity (correlations) but nothing can be said about effective connectivity (Friston, personal communication).

3.6.2 SEM and ill-conditioned models.

As Structural Equation Modelling (SEM) can be seen as an extension of the General Linear Model (GLM) it shares the same limitation to assess ill-conditioned models. In practice the SEM's community avoid the problem of empirical under-determination assuming that the error term in the model is not only *measurement noise* but also it is *system noise* (which is propagated through the network). The justification to the assumption of system noise is that every model is a simplification of the real system and there are different sources of information which are not represented in the model (nodes not included in the model). The conceptualisation of the error term as a system noise can be thought of as an "injection" of extra information, which lets the model to have a unique solution, i.e. the system of equations is solved by a unique set of parameters (values).

However this assumption is not valid in fMRI data since the noise in the time series is dominated by measurement noise (e.g. variation in the signal due to the limitations of the equipment); and if any system noise exists, it is not gaussian noise because the system is responding to designed inputs that have a large deterministic component (Friston, personal communication). Consequently, the lack of "noise propagation" is that the model has many solution and the values of the parameter are highly dependant on the starting values used in the optimisation technique (Bollen,1989).

For example, let us apply standard MonteCarlo simulation techniques used in SEM to generate data for the model presented in figure 3.14. The time series were generated in a similar way that those showed in figure 3.13. The parameters used in the simulation are shown in figure 3.16 (a). SEM software (AMOS 4) was used to identify the parameters of the model. Since error terms were propagated, SEM was able to identify the original parameters, see figure 3.16 (b).
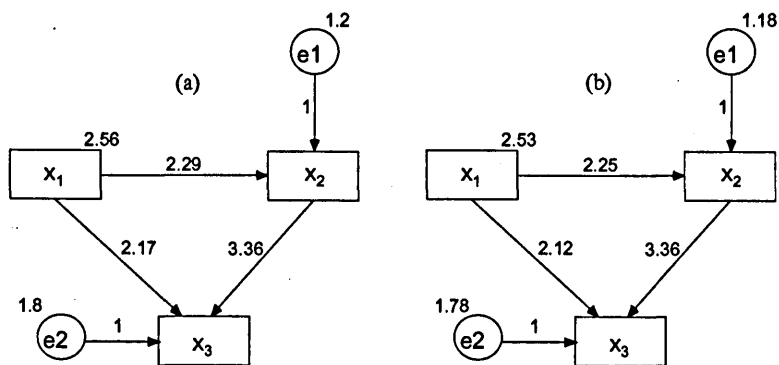
Figure 3.16 System noise. (a) Original model used to generate the data. The error terms (gaussian noise) where propagated through the network. (b) Model estimated using SEM software. The probability level was > 0.1.

However generating data for the same model without noise propagation (a more realistic model for fMRI data), the parameter estimation is empirically under-determined because the solution reached is determined by the starting values used by the optimisation technique. As it can be appreciated from the data simulation without noise propagation, a model different to that one used to generate the data was able to reach statistical significance to represent the data, see figure 3.17 (a).

On the other hand, fixing one of the parameter values (θ_3=3.36) to bias the search to converge to the "true" solution, the parameter estimation is able to find the original model, see figure 3.17 (b). This example shows the instability of the SEM's parameter estimation under ill-conditioned models.
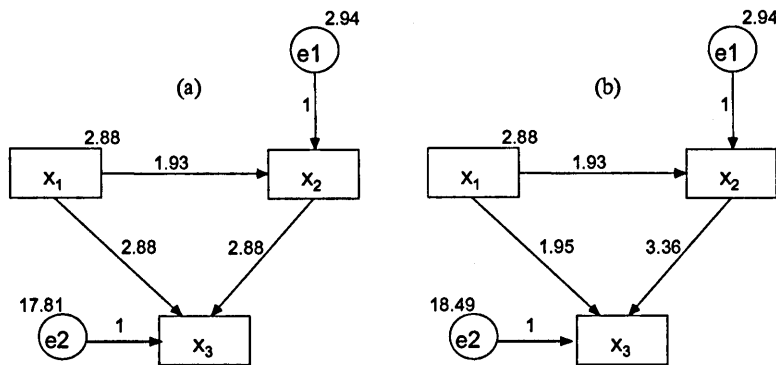
Figure 3.17 Measurement noise. Data for the model presented in figure 3.16a was generated without noise propagation. (a) Model estimated using SEM software. (b) The model was forced to find the original model. Both models reached a probability level > 0.1.

One of the most common applications of SEM in neuroimaging has been the comparison of the parameters in a given model under different experimental conditions or between groups (see section 2.6 in chapter 2). However, as it was showed above, differences in parameters might not necessarily mean changes of effective connectivity.

3.6.3 NARMAX and ill-conditioned models.

Although the Non-linear AutoRegressive Moving Average with eXogenous variables (NARMAX) approach provides a better representation than SEM to model non-linear dynamic systems like fMRI data, it has similar limitations to deal with ill-conditioned models. Using the NARMAX approach singularity is not a problem, because the inverse of X^TX is computed iteratively in the orthogonalisation procedure. However, the system identification is affected at the time to select the orthogonal components in the covariance matrix.

For example, let us consider the data generated for the model shown in the figure 3.17. (no noise propagation). Then, after using NARMAX algorithm to identify the system, it suggested the model shown in figure 3.18 (a). As it can be seen in the model, the orthogonalisation procedure deleted the arc form x_1 to x_3. The explanation for that is that one predictor (x_2) is enough to explain most of the variance in x_3.

However using a more relaxed criterion to decide which terms of X are redundant (and have to be deleted), i.e. how similar (correlated) two terms should be to be considered collinear. The NARMAX algorithm reached the same solution as SEM. See figure 3.18 (b).
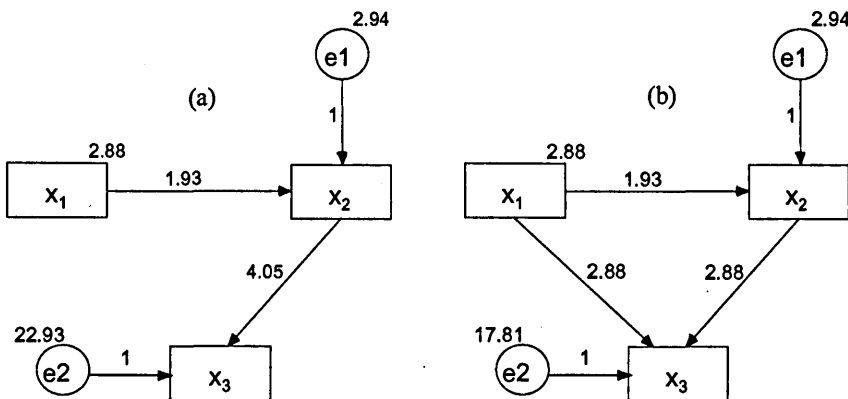
Figure 3.18 System identification of ill-posed models using NARMAX. NARMAX algorithm was used to identify the model shown in figure 3.17. The algorithm was run using a dynamic order=0 and polynomial order=1. (a) The orthogonalisation criteria to delete terms was set up to 0.5, i.e. when two terms have correlation bigger than 0.5 .The convergence criteria reported that x_2 was able to explain 90.73 % of the variance in x_2. (b) The orthogonalisation criteria was set up to 0.85.

As it can be appreciated from the previous example, although NARMAX approach "solves" the ill-conditioned problem through the ortogonalisation of the covariance matrix, the cost of this solution is not acceptable in this application. By definition, effective connectivity is a confirmatory analysis, i.e. it relies on the anatomical model. So, the structure of the model can not be changed on the sake of mathematical consistency.

NARMAX approach can be thought of as a forward regression orthogonal algorithm which brings SEM towards a non-linear dynamic modelling technique. However under ill-conditioned models, NARMAX is good for prediction but not for explanation (McIntosh personal communication). That is to say, NARMAX approach is good to develop exploratory analysis (find a model which fits the observed data) but not necessarily to make a confirmatory analysis (identify the "true" parameters). So, in ill-conditioned models the parameters estimated by NARMAX are as meaningless as those estimated by SEM.

As a summary it can be said that ill-conditioned models are those models which have not a unique solution because nodes are highly correlated and predictors of dependent variables are usually redundant. Effective connectivity models using fMRI data easily becomes ill-conditioned since cortical activations are produced by designed stimuli that have a large deterministic component.

Under ill-conditioned models it is not possible to make any conclusion about effective connectivity since the solution of this models are very unstable and it is not possible to know the "true" parameters which produced the data in the physical model. Neither SEM nor NARMAX approaches are able to solve this identification problem because there is not enough information in the system to find a unique solution. So, the parameters identified in this kind of models are meaningless. These models are empirically under-determined and no reliable conclusions can be made about effective connectivity, conclusions can be formulated only in term of functional connectivity (correlations).

Approaches based on GLM (from which SEM and NARMAX are part) are not appropriate to deal with effective connectivity ill-conditioned problems since they are good for prediction but not for explanation. Although some alternative representations like state spaces or Bayesian modelling have been recently explored to develop connectivity analysis, more research has to be done to create better representations of effective connectivity models using fMRI data (Friston 2002; Friston, Harrison et al. 2002).

Another open problem in effective connectivity analysis using fMRI time series is that although empirically identified models have unique solutions, the meaning of the parameters remains uncertain. There is not clear understanding about the relationship between neural activity and Blood Oxygenated Level Dependent (BOLD) changes. So, it is not reliable to assume that parameters of effective connectivity are the same at both levels.

For example, hemodynamic responses from different regions on the brain are characterised by different parameters, e.g. neuronal efficacy, autoregulation, resting oxygen extraction, etc. (Friston, Mechelli et al. 2000; Zheng, Martindale et al. 2002). Thus, the same input can produce different hemodynamic responses in different regions (fig. 3.19).
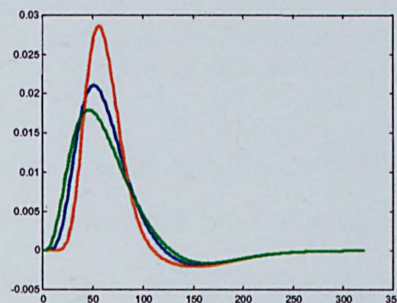
Figure 3.19 Hemodynamic responses

Therefore, although at a neural level the relationships in the model could be linear, at a BOLD level the model is inherently non-linear. This implies that the parameters identified at BOLD level would not be the same that those in the physical system (neuronal level), see figure 3.20.
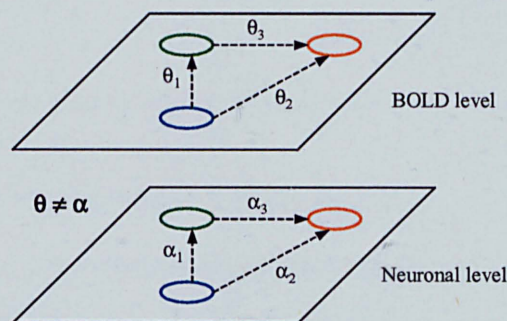
Figure 3.20 Different levels of representation for effective connectivity models.

One possible solution to estimate the parameters at low level is to infer the neuronal activity from the BOLD response, i.e., to convert hemodynamic response into neurone spikes. Then, the estimation of the model can be made at a neuronal level (fig. 3.21).
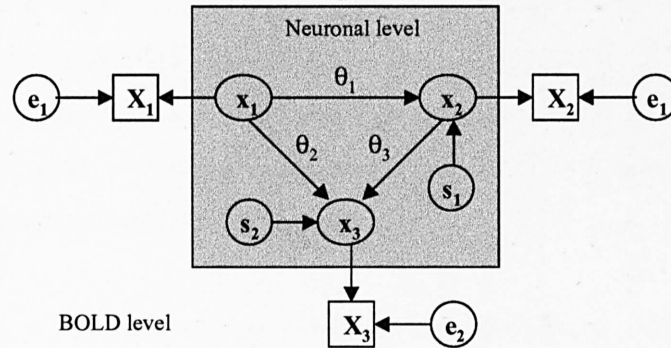
Figure 3.21 Model representation at BOLD and neuronal levels. BOLD measurements are represented as squares (X). The error term associated at this level is considered measurement noise (e). Neuronal activity is represented as ellipsis (x). The error terms associated at this level are considered system noise (s).

There are well known forward methods to estimate BOLD response from neural activity. These methods take in account the physiological properties of the cortical region from which to estimate the BOLD response. For example, "the Balloon model" proposed by Friston *et al* (Friston, Mechelli et al. 2000; Zheng, Martindale et al. 2002) estimates the BOLD signal as a non-linear function of volume (v), deoxyhemoglobin (q) and oxygen extraction fraction by the capillary bed (E_o), see figure 3.21.
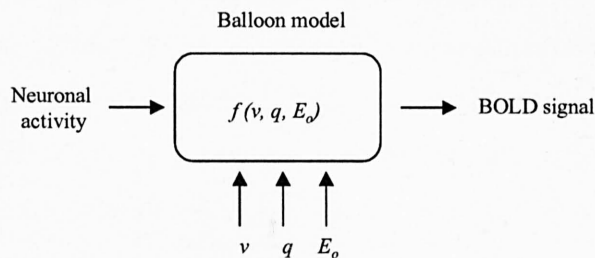
Figure 3.21 Balloon model.

However, the estimation of neuronal activity from BOLD signal is not trivial and by the moment it does not exist a methodology to develop this "inversion" properly (Zheng, Martindale et al. 2002).

Another advantage of this multilevel analysis is that error terms can be modelled discriminating between measurement noise (e) and system noise (s). Measurement noise can be estimated as the remaining part of the variance (at BOLD level) that can not be converted into neuronal spikes. Then, any remaining error at neuronal level should be considered system noise. We hypothesise that system error should to be propagated through the network. It is important to note that the representation of the model at different levels would permit not only to have a better understanding of the interactions at low level but also to specify the problem in agreement with SEM assumptions.

The necessity of the analysis of effective connectivity at different levels of representation was also addressed by Friston *et al* (Friston 2002; Friston, Harrison et al. 2002) almost at the same time than us. In accordance with our observations, they concluded that the success of this approach depends on the "inversion" of the forward models, i.e. the estimation of neuronal activity from BOLD signals (Friston personal communication).

NOTE: It is important to clarify that effective connectivity analysis was only developed on the data acquired in experiment 1 (chapter 4) because the experimental stimulation used on it was specifically designed to test the hypothesis of the involvement of V5 region in the processing of stereo information. Due to finding that V5 was not sensitive to stereo disparities it was possible to apply SEM and NARMAX methodologies since the patterns of activation were not correlated (the model was not ill-conditioned). However, on the subsequent experiments (chapters 5, 6 and 7) our efforts were concentrated on the identification of the cortical regions involved in stereo disparity processing and the experimental stimuli were not designed to develop any analysis of effective connectivity.

3.7 Conclusions.

In the present chapter a Non-linear Auto-Regressive Moving Average with eXogenous variables (NARMAX) algorithm was introduced as an alternative approach to represent models of effective connectivity using fMRI time series. This approach can be thought of as a dynamic non-linear system identification technique which overcomes some of the limitations of Structural Equation Modelling (SEM). SEM is in principle appropriate to represent static linear systems, whereas fMRI time series are inherently non-linear dynamic systems (Friston, Frith et al. 1995; Friston, Josephs et al. 1998).

BOLD hemodynamic response is a non-linear transformation of the stimulation input, and interactions between regions are non-linear as well. Then, non-linearity is an important characteristic to contemplate in models of effective connectivity.

Although in SEM it is possible to introduce moderator variables (interaction terms) to represent non-linear relationships, the advantage of NARMAX approach is that it is able to identify those terms, in other words, NARMAX can be considered as a method not only for parameters estimation but also for model identification.

Hemodynamic response latencies differ between cortical regions and the assumption of immediate responses is not convenient. A more relaxed assumption is to let the model accommodate lags in time, i.e. to assume a causal relationship in which the input could happen in the past. This temporal property is not considered in SEM since it was created for the modelling of static systems. On the other hand, time is an important characteristic in the analysis of fMRI time series. The hemodynamic response in a region, not only depends on the inputs but also on its own activity in the past. This autoregressive property is incorporated in NARMAX approach as part of the polynomial expansion.

It is important to notice that although Volterra series are as powerful as NARMAX in representing non-linear dynamical systems, Volterra representation is useful to model the behaviour of a *single region* (MISO), whereas NARMAX representation offers a framework to represent interactions between multiple regions (MIMO) (Billings, Chen et al. 1989; Friston, Josephs et al. 1998; Friston and Buchel 2000).

Another difference between SEM and NARMAX approaches is how they interpret the noise (error) term. In SEM the noise term is considered "system noise", that is to say, the error is propagated through the network, whereas in NARMAX it is considered "measurement noise", i.e. it is uncorrelated with the regions included in the model (Bollen 1989).

The definition of the error term as measurement noise is much more appropriate for fMRI data, because system noise is dominated by the huge amount of measurement noise characteristic of fMRI time series (Friston, personal communication).

Finally, the parameter estimation procedure used by NARMAX is more efficient than that used by SEM in the way the observed covariance matrix is inverted. When two or more regions are very similar, they create a collinearity problem, in other words, the inverse of the covariance matrix cannot be computed. A typical mathematical error produced by the parameter estimation procedure of SEM is the "singularity" of the covariance matrix. By the operational definition of NARMAX approach, it is not a problem due to the orthogonal decomposition of the inputs (Billings and Voon 1986b; Chen, Billings et al. 1989).

As a conclusion, it can be said that SEM was a proper technique for its first applications to effective connectivity analysis using static data (2-deoxyglucose autoradiography or PET) (McIntosh and Gonzalez-Lima 1992; McIntosh and Gonzalez-Lima 1994a). However, the application of this technique to fMRI time series requires a modification of the technique to satisfy the non-linear dynamic properties of fMRI data.

The Non-linear Auto-Regressive Moving Average with eXogenous variables (NARMAX) approach described in the present chapter can be thought of as an attempt to bring SEM towards a non-linear dynamic system modelling technique which permits a more appropriate representation of effective connectivity models using fMRI time series. However, although NARMAX approach provides a better representation than SEM to model non-linear dynamic systems, it has similar limitations to deal with ill-conditioned models.

Chapter 4

Experiment 1

Stereo Motion Interactions

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 stereopsis (V3B) and one more
responsive to both stimuli (V3A). 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.

4.1 Introduction.

Although many psychophysical studies have investigated how the human brain computes stereoscopic information (Tyler 1973; Burt and B. 1980), it is not certain 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 (Cumming and Parker 1999). MT/V5 in monkeys shows a columnar organisation tuned for disparity (DeAngelis and Newsome 1999). MT/V5 in human brains has been widely reported as a motion sensitive area (McKeefry, Watson et al. 1997; Tootell, Mendola et al. 1997; Smith, Greenlee et al. 1998a; Braddick, O'Brien et al. 2001). Given the similarity between the visual system of the monkey and the human (Gonzalez and Perez 1998), it is not unreasonable to think that V5 in human brains is also involved in the processing of stereo information.

However some studies of monkeys with lesion in V4 and V5 regions reported no loss of performance in a stereoscopic depth task (Schiller 1993). This could be attributed because either V5 is not related to stereo disparities processing or because there is another region (beyond V5) more sharply tuned to stereoscopic information. The aim of the present study is not only to investigate the cortical areas involved in the processing of stereo information but also to show how these areas interact with the V5 region.

4.2 Experiment Design.

All subjects were given 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 21s., giving 7 multislice volumes per condition (TR=3 s.).

A dummy condition of a blank screen was presented during the first 15s (5 scans) of each run to eliminate magnetic saturation. To avoid habituation the conditions were counterbalanced using a Latin Squares design (figure 4.1).
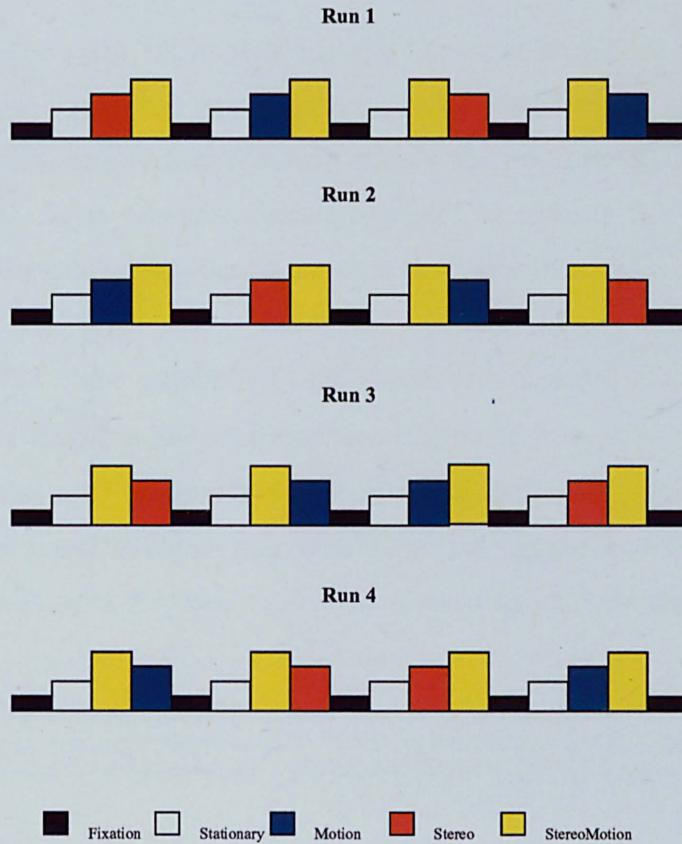
Figure 4.1 Stimulation Sequences.

The motion stimulus was radial to facilitate fixation following by Buchel's study (Buchel and Friston 1997; Buchel, Josephs et al. 1998). The subjects were instructed to fixate a point (0.3 deg. of radius) in the middle of the screen (circular field of view 13 deg.) and foveate while presented with the visual stimuli. There were five conditions of visual stimulation (fig. 4.2).

a) **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.

b) **Stationary**: Two hundred and fifty dots (with radio 0.1 deg.) were randomly positioned within the circular field of view (mean dot density 8 dot deg^{-2} at the centre and 1 dot deg^{-2} at the edges) , the aim of this condition was to activate the visual cortex areas sensitive to the luminance produced by the dots.

c) **Motion**: The same set of dots moving (constant speed 6.8 deg s^{-1}) radially, changing from expansion to contraction every 3 s. The dot density was kept 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. Radial motion was used because it facilitates fixation.

d) **Stereo**: The same number of dots positioned in depth (red/green anaglyph stereogram) forming a 3D cone structure (disparity from 0 to ±1.5 deg.). A 3D cone shape was used to provide a wide range of disparities to stimulate binocular neurons. To counterbalance this condition, the appearance of the cone was changing from concave to convex, but maintained constant during each epoch.

e) **StereoMotion**: The above Stereo and Motion stimuli were combined. This visual stimulus was designed to activate stereo and motion sensitive areas at the same time.

4.3 Methods.

4.3.1 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.
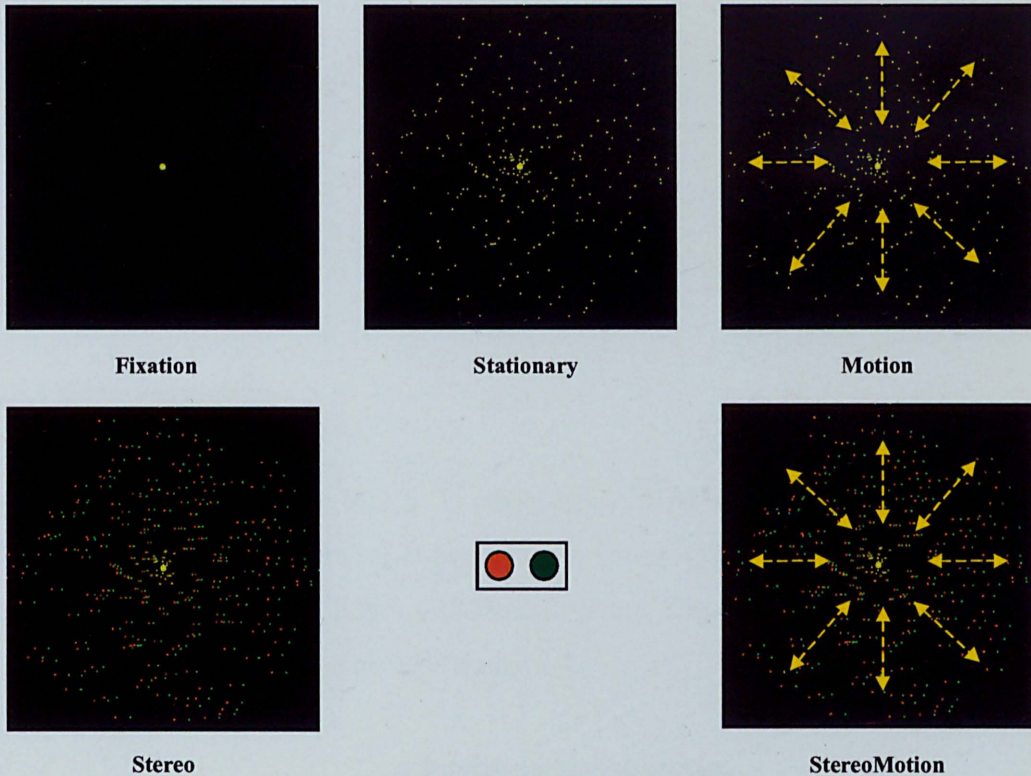
Figure 4.2 Visual stimuli. Subjects viewed the stimuli through the red/green glasses to present anaglyph stereo stimuli. Two directional arrows represent the expansion contraction radial motion.

4.3.2 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 just outside the magnet (fig. 4.3). The viewing distance was 2.4 m. Stimuli were projected on to the screen using an EPSON (EMP-7300) projector driven by a G3 Mac running Psychophysics Tool Box ver. 2.44 (Brainard 1997; Pelli 1997) under MATLAB ver. 5.3.

The mean luminance of the image was 30 cd/m^2. Although the stimuli were displayed at a video frame rate of 60 Hz, the image was only updated on every 10th frame, producing an effective frame rate of 6 Hz.
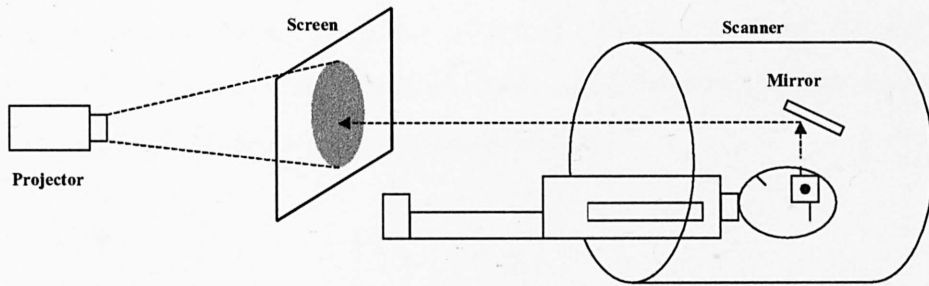
Figure 4.3 Stimulus display.

4.3.3 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 (figure 4.4).
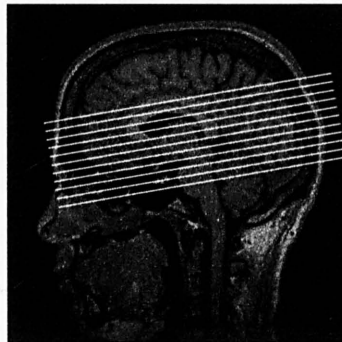
Figure 4.4 Data acquisition.

4.3.4 Data analysis.

The data was pre-processed and analysed using SPM99 (Wellcome Department of Cognitive Neurology). The first five scans of each run were discarded in order to exclude magnetic saturation artefacts. All volumes were slice timed with respect to the top slice. Motion correction was made taking as a reference the first volume of the first run. The data was normalised in the MNI (Montreal Neurological Institute) stereotaxic space and smoothed using a 6 mm FWHM (full width at half maximum) isotropic Gaussian kernel (WDCN 1997).

Data analysis was performed using a boxcar design matrix of the different conditions (fixation, stationary, motion, stereo and stereomotion) convolved with a canonical hemodynamic response function (fig. 4.5).
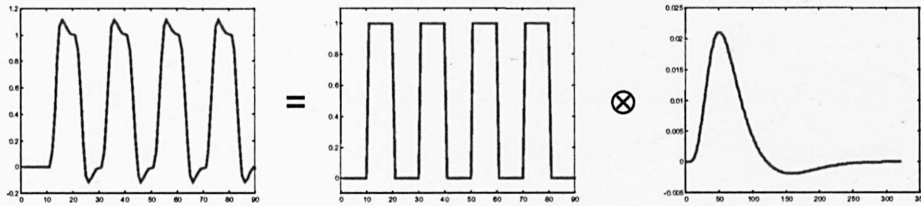
Figure 4.5. Hemodynamic model. The hemodynamic model used for regression was created by convolving ($\otimes$) the box car sequences of stimulation with the standard hemodynamic response function calculated by spm_hrf(3).

Specific effects were tested by applying the corresponding linear contrast to the parameters obtained applying General Linear Model (GLM) using the design matrix shown in figure 4.6 (see also experiment design).
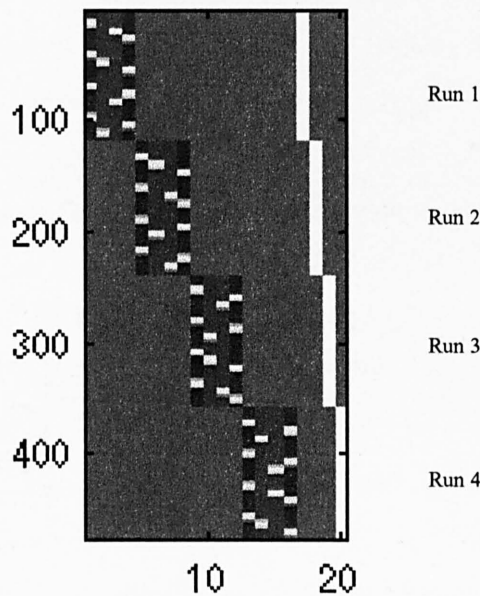
Figure 4.6 Design matrix. The design matrix used in the present analysis is shown in the conventional SPM display. Each block represents one sequence (run). Each column of the block represents one of the experimental conditions in the following order (stationary, motion, stereo and stereomotion). The base line condition (fixation) is represented as the complement of the sum of all the experimental conditions. The white squares shows the periods in which the related condition was presented. The four columns at the end of the design matrix represent each of the independent sequences. Rows represents scans (476 scans = 7 scans per condition * 17 conditions per sequence * 4 sequences).

Voxel values were normalised to the Global Mean Intensity (GMI) using global effects (scaled), which scales each image within a session to 100/(mean of GMIs of the session). The data was temporally smoothed using an *hrf* as a low-pass filter (WDCN 1997).

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), (WDCN 1997). Group analysis (random effects) was not developed due to the small number of subjects scanned (Friston, Holmes et al. 1999).

4.4 Results.

The statistical difference of the neural activity under individual experimental conditions was tested using the SPM contrast approach. The *contrast* of two experimental conditions is constructed by the subtraction of the individual regression parameters obtained by the linear regression (GLM), in other words, the contrast Motion against Stationary (Motion-Stationary) tests the statistical validity of the hypothesis that the activation under Motion condition is not bigger than the activation under Stationary condition (WDCN 1997). It is important to notice that the contrasts Motion against Stationary, and Stationary against Motion, have different meanings.

As part of the statistical analysis made by SPM software, a window of results is presented. This window contains four main sections: The first one shows the activation in a 3 view (axial, coronal, sagital) glass brain. A cursor (<) points to the voxel of the location as shown in the table in figure 4.7. The second section of the window shows the contrast design matrix used in the analysis. The bars above each column of the design matrix refer to the conditions used in the analysis. The bars above the show line means that the activity in this condition is bigger than the condition in which the bar appears below the mean line.

The third section of the window shows the statistical parameters of each activation, for example, location in mm., cluster size, statistical significance (*p value*), Z scores, etc. Finally, the window shows a summary of the statistical criteria used in the analysis, e.g., degrees of freedom, voxel size, height threshold, etc. A typical SPM analysis output is shown in the figure 4.7.

In the present study, six analyses were performed in order to identify the areas involved in each particular stimulus. For reasons of space, individual analysis outputs are not presented, instead the results are reported as a synopsis through the use of condensed statistical tables which shows a summary of the relevant statistical parameters of the regions of activation. In the same manner, the statistical map of the most sensitive subject is presented. It is important to clarify that the tables shown in those summaries refer to right and left hemispheres of the subject, for example, the left and right hemispheres of subject one are refered to as 1R and 1L respectively. Subjects or hemispheres not included in a particular table means that not significant activation was found.

Due to technical restrictions, it was not possible to develop retinotopic mapping to identify the visual areas activated in our studies, instead, the regions of activation identified through the statistical analysis were mapped to anatomical locations using as a reference their Talairach coordinates. The labels assigned to each region were given, matching the anatomical location reported by other authors. So although the labelling was made as precise as possible using Talairach space, it is important to acknowledge that it was not possible to distinguish V3, V3A andV3B reliably and the labels assigned are best guesses.

Finally, it is important to remark that the analysis of effective connectivity was only developed in this study and not in the subsequent experiments for the reasons explained in the Chapter # 3 (section 3.6).
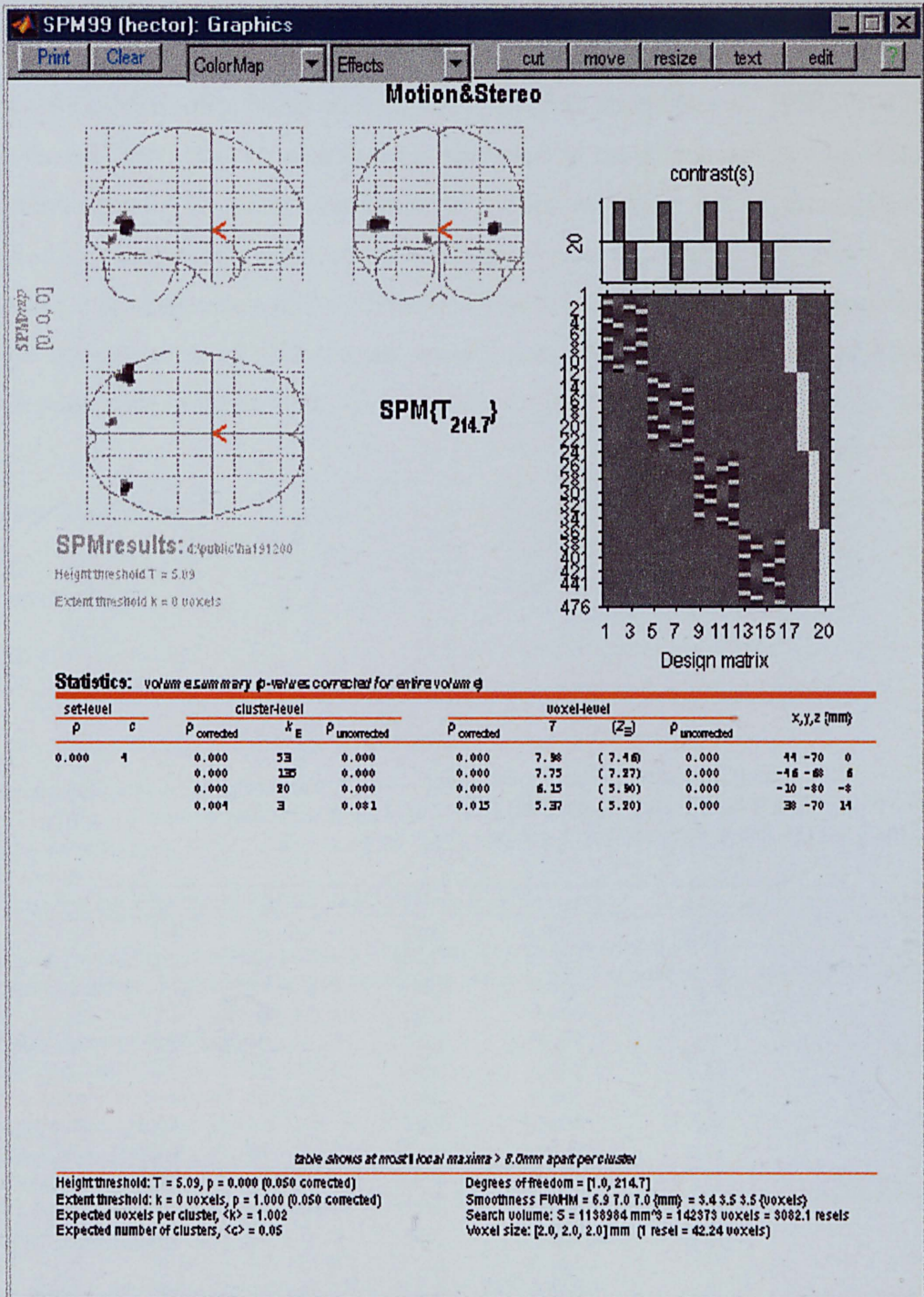
Figure 4.7 SPM 99 output display.

4.4.1 Motion against Stationary.

This contrast compares the activation caused by the moving dots against that caused by stationary dots. In accordance with previous research (Watson, Myers et al. 1993; McKeefry, Watson et al. 1997; Buchel, Josephs et al. 1998; Friston and Buchel 2000), the V5 area showed activation to the directional motion. Also the V3A showed consistent activation to motion stimulus. All of the subjects showed consistent activation on motion sensitive areas. Figure 4.8 shows the statistical parametric map of the V5 and V3A activation, it also shows the location of the activation in V5. To see the specific location of the activation in V3A consult the table in figure 4.12.

V5
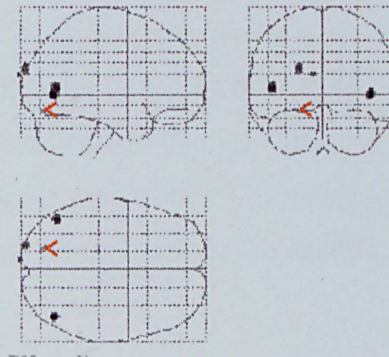
Subject	Location	Z-Score	P corrected	Cluster size
1 R	56, -64, -6	(Inf)	0.000	151
1 L	-48, -74, -4	(Inf)	0.000	239
2 R	44, -82, 6	(Inf)	0.000	353
2 L	-48 -82, 4	(Inf)	0.000	396
3 R	48, -66, -6	7.58	0.000	99
3 L	-40, -84, -6	(Inf)	0.000	122
4 R	40, -76, -2	(Inf)	0.000	223
4 L	-38, -72, 2	7.60	0.000	128
5 R	48, -78, 0	(Inf)	0.000	824
5 L	-42,-84,6	(Inf)	0.000	215
6 R	42, -64, 4	(Inf)	0.000	474
6 L	-46, -74, 8	(Inf)	0.000	333
7a R	46, -70, 0	(Inf)	0.000	610
7a L	-46, -68, 8	(Inf)	0.000	438
7b R	44, -70, 0	(Inf)	0.000	85
7b L	-46, -68, 6	(Inf)	0.000	186

Figure 4.8 Motion against Stationary contrast. The statistical map shows the areas sensitive to motion, the activation include V5 and V3A regions. The table shows the locations of V5 region. To see the locations of V3A look at the figure 4.12 in contrast 4.4.5.

4.4.2 Stereo against Stationary.

The activation found by the contrast of the stereo cone against the stationary dots was not as consistent as in the previous contrast. This contrast reveals two main areas, one only sensitive to stereo and one sensitive to stereo and motion, i.e. there was an overlap with the region sensitive to motion (V3A). Five of the subjects showed consistent activation in V3A. Four of the subjects showed consistent activation in a small area only sensitive to stereo stimulus. In accordance with the theory of the right hemisphere dominance in stereo information processing (Nishida, Hayashi et al. 2001), the four subjects whose showed activation in this area had activation in the right hemisphere. The anatomical location of the region which was sensitive only to stereo stimulus refers to the visual area V3B. The literature reports this area as sensitive to stereo disparities (Backus, Fleet et al. 2001), kinetic boundaries (Orban, Dupont et al. 1995; Van Oostende, Sunaert et al. 1997) and to second order motion (Smith, Greenlee et al. 1998a). Figure 4.9 shows the statistical parametric map of the V3B and V3A activation, it also shows the location of the activation in V3B. To see the specific location of the activation in V3A consult the table in figure 4.12.

V3B
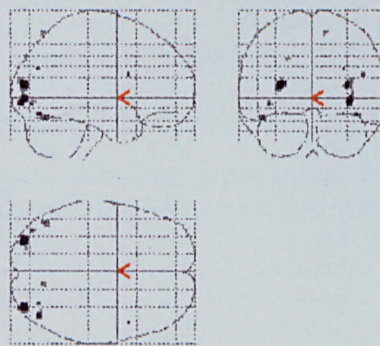
Subject	Location	Z-Score	P corrected	Cluster size
4 R	44, -88, 10	6.09	0.000	245
4 L	36, -96, 0	4.80	0.098	76
6 R	40, -88, 8	5.35	0.006	7
7a R	36, -88, 0	(Inf)	0.000	294
7a L	-36, -88, 2	(Inf)	0.000	243
7b R	34, -88, -2	6.34	0.000	71

Figure 4.9 Stereo against Stationary contrast. The statistical map shows the areas sensitive to stereo, the activation include V3B and V3A regions. The table shows the locations of V3B region. To see the locations of V3A look at the figure 4.12 in contrast 4.4.5.

4.4.3 Motion against Stereo.

The assessing of the existence of one individual region sensitive only to motion was tested using the contrast of moving dots versus the stereo cone. This contrast can be thought of as the part of motion sensitive areas that are not sensitive to stereo stimuli.
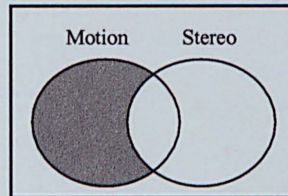
This contrast suggests, in all of the subjects, that the V5 regions is sensitive to motion (fig. 4.10).

V5
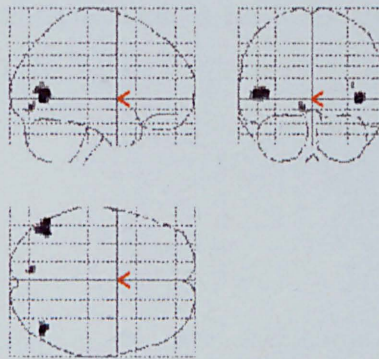
Subject	Location	Z-Score	P corrected	Cluster size
1 R	54, -68, -8	6.12	0.000	24
1 L	-48, -74, -6	7.40	0.000	92
2 R	-48, -76, 10	7.01	0.000	223
2 L	46, -84, 0	7.22	0.000	204
3 R	46, -68, -6	6.49	0.000	66
3 L	-40, -84, -6	7.20	0.000	63
4 R	40,-76, -2	6.36	0.000	105
4 L	-40, -72, 0	7.14	0.000	66
5 R	48, -78, 2	(Inf)	0.000	117
5 L	-42, -84, 6	6.77	0.000	76
6 R	42, -64, 4	(Inf)	0.000	140
6 L	-46, -74, 8	6.52	0.000	125
7a R	44, -70, 0	7.46	0.000	56
7a L	-46, -68, 6	7.27	0.000	135
7b R	42, -68, 2	5.7	0.001	8
7b L	-46, -68, 8	7.10	0.000	86

Figure 4.10 Motion against Stereo contrast. The statistical map shows the region V5 sensitive only to motion.

4.4.4 Stereo against Motion.

The assessing of the existence of one individual region, sensitive only to stereo was tested using the contrast of stereo cone versus moving dots. This contrast can be thought of as the part of stereo sensitive areas that are not sensitive to motion stimuli.
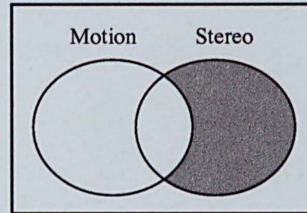
The results of this contrast was not as consistent as the previous one. Only two subjects showed consistent activation in the V3B area, one of them only in the right hemisphere (fig. 4.11).

V3B
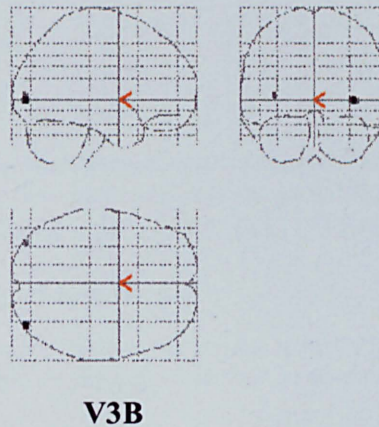
Subject	Location	Z-Score	P corrected	Cluster size
6 R	46, -84, 4	5.52	0.003	3
7a R	38, -88, 0	5.03	0.033	1
7a L	-36, -88, 2	5.58	0.002	6

Figure 4.11 Stereo against Motion contrast. The statistical map shows the region V3B sensitive only to stereo.

4.4.5 Conjunction analysis of Motion and Stereo.

Previous contrasts suggest the existence of two independent regions, one sensitive to motion and one sensitive to stereo. A combined contrast to identify the regions sensitive to both stimuli is given by the intersection between these two groups of neurones.
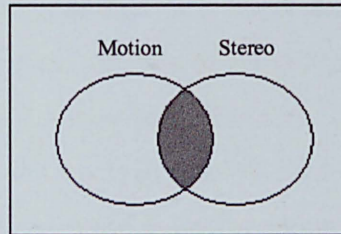
The statistical map of the regions (V3A) sensitive to stereo and to motion is shown in figure 4.12.

V3A
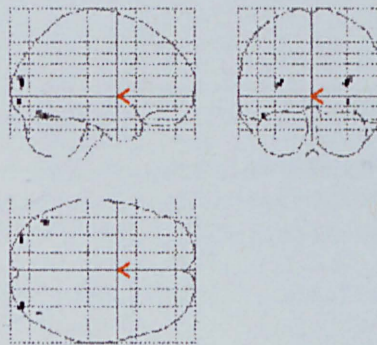
Subject	Location	Z-Score	P corrected	Cluster size
4 R	30, -88, 22	(Inf)	0.000	13
4 L	-28, -88, 24	5.13	0.023	2
5 R	14, -94, 30	7.53	0.000	114
5 L	-10, -98, 28	6.97	0.000	10
6 R	28, -92, 22	6.91	0.000	45
6 L	-24, -86, 32	5.02	0.031	3
7a R	32, -92, 16	6.66	0.000	66
7a L	-24, -96, 22	6.37	0.000	27
7b R	32, -90, 14	6.19	0.000	33
7b L	-24, -96, 24	7.73	0.000	26

Figure 4.12 Conjunction analysis of Motion and Stereo. The statistical map shows the region V3A sensitive to motion and to stereo.

4.4.6 StereoMotion against Motion and Stereo.

The present contrast was designed to test if the region sensitive to stereo and motion become more active when motion and stereo stimulus are presented at the same time rather than when the individual stimuli are presented in isolation. In other words, this contrast looks for a non-linear relationship in the joint stimulation which may suggest that this joint stimulation produces a bigger response to that produced by the average of the individual stimulations. Given that the location of the stereo-motion sensitive area (V3A) was identified in the previous contrast, the present analysis was restricted to those voxels positioned under a sphere of radius 8mm around the locations reported in figure 4.12. This analysis is also called Small Volume Correction (WDCN 1997). The results of the SVC analysis suggest that V3A is more responsive when stereo and motion are presented at the same time (fig. 4.13).

V3A

Subject	Location	Z-Score	P corrected
4 R	30, -88, 22	6.16	0.000
4 L	-28, -88, 24	4.99	0.000
5 R	14, -94, 30	5.79	0.000
5 L	-10, -98, 28	5.02	0.000
6 R	28, -92, 22	3.18	0.071
6 L	-24, -86, 32	Not	significant
7a R	32, -92, 16	3.55	0.028
7a L	-24, -96, 22	4.46	0.002
7b R	32, -90, 14	6.04	0.000
7b L	-24, -96, 24	6.08	0.000

Figure 4.13 StereoMotion against Stereo + Motion.

The results found in the statistical analysis of the activations observed in the present experiment can be summarised as: motion stimuli activates V5 region in 8 out of 8 of the scanned subjects. Stereo condition activates V3B region in 4. However, in the contrast Stereo against Motion, only two subjects showed significant activation in the V3B region. Furthermore, in five subjects, V3A was sensitive to both stereo and motion stimuli, showing a bigger response when the stimuli were presented together.

Using the anatomical data of the most sensitive subject, a summary of the main activation is shown in figure 4.14
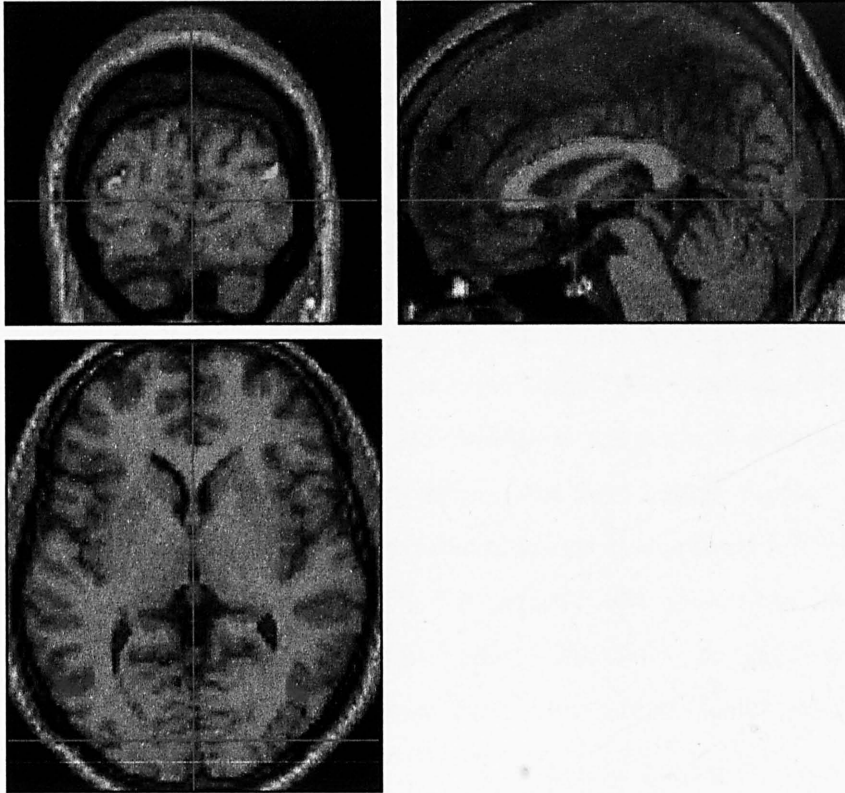
Figure 4.14. Axial, sagittal and coronal views of activations at point [0 –90 2]$_{MNI}$. The anatomical images are used to show the location of the regions sensitive to: Motion V5 (blue), Stereo V3B (red) and Stereo and Motion V3A (yellow).

4.5 Effective Connectivity Analysis.

Having identified the cortical regions sensitive to each stimulus, the next question to answer is how to explain the relationships between regions?. The answer to this question can be formulated as an effective connectivity analysis. The two main parts of an effective connectivity analysis are the anatomical model (*which areas are connected*) and the mathematical model (*how these areas are connected*). Based on the observed data, the following section proposes and assesses a hypothesis of effective connectivity.

4.5.1 Anatomical model.

From studies in monkey it is known that the visual system can be functionally divided into two main streams. The ventral stream which projects over the occipito-temporal cortex and the dorsal stream which projects over the occipito-parietal cortex. The ventral stream ('*what*') is generally associated with object information processing and the dorsal stream ('*where*') is associated with spatial and three dimensional information processing (Mishkin, Ungerleider et al. 1983; Creem and Proffitt 2001).

Moreover, the dorsal stream can be subdivided into two branches, both of them starting in the primary visual cortex and ending in the parietal cortex (Zeki and Shipp 1988; Felleman and Van Essen 1991). The first branch projects from the primary visual cortex to the superior parietal cortex through the V5/V5A visual areas (Podzebenko, Egan et al. 2002). The second one goes from the primary visual cortex to the intraparietal cortex through the V3/V3A visual areas. This stream is related with the perception of three dimensional structures (Faillenot, Decety et al. 1999).

Although there is no anatomical evidence of the projection from V5 to V3A, the hypothetical model of effective connectivity shown in figure 4.15 is proposed to explain the cortical interaction observed in the present experiment. It is important to note that because only eighteen slices, covering most of the visual cortex, were used to acquire data, most of the parietal cortex was missed and nothing can be deduced about its activations. Therefore, parietal regions are not included in the proposed model.
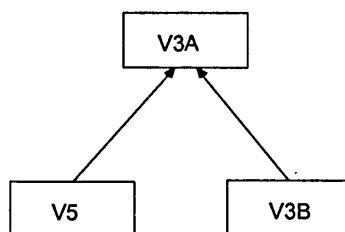
Figure 4.15 Anatomical model.

4.5.2 Mathematical model.

In order to assess this hypothesis of effective connectivity (Friston, Holmes et al. 1994), path analysis was made using the time series from the regions involved. For that reason, the data of the most sensitive subject (subject 7aR) was used for the present analysis. The first principal component time series of a sphere 8mm radius around the voxel with highest activation of each cortical region is shown in figures 4.17 and 4.18. The NARMAX algorithm was applied to identify the interaction terms (i.e. non linear relationships) among regions. Using an autoregressive linear model of dynamic order=0 and polynomial order=2, the forward regression orthogonal algorithm identified the model shown in figure 4.16. An interactive term (V5*V3B) was suggested by NARMAX to explain the response activation of the joint activity of V5 and V3B.

After the expansion of the model to include the interaction term (V5*V3B) suggested by NARMAX, the parameters of the model were identified using the Analysis of Moment Structures software package (AMOS ver. 4.0). As an indicator of correctness, the probability level achieved was near to one.
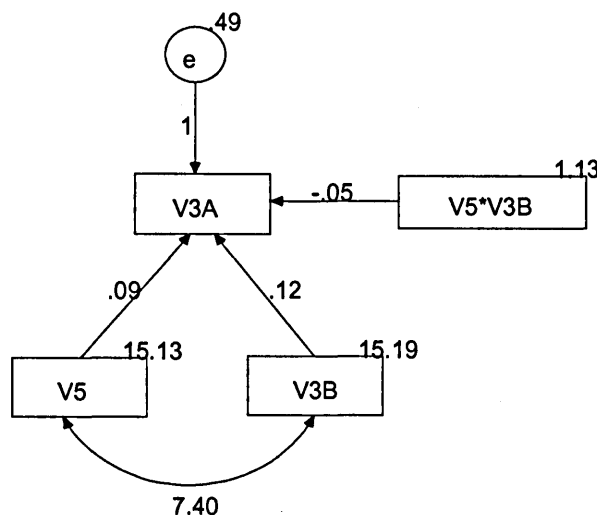
Figure 4.16 Effective connectiviy model. The model identified using the NARMAX algorithm was evaluated using SEM (AMOS 4.0). The model fitted the data (x^2=0.2, dof=2, p > 0.1)
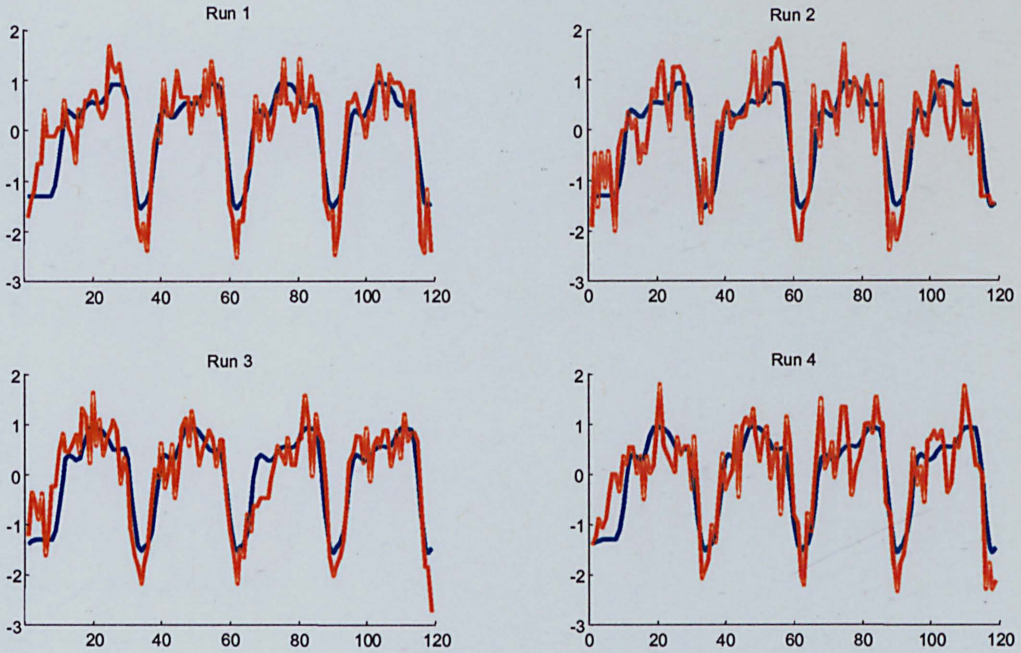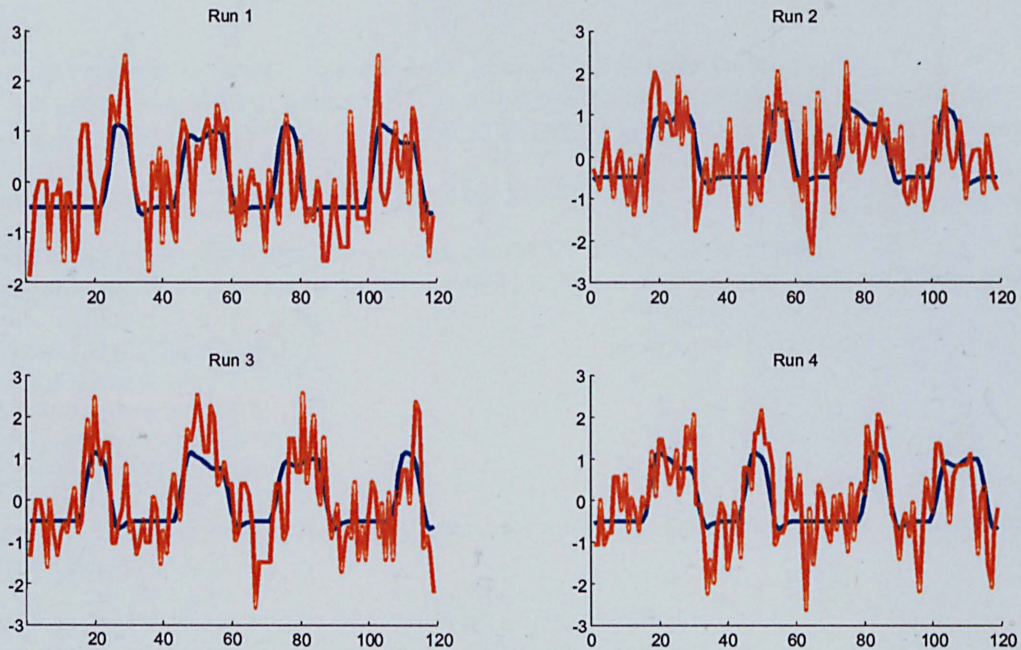
Figure 4.17 Time series. Blue time series corresponds to the hemodynamic model estimated from the experimental stimulation paradigm. The time series were normalised to mean zero and standard deviation one.

V3B

Talairach (38 -85 4)

V3A

Talairach (32 -88 19)
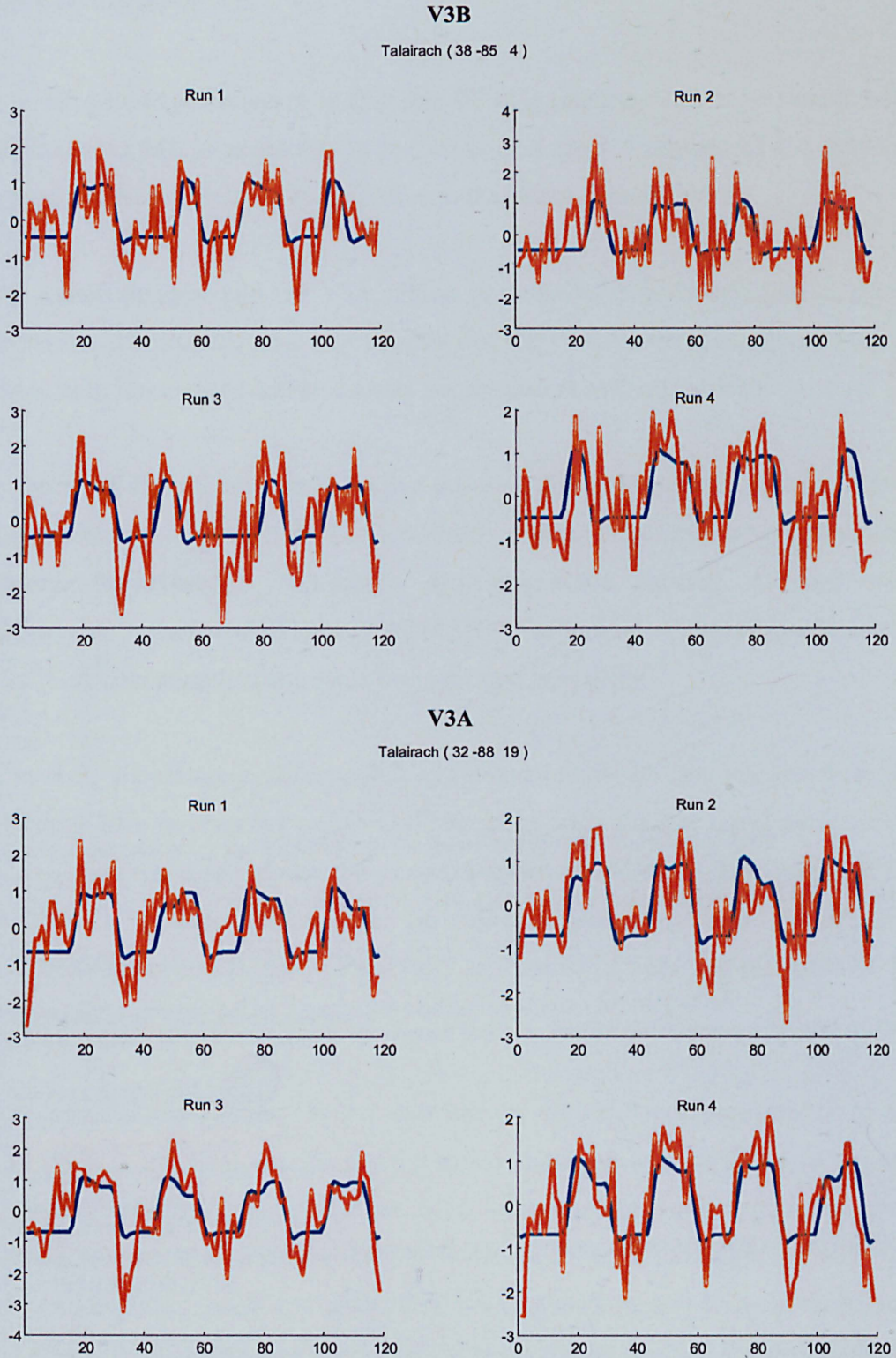
Figure 4.18 Time series. Blue time series corresponds to the hemodynamic model estimated from the experimental stimulation paradigm.

4.6 Conclusions

A conclusion of our research is that area V5 in humans appears to be functionally dissimilar to MT in monkeys, in the sense that the V5 regions of the subjects scanned in this study were not sensitive to the stereo stimulus used.

Our results suggest that the V3A region is sensitive to both motion and stereo disparities. Our experiments also suggest that the V3B region is sensitive to stereo disparities, however its activation was not consistent over all subjects.

A model of effective connectivity that accounts for the functional architecture of stereo vision was presented. The non-linear relationships under the model were identified applying a Non-linear Auto-Regressive moving Average with eXogenous variables algorithm (NARMAX). The validity of the proposed model was tested statistically using structural equation modelling.

However, the structure of the model was derived from the data and not from the anatomy. This represents a problem of circularity since the functional model relies on the anatomical model and the structure of the model can not be modified for the sake of a better goodness of fit. Nevertheless, this analysis of effective connectivity was useful to exemplify the application of SEM and NARMAX to a real data set of fMRI time series.

One possible explanation for the inter subject variability or unreliability of the activation in the V3B area during the stereo condition was attributed to the fact that all the subjects reported that the 'intensity' of the stereo precept declined over the 21s. presentation time, however no evidence of it was found in the data. We have no explanation for this 'adaptation' but in the study that follows we attempt to prevent its occurrence by periodically alternating the depth of the stimulus from convergent to divergent disparities.

Chapter 5

Experiment 2

SereoMotion in Depth

Abstract.

Functional magnetic resonance imaging was used to identify the stereo sensitive regions of the brain. Red/green random dot anaglyph stereograms were used to create four squares lying in two different planes. The squares were periodically alternating in depth from convergent to divergent disparities. Although subjects reported seeing the squares moving in depth, the inter-subject variability of the activations suggests that the stereo stimulus was dominated by the control conditions. However, consistent with the results of the previous experiment, evidence of sensitivity of the areas V3B and V3A to stereo stimulus was found.

5.1 Introduction.

The results of the previous experiment suggested the existence of a cortical region sensitive to stereo disparities. According to its anatomical location, this cortical region lays more posterior to V5 region and corresponds to the location of V3B. The V3B region has been mainly associated with the processing of kinetic boundaries and second order motion (Orban, Dupont et al. 1995; Dupont, De Bruyn et al. 1997; Van Oostende, Sunaert et al. 1997; Smith, Greenlee et al. 1998a; Seghier, Dojat et al. 2000).

From the previous experiment the hypothesis is that V3B and V3A are both sensitive to stereo disparities. The main goal of the present study is to re-examine the cortical regions sensitive to stereo disparities, see figure 5.1.

V3B

R [38, -88, 0]
L [-36, -88, 2]

V3A

R [32, -92, 16]
L [-24, -96, 22]
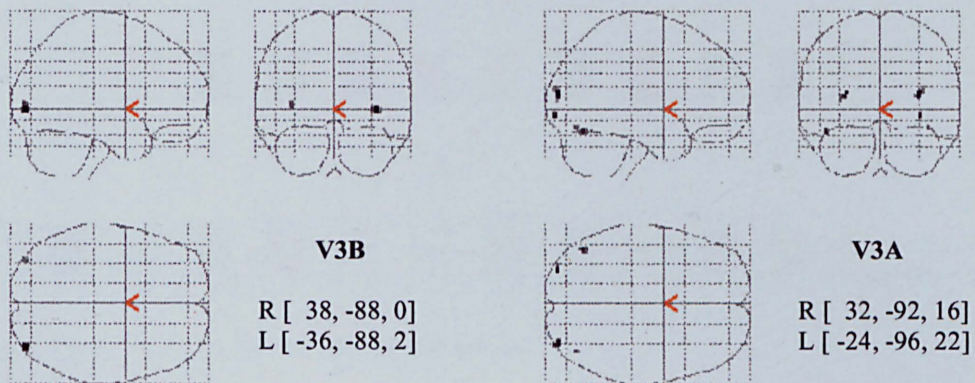
Figure 5.1 Stereo sensitive regions.

Although the activations were not consistent over all subjects, a possible explanation of the inter-subject variability is that all the subjects reported that the 'intensity' of the stereo perception declined over the 21s presentation time, however no evidence of it was found in the data. We have no explanation for this 'adaptation'. In an attempt to prevent this happens, in this experiment the disparities of the stimulus were changed over time by periodically alternating the depth of the stimulus from convergent to divergent disparities.

5.2 Experiment Design.

All subjects were given 4 sequential scans each lasting 9 min. (25 epoch) with a 5 min interscan interval to permit subjects to rest. One hundred and eighty image volumes were obtained in each run. Each condition lasted 21s., giving 7 multislice volumes per condition (TR=3 s.). A dummy condition of a blank screen was presented during the first 15s. (5 scans) of each run to eliminate magnetic saturation. To avoid any order effects the conditions were counterbalanced using a Latin Squares design (figure 5.2).

Run1

Run2

Run3

Run4

▮ StereoMotion in depth with random motion.	▮ Stationary.	
▮ Random motion.	▯ Lateral motion.	
▮ StereoMotion in depth.	▮ Stereo still.	
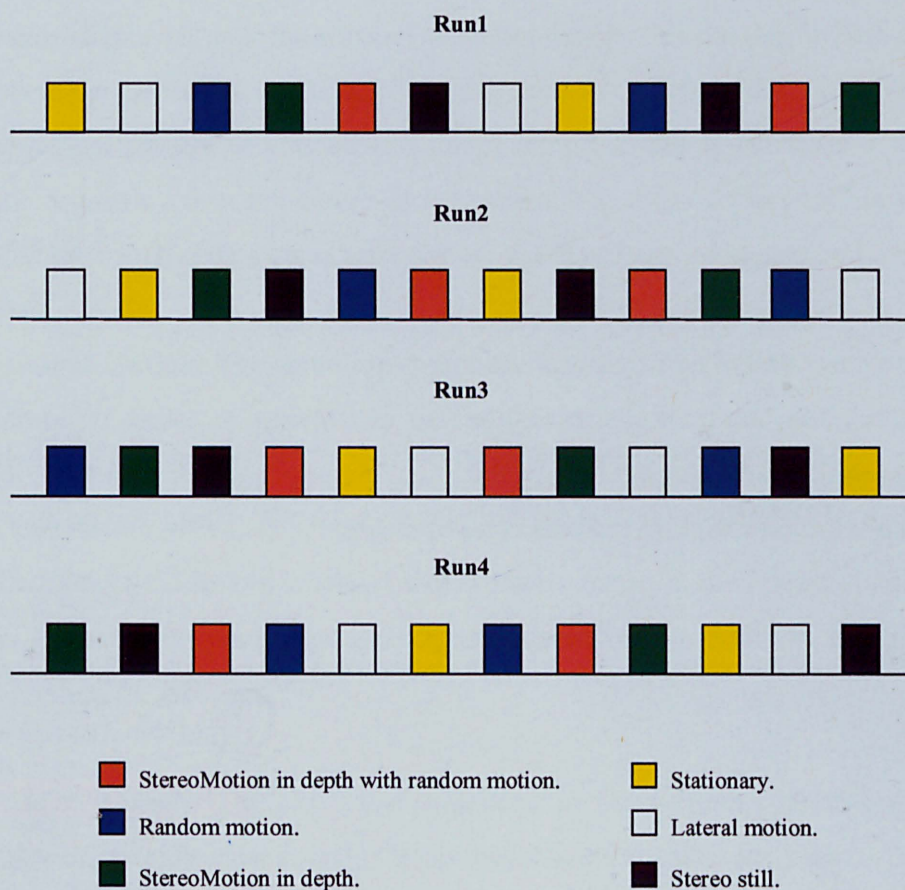
Figure 5.2 Stimulation Sequences.

The subjects were instructed to fixate a point (0.3 deg. of radio) in the middle of the screen (circular field of view 13 deg.) and foveate while presented with the visual stimuli. There were seven conditions of visual stimulation (figure 5.3).

a) **Rest**: From the experience of the previous experiment, in which some subjects reported 'tired eyes' and sometimes tears. The rest condition was designed as a relaxation period. In the first 15 seconds of this condition a blank screen was displayed and the subjects were instructed to blink or even to close their eyes for a couple of seconds. In the second stage of the condition (from second 15 to 21), a fixation point was displayed in the centre of the screen. The appearance of the fixation point was used to attract the attention of the subject and to get him/her prepared for the coming stimulus. This rest condition was taken as a base line.

b) **Stationary**: One thousand and twenty four dots (with radio 0.1 deg.) were randomly displayed over the screen (mean dot density 1.5 dot deg^{-2}). The absence of dots in the central axes (0.8 deg. wide) of the screen defined four doted textured squares. Each square was located in one quadrant of the visual field. A fixation point was displayed in the centre of the screen. The aim of this condition was to activate the visual cortex areas sensitive to the luminance produced by the dots.

c) **Random motion**: The condition stationary was modified letting each dot move randomly. In order to maintain a uniform distribution of the dot density, the random motion was implemented by specifying a reference point for each dot, then the position of the dots between frames was determined randomly under the area defined by a circle of radius 1 deg. with its centre in the reference locations. This stimulus was designed as a control condition for StereoMotion in depth with random motion condition.

d) **Lateral motion**: The four squares defined in the stationary condition moved laterally in the following manner. When squares in the quadrant 1 and 3 move to the right, squares in the quadrant 2 and 4 move to the left and vice versa. The speed of displacement was set constant to 0.53 deg. s^{-1}. This stimulus was designed as a control condition for StereoMotion in depth condition.

e) **Stereo still**: This condition was designed to stimulate binocular neurones. Four squares were displayed at two different depths (red/green anaglyph stereogram).

Squares in quadrants 1 and 3 lay in the same plane behind the fixation point (0.172 deg.). Squares in quadrants 2 and 4 lay in the same plane in front of the fixation point (-0.172 deg.).

f) **StereoMotion in depth**: The pairs of squares laying in two different planes changed their depth smoothly from -0.172 to 0.172 deg. and vice versa (0.53 deg. s^{-1}). The variation of disparities over time was introduced to avoid any habituation effects occurred in the previous experiment. In the training sessions some of the subjects reported that they perceived the squares moving laterally instead of seeing the squares moving in depth. This visual effect was attributed to the fact that some individuals are 'motion dominant'. In other words, for motion dominant subjects the change in position of the stereoscopic pair of dots (disparity) was stronger that the binocular fusion itself. Although this problem can be avoided by randomly locating the position of the dots between frames (see next condition), this condition was kept in order to investigate this phenomenon.

Because motion in depth not only implies the processing of stereo disparities but also the computation of the motion gradient, this condition is expected to activate motion sensitive regions. We hypothesised that the subtraction of Lateral motion condition (d) from StereoMotion in depth condition should reveal the stereo sensitive regions.

g) **StereoMotion in depth with random motion**: This condition is a modification of the previous one (f). The introduction of random motion tends to eliminate the effect of lateral motion, since no directional clues are involved in the stimulus. With the subtraction of random motion condition (c), from StereoMotion in depth with random motion, the stereo sensitive regions were expected to be identified.
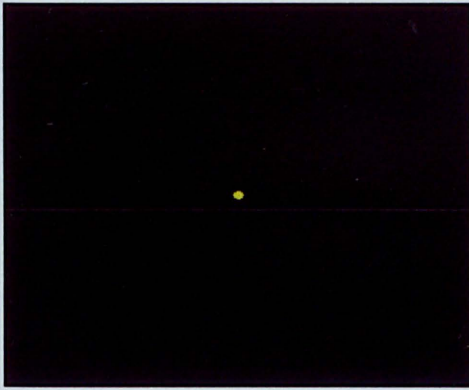
Rest
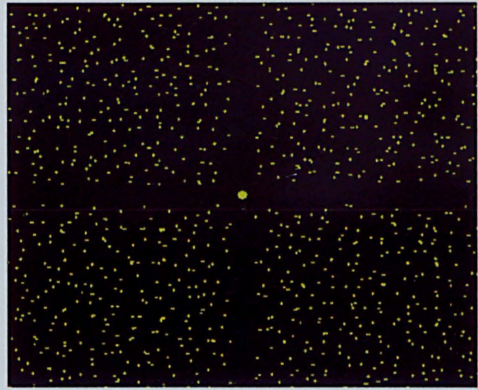
Stationary
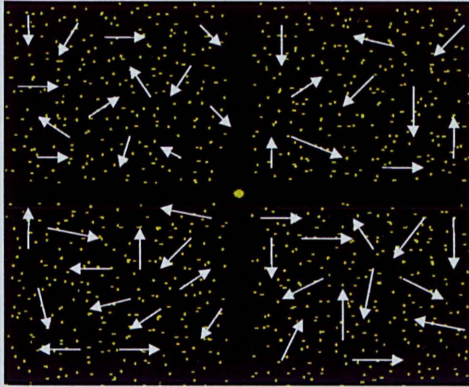
Random Motion
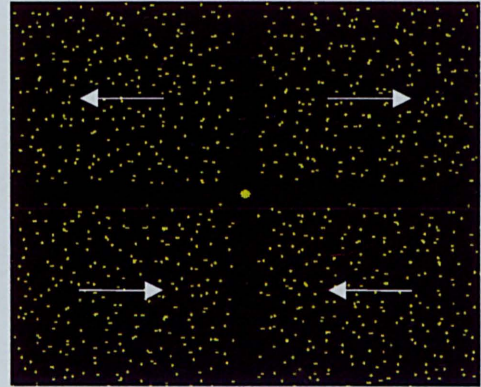
Lateral Motion
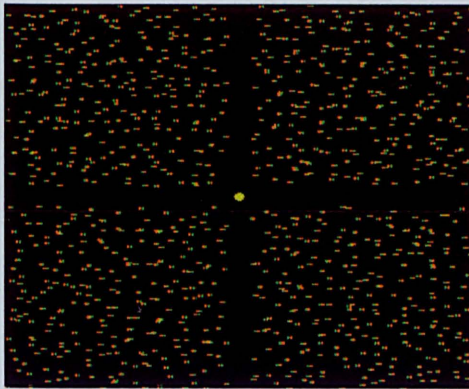
Stereo Still
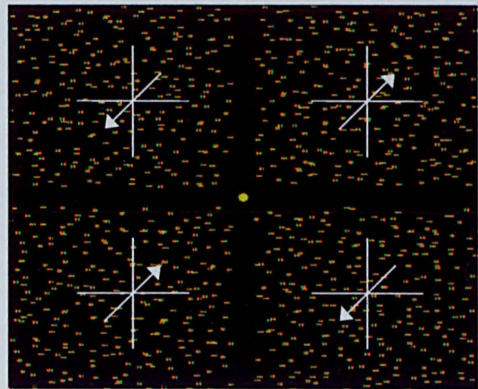
StereoMotion in Depth

Figure 5.3 Visual stimuli. Subjects viewed the stimuli through the red/green glasses to present anaglyph stereo stimuli. The white arrows in the displays represent the direction in which the dots were moving .

5.3 Methods.

5.3.1 Subjects.

Eight healthy right-handed volunteers (3 female,5 male) aged from 20 to 60 years participated in the present study. Five of the subjects were the same as those scanned in the previous experiment. The subjects had a preliminary session outside the magnet to get familiar with the visual stimulation. All subjects gave informed written consent.

5.3.2 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 just outside 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 G3 Mac running Psychophysics Tool Box ver. 2.44 (Brainard 1997; Pelli 1997) under MATLAB ver. 5.3.

The mean luminance of the image was 18 cd/m^2. Although the stimuli were displayed at a video frame rate of 60 Hz, the image was only updated on every 10th frame, producing an effective frame rate of 6 Hz.

5.3.3 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.

5.3.4 Data analysis.

The data was pre-processed and analysed using SPM99 (Wellcome Department of Cognitive Neurology). The first five scans of each run were discarded in order to exclude magnetic saturation artefacts. All volumes were slice timed with respect to the first slice. Motion correction was made taking as a reference the first volume of the first run. The data was normalised in the MNI (Montreal Neurological Institute) stereotaxic space and smoothed using a 6 mm FWHM (full width at half maximum) isotropic Gaussian kernel (WDCN 1997).

Data analysis was performed using a boxcar design matrix of the different conditions convolved with the hemodynamic response function. Specific effects were tested by applying the corresponding linear contrast to the parameters obtained applying General Linear Model using the design matrix shown in figure 5.4 (see also experiment design).
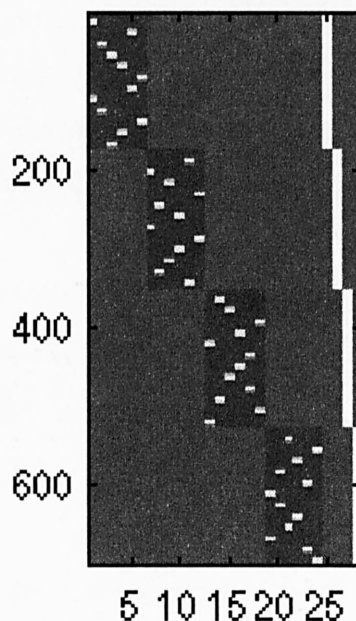
Figure 5.4 Design matrix. The design matrix used in the present analysis is shown in the conventional SPM display. Each block represents one sequence (run). Each column of the block represents one of the experimental conditions in the following order (stationary, random motion, stereo-motion in depth, stereo-motion in depth with random motion, lateral motion and stereo still). The white squares shows the periods in which the related condition was presented. The four columns at the end of the design matrix represent each of the independent sequences. Rows represents scans (700 scans = 7 scans per condition * 25 conditions per sequence * 4 sequences).

Voxel values were normalised to the Global Mean Intensity (GMI) using global effects (scaled), which scales each image within a session to 100/(mean of GMIs of the session). The data was temporally smoothed using an *hrf* low-pass filter (WDCN 1997).

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), (WDCN 1997). Group analysis (random effects) was not developed due to the small number of subjects scanned (Friston, Holmes et al. 1999).

5.4 Results.

In the present study, nine analyses were performed in order to identify the areas involved in each particular stimulus. For reasons of space, individual analysis outputs are not presented, instead the results are reported as a synopsis through the use of condensed statistical tables which shows a summary of the relevant statistical parameters of the regions of activation. In the same manner, the statistical map of the most sensitive subject is presented.

Note: Because the stereo still condition was included in the stimulation paradigm after the first scan session, the contrasts involving this condition are not available for the first three subjects.

5.4.1 Random motion against Stationary.

This contrast compares the activation produced by the dots moving randomly with that produced by the stationary dots. Consistent with other studies (McKeefry, Watson et al. 1997; Smith, Greenlee et al. 1998a), random motion activates V1 and V5 regions. The activation was more consistent in V5 (7 subjects) than in V1 (5 subjects). There was neither consistent bilateral activation nor consistent hemisphere activation (fig. 5.5).

V1
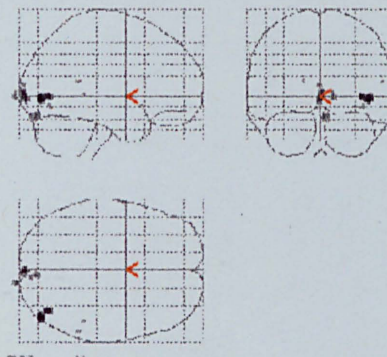
Subject	Location	Z-Score	P corrected	Cluster size
1 R	14, -98, -4	6.41	0.000	106
1 L	-14, -94, -10	6.30	0.000	359
2 R	12, -92, 6	6.53	0.000	202
3 R	2, -96, 2	6.10	0.000	34
5 L	-12, -98, 10	5.72	0.001	18
7 R	6, -82, -6	5.75	0.000	37

V5

Subject	Location	Z-Score	P corrected	Cluster size
1 R	46, -70, 2	7.24	0.000	39
1 L	-48, -68, 4	5.71	0.001	21
2 R	40, -76, 0	5.40	0.005	4
3 R	46, -80, -2	6.68	0.000	58
4 R	42, -72, -2	6.19	0.000	54
4 L	-46, -76, 2	6.02	0.000	27
5 R	44, -74, 2	7.51	0.000	93
5 L	-32, -94, -2	6.04	0.000	20
6 L	-48, -76, 0	5.11	0.022	5
8 R	44, -84, 2	(Inf)	0.000	105
8 L	-50, -76, 10	7.01	0.000	241

Figure 5.5 Random motion against Stationary contrast. The statistical map shows the areas sensitive to random motion. The activation includes V1 and V5 regions. The table shows the locations of the activations and the statistical parameters.

5.4.2 Lateral motion against Stationary.

This contrast compares the activation produced by the dots moving laterally with that produced by the stationary dots. Consistent with other studies (Watson, Myers et al. 1993; McKeefry, Watson et al. 1997; Uusitalo, Virsu et al. 1997; Sunaert, Van Hecke et al. 1999; Friston and Buchel 2000), lateral (directional) motion activates V5 region in both hemispheres. It is important to note that although random motion activated V5, lateral motion produces a stronger activation in terms of activated area (fig. 5.6). Contrary to our expectation, none of the subjects showed activation in the V3A region.

V5
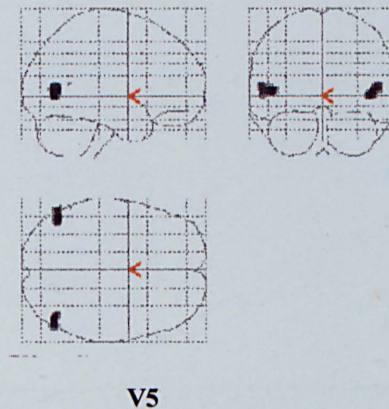
Subject	Location	Z-Score	P corrected	Cluster size
1 R	46, -70, 2	(Inf)	0.000	120
1 L	-48, -66, 6	(Inf)	0.000	138
2 R	44, -84, 10	6.08	0.000	12
4 R	50, -74, 2	(Inf)	0.000	135
4 L	-42, -74, 0	5.37	0.006	8
5 R	42, -74, 0	(Inf)	0.000	212
5 L	-36, -84,4	7.54	0.000	168
6 R	44, -72, 6	5.96	0.000	42
6 L	-48, -70, 6	5.04	0.029	2
7 L	-44, -70, -6	7.84	0.000	105
8 R	46, -84, 2	(Inf)	0.000	328
8 L	-52, -74, 4	(Inf)	0.000	363

Figure 5.6 Lateral motion against Stationary contrast. The statistical map shows the areas sensitive to directional motion. The activation includes V5 region. The table shows the locations of the activation and the statistical parameters.

5.4.3 Stereo still against Stationary.

The present contrast compares the activation during the presentation of the squares in two different planes with the activation during the squares lying in one plane (zero disparities). With this contrast, the activation of the stereo sensitive regions was expected. Only one subject showed significant activation in this contrast, showing activation in V3B and V3A regions (fig. 5.7). Although the activations showed by the subject 4 were not significant at p corrected < 0.05, they were significant at this level using a small volume correction (WDCN 1997). It is important to remember that this contrast is available only for 5 subjects.

V3B
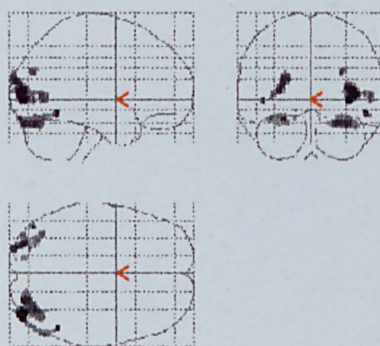
Subject	Location	Z-Score	P corrected	Cluster size
1 R	40, -88, 6	7.18	0.000	244
1 L	-28, -86, 10	6.37	0.000	122
2	Contrast	not	available	
4 R	44, -74, -2	4.53	0.225	146
4 L	-26,-92, -14	4.21	0.586	29
5	Contrast	not	available	
6	Contrast	not	available	

V3A

Subject	Location	Z-Score	P corrected	Cluster size
1 R	28, -78, 24	5.74	0.001	28
1 L	-24, -96, 20	6.28	0.000	122
2	Contrast	not	available	
4 R	30, -82, 24	4.20	0.607	29
4 L	-22, -90, 26	3.93	0.910	35
5	Contrast	not	available	
6	Contrast	not	available	

Figure 5.7 Stereo still against Stationary contrast. The statistical map shows the areas sensitive to stereo. The activation includes V3B and V3A regions. For subjects 2,5,6 stereo still condition was not available.

5.4.4 StereoMotion in depth against Lateral motion.

This contrast compares the activation during the presentation of the squares moving in depth with the activation during the squares moving laterally. Because the condition StereoMotion in depth involves directional motion (V5,V3A) and stereo disparities (V3A,V3B), with this contrast, the activations of the stereo sensitive region V3B was expected. Only one subject showed significant activation in this contrast (fig. 5.8). V3B was activated bilaterally and V3A showed activation in the left hemisphere. It is important to clarify that subject 1 was the same subject who showed a strong activation in the V3B under stereo stimulation in the previous experiment.

V3B
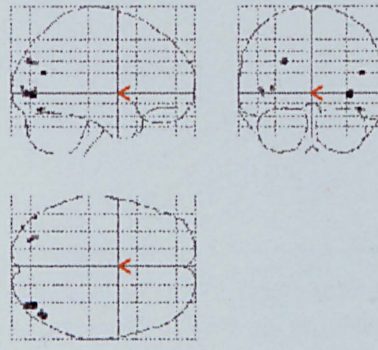
Subject	Location	Z-Score	P corrected	Cluster size
1 R	36, -80, -2	5.98	0.000	38
1 L	-38, -90, 6	5.20	0.011	4

V3A

Subject	Location	Z-Score	P corrected	Cluster size
1 L	-26, -82, 30	5.35	0.005	14

Figure 5.8 StereoMotion in depth against Lateral motion. The statistical map shows the areas sensitive to stereo. The activation includes V3B and V3A regions.

5.4.5 StereoMotion in depth against Stereo still.

This contrast compares the activation during the presentation of the squares moving in depth with the activation produced by the squares in two different planes. Because the condition StereoMotion in depth involves directional motion and stereo disparities (V5, V3A V3B), with this contrast, the activation of the V5 motion sensitive region was expected. All the subjects showed significant activation (fig. 5.9).

V5
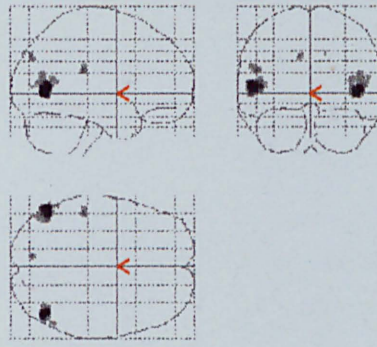
Subject	Location	Z-Score	P corrected	Cluster size
1 R	44, -68, 2	(Inf)	0.000	222
1 L	-54, -66, 6	(Inf)	0.000	259
2	Contrast	not	available	
3 R	40, -70, -4	6.11	0.000	14
4 R	48, -66, 6	6.14	0.000	41
4 L	-44, -72, 0	6.94	0.000	32
5	Contrast	not	available	
6	Contrast	not	available	
7 L	-44, -70, 6	5.89	0.000	19
8 R	40, -80, 4	(Inf)	0.000	237
8 L	-50, -76, 6	6.95	0.000	127

Figure 5.9 StereoMotion in depth against Stereo still contrast. The statistical map shows the areas sensitive to directional motion. The activation includes V5.

5.4.6 StereoMotion in depth with random motion against Random motion.

This contrast compares the activation during the presentation of the squares moving in depth with dots moving randomly against the activation during the dots moving randomly. Because the condition StereoMotion in depth with random motion involves random motion (V1 and V5), directional motion (V5 and V3A) and stereo disparities (V3A and V3B), with this contrast, the activation of the stereo sensitive region V3B was expected. Only one subject showed significant activation in this contrast. V3B showed significant activation in right hemisphere (fig. 5.10).

V3B
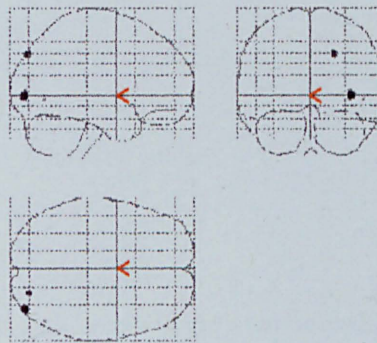
Subject	Location	Z-Score	P corrected	Cluster size
1 R	38, -86, 0	6.56	0.000	33

Figure 5.10 StereoMotion in depth with random motion against Random motion contrast. The statistical map shows the areas sensitive to stereo stimulus. The activation include V3B region.

5.4.7 StereoMotion in depth with random motion against Lateral motion.

This contrast compares the activation during the presentation of the squares moving in depth with dots moving randomly against the activation during the dots moving laterally. Because the condition StereoMotion in depth with random motion involves random motion (V1 and V5), directional motion (V5 and V3A) and stereo disparities (V3B and V3A), with this contrast, the activations from areas sensitive to random motion (V1) and to stereo disparities (V3B) were expected. Four subjects showed significant activation in V1. However none of them showed activation in V3B (fig. 5.11).

V1
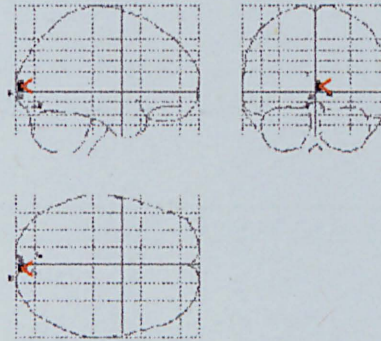
Subject	Location	Z-Score	P corrected	Cluster size
1 R	14,-96, -4	7.08	0.000	178
1 L	-12,-104, -6	5.79	0.001	25
2 R	4, -90, -6	6.74	0.000	214
3 R	4, -96, 6	6.21	0.000	25
6 R	2, -88, -6	5.58	0.002	36

Figure 5.11 StereoMotion in depth with random motion against Lateral motion contrast. The statistical map shows the areas sensitive to random motion stimulus. The activation includes V1 region.

5.4.8 StereoMotion in depth with random motion against Stereo still.

This contrast compares the activation during the presentation of the squares moving in depth with dots moving randomly against the activation during squares lying in two different planes. With this contrast, the activations from areas sensitive to random motion (V1) and to directional motion (V5 and V3A) were expected. As expected V1, V5 were activated by the motion stimuli (fig. 5.12).

V1

Subject	Location	Z-Score	P corrected	Cluster size
1 R	12, -98, -6	5.70	0.000	46
1 L	-12,-102, -4	5.33	0.006	10
2	Contrast	not	available	
3 R	4,-96, 4	5.62	0.002	47
3 L	-4, -96, -4	6.62	0.002	40
5	Contrast	Not	available	
6	Contrast	not	available	

V5

Subject	Location	Z-Score	P corrected	Cluster size
1 R	44, -68, 2	(Inf)	0.000	105
1 L	-52, -66, 6	(Inf)	0.000	110
2	Contrast	not	available	
3 R	44, -80, -2	7.26	0.000	94
4 R	50, -68, 6	6.95	0.000	107
4 L	-46, -74, 0	6.42	0.0	49
5	Contrast	not	available	
6	Contrast	not	available	
8 R	44, -82, 6	7.56	0.000	36
8 L	-50, -76, 6	6.05	0.000	58

V3A

Subject	Location	Z-Score	P corrected	Cluster size
2	Contrast	not	available	
3 L	-20,-100,16	5.98	0.000	46
5	Contrast	not	available	
6	Contrast	not	available	
8 R	-18, -98, 22	5.27	0.011	5

Figure 5.12 StereoMotion in depth with random motion against stereo still contrast. The activation includes V1, V5 and V3A regions.

5.4.9 Rest against Action.

Although this contrast was not originally planned in this study, its unexpected results are worth reporting as a significant finding. This contrast compares the neural activity during the rest period with that under visual activity. For this purpose a dummy condition composed of a sum of all the conditions (stationary, random motion, lateral motion, etc.) was created. This dummy condition was called *Action*.

Agreeing with previous reports (Corbetta, Akbudak et al. 1998; Coghill, Sang et al. 1999; Taylor, Schmitz et al. 2000), a consistent activation between all the subjects was found in occipital, temporal and parietal areas. This finding suggests that some parts of the brain become active when the subjects are inactive. No general accepted theory exists to explain this phenomenon. However, it is believed that during rest periods the brain moves into an attentional mode, i.e. it is checking for new stimuli in order to be ready to react to any event.

For instance, Raichle (Raichle, MacLeod et al. 2001) defined this mental state as *the default mode of brain function* in which "the posterior cingulate and medial parietal cortices may well be the 'sentinels', which, when beam of light move over them, cry 'who goes there' and call the fovea to the spot".

Figure 5.13 shows the regions of activation for the Rest against Action (red) and for the Action against Rest (blue) contrasts.
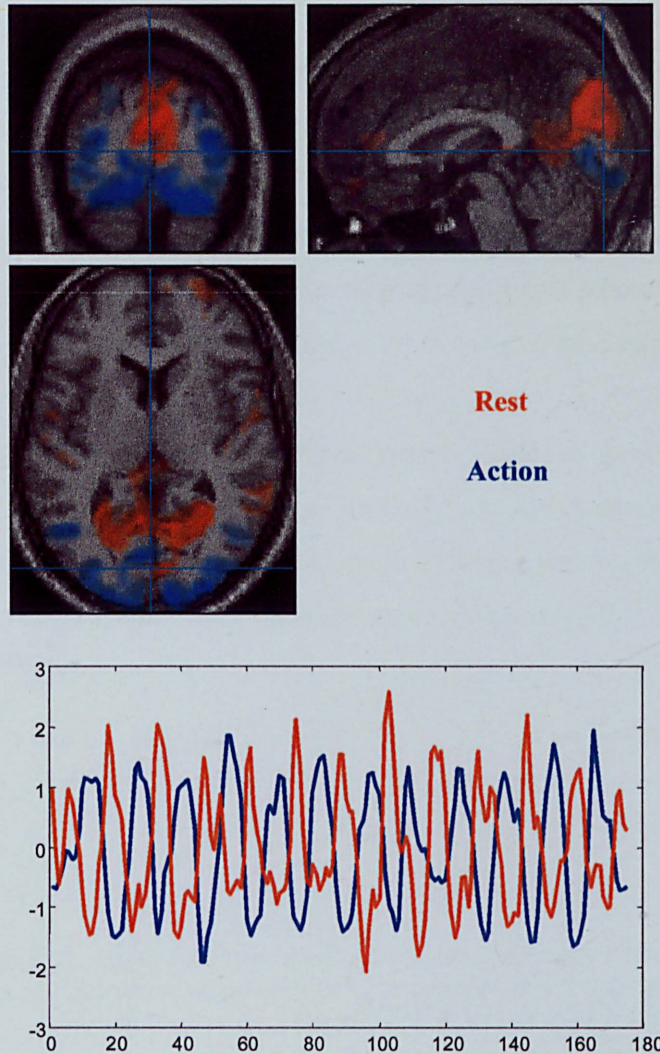
Figure 5.13 Rest against Action. At the top, axial, coronal and sagital views of the structural data are used to show the location of the activations. The areas of activation during rest and action conditions are shown in red and blue respectively. At the bottom, the characteristic mean time series of each group of activations are displayed. As it can be appreciated in the picture, the activation of the regions is anti-correlated in time.

5.5 Conclusions

In the present study, three main experimental conditions were designed to stimulate binocular neurones. Stereo still, Stereomotion in depth and Stereomotion in depth with random motion conditions were contrasted with their corresponding control conditions, revealing the following results.

Stereo still against Stationary contrast revealed bilateral activation in the expected visual areas (V3B and V3A) in two of the five subjects. Stereomotion in depth against Lateral motion activated V3B bilaterally and V3A in the left hemisphere in one of the subjects. Stereomotion in depth with random motion against the random motion contrast showed activation in V3B in the right hemisphere in one subject. It is important to clarify that the subject mentioned above was the same who showed the highest activation in these areas in the previous experiment.

These activations suggest that the stereo region V3B is sensitive to stereo disparities, and it is functionally and spatially independent of V5 (motion sensitive) area (fig. 5.14). However not enough evidence among the subjects was found to confirm this hypothesis of the stereo sensitivity of V3B.
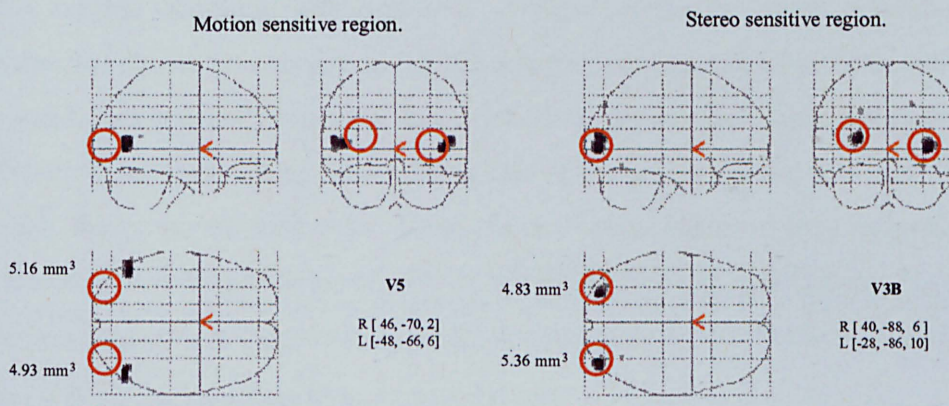
Figure 5.14 Functional defined regions. Regions of activation for motion and stereo stimuli were identified using the data of the most sensitive subject. Left panel shows the location of the motion sensitive visual area (V5). Right panel shows the location of the stereo sensitive visual area (V3B). Red circles helps to show the different location of each region. The distance between centres of each cluster are 28 mm. and 18 mm. for the left and right activation respectively.

Although the subjects reported to perceiving the squares moving in depth, the inter-subject variability of these results suggest that the stereo stimulus was very weak and the control conditions dominate the activations. The question which arises is how to increase the strength of stereo stimuli?.

It has been demonstrated that attentional requirements can selectively modulate the response of a given stimulus (Corbetta, Miezin et al. 1991; Buchel, Josephs et al. 1998; Friston and Buchel 2000). Therefore, one alternative to increase the strength of the response to the stereo stimulus is to introduce an attentional requirement, i.e. design an specific task in which the subject has to do something more that just look, for example, detecting a feature in the display.

In the study presented in the next chapter an attentional requirement was included in the paradigm through the use of an square (lying in front of the background) moving from left to right. The subjects were instructed to perform pursuit eye movement to follow the path of the square with their eyes.

Finally, an unexpected finding of this study was the identification of a cortical region on the occipital, temporal and parietal areas, in which activation is correlated with the rest condition. Although this phenomenon has been reported previously, no generally accepted theory exists to explain its causes. One possible explanation for this finding in our particular study is that during rest periods the subjects freely moved their eyes. Then, the activation observed may be produced by the network of functional areas which control eye movements (Darby, Nobre et al. 1996; Corbetta, Akbudak et al. 1998; Petit and Haxby 1999; Nobre, Gitelman et al. 2000).

Unfortunately, this hypothesis could not be tested because the areas involved in the eye movement network (frontal eye field, supplementary eye field, intraparietal sulcus and parietal eye field) were not included in the data we acquired in our experiment. Because the nature of this finding is out of the scope of this thesis, the exploration of this activation is proposed as a future project.

Chapter 6

Experiment 3

Global Stereo Tracking

Abstract.

This chapter explores the sensitivity of binocular neurones to a global stereo tracking stimuli. Red/green random dot anaglyph stereograms were used to define a square (lying in front of the background) moving from left to right and vice versa. A similar display (square defined by luminance) was used as a control condition. The subjects were instructed to perform pursuit eye movement to follow the path of the square with their eyes. Consistent activations in V3A, V3B and parietal regions were revealed by the stereo stimuli. The activation of V3B was not generally consistent with the profile reported by other authors. However, we attribute its activation to the stereoscopic motion component induced by the global stereo tracking task.

6.1 Introduction.

Results from our previous experiments suggest the existence of two cortical regions sensitive to stereo disparities (V3A and V3B). In agreement with other studies, V3A was sensitive not only to stereo stimuli, but also to motion stimuli (Tootell, Mendola et al. 1997; Hasnain, Fox et al. 1998; Mendola, Dale et al. 1999). Contrary to the consistency of V3A region with respect to its functional profile as reported in the literature, only one study (Backus, Fleet et al. 2001) reports V3B as a region sensitive to stereo disparities, whereas some studies associate this region with the processing of kinetic boundaries (boundaries between adjacent areas in which motion is in the opposite direction) and also to second order motion (motion than cannot be identified by tracking the position of spatial structures) (Orban, Dupont et al. 1995; Dupont, De Bruyn et al. 1997; Van Oostende, Sunaert et al. 1997; Smith, Greenlee et al. 1998a; Seghier, Dojat et al. 2000). The conclusions made by this study supports the first theory, since neither kinetic boundaries nor second order motion were used in our stimuli experiments.

On the other hand, the inter-subject variability of our experiments suggests that although the stereo activation was anatomically consistent in two subjects, it was too weak to be identified consistently over all subjects. One possible way to increase the response is to introduce an attentional requirement in the experimental conditions. It has been demonstrated that attentional requirements can selectively modulate the response to a given stimulus (Corbetta, Miezin et al. 1991; Buchel, Josephs et al. 1998). For example, using radial motion, Friston *et al* (Friston and Buchel 2000) showed how the activity in V5 was increased when the subjects were instructed to detect changes in the speed of the motion rather than when they were watching the stimulus passively.

Our ideal in the design of the present experiment was to increase the response of the stereo sensitive regions by requiring the subjects to perform a task of Global Stereo Tracking (GST). The basic idea behind this stimulus was to define a square (lying in front of the background) moving from left to right and vice versa.

The subjects were instructed to perform pursuit eye movement to follow the path of the square with their eyes. The advantage of this paradigm is that, it is an active rather than a passive stimulus in which the success of the task depends on the continuous perception of the square, i.e. to follow the path of the square, it is necessary to involve the stereoscopic processing system.

6.2 Experiment Design.

Subjects were given 4 sequential scans each lasting 5.15 min. (10 epoch) with a 5 min. interscan interval to permit subjects to rest. One hundred image volumes were obtained in each run. Each condition lasted 30s., giving 10 multislice volumes per condition (TR=3s.). A dummy condition of a blank screen was presented during the first 15s. (5 scans) of each run to control for magnetic saturation effects. The stimulation sequence is shown in figure 6.1.
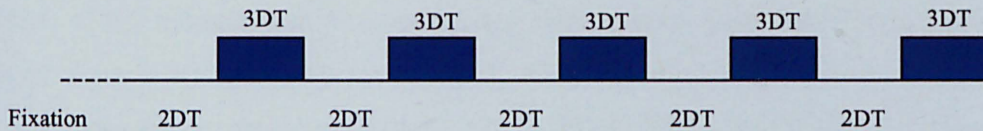
Figure 6.1 Stimulation sequence. Dotted line represents the dummy condition displayed during the first 15 s. to avoid magnetic saturation. Two dimensional tracking (2DT) was used as a base line. Blue blocks represent the three dimensional tracking condition (3DT).

The display contained one thousand and twenty four dots (with radio 0.1 deg. and zero disparity) distributed over the screen (mean dot density 1.5 dot deg^{-2}). The subjects were instructed to fixate on the right superior corner of a square (5.23 deg. side long) moving laterally in the screen (13 deg. field of view). The square was moving from left to right and vice versa at 2.19 deg.sec^{-1}, each time that the square reached one edge of the screen it changed its direction. Dynamic random noise was used in order to remove the motion cues introduced by the chance in disparity (Hanazawa, Kawashima et al. 2000). There were two modalities to define the square, each one represents one experimental condition (fig. 6.2).

a) **Two Dimensional Tracking (2DT)**: The square was luminance defined, its luminance (8.56 cd/m^2) was lower that the background (18 cd/m^2). This condition was used as a base line.

b) **Three Dimensional Tracking (3DT)**: The square was depth defined (red/green anaglyph stereogram), it laid at the front (-0.3 deg.) of the background (zero disparity). It is important to clarify that the square was moving in the plane X/Y, not in plane Z (motion in depth).

6.3 Methods.

6.3.1 Subjects.

Ten healthy subjects, nine right-handed and one left-handed volunteers (7 female, 3 male) aged from 20 to 30 years participated in the present study. The stereo acuity of the subjects was measured using stereo vision test (RANDOT SO-002), all of them were below 40 sec of ARC. The subjects had a preliminary session outside the magnet to become familiar with the visual stimulation. All subjects gave informed written consent.

6.3.2 Stimulus presentation.

Subjects lay on their backs 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 just outside 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 G3 Mac running Psychophysics Tool Box ver. 2.44 (Brainard 1997; Pelli 1997) under MATLAB ver. 5.3. The stimuli were displayed at a video frame rate of 60 Hz.

Two dimensional tracking
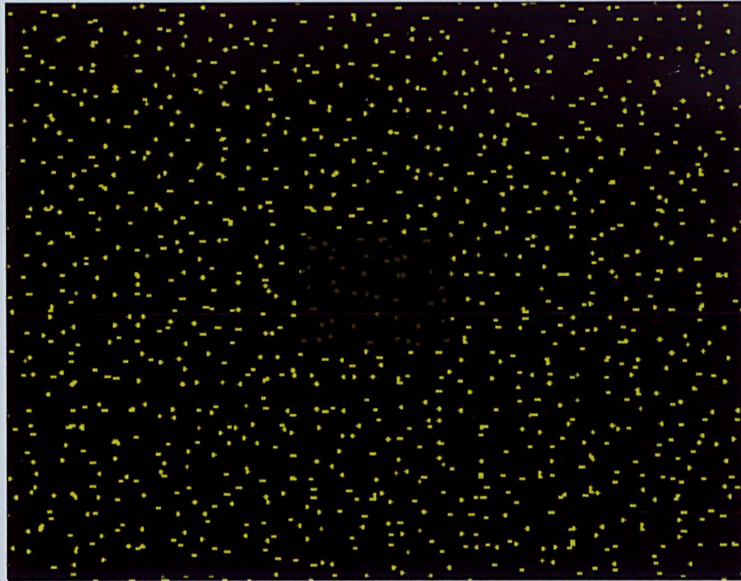
Three dimensional tracking
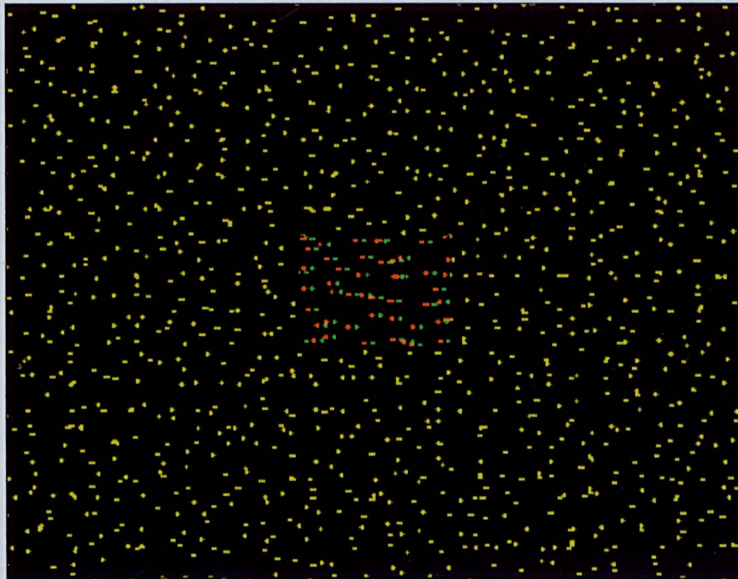
Figure 6.2 Visual stimuli. Subjects viewed the stimuli through the red/green glasses to present anaglyph stereo stimuli. In order to exemplify the stimulus used, the central position of the square is displayed in the frames above.

6.3.3 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.). Because some studies suggest the involvement of parietal regions in the processing of stereo information (Gulyas and Roland 1993; Hanazawa, Kawashima et al. 2000; Nishida, Hayashi et al. 2001), the area of acquisition was expanded to cover these regions. Thirty two slices covering the whole brain were acquired (fig. 6.3).

It is important to clarify, that no activations in parietal regions are expected due to either eye movements or attention modulations (Darby, Nobre et al. 1996; Corbetta, Akbudak et al. 1998; Petit and Haxby 1999; Nobre, Gitelman et al. 2000), because the control condition was designed to counterbalance them.
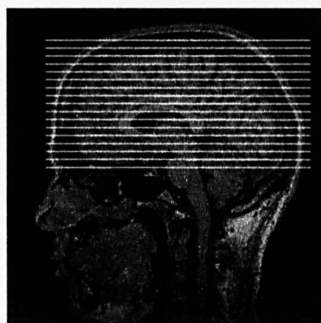
Figure 6.3 Data acquisition.

6.3.4 Data analysis.

The data was pre-processed and analysed using SPM99 (Wellcome Department of Cognitive Neurology). The first five scans of each run were discarded in order to exclude magnetic saturation artefacts. All volumes were slice timed with respect to the first slice. Motion correction was made taking as a reference the first volume of the first run. The data was normalised in the MNI (Montreal Neurological Institute) stereotaxic space and smoothed using a 6 mm FWHM (full width at half maximum) isotropic Gaussian kernel (WDCN 1997).

Data analysis was performed using a boxcar design matrix of the different conditions (2DT and 3DT) convolved with the hemodynamic response function. Specific effects were tested by applying the corresponding linear contrast to the parameters obtained applying General Linear Model using the design matrix shown in figure 6.4 (see also experiment design). 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), (WDCN 1997). Group analysis (random effects) was not developed due to the small number of subjects scanned (Friston, Holmes et al. 1999).
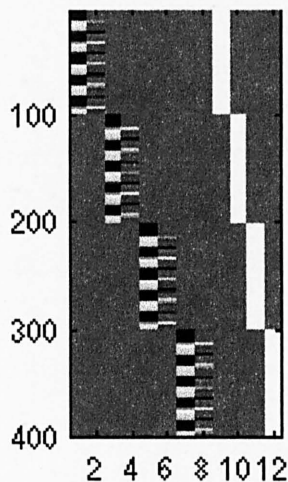
Figure 6.4 Design matrix. The design matrix used in the present analysis is shown in the conventional SPM display. Each block represents one sequence (run). Each column of the block represents one of the experimental conditions in the following order: 3DT, first derivative. The base line condition (2DT) is represented as the complement of the experimental conditions. The white squares shows the periods in which the related condition was presented. The four columns at the end of the design matrix represent each of the independent sequences. Rows represents scans (400 scans = 10 scans per condition * 10 conditions per sequence * 4 sequences).

6.4 Results.

In the present study, two analyses were performed in order to identify the areas involved in each particular stimulus. For reasons of space, individual analysis outputs are not presented, instead the results are reported as a synopsis through the use of condensed statistical tables which show a summary of the relevant statistical parameters of the regions of activation. In the same manner, the statistical map of the most sensitive subject is presented.

6.4.1 Three dimensional tracking against Two dimensional tracking.

This contrast compares the activation produced by the moving square defined by depth with that produced by the moving square defined by luminance. Consistent with our expectations V3A, V3B and parietal regions were activated (fig. 6.5). Although we could not track the eye movements to assess how well the subjects were developing the task, we do not believe that the observed activation were produced by saccadic eye movements because after both, the training session and the scan session, the subjects reported easily following the path of the square in 2DT and 3DT conditions.

On the other hand, studies in eye movements suggest the involvement of parietal areas in the execution of eye movements but not the involvement of V3A and V3B regions (Darby, Nobre et al. 1996; Corbetta, Akbudak et al. 1998; Petit and Haxby 1999; Nobre, Gitelman et al. 2000).
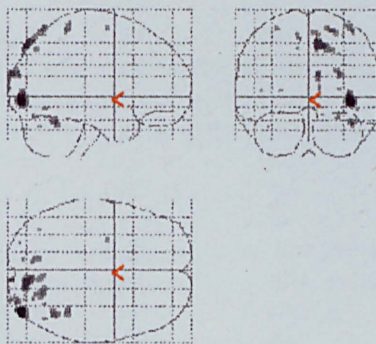
Figure 6.5 Three dimensional tracking against Two dimensional tracking. The statistical map shows the areas sensitive to stereoscopic information. The activation includes V3A, V3B and parietal regions.

Area V3A was activated bilaterally in 2 subjects, and in the right hemisphere of 5 subjects (table 6.1). The right hemisphere dominance of the stereo activation in V3A is supported by the theory of right cerebral dominance in stereo vision (Howard and Rogers 1995).

Subject	Location	Z-Score	P corrected	Cluster size
1R	32, -96, 16	5.65	0.002	21
2R	30, -92, 24	6.20	0.000	16
3R	28, -88, 16	5.09	0.029	3
4 R	34, -94, 16	5.62	0.003	4
7R	34, -92, 22	6.25	0.000	8
8R	36, -86, 16	(Inf)	0.000	45
8L	-28, -96, 16	(Inf)	0.000	62
10R	32, -94, 16	(Inf)	0.000	116
10L	-28, -98, 14	(Inf)	0.000	546

Table 6.1 Stereo sensitive region: V3A. The table shows the locations of the activations and the statistical parameters.

Consistent with our previous results, area V3B was activated bilaterally in 5 subjects, in the right hemisphere of 3 subjects and in the left hemisphere of 1 subject (table 6.2). Since the pattern of activation was similar for both regions, it suggest a functional relationship between the activation on V3A and V3B. i.e. (see table 6.4).

Subject	Location	Z-Score	P corrected	Cluster size
1R	34, -84, -6	6.25	0.000	36
1L	-26, -94, 2	5.80	0.001	25
2R	40, -88, -4	6.98	0.000	38
3R	38, -88, -4	5.29	0.011	6
3L	-28, -96, -10	5.54	0.003	4
4 R	40, -88, 0	5.81	0.001	16
6 R	28, -104, 2	7.03	0.000	93
7 R	36, -92, 0	(Inf)	0.000	105
7L	-28, -102, 2	(Inf)	0.000	37
8R	36, -92, 4	(Inf)	0.000	35
8L	-26, -100, 6	(Inf)	0.000	176
9 L	-28, -100, -2	4.54	0.000	51
10R	38, -90, 2	5.88	0.000	3
10L	-28, -100, 2	7.84	0.000	7

Table 6.2 Stereo sensitive region: V3B. The table shows the locations of the activations and the statistical parameters.

Consistent with other studies, parietal activation were found in this contrast. The precuneus was activated in the right hemisphere of 3 subjects and in the left hemisphere of 2 subjects (Gulyas and Roland 1993; Fortin, Ptito et al. 2000). The superior parietal was activated in the right hemisphere of 2 subjects (Nishida, Hayashi et al. 2001),(table 6.3).

Subject	Location	Z-Score	P corrected	Cluster size
2R	16, -90, 40	6.53	0.000	19
3 L	-12, -90, 42	5.84	0.001	11
4 R	8, -82, 50	6.96	0.000	145
7R	24, -56, 64	(Inf)	0.000	104
8R	32, -64, 62	(Inf)	0.000	43
8L	-22, -74, 42	(Inf)	0.000	44
10R	14, -88, 44	(Inf)	0.000	224

Table 6.3 Stereo sensitive regions: Superior parietal lobule (grey rows) and precuneus. The table shows the locations of the activations and the statistical parameters.

6.4.2 Two dimensional tracking against Three dimensional tracking.

The subtraction of the activations produced by following the square defined by depth from the activation produced by the square defined by luminance, did not revealed any consistent activation, i.e. although statistically significant activations were found for this contrast, all of them were spread over the brain without any consistent location.

6.5 Summary of activations.

The following tables compare the functional profiles and the average Talaraich coordinates of the V3A and V3B areas founded in the present experiment with those reported by other authors.

V3A

Reference	x	y	z	Sensitive to
Tootell, 1997	±14	-84	19	Directional motion.
Buchel, 1998	-30	-90	9	Directional motion.
	36	-84	18	
Hasnain, 1998	-25.7 (2.9)	-83.8 (4.9)	11.0 (6.1)	Defined by
	17.4 (9.3)	-84.1 (6.1)	15.9 (3.4)	retinotopic mapping.
Mendola, 1999	-21.2 (9.3)	-89.1 (4.3)	16.5 (11.1)	Stereopsis and
	22.7 (10)	-88.3 (5.6)	16.5 (10.5)	illusory contours.
Backus, 2001	*No given*	*by the*	*author*	Stereopsis.
Acosta, 2001	-27.7 (1.5)	-93.2 (1.41)	18.48 (1.41)	Directional motion
(Present study)	32 (2.94)	-88 (3.54)	20.9 (3.46)	and Stereopsis.

V3B

Reference	x	Y	z	Sensitive to
Orban, 1995	-25	-88	-1	Kinetic boundaries.
Dupont, 1997	-28	-94	-4	Kinetic contours, shape
	34	-88	0	and motion.
VanOostende,1997	±31	-91	0	Kinetic boundaries.
Smith , 1998a	±26 (8)	-89 (8)	-2 (8)	Second order motion.
Backus, 2001	*No given*	*by the*	*author*	Stereopsis.
Acosta, 2001	-27 (1.03)	-95.58 (3)	4.7 (5.51)	Stereopsis, stereoscopic
(Present study)	35.8 (3.91)	-87.94(5.9)	3.22 (4.42)	motion.

Although the design of the present experiment did not contain a contrast to identify the V5 region, the data of the previous experiment was used to identify its mean location. The following table compares our results with those presented in other studies.

V5

Reference	x	Y	z	Sensitive to
Orban, 1995	-44	-67	3	Directional motion.
	40	-70	4	
VanOostende,1997	-42	-78	4	Directional motion.
	41	70	2	
Buchel, 1998	-48	-72	-9	Directional motion.
	42	-78	-12	
Hasnain, 1998	-39 (2.6)	-71.9 (4.2)	-0.8 (4.1)	Retinotopic mapping.
	41.3 (4.7)	-64.8 (7)	-1.6 (6.2)	
Smith , 1998a	±46 (7)	-70 (7)	4 (7)	First and second order motion.
Mendola, 1999	-45.9 (7.5)	-70.1 (5.2)	1.8 (8.2)	Directional motion.
	45.5 (8.1)	-65.9 (7.9)	-0.9 (6.5)	
Acosta, 2001	-44.5 (5.6)	-70.6 (6.1)	5.67 (4.6)	Directional motion.
(Previous study)	44.8 (2.7)	-73.84 (6.1)	5.79 (4.2)	

From the average location of the areas V3B and V5, it is possible to conclude that the results of the present experiment support our hypothesis for the existence of a stereo sensitive area located at almost the same horizontal level as V5 but more posterior and more medial (fig. 6.6).
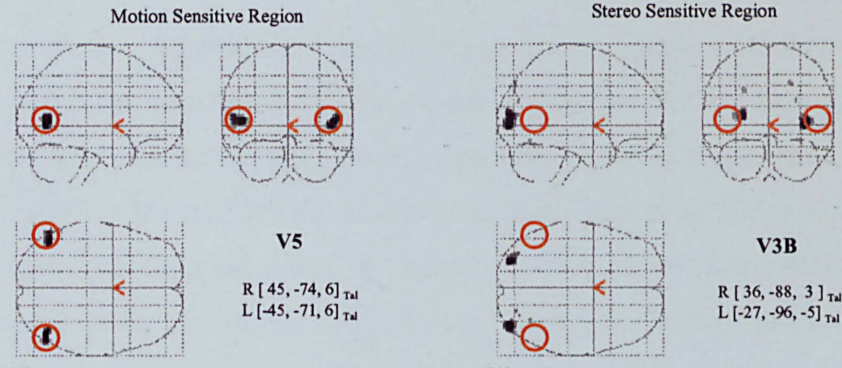
Figure 6.6 Anatomical location of V3B and V5 regions. Left panel shows the location of the motion sensitive visual area (V5). Right panel shows the location of the stereo sensitive visual area (V3B). Red circles helps to show the different location of each region.

Our results supports the hypothesis of right hemisphere dominance in stereoscopic vision, however, there was no evidence in the data that suggested any relationship between gender, eye dominance or hand dominance with the activations found in our experiments. Table 6.4 summarises the results of this study.

Subject	Gender	Dom. Eye	Dom. hand	V3A	V3B	BA7
1	F	L	L	R	R,L	X
2	F	R	R	R	R	R
3	F	R	R	R	R,L	L
4	F	L	R	R	R	R
5	F	R	R	X	X	X
6	F	R	R	X	R	X
7	M	R	R	R	R,L	R
8	M	R	R	R,L	R,L	R,L
9	F	R	R	X	L	X
10	M	R	R	R,L	R,L	R

Table 6.4 Summary of results. The characters L, R and X in the table refers to left, right and no-activation respectively.

6.6 Conclusions.

The results of this experiment revealed three main regions sensitive to the stereoscopic information, V3A, V3B and parietal regions. Consistent with other studies, V3A and parietal regions showed sensitivity to stereo disparities (Gulyas and Roland 1993; Mendola, Dale et al. 1999; Backus, Fleet et al. 2001). However, only one study has reported the V3B region as being sensitive to stereoscopic information (Backus, Fleet et al. 2001). In our results, the preferential activation of V3B to the stereo stimuli cannot be explained either by different patterns of eye movements or by stronger attentional engagement since the control condition was designed to provide identical requirements for these parameters. On the other hand, we do not believe that the activation on V3B can be produced by the segmentation in the three dimensional space required by the global stereo tracking task, since fMRI studies in perception of object shape suggest that the activation of cortical regions (Lateral Occipital Complex) involved in the analysis of object structure is independent of the cues (luminance, colour, depth) that define the shape (Kourtzi and Kanwisher 2000).

A possible explanation for the activation of V3B in our data (and not in other stereoscopic studies) can be postulated by looking at the differences in the experimental stimulus used. All those stereoscopic studies have used static stereo stimulus, whereas in ours, it is a dynamic stimulus, i.e., our stereo stimuli not only requires the perception of square defined by depth (form task), but also requires the tracking of it over time (motion task).

Motion can be detected using different sources of information, for example, luminance, colour, texture or depth. Depending on its operational definition, motion perception can be categorised as first or second order motion. First order motion can be identified by tracking the position of spatial structures (one point moving in space), whereas it is not possible with second order motion (motion defined by spatio-temporal changes of horizontal disparities, also referred to as stereoscopic -cyclopean- motion).

Based on this we hypothesise that the V3B region was activated by the stereoscopic motion component of the global stereo tracking task. Supporting this hypothesis, Patterson's studies (Patterson, Donnelly et al. 1997) showed how speed discrimination can be done under conditions of minimal position information. This suggests that stereoscopic motion is processed in a different way, perhaps by a different mechanism than that which processes first order motion (Smith and Scott-Samuel 1998b). Moreover, in an fMRI study of second order motion, Smith *et al* (Smith, Greenlee et al. 1998a) found a cortical region sensitive to motion defined by contrast rather than by luminance. This cortical region was previously identified and named Kinetic Occipital (KO) for it's sensitivity to kinetic defined contours. Smith called it V3B for the sake of consistency with the classical alpha-numerical classification of the visual areas.

The comparison of the anatomical location of these regions, shows that they are very similar or probably the same (see table in section 6.5). This suggests that the activation found in our experiment, matches both functionally and anatomically the V3B region. Thus, it is not unreasonable to hypothesise that the V3B region is sensitive to second order motion as defined by contrast and also by stereoscopic motion.

In the next chapter, a second experiment introducing a stereoscopic motion component is presented to test the consistency of our hypothesis. The remaining question is why this region was activated in two of the eight subjects in the first experiment since the stereo stimulus (stereo cone) did not contain any second order motion at all?. However that activation could not be considered as consistent (because it was present in only two subjects) and more research has to be done to investigate this phenomenon.

Finally, it is important to remember that the effective connectivity model proposed in chapter 4 (experiment 1), was based on the assumption that V3B was involved in the processing of pure stereo stimuli, therefore, that model should be reinterpreted in the light of these results (see section 8.1.2 of the next chapter).

Chapter 7

Experiment 4

Shape Discrimination from Stereopsis

Abstract.

The objective of the present experiment is to explore not only the cortical regions involved in pure stereoscopic vision but also the behaviour of the V3B region identified in our previous experiments. Red/green random dot anaglyph stereograms were used to define a *pacman* shape (lying in front of the background). A similar display (*pacman* defined by luminance) was used as a control condition. The figure changes to one of four possible positions every second. The subjects were instructed to press a button when they identified a certain position of the figure. The advantages of this stimulus over the previous one are that it avoids eye movements and provides a mechanism to assess how well the subjects are developing the task. The results support our hypothesis of the engagement of V3A and precuneus in the processing of stereo disparities, and the possible engagement of V3B in processing stereoscopic motion. These results also suggest that V1 region is sensitive to near zero disparities.

7.1 Introduction.

The results of the previous experiment revealed three main regions sensitive to the stereoscopic information. Consistent with other studies, V3A and parietal regions showed sensitivity to stereo disparities (Gulyas and Roland 1993; Mendola, Dale et al. 1999; Backus, Fleet et al. 2001). However, the use of the Global Stereo Tracking (GST) task showed a cortical region that has not been usually shown in the results reported by other authors. We hypothesised that this region was activated by the stereoscopic motion induced by the GST task, i.e., this region is sensitive to second order motion defined by spatio-temporal changes of horizontal disparities (stereoscopic -cyclopean- motion).

Supporting this hypothesis, Smith's study (Smith, Greenlee et al. 1998a) reports a cortical region sensitive to second order motion defined by contrast which matches our region both functionally and anatomically. We concluded that these regions are very similar or probably the same as that referred to as V3B. Thus, we hypothesised that the V3B region is sensitive to second order motion as defined by contrast and also by stereoscopic motion.

It is important to remember that our ideal in the design of the previous experiment was to increase the response of the stereo sensitive regions by the introduction of an attentional task (GST). The subjects were instructed to perform pursuit eye movement to follow the path of a square with their eyes. The advantage of this paradigm is that it is an active rather than a passive stimulus, in which the success of the task depends on the continuous perception of the square, i.e. to follow the path of the square, it is necessary to involve the stereoscopic processing system.

The objective of the present experiment is to explore not only the cortical regions involved in pure stereoscopic vision but also the behaviour of the V3B region identified in our previous experiments. As in the GST task, the present stimulus was designed to maintain the attention of the subject through the use of an active task.

A *"pacman"* shape defined either by luminance (2D) or by depth (3D) was displayed at the centre of the screen. The figure changes to one of four possible positions every second. The subjects were instructed to press a button when they identified a certain position of the figure. The advantages of this task over the GST are that it avoids eye movements and provides a mechanism to assess how well the subjects are performing the task.

7.2 Experiment Design.

Subjects were given 4 sequential scans each lasting 5.15 min. (10 epoch) with a 5 min. interscan interval to permit subjects to rest. One hundred image volumes were obtained in each run. Each condition lasted 30s., giving 10 multislice volumes per condition (TR=3s.). A dummy condition of a blank screen was presented during the first 15s. (5 scans) of each run to control for magnetic saturation effects. The stimulation sequence is shown in figure 7.1.
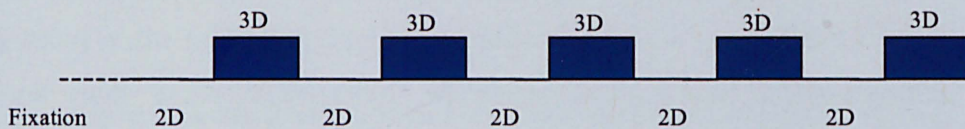
Figure 7.1 Stimulation Sequence. Dotted line represents the dummy condition displayed during the first 15 s. to avoid magnetic saturation. Luminance task (2D) was used as a base line. Blue blocks represent the depth task condition (3D).

The display contained four thousand seven hundred and sixty one dots (radius 0.04 deg.) distributed over the screen (mean dot density 7 dot deg^{-2}). A *pacman* shape (radius 4.3 deg.) was displayed at the centre of the screen. The *pacman* was changing randomly in one of four possible positions (up, down, left, right) every second (see figure 7.2). The change in position was constrained to avoid the fact that one position was displayed for more than one second. In order to remove possible shape cues introduced by the red/green stereoscopic pair of dots in the stereo condition, the positions of the dots were changed between frames (in both conditions) to produce the effect of dynamic random noise (Hanazawa, Kawashima et al. 2000).
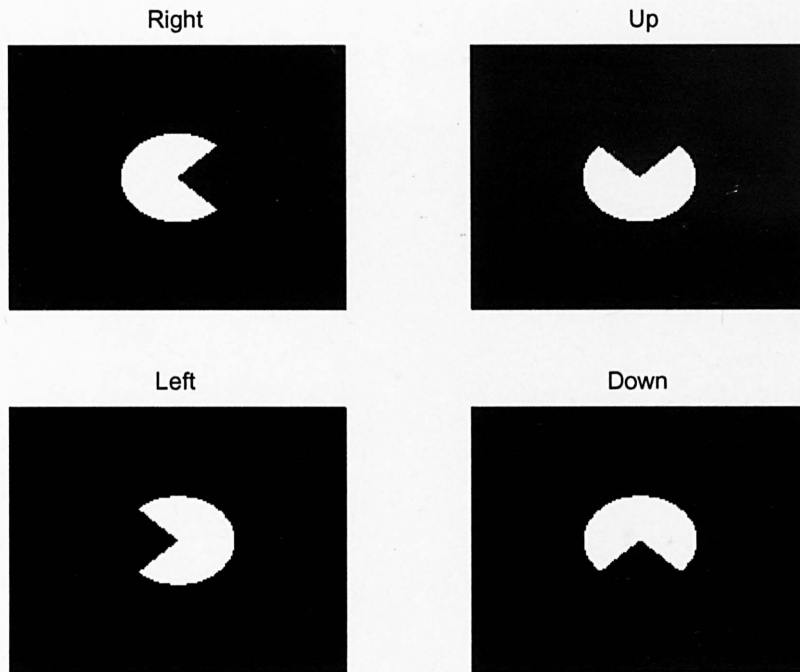
Figure 7.2 Pacman positions. The pacman figure was changing its position every second. Four positions were defined: right, left, up and down.

The subjects were instructed to fixate a point (0.3 deg. of radio and zero disparity) in the middle of the screen (circular field of view 13 deg.) and press a button (with her/his right hand) when the *pacman* was in up position. The response on the button box was recorded to assess the performance of the subjects. There were two modalities to define the *pacman*, each one represents one experimental condition (fig. 7.3).

a) **Luminance (2D)**: The *pacman* was luminance defined, its luminance (9.29 cd/m^2) was lower that the background (17.7 cd/m^2). Both the background and the *pacman* were laid in the same plane (zero disparity). This condition was used as a base line.

b) **Depth (3D)**: The *pacman* was depth defined (red/green anaglyph stereogram), and laid at the front (-0.076 deg. of disparity.) of the background (0.076 deg. of disparity). The *pacman* and the background were displayed in front and behind the fixation point respectively in order to remove possible shape cues introduced by the red/green stereoscopic pair of dots in the stereo condition.

Luminance
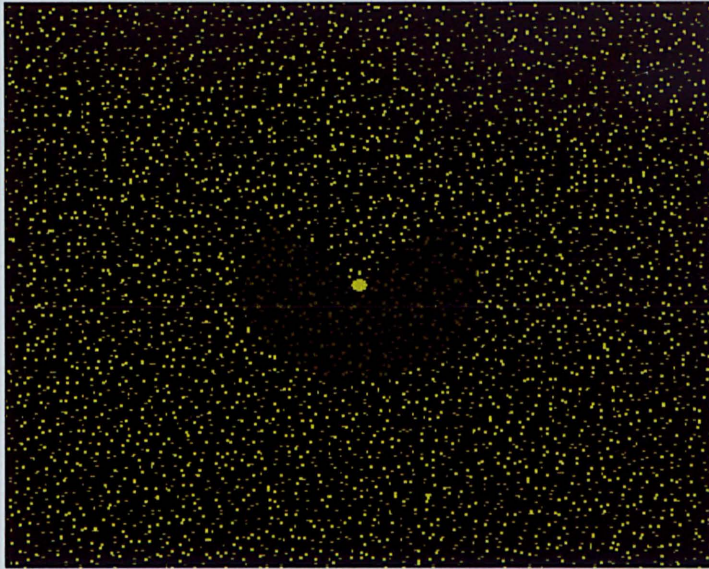
Depth
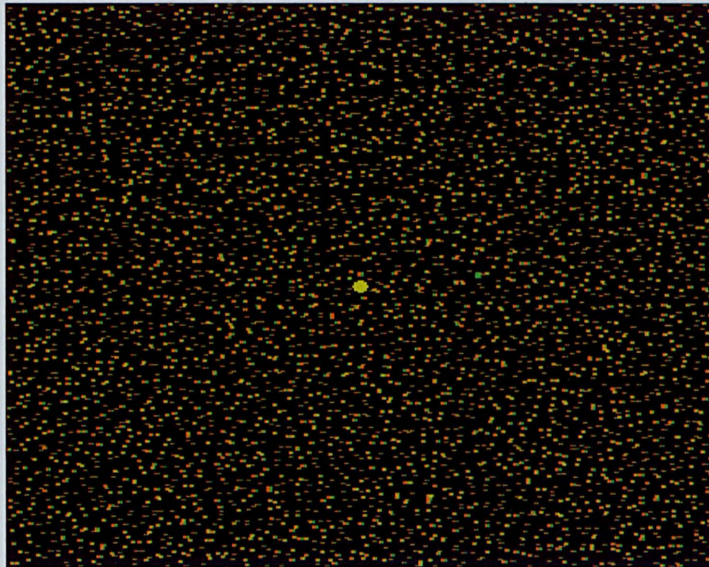
Figure 7.3 Visual stimuli. Subjects viewed the stimuli through the red/green glasses to present anaglyph stereo stimuli. In order to exemplify the stimulus used, the up position of the *pacman* is displayed in the frames above.

7.3 Methods.

7.3.1 Subjects.

Five healthy right-handed subjects (2 female, 3 male) aged from 20 to 30 years participated in the present study. The stereo acuity of the subjects was measured using stereo vision test (RANDOT SO-002), all of them were below 40 sec of ARC. The subjects had a preliminary session outside the magnet to become familiar with the visual stimulation. All subjects gave informed written consent.

It is important to clarify that for reasons of financial restrictions, in the present study only five subjects were scanned.

7.3.2 Stimulus presentation.

Subjects lay on their backs 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 just outside 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 G3 Mac running Psychophysics Tool Box ver. 2.44 (Brainard 1997; Pelli 1997) under MATLAB ver. 5.3. The stimuli were displayed at a video frame rate of 60 Hz.

7.3.3 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.). As performed in the previous experiment (see section 6.3.3), thirty-two slices covering the whole brain were acquired.

7.3.4 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 effects. All volumes were slice timed, motion corrected and normalised in the MNI (Montreal Neurological Institute) stereotaxic 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 (2D and 3D) convolved with the hemodynamic response function.

Specific effects were tested by applying the corresponding linear contrast to the parameters obtained applying General Linear Model using the design matrix shown in figure 7.4 (see also experiment design). The statistical parametric maps (SPMs) were then interpreted by referring to the probabilistic behaviour of Gaussian random fields. Although for some subjects the activation in V3A and V3B was not significant at p corrected < 0.05, they were significant at this level using a small volume correction (WDCN 1997).
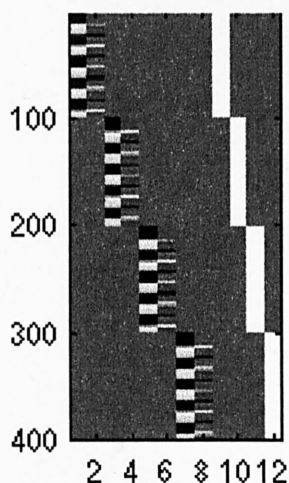
Figure 7.4 Design matrix. The design matrix used in the present analysis is shown in the conventional SPM display. Each block represents one sequence (run). Each column of the block represents one of the experimental conditions in the following order: 3D, first derivative. The base line condition (2D) is represented as the complement of the experimental conditions. The white squares shows the periods in which the related condition was presented. The four columns at the end of the design matrix represent each of the independent sequences. Rows represents scans (400 scans = 10 scans per condition * 10 conditions per sequence * 4 sequences).

7.4 Results.

In the present study, two analyses were performed in order to identify the areas involved in each particular stimulus. For reasons of space, individual analysis outputs are not presented, instead the results are reported as a synopsis through the use of condensed statistical tables which show a summary of the relevant statistical parameters of the regions of activation. In the same manner, the statistical map of the most sensitive subject is presented.

The performance of the task was assessed integrating all the occasions in which the subjects pressed the button at the right time (when the *pacman* was in up position). Then, subtracting the occurrences when the subjects pressed the button at wrong times (when the *pacman* was not in up position). The final score was computed dividing this difference by the number of occurrences of the *pacman* in the up position. In both, the training session and the scan session, all of the subjects scored at least 95% of accuracy in the performance of the discrimination task.

7.4.1 Depth against Luminance.

This contrast compares the activation produced by the *pacman* defined by depth with that produced by the *pacman* defined by luminance. Consistent with our expectations V3A, V3B and precuneus were activated. Activations were also found in striate cortex (V1), we attribute this activation to the fact that small disparities were used in the depth condition, i.e. there is evidence that V1 is tuned to near zero disparities (Poggio, Gonzalez et al. 1988; Backus, Fleet et al. 2001). The statistical parametric map of one of the subjects is shown in figure 7.5.

Area V3A was activated bilaterally in 3 subjects, and in the right hemisphere of 2 subjects (table 7.1). The right hemisphere dominance of the stereo activation in V3A is supported by the theory of right cerebral dominance in stereo vision (Howard and Rogers 1995).
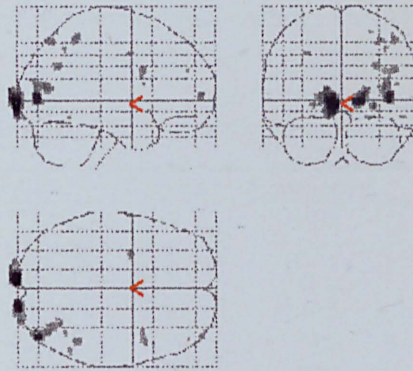
Figure 7.5 Depth against Luminance. The statistical map shows the areas sensitive to stereoscopic information. The activation includes V1, V3A, V3B and precuneus.

Subject	Location	Z-Score	P corrected	Cluster size
1L	-24, -98, 16	5.30	0.000	1
1R	36, -88, 18	5.29	0.000	1
2R	28, -78, 24	3.11	0.077	1
3R	34, -90, 20	4.67	0.001	47
4L	-10,-104,16	3.77	0.02	2
4R	36,-90, 16	5.99	0.000	1
5L	-22, -98, 18	7.21	0.000	52
5R	34, -94, 12	5.63	0.000	9

Table 7.1 Stereo sensitive region: V3A. The table shows the locations of the activations and the statistical parameters.

Consistent with our previous results, area V3B was activated bilaterally in 1 subject, and in the right hemisphere of 4 subjects (table 7.2). Since the pattern of activation was similar for both regions, it suggests a functional relationship between the activation on V3A and V3B.

Subject	Location	Z-Score	P corrected	Cluster size
1R	42, -82, 4	(Inf)	0.000	149
2R	38,-90, 2	3.96	0.056	39
3R	40, -80, 0	4.58	0.001	21
4L	-26, -100, 4	7.59	0.000	28
4R	36, -94, 6	(Inf)	0.000	103
5L	-36, -92,4	5.47	0.000	7
5R	36, -88, -4	6.55	0.000	21

Table 7.2 Stereo sensitive region: V3B. The table shows the locations of the activations and the statistical parameters.

The precuneus was activated bilaterally in 1 subject and in the right hemisphere of 4 subjects (table 7.3).

Subject	Location	Z-Score	P corrected	Cluster size
1R	32, -72, 52	7.06	0.000	1
2L	-26, -74, 46	4.67	0.003	4
2R	28, -66, 54	6.01	0.000	96
3R	14, -78, 52	3.90	0.015	5
4R	28, -64, 54	4.30	0.003	8
5R	14, -74, 54	5.47	0.000	1

Table 7.3 Stereo sensitive region: Precuneus (BA7). The table shows the locations of the activations and the statistical parameters.

Consistent with other studies which used stimulus with small disparities, activations in V1 were found in this contrast (Backus, Fleet et al. 2001). The primary visual cortex was activated bilaterally in 2 subjects, in the left hemisphere of 2 subjects and in the right hemisphere of 1 subject (table 7.4).

Subject	Location	Z-Score	P corrected	Cluster size
1L	-8, -102, -2	(Inf)	0.000	236
1R	16, -100, 0	(Inf)	0.000	149
2L	-8, -106, -2	6.20	0.000	59
3L	-6, -104, -6	5.27	0.000	55
4L	-14, 106, 0	5.83	0.000	15
4R	16, -100, 6	6.65	0.000	214
5R	10, -100, 4	5.06	0.001	1

Table 7.4 Stereo sensitive regions: V1. The table shows the locations of the activations and the statistical parameters.

7.4.2 Luminance against Depth.

The subtraction of the activations produced by the *pacman* defined by depth from the activation produced by the *pacman* defined by luminance, did not revealed any consistent activation, i.e. although statistically significant activations were found for this contrast, all of them were spread over the brain without any consistent location. This inter-subject variability was considered as noise.

7.4.3 Motor response.

Although the identification of the motor regions was not the purpose of the present study, the activation produced by the finger movements (when the subjects press the button to confirm that they perceived the *pacman* in up position) were identified in the present contrast.

The responses recorded by the button box were convolved with the canonical hemodynamic response function to use it as a column in the GLM design matrix. As expected, this contrast revealed activations in the primary motor cortex contralateral to the hand used in the motor task. The statistical parametric map of one of the subjects (which used the left right hand to press the button) is shown in figure 7.6.
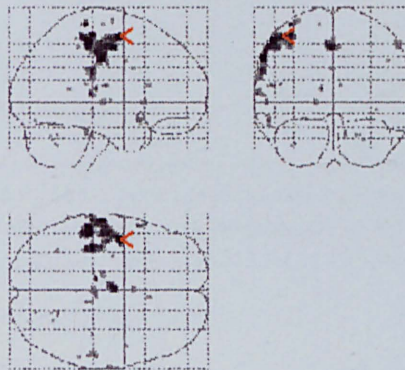
Figure 7.6 Motor response. The statistical map shows the areas involved in the motor activity. The activation was found in the primary motor cortex, contralateral to the hand used to press the button.

Three main regions were activated in the left primary motor cortex: Precentral Gyrus, Postcentral Gyrus (BA3), and Cyngulate Gyrus (table 7.5). It is important to clarify that because this contrast was not properly counterbalanced (to permit the signal to decay), the statistical significance of the response was weak and not valid at $P < 0.05$ corrected. So, it was assessed with criteria of $P < 0.001$ uncorrected.

Region	Location	Z-Score	P uncorrected
Precentral Gyrus.	-44,-4, 58	4.94	0.000
Postcentral Gyrus(BA3).	-62, -20, 38	4.80	0.000
Cyngulate Gyrus.	-2, -10, 48	4.59	0.000

Table 7.5 Motor cortex: Precentral Gyrus, Postcentral Gyrus (BA3), and Cyngulate Gyrus. The table shows the locations of the activations and the statistical parameters.

Our results supports the hypothesis of right hemisphere dominance in stereoscopic vision, however, there was no evidence in the data that suggested any relationship between gender, eye dominance or hand dominance with the activations found in our experiments. Table 7.6 summarises the results of this study.

Subject	Gender	Dom. Eye	Dom. hand	V1	V3A	V3B	BA7
1	F	R	R	R,L	R,L	R	R
2	F	L	R	L	R	R	R,L
3	M	R	R	L	R	R	R
4	M	R	R	R,L	R,L	R,L	R
5	M	L	R	R	R,L	R	R

Table 7.6 Summary of results.

7.5 Conclusions.

In accordance with other studies on stereoscopic perception, our results showed inter-subject consistent activations in the regions: V1, V3A, V3B and precuneus. Given that V1 is tuned to near zero disparities (Poggio, Gonzalez et al. 1988; Backus, Fleet et al. 2001), the activation observed in the V1 region (and not observed in our previous experiments) was attributed to the small disparities (±0.076 deg.) used in the depth stimuli.

The activation of the V3B region supports our hypothesis of the sensitivity of the V3B region to stereo disparities, particularly in the processing of motion defined by stereopsis.

Although the stereoscopic motion component induced by the rotatory *pacman* was not the *optimal stimuli* to activate the V3B region (see section 8.3.2 of the next chapter), we believe that it was strong enough to activate the stereoscopic motion sensitive region of all the subjects (see table 7.2).

Figure 7.7 shows the time series of the right V3B region of each subject. The grey columns represent the periods in which the *pacman* defined by depth was displayed, whilst the white columns represent the base line condition (*pacman* defined by luminance). In average, the amplitude of the signals was 1% over the mean activation, with frequency equal to that used in the experimental stimulation (0.0167 Hz).
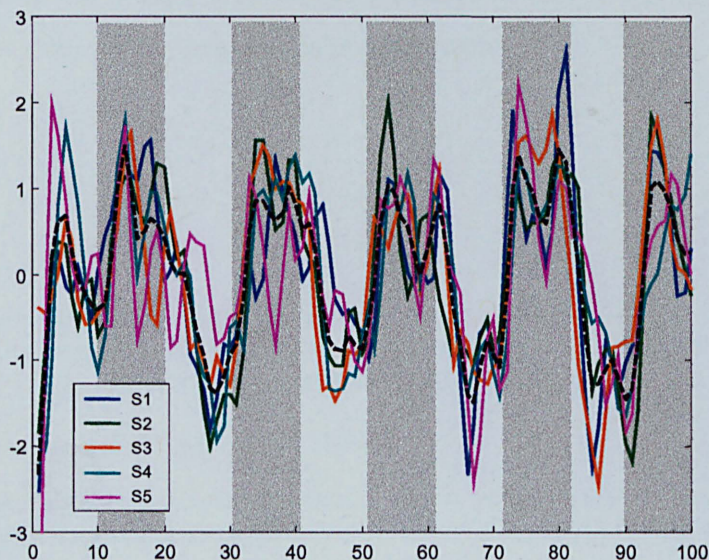
Figure 7.7 V3B time series. The first principal component time series of a sphere 8mm radius around the voxels specified in table 7.2 (right V3B region) were averaged over sessions (runs) to create a single time series (100 time points) per subject (S). Each time series represents the neural activity over time (scans) of the right V3B region in each subject. The time series were normalised to mean zero and standard deviation one. The dashed black time series represents the average of the subjects. Grey columns show the stereo stimulation periods.

The experimental design of the present study improves two aspects of the previous experiment. First, the activation on V3B region cannot be related with eye movements because in this study the subjects were instructed to fixate in a static feature.

Second, the only cue available to develop the discrimination task was the stereo disparities, thus, the correct discrimination of the shape was an indicator of depth perception, i.e., the identification of the "up position" of the *pacman* implied a proper stereoscopic fusion.

In conclusion, the results of this experiment support our hypothesis of the engagement of V3A and precuneus in the processing of stereo disparities and the possible engagement of V3B in the processing of stereoscopic motion. These results also suggest that V1 region is sensitive to near zero disparities. The remaining question is why: using stimulus that did not contain any stereoscopic motion at all (our first two experiments and Backus' study) some subjects showed activation in the V3B region?. One possible explanation is that it may be produced by illusory stereoscopic motion effects (Zeki, Watson et al. 1993), however more research has to be done to investigate this phenomenon.

Chapter 8

Discussion and Future Work

Abstract.

The present chapter summarises the results of the
studies made in human and non-human primates
to understand the functional anatomy of
stereoscopic vision. The findings of our
experiments are discussed in the context of other
related studies, and conclude that V3A and
precuneus seems to be involved in pure
stereoscopic processing whereas V3B might be
particularly sensitive to stereoscopic motion.
Also, a new technique for a more appropriate
representation of effective connectivity models
using fMRI time series is introduced as the
NARMAX approach. This representation attempts
to bring SEM towards a non-linear dynamic
system modelling technique. Finally, the direction
for the new studies arising from this thesis is
described with ideas for future research.

8.1 Functional anatomy of stereoscopic vision.

Stereoscopic vision is one of the most important sources of depth information for species with frontally located eyes. Although there are other sources of three-dimensional cues, such as shading, texture, motion or perspective; binocular disparities are particularly important because they provide a mechanism to estimate not only the form but also the position of objects in space.

Many studies have investigated the physiological mechanisms from which the brain transforms the retinal disparities into three-dimensional representations of the world. Given the invasive nature of the techniques to assess neural activity, most of these studies have been in non-human primates and the research on humans has been restricted to psychophysical studies.

Recently, modern non-invasive neuroimaging techniques have allowed the investigation of the functional anatomy of the human brain. These new techniques identify the sites of neural activity using exogenous indicators such as fluorodeoxyglucose (PET) or oxygen concentration on tissue (fMRI). However, the understanding of the relationship between these indicators and the 'effective' neural activity at synaptic level remains incomplete.

Although many PET and fMRI studies have explored the functional properties of the visual cortex, few have explored the functional anatomy of stereoscopic vision. Most of these studies seem to be pilot work; and they show inconsistency regarding the areas sensitive to stereo disparities. It seems to be that the stereo disparity processing is widespread over a network of cortical regions rather than being confined to a specific cortical region. Furthermore, there is no general agreement about the cortical regions selective to stereo disparities or the specific role that each of these has in the perception of depth. The regions most common reported as disparity selective are: V1, V2, V3A, V3B, and precuneus (BA7).

One possible explanation for the variability of the results reported by different laboratories is the lack of homogeneity in the design of the stereo stimulus used. It is relatively easy to define the concept of stereo disparity, but for practical purposes, its implementation can be made in different ways. For example; the technique used to generate the RDS (anaglyph, polaroid, twin monitors, etc), or the parameters used to define the stereo pair of features (feature size, disparity size, density, etc). In order to avoid contradictory conclusions, the interpretation of results should be carefully analysed in the context of the experimental parameters used to stimulate disparity selective neurones.

In conclusion, although there may be some advances in the identification of the cortical regions involved in stereoscopic vision, the accurate understanding of the functional anatomy of stereoscopic vision is uncertain and more research is needed to investigate the neural basis of depth perception from stereoscopy. In the next section the main findings of our experimental work are discussed.

8.1.1 Summary of experimental results.

Four experiments, using red/green anaglyph to present stereo stimuli, were conducted to explore the cortical regions involved in stereoscopic vision. In the first experiment the sensitivity of the V5 region to motion and stereo disparities (as reported in non-human primates) was tested using radial motion and radial disparity (stereo cone) respectively. The V5 region was activated by motion, however no evidence was found of the sensitivity of this region to stereo disparities (Acosta-Mesa, Mayhew et al. 2001).

In accordance with previous studies, our results revealed the sensitivity of V3A to motion and stereo stimulus (Buchel, Josephs et al. 1998; Mendola, Dale et al. 1999; Sunaert, Van Hecke et al. 1999; Paradis, Cornilleau-Peres et al. 2000; Backus, Fleet et al. 2001; Braddick, O'Brien et al. 2001). A non-linear interaction between the effects of the stereo and motion stimulus was identified in the V3A region.

Interestingly, the V3B region (not commonly reported as stereo sensitive) was activated by the stereo stimuli in 2 of 8 subjects. Although the activation of V3B region was not consistent between subjects, a model of effective connectivity was proposed to explain the observed activations (see section 8.1.2). Since the scanned area was restricted to the occipital cortex, parietal regions were not included in the analysis.

A possible explanation of the inter-subject variability on the activation of V3B region can be attributed to the fact that all the subjects reported that the 'intensity' of the stereo precept declined over the presentation time, although no evidence of this was found in the data. We have no explanation for this 'adaptation'. In an attempt to prevent this occurring, in the second experiment, the disparities of the stimulus were changed over time by periodically alternating the depth of the stimulus from convergent to divergent disparities. The new stimuli consisted of two pairs of squares lying in two different planes, changing their depth smoothly from convergent to divergent disparities.

In accordance with our previous results, areas V3A and V3B were activated by the stereo stimuli. However, the activation was not consistent between the subjects. Although the subjects reported perceiving the squares moving in depth, the inter-subject variability of these results suggests that the stereo stimulus was very weak and the control conditions dominated the activations. Although the results of this experiment did not provided conclusive evidence of the stereo sensitivity of V3B, they support our hypothesis of its involvement in the processing of stereo disparities and suggest that the kind of stereo stimuli used was not optimal to stimulate the V3B region.

In order to explore the role of V3B region in the processing of stereoscopic information, the main goal of the third experiment was to introduce an attentional demand that ensured that the subjects concentrated their attention on the stereoscopic component of the stimuli.

This attentional requirement was included in the paradigm through the use of a square (defined either by luminance or by depth) moving from left to right. The subjects were instructed to perform pursuit eye movement to follow the path of the square with their eyes. Because some studies suggest the involvement of parietal regions in the processing of stereo information, the area of acquisition was expanded to cover the whole brain.

The results of this experiment revealed a consistent activation in three main regions sensitive to the stereoscopic information, V3A, V3B and precuneus. It is important to emphasise that, although for technical reasons eye movements were not measured, the preferential activation of these regions to the stereo stimuli cannot be explained either by different patterns of eye movements or by stronger attentional engagement since the control condition was designed to provide identical requirements for these parameters.

On the other hand, we do not believe that the activation of V3B can be produced by the segmentation in the three dimensional space required by the global stereo tracking task (GST), since fMRI studies in perception of object shape suggest that the activation of cortical regions (Lateral Occipital Complex) involved in the analysis of object structure is independent of the cues (luminance, colour, depth) that define the shape (Mendola, Dale et al. 1999; Kourtzi and Kanwisher 2000).

The explanation for the activation of V3B in our data (and not in other stereoscopic studies) could be due to the differences in the experimental stimulus used. Most of the stereoscopic studies have used static stereo stimulus, whereas in ours, it is a dynamic stimulus, i.e., our stereo stimuli not only requires the perception of the square defined by depth (form task), but also requires the tracking of it over time (motion task). Based on this, we hypothesised that V3B region could have being activated by the stereoscopic motion induced by the GST task, i.e., this region is sensitive to second order motion defined by spatio-temporal changes of horizontal disparities (stereoscopic -cyclopean- motion).

In relation of this hypothesis, Smith's study (Smith, Greenlee et al. 1998a) reported a cortical region sensitive to second order motion defined by contrast which matches our region both functionally and anatomically. We concluded that these regions are very similar or probably the same as that referred to as V3B. Thus, we hypothesised that the V3B region is sensitive to second order motion as defined by contrast and also by stereoscopic motion.

The objective of the last experiment was designed to explore not only the cortical regions involved in pure stereoscopic vision but also to test the sensitivity of V3B to stereo disparities. As in the GST task, this stimulus was designed to maintain the attention of the subject through the use of an active task.

A *"pacman"* shape defined either by luminance or by depth was displayed at the centre of the screen. The figure changes to one of four possible positions every second. The subjects were instructed to press a button when the figure was in a certain position. The advantages of this task over the GST are that it avoids eye movements and provides a way to assess how well the subjects are performing the task.

In accordance with other studies on stereoscopic perception, our results showed that all the subjects had consistent activations in the regions: V1, V3A, V3B and precuneus. The activation observed in the V1 region (and not observed in our previous experiments) was attributed to the small disparities used in the depth stimuli (Poggio, Gonzalez et al. 1988; Backus, Fleet et al. 2001). The activation of the V3B region supports our hypothesis of the sensitivity of the V3B region to stereo disparities, particularly in the processing of motion defined by stereopsis.

Although the stereoscopic motion component induced by the rotatory *pacman* was not the *optimal stimuli* to activate the V3B region (see section 8.3.2), it was strong enough to activate the stereoscopic motion sensitive region of all the subjects. The activation on V3B region cannot be related with eye movements because in this study the subjects were instructed to fixate on a static feature.

On the other hand, the only cue available to develop the discrimination task was the stereo disparities, thus, the correct discrimination of the shape was an indicator of depth perception. Our results supports the hypothesis of right hemisphere dominance in stereoscopic vision. However, there was no evidence in the data that suggested any relationship between gender, eye dominance, or hand dominance with the activations discovered in our experiments. To summarise, our results support the hypothesis of the engagement of V3A and precuneus in the processing of stereo disparities. Contrary to the results reported in studies with monkeys, our experiments did not reveal any evidence of the sensitivity of V5 to stereo disparities. Finally, our results showed a region which functional profile and anatomical location matched the V3B region. We hypothesised that this regions was activated by stereoscopic motion. See figure 8.1.
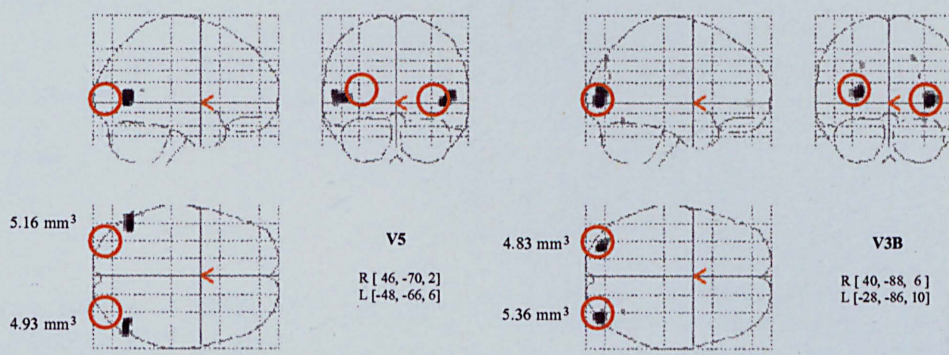
Figure 8.1 Functional defined regions. Regions of activation for motion and stereo stimuli were identified using the data of the most sensitive subject. Left panel shows the location of the motion sensitive area (V5). Right panel shows the location of the stereo sensitive visual area (V3B). Red circles helps to show the different location of each region. The distances between centres of each cluster are 28 mm. and 18 mm. for the left and right activation respectively.

The remaining question is why: when using stimulus that did not contain any stereoscopic motion at all (our first two experiments and Backus' study) some subjects showed activation in the V3B region?. One possible explanation is that it may be produced by illusory motion effects (Zeki, Watson et al. 1993), i.e, the perception of motion which is not physically present in the visual stimulus (e.g. *Enigma* picture), however more research has to be done to investigate this phenomenon.

8.1.2 Connectivity model of stereoscopic vision.

Based on neuro-anatomical studies in monkeys and in the activation observed in our first pilot experiment (chapter 4), an effective connectivity model was proposed to explain the relationships between the motion and the stereo selective regions. The model was defined assuming that V3B was engaged in the processing of pure stereoscopic disparities. The V5 and V3B regions were considered the main centres of motion and stereo information respectively. The dual sensitivity of V3A (motion and stereo) was explained by considering this region as a second stage of the processing in which an integration of the two sources of information is made (fig 8.2).
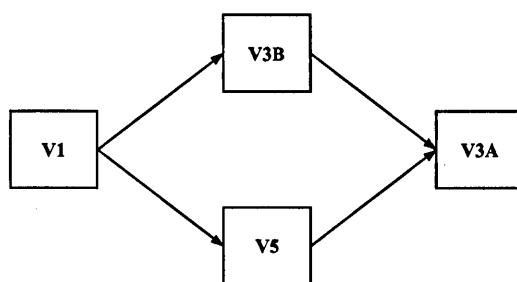
Figure 8.2 Effective connectivity model of stereoscopic vision proposed in the pilot study.

However, in the light of the results observed in the last experiments, which suggest that V3B is engaged in processing stereoscopic motion, this model has to be reinterpreted. Firstly, since the activation of V3A is correlated with the presence of stereo disparities, and not directly engaged in stereoscopic motion, the main centre of stereo processing should be V3A (Backus, Fleet et al. 2001). Because stereoscopic motion requires that disparity information be processed prior to motion information and V3B is preferentially activated by stereoscopic motion, it can be hypothesised that V3B receives inputs from V3A. That is to say, in order to detect the stereoscopic motion in V3B, it is necessary to solve the correspondence problem in V3A.

Secondly, in our last two experiments, the contrasts which compared luminance against depth and depth against luminance, did not reveal any activation in V5. This fact may have two explanations: V5 was not activated in any of the two conditions or V5 was equally activated by each one. The likelihood of the first explanation is minimal because the stimulus produced a clear perception of rotatory motion. Thus, although there is no direct evidence, we hypothesise that V5 is engaged in stereoscopic motion as much as it is in first order motion. In the light of these findings, the effective connectivity model proposed in the pilot study was modified to be consistent with our conclusions (fig. 8.3).
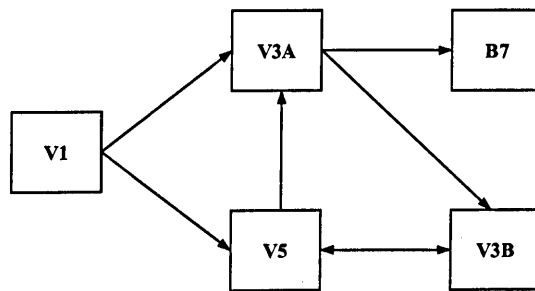
Figure 8.3 Effective connectivity model of stereoscopic vision (modified to account for stereoscopic motion cues).

This new model accounts for the hypothesis of the involvement of V3B in the processing of stereoscopic motion. It also suggest a possible interaction between the first order motion sensitive region (V5) and the second order sensitive region (V3B). The stereo sensitive region observed in the parietal region (B7) was also included in the model according to monkey's anatomical map, although its functional role in stereoscopic vision is unknown.

It is important to emphasise that this model is a hypothetical attempt to explain the relationships of the activation observed in our experiments, so research must be done to explore its validity. The next section proposes an experimental paradigm to investigate this hypothesis in future work.

8.1.3 Future work.

The results obtained in our experiments suggest that V3B is engaged in the processing of second order motion defined by stereo disparities. However, as the main objective of our studies was to investigate the cortical areas selective to stereoscopic information, no specific experiment was designed to test the hypothesis of the selectivity of the V3B region to stereoscopic motion.

Although both tasks used in the last two experiments (global stereo tracking and the form discrimination) implied in some way the processing of motion defined by spatio-temporal changes of horizontal disparities, neither of them maximises the directional component of the motion stimulus. Classical studies of motion selective regions have used radial motion to excite a big range of direction selective neurones.

A more appropriate stimuli to explore the selectivity of V3B region to stereoscopic motion, is to use concentric rings moving radially from expansion to contraction. The use of concentric rings ensures the stimulation of a larger population of directional motion sensitive neurones and facilitates fixation. On the other hand, the change from expansion to contraction prevents any habituation effect. In order to provide a mechanism to maintain the attention of the subject, one can develop a specific task (e.g. press a button) when the direction of the motion is altered. Four conditions are proposed to identify the region involved in stereoscopic motion.

a) **Stationary:** One thousand and twenty four dots (zero disparity) randomly distributed over the screen. The subject is instructed to fixate on the point located in the centre of the screen. This condition is used as a base line.

b) **Luminance:** Four concentric rings defined by luminance (half luminance with respect to the background) are displayed.

c) **Depth:** Four concentric rings defined by depth (red/green anaglyph stereogram) positioned at the front (-0.15 deg.) of the background (0.15 deg.). The rings and the background are displayed in front and behind the fixation point (zero disparity).

In order to remove possible shape cues introduced by the red/green stereoscopic pair of dots in the stereo condition, the positions of the dots are changed between frames (in both conditions) to produce the effect of dynamic random noise (Hanazawa, Kawashima et al. 2000). It is important to notice that the rings are moving in the plane X/Y, not in plane Z (motion in depth)

d) **LuminancewithDepth:** The rings are defined by both stimulus (Luminance and Depth stimulus), the purpose of this condition is to assess the effect of the interaction of first and second order motion over those regions sensitive to both stimuli.

In order to assess the activation produced for each specific condition, four main contrast are proposed:

i) **Luminance against Stationary :** The subtraction of the stationary dots from the moving rings defined by luminance should identify those regions involved in the processing of first order motion (V5 and V3A).

ii) **Depth against Stationary:** The subtraction of the stationary dots from the moving rings defined by stereo disparities should identify those regions involved in the processing of stereoscopic motion (V5, V3A ,V3B and BA7).

iii) **Depth against Luminance:** This contrast compares the activation produced by the rings defined by depth with that produced by the rings defined by luminance. According to our hypothesis of the stereoscopic motion selectivity, with this contrast we expect to identify the V3B and BA7 regions.

iv) **LuminancewithDepth against Luminance and Depth:** This contrast is intended to identify those regions which become more active when luminance and depth stimuli are presented at the same time, rather than when the individual stimuli are presented in isolation. Assuming that V5 is engaged in first order motion (luminance) and second order motion (stereoscopy), the joint stimulation of both conditions is expected to result in an increase in the response.

8.1.4 Second order motion and kinetic boundaries.

One interesting extension in the study of the V3B region is the exploration of the relationship between stereoscopic motion and kinetic boundaries detection. As reported in the literature, the V3B regions is also called Kinetic Occipital (KO) for its selectivity to illusory boundaries defined by features moving in opposite direction (Orban, Dupont et al. 1995; Dupont, De Bruyn et al. 1997; Van Oostende, Sunaert et al. 1997). These results suggest that V3B is involved in the processing of high order properties of motion stimuli. However, the specific role that this region has in the perception of kinetic boundaries is not known.

8.2 Effective connectivity modelling.

The traditional theory of the existence of functional segregated areas, highly specialised in specific cognitive tasks, has been replaced by the concept that brain functionality is the result of the interaction between several regions. In this new approach the functional profile of a cerebral region is determined by its relationship with other regions (McIntosh 2000).

Modern approaches to the study of the functional anatomy of cognitive process have been focused on the analysis of the interaction between cortical regions under different experimental conditions, also called *Effective connectivity* analysis. *Effective connectivity* is "the influence one neural system exerts over another" (Friston, Frith et al. 1993).

This concept can be thought of as the neural connectivity at synaptic level. "*Effective connectivity* should be understood as the simplest possible circuit-diagram that would replicate the observed time relationship between the recorded neurones"(Aertsen and Preissl 1991). Effective connectivity analysis has been focused on finding changes in connectivity, through the evaluation of an anatomically defined model under different experimental conditions.

The "classical" application of effective connectivity analyses have been made using Structural Equation Modelling (SEM). Although SEM represents a useful tool to assess models that define relations among variables, some of its basic assumptions are not totally compatible with fMRI data. However, it can be extended to account for non-linear dynamic systems like fMRI time series.

8.2.1 Structural equation modelling.

Although SEM is not new, the application of this technique to quantify functional relationships among neural regions from fMRI or PET neuroimaging data is a new area of research. In neuroimaging, the connections among the variables representing brain regions are anatomically constrained and the measurement of each variable can be made directly with functional imaging. These features have made SEM a useful tool for models of effective connectivity.

However, the results obtained with SEM have to be carefully interpreted because its meaning is highly dependent on the assumptions made in the analysis, for example:

i) A small number of observations (N) will over-estimates the goodness of fit of the model because the chi-square value is a function of the ML estimator times N. i.e. the same model can be a good or bad representation of the data depending on the size of the time series used in the analysis (Bullmore, Horwitz et al. 2000).

ii) The fixation of the error terms can bias the parameter estimation at the same time that it biases the goodness of fit due to the increases in the degrees of freedom.

iii) The collinearity between pre-cursors of the same node, produces empirically under-determined models, because the propagation of the error system (generally assumed by SEM modellers) does not exist in fMRI time series or is relatively much smaller than measurement noise. Thus, although the model can be theoretically identified, it could be empirically under-determined.

iv) The differences in signal to noise ratios between the data sets used in staked model analysis can be enough to conclude the difference between models. On the other hand, the classical stacked analysis *assumes* that the data sets used in the analysis have a significant goodness of fit on the model (Goncalves, Hall et al. 2001). It implies that the difference of two models can be statistically significant although the individual goodness of fits of each data set is not assessed.

This lack of standardisation in the application of SEM to perform analysis of effective connectivity makes difficult the interpretation of the results reported in neuroimaging.

Another important aspect to consider in the analysis of effective connectivity models using fMRI time series is the inherently dynamic nature of these systems. SEM has been widely applied to static systems, in fact, it was the correct technique for its first applications in effective connectivity analysis using static data (2-deoxyglucose autoradiography or PET) (McIntosh and Gonzalez-Lima 1992; McIntosh and Gonzalez-Lima 1994a). However, the application of SEM to fMRI time series requires an extension of the technique to satisfy the non-linear dynamic properties of this relatively new application. Such extension is proposed using a Non-linear Autoregressive Moving Average with eXogenous inputs algorithm (NARMAX).

8.2.2 NARMAX.

The Non-linear Auto-Regressive Moving Average with eXogenous variables (NARMAX) algorithm was introduced as a new approach to represent models of effective connectivity using fMRI time series. It can be thought of as a dynamic non-linear system identification technique which overcomes some of the limitations of Structural Equation Modelling (SEM). The main advantages of this extension can be summarised as :

i) Unsupervised identification of interaction terms. Although in SEM it is possible to introduce moderator variables to represent non-linear relationships, the polynomial expansion made by NARMAX permits the automatic identification of these hidden variables, thus, this approach can be seen as method not only for parameters estimation, but also for model identification.

ii) An appropriate treatment of noise term. The treatment of noise terms is another significant difference between SEM and NARMAX approaches. Whereas for SEM term noise is considered "system noise" (it is propagated through the model), in NARMAX the error term is considered "measurement noise" (it is uncorrelated with the regions included in the model). The definition of the error term as measurement noise is much more suitable for fMRI data, because system noise is dominated by the huge amount of measurement noise characteristic of fMRI time series.

iii) Avoidance of singularity on the covariance matrix. Under NARMAX approach, the collinearity problem in the data does not produce mathematical indetermination in the parameter estimation procedure.. The orthogonal decomposition developed by the algorithm ensures an invertible covariance matrix.

iv) Representation of dynamic systems. SEM is in principle appropriate to represent static linear systems, whereas fMRI time series are inherently non-linear dynamic systems. The capability of NARMAX to accommodate lagged inputs and autoregressive interactions allows for a better representation of the non-linear dynamic behaviour of fMRI data.

Although Volterra series are as powerful as NARMAX in representing non-linear dynamical systems, the Volterra representation is useful to model the behaviour of a *single region* (MISO), whereas NARMAX representation offers a framework to represent interactions between multiple regions (MIMO).

In summary, the NARMAX approach can be thought of as an attempt to bring SEM towards a non-linear dynamic system modelling technique which permits a more appropriate representation of effective connectivity models using fMRI time series.

8.2.3 Future work.

It is important to remember that the fMRI BOLD response is only an indicator of neural activity, however, the relationship between oxygen concentration in tissue and "neural activity" is not well understood. The misinterpretation in this relationship may have a direct implication in the assessing of effective connectivity since the evaluation of the model at BOLD level could differ from the representation at lower levels (i.e. synaptic).

Neural activity can be assessed at different levels of abstraction, for example, synaptic action potentials, blood flow, blood volume, oxyhemoglobin, etc. An interesting extension on the effective connectivity analysis can be to assess the behaviour of the model at different levels of representation. It can be made using either different neuroimaging techniques or models to infer neural activity from high levels, e.g. BOLD (Buxton, Wong et al. 1998; Zheng, Martindale et al. 2002).

The basic idea for the multi-level representation is to create independent models for BOLD (**y**) and neuronal (**x**) levels (fig. 8.4). The activity in **y** level is a function of the hidden (non-measurable) level **x** and the activation in **x** level is a function of the stimulus (u).

$$x = f_1(u, \alpha)$$

$$y = f_2(x, \gamma)$$

Then, the identification consists in solving the system to find the parameters (α) on **x** level. Although this problem has been recently addressed Friston *et al* (Friston 2002), this represent a new area of investigation and further studies are needed to explore its plausibility.
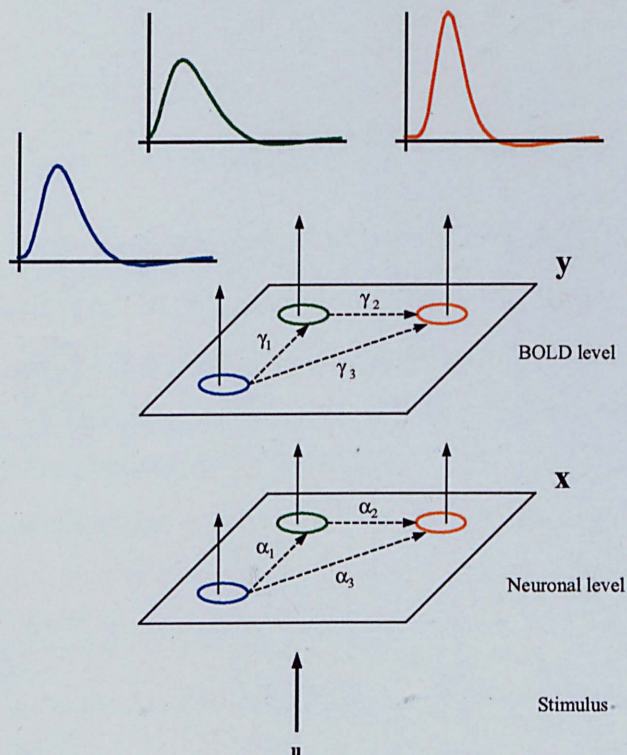
Figure 8.4 Levels of representation for effective connectivity models. The same anatomical model is shown in a different level of representation. The measurements (**y**) at BOLD level are indicators of the activations at neural level (**x**) produced by the stimulus (u). The research question is how the connectivity parameters (α and γ) behaves between levels.

Appendix

CODE

Abstract.

This appendix describes the code developed to implement two approaches for the analysis of effective connectivity models, Structural Equation Modelling (SEM) and Non-linear AutoRegressiove Moving Average with eXogenous variables (NARMAX) algorithm. The code was written in MATLAB 5.3 for Windows NT 4.0.

A.1 Neural structural equation modelling tool.

The prototype of the Structural Equation Modelling tool developed to investigate models of effective connectivity using fMRI time series was implemented following the procedure described in Bollen's book (Bollen 1989). The code was written in MATLAB 5.3 for Windows NT 4.0. The prototype takes as input the fMRI time series from a matrix structure of the form (time_points, cortical_region). There are two fitting function available, Maximum Likelihood (ML) or Unweighted Least Squares (ULS). A routine that moves the phase of the time series is provided to accommodate lags in responses (fig. A.1).
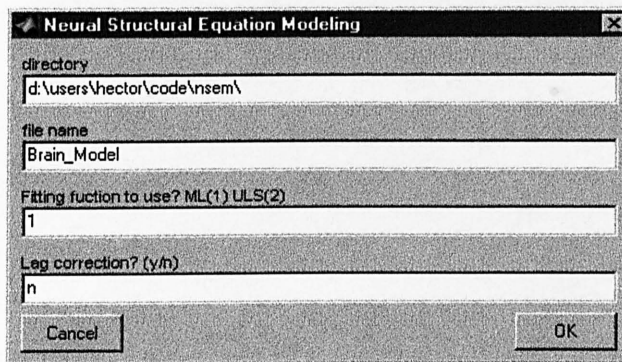
Figure A.1 SEM tool.

The connectivity structure is defined through a graphical interface which displays a statistical map of the regions of interest (fig. A.2a). The interface lets the user specify the effective connectivity pathways using the "mouse" (fig. A.2b). As output, the SEM tool displays the estimated lags, the connectivity coefficients, and the goodness of fit of the model (fig. A.2c).

The code of the different routines used to implement the neural structural equation modelling tool is listed below. For the sake of legibility, the commands engaged in the formatting of the outputs were deleted. It is important to clarify that the present implementation only works for recursive models and its extension to account for non-recursive model requires additional routines.
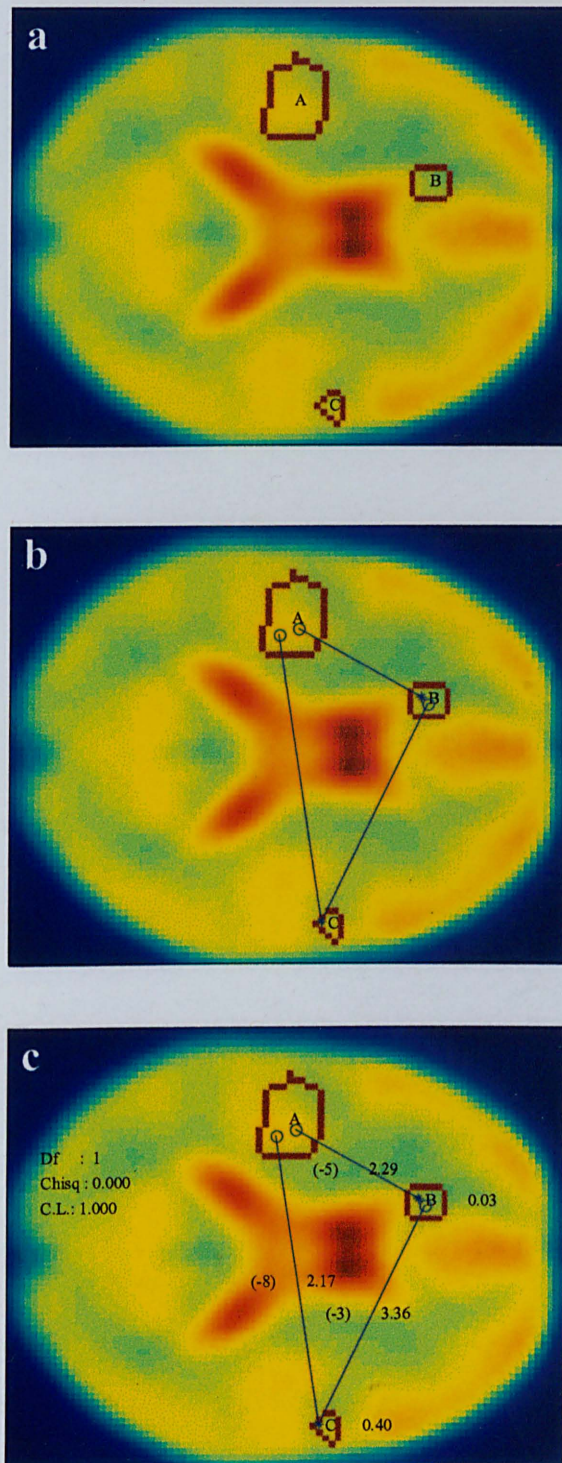
Figure A.2 SEM tool interface. a) The regions of activation are displayed in the statistical map. b) The connectivity paths are defined using the mouse, first click defines the cause (o) and the second click defines the effect (*). c) The connectivity coefficients, lags estimation, and goodness of fit (χ^2) are displayed over the statistical map.

% Main function

```
Prompt = {'directory',...
    'file name',...
    'Fitting fuction to use? ML(1) ULS(2)  ',...
    'Lag correction? (y/n)',...
    };
    Lines = 1;
    dir = [pwd '\'];
DefAns = {dir,...
    'Brain_Model',...
    '1',...
    'n',...
    };
answer = inputdlg(Prompt, 'Neural Structural Equation Modeling', Lines, DefAns);

% Reading data
dir = [answer{1} '\'];
infile = answer{2};
method = eval(answer{3});
lagc = answer{4};
load([dir infile]);

% Displaying structural data
[samples vars]=size(block);
areas2=reshape(areas,h,w);
edges=edge(areas2);
edges=reshape(edges,h*w,1);
ind=find(edges > 0);
img(ind)=max(img);
imagesc(reshape(img,h,w)),colorbar
title('Define path diagram.')
ans=[];
i=1;
hold on

% Defining connectivity structure
while isempty(ans)
  [y x]=getline;
  arcs(i,3)=floor(mean(y))+2;
  arcs(i,4)=floor(mean(x));
  title('Cause (o)        Effect (*)        Temporal violation (---)')
  x=fix(x);
  y=fix(y);
  arcs(i,1:2)=[areas2(x(1),y(1)) areas2(x(2),y(2))];
  mode='-';
  if (lag_detection(block(:,areas2(x(1),y(1))),block(:,areas2(x(2),y(2)))) > 0)
      mode=':';
  end
  line(y,x,'linestyle',mode);
  plot(y(1),x(1),'o')
  plot(y(2),x(2),'*')
  i=i+1;
  ans=input( 'Press <Return> to define another path or any letter to finish. ','s');
end
hold off
cont_exo=1;
cont_end=1;
```

```
for i=1:vars
    if ((ismember(i,arcs(:,1))) & (~(ismember(i,arcs(:,2)))))
    exogenous(cont_exo)=i;
        cont_exo=cont_exo+1;;
    else
    if ((~(ismember(i,arcs(:,1)))) & (~(ismember(i,arcs(:,2)))))
        o=0;
    else
    endogenous(cont_end)=i;
        cont_end=cont_end+1;;
    end
    end
end
guideline=0.5; % Starting values estimation (0.9) Moderate(0.4) Weak (0.2)
cont_series=1;
coef=1;
% Defining implied matrix (S)
BETA=zeros(length(endogenous),length(endogenous));
for i=1:length(endogenous)
    data(:,cont_series)=block(:,endogenous(i));
    cont_series=cont_series+1;
    ind=find(arcs(:,1) == endogenous(i));
    for j=1:length(ind)
        param(coef)=(std(block(:,arcs(ind(j),2))))/std(block(:,endogenous(i))))*guideline;
        aux(coef,1)=1;
    aux(coef,2)=find(endogenous==arcs(ind(j),2));
    aux(coef,3)=find(endogenous==endogenous(i));
    arcs(find(arcs(:,1)==endogenous(i) & arcs(:,2)==arcs(ind(j),2)),5)=coef;
    BETA(find(endogenous==arcs(ind(j),2)),find(endogenous==endogenous(i)))=coef;
    coef=coef+1;
    end
end
clear ind;
ALPHA=zeros(length(endogenous),length(exogenous));
for i=1:length(exogenous)
    ind=find(arcs(:,1) == exogenous(i));
    for j=1:length(ind)
        param(coef)=(std(block(:,arcs(ind(j),2))))/std(block(:,exogenous(i))))*guideline;
    aux(coef,1)=2;
    aux(coef,2)=find(endogenous==arcs(ind(j),2));
    aux(coef,3)=find(exogenous==exogenous(i));
    arcs(find(arcs(:,1)==exogenous(i) & arcs(:,2)==arcs(ind(j),2)),5)=coef;
        ALPHA(find(endogenous==arcs(ind(j),2)),find(exogenous==exogenous(i)))=coef;
    coef=coef+1;
    end
end
PSI=zeros(length(endogenous),length(endogenous));
for i=1:length(endogenous)
    param(coef)=var(block(:,endogenous(i)))*guideline;
    aux(coef,1)=3;
    aux(coef,2)=i;
    aux(coef,3)=i;
    noises(i,1)=endogenous(i);
    noises(i,2)=coef;
    PSI(i,i)=coef;
    coef=coef+1;
end
```

```
PHI=zeros(length(exogenous),length(exogenous));
for i=1:length(exogenous)
    data(:,cont_series)=block(:,exogenous(i));
    cont_series=cont_series+1;
    PHI(i,i)=var(block(:,exogenous(i)));
end

% Construct observerd covariance matrix (S)
if lagc=='y'
    [S,Mlags]=lag_correction(data);
else
    S=cov(data);
end

opt = optimset;
opt.MaxFunEvals = 1e10;
opt.MaxIter =1e10;
opt.TolFun=1e-5;
opt.TolX=1e-5;

% Parameter Identification
[param,f]=fminsearch('fit_function',param,[],BETA,ALPHA,PSI,PHI,S,method,aux);

% Chi square
chi_obs=abs(samples*f);
% Degrees of freedom
df=0.5*vars*(vars+1)-length(param);
% P value
pvalue=chi2(chi_obs,df);
```

% Fitting function
```
% This function is used by the MATLAB optimisation tool box to estimate the parameters

function res=model(param,BETA,ALPHA,PSI,PHI,S,method,aux)

for i=1:length(param)
   if (aux(i,1)==1)
     BETA(aux(i,2),aux(i,3))=param(i);
   else
     if (aux(i,1)==2)
       ALPHA(aux(i,2),aux(i,3))=param(i);
     else
       PSI(aux(i,2),aux(i,3))=param(i);
     end
   end
end
% Constructing implied matrix

%BETA+PSI
[rb cb]=size(BETA);
I=eye(rb,cb);
B2=inv(I-BETA);
s(1:rb,1:cb)=B2*(ALPHA*PHI*ALPHA'+PSI)*B2';
%ALPHA
[ra ca]=size(ALPHA);
s(1:ra,cb+1:cb+ca)=B2*ALPHA*PHI;
%s(rb+1:rb+ra,1:ca)=PHI*ALPHA'*B2';
s(rb+1:rb+ca,1:ra)=PHI*ALPHA'*B2';
```

```
%PHI
[rp cp]=size(PHI);
s(rb+1:rb+rp,cb+1:cb+cp)=PHI;

% Fitting functions

if (method==1) % Maximum likelihood
  p1=log(det(s));
  p2=trace(S*inv(s));
  p3=log(det(S));
  res=p1+p2-p3-(ra+ca);
else % Unweighted least Squares
  res1 = S-s;
  res2 = res1.*res1;
  res =0.5*trace(res2);
end
```

% Lag detection function
% This function verifies if a move of phase maximises the correlation between regions

```
function [l,c]=lag_detection(serie1,serie2)

serie1=serie1-mean(serie1);
serie2=serie2-mean(serie2);
len=length(serie1);

for i=1:30
  covs(i+30)=(serie1(1:len+1-i,1)'*serie2(i:len,1))/(len-1);
  covs(i)=(serie1(i:len,1)'*serie2(1:len+1-i,1))/(len-1);
end

l=find(abs(covs) == max(abs(covs)));
l=l(1,1);
c=covs(l);
l=l-1;
```

% Lag correction function
% This function moves the phase of the time series to maximise the correlation between regions

```
function [S,Mlags]=lag_correction(data)

[rows cols]=size(data);
S=zeros(cols,cols);
Mlags=zeros(cols,cols);

data=data-repmat(mean(data),rows,1);

for vars=1:cols
 for id=1:cols
  if id < vars
    for i=1:10
      covs(i+10)=(data(1:rows+1-i,vars)'*data(i:rows,id))/(rows-1);
      covs(i)=(data(i:rows,vars)'*data(1:rows+1-i,id))/(rows-1);
    end
    l=find(abs(covs) == max(abs(covs)));
    l=l(1,1);
```

```
    c=covs(l);
    l=l-1;
    if (l > 10)
       l=(l-10)*-1;
    end
    S(vars,id)=c;
    Mlags(vars,id)=l;
  end
 end
end
S=S+S'+cov(data).*eye(cols,cols);
Mlags=Mlags-Mlags';
```

% Chi-square p-value estimation
% This function estimates the p-value for a given chi-square value at a specific degrees of freedom.

```
function res = chi2(x,n)
% chi2(chi-square,degrees_o_freedom)

p=exp(-0.5*x);
if (mod(n,2) == 1)
  p=p*sqrt(2*x/pi);
end
k=n;
while(k>=2)
  p=p*x/k;
  k=k-2;
end
t=p;
a=n;
while(t>(0.000001*p))
  a=a+2;
  t=t*x/a;
  p=p+t;
end

res=1-p;
```

A.2 NARMAX algorithm.

The present NARMAX implementation was made following the algorithms described on Korenberg , Chen and Billings' papers (Korenberg, Billings et al. 1988; Billings, Chen et al. 1989; Chen, Billings et al. 1989). The algorithm described in 3.4.1 was codified in MATLAB 5.3 for Windows NT 4.0. The main inputs of this tool are the fMRI time series of the regions of interest and their connectivity structure specified by the user. The general settings for the system identification and parameter estimation (see chapter 3) are requested in the input window (fig. A.3).

Figure A.3 NARMAX tool.

The outputs of the algorithm are the connectivity coefficients, the lags between activations, and the high order interaction terms (non-linear terms). For an example of the output consult chapter 3. It is important to clarify that the present implementation of NARMAX assumes that the error terms are random noise (measurement noise), so no further analysis is made to assess correlated errors. The code of the different routines used to implement the NARMAX algorithm is listed next. For the sake of legibility, the commands engaged in the formatting of the outputs were deleted.

% Main function

```
global Colinear_Criteria Explain_Criteria AR

Prompt = {'Directory',...
    'File name',...
    'Dynamic order ',...
    'Polynomial order (1,2 or 3)',...
    'Standardized Estimates? (y/n)',...
    'Minimum explaned variance permited (%)',...
    'Angle to consider orthogonal (>) ',...
    'Autoregressive (y/n) ',...
    };
    Lines = 1;
    dir = [pwd '\'];
    DefAns = {dir,...
    'Example4',...
    '0',...
    '2',...
    'n',...
    '1',...
    '45',...
    'n',...
    };
answer = inputdlg(Prompt, NARMAX', Lines, DefAns);

dir = [answer{1} '\'];
infile = answer{2};
Max_lag = eval(answer{3});
Max_pol = eval(answer{4});
Stand_E = answer{5};
Explain_Criteria = eval(answer{6})/100;
Colinear_Criteria=eval(answer{7});
AR=answer{8};

load([dir infile]);

if Stand_E == 'y' % Normalization
  Data=Data-repmat(mean(Data),max(size(Data)),1); % Mean
  Data=Data./repmat(std(Data),max(size(Data)),1); % Std
end

vars=length(strs1);
str_pos=num2str((1:vars));

blank=cellstr('£'); %Fill predictors
for j=1:length(strs1)
    str_aux2='';
  for i=1:length(strs1)
    if ~strcmp(strs1(j),strs1(i))
      str_aux2=strcat(str_aux2,strs1(i),blank);
    end
    end
  str_aux2=char(strrep(str_aux2,'£',' '));
  DefAns(j)=cellstr(str_aux2(1:length(str_aux2)-1));
end

answer = inputdlg(strs1, 'Enter list of predictors per Variable',Lines,DefAns);
```

```
for i=1:vars
  Dv=Data(:,i);
  str_Dv=strs1(i);
  exps=char(answer(i));
if length(exps) > 0

  % In answer % Look at which position correspond in Data.
  clear keep srt_aux answer2
  con_keep=1;
  j=1;
  k=1;

  for ii=1:length(exps)
    if ~strcmp(exps(ii),' ')
      srt_aux(j)=exps(ii);
      j=j+1;
    else
      answer2(k)=cellstr(srt_aux);
      clear srt_aux;
      j=1;
      k=k+1;
    end
  end
  answer2(k)=cellstr(srt_aux);

for j=1:length(answer2)
    for jj=1:length(strs1)
      %answer2(j)
      %strs1(jj)
      if strcmp(answer2(j),strs1(jj))
        keep(con_keep)=jj;
        con_keep=con_keep+1;
      end
    end
  end
  factors=Data(:,keep);
  strs2=strs1(keep);
  ncauses=length(strs2);
  Causes=factors;
  Var_names=strs2;

else
  ncauses=0;
end

  fprintf(fid,' %s\t\t',char(str_Dv));

if ncauses > 0

    Var_names(length(Var_names)+1)=str_Dv;
    Causes(:,length(Var_names))=Dv;

    clear Parameters lab_fac Explain val_fac

    [Parameters,lab_fac,Explain,val_fac,sem]=fnarmax(Causes,Dv,Max_lag,Max_pol,Var_name
    s);
    Rv=val_fac*Parameters;
    noise=Dv-Rv;
```

```
    NoiseMat(:,i)=noise;
    Rvs(:,i)=Rv;
fprintf(fid,'Endogenous\t %3.3f\t\t',var(noise));

  else
    fprintf(fid,'Exogenous\t %3.3f\n',var(Dv));
  end
```

% NARMAX Algorithm
% Apply two main steps of the algorithm

```
function [Parameters,lab_fac,Explain,val_fac,sem]=fnarmax(factors,Dv,Max_lag,Max_pol,strs);

global Colinear_Criteria Explain_Criteria AR
% -------------------- Generating Design_Matrix
[DM labels]=fdesmat(factors,Max_lag,Max_pol,strs);
nfacts=min(size(factors)); %labels
% -------------------- Orthogonalisation Algorithm
[Parameters,goods,Explain,val_fac,sem]=fMGramSchmidt(DM,Dv,nfacts);

lab_fac=labels(goods,:);
ind=find(goods <= min(size(factors)));
```

% Design matrix generation
% This function generates predictors through the polynomial expansion

```
function [Design_Matrix,labels]=fdesmat(factors,Max_lag,Max_pol,fac_str);

global Colinear_Criteria Explain_Criteria AR

N=max(size(factors));
nf=min(size(factors));

cont=1;
if strcmp(AR,'y')
  for j=-Max_lag:-1
      Design_Matrix(:,cont)=putlag(1,factors(:,nf),j);
      str_tem=strcat(fac_str(nf),'(t-',num2str(abs(j)),')');
      lg=length(str_tem);
      str_labels(cont,1:lg)=str_tem;
    cont=cont+1;
    end
end

for i=1:nf-1
  for j=-Max_lag:0 %Max_lag
    Design_Matrix(:,cont)=putlag(1,factors(:,i),j);
    if j > 0
      str_tem=strcat(fac_str(i),'(t+',num2str(abs(j)),')');
    else
        str_tem=strcat(fac_str(i),'(t-',num2str(abs(j)),')');
    end
    lg=length(str_tem);
    str_labels(cont,1:lg)=str_tem;
    cont=cont+1;
  end
end
```

```
if Max_pol > 1
l=min(size(Design_Matrix));
gabbo=1;
for i=1:l
  for j=1:l
    if j >= i
      residual=Design_Matrix(:,i).*Design_Matrix(:,j);

      residual=residualise(Design_Matrix(:,i),Design_Matrix(:,j));

      Design_Matrix(:,cont)=residual;
      str_tem=strcat(str_labels(i,:),str_labels(j,:));
      lg=length(str_tem);
      str_labels(cont,1:lg)=str_tem;
      cont=cont+1;
    end
  end
end
end


if Max_pol > 2
l2=min(size(Design_Matrix));
for i=1:l
  for j=l+1:l2
      residual=Design_Matrix(:,i).*Design_Matrix(:,j);

      residual=residualise(Design_Matrix(:,i),Design_Matrix(:,j));

      Design_Matrix(:,cont)=residual;
      str_tem=strcat(str_labels(i,:),str_labels(j,:));
      lg=length(str_tem);
      str_labels(cont,1:lg)=str_tem;
      cont=cont+1;
  end
end
end
labels=str_labels(1:cont-1,:);
```

% Residualisation
```
% This function residualise the product of x * y

function residual=residualise(x,y)

residual=x.*y;

alpha=covar(residual,x)/var(x); % First orthogonalization
residual=residual-(alpha*x);

% Make orthogonal the second term with the previous one.
aux=y;
alpha=covar(aux,x)/var(x);

var_aux=var(aux);
Simil=((alpha*alpha)*var_aux)/var(x);

aux=aux-(alpha*x);
alpha=covar(residual,aux)/var(aux);
    residual=residual-(alpha*aux);
```

```
% Orthogonalisation algorithm
% Modified Gram-Schmidt Algorithm

function [param,Goods,Explain,val_fac,sem]=fMGramSchmidt(P,Z,nfacts)

global Colinear_Criteria Explain_Criteria
s=1;
[N M]=size(P);
Sdm=M;
A_aux=zeros(Sdm,Sdm);
goods=(1:Sdm);
PP=P;
ZZ=Z;
Cn=1;
% Select the fist component

for i=1:M % Find the most informative  component in P
   Gi(i)=covar(P(:,i),Z)/var(P(:,i));
   expl_i(i)=((Gi(i)*Gi(i))*var(P(:,i)))/var(Z);
end
ind=find(expl_i == max(expl_i));

if length(ind) > 1
   ind=ind(1,1);
   display('Two or more with equal score')
end
Aux_P=P(:,s);
P(:,s)=P(:,ind); % Most informative element
P(:,ind)=Aux_P;
G(s)=Gi(ind);

Expl(s)=expl_i(ind);

Aux_g=goods(s);
goods(s)=goods(ind);
goods(ind)=Aux_g;

A=A_aux(1:s,1:s)+eye(s,s);
param=inv(A)*G';
Goods=goods(1:s);
New_Z=PP(:,Goods)*param;
Noise(Cn)=var(ZZ-New_Z);

s=s+1;

% Look for the rest of the orthogonal components.

flag1=1;
while s <= M & flag1 == 1
   clear Gi expl_i alpha
   flag2=0;
   for k=s:M  % Substract previous orthogonal component from the rest
      alpha(k)=covar(P(:,k),P(:,s-1))/var(P(:,s-1));
      angle=acos(corr(P(:,k),P(:,s-1)))/0.0175;
         if angle > Colinear_Criteria% 0.5 % Avoid matrix become singular
             P(:,k)=P(:,k)-(alpha(k)*P(:,s-1));
             var_Pk=var(P(:,k));
             If var_Pk == 0
```

```
            display('Increase orthogonal criteria')
          expl_i(k)=-1000;
              flag2=1;
      else
              Gi(k)=covar(P(:,k),Z)/var_Pk;
              expl_i(k)=((Gi(k)*Gi(k))*var_Pk)/var(Z);
          end
    else
      expl_i(k)=-1000;
      flag2=1;
    end
  end
 ind=find(expl_i == max(expl_i));
 if length(ind) > 1
   ind=ind(1,1);
   display('Two or more with equal score')
 end


 if expl_i(ind) >= Explain_Criteria

   Expl(s)=expl_i(ind);
   Aux_P=P(:,s);
   P(:,s)=P(:,ind);
   P(:,ind)=Aux_P;
   G(s)=Gi(ind);

   AA_aux=A_aux(:,s);
   A_aux(:,s)=A_aux(:,ind);
   A_aux(:,ind)=AA_aux;

   Aux_A=alpha(s);
   alpha(s)=alpha(ind);
   alpha(ind)=Aux_A;
   A_aux(s-1,:)=A_aux(s-1,:)+alpha;
   Aux_g=goods(s);
   goods(s)=goods(ind);
   goods(ind)=Aux_g;

   Aux_exp=expl_i(s);
   expl_i(s)=expl_i(ind);
   expl_i(ind)=Aux_exp;

   if flag2 == 1
     [P,A_aux,goods,M,alpha]=del_colinear(expl_i,P,A_aux,goods,M,alpha);
   end

A=A_aux(1:s,1:s)+eye(s,s);
param=inv(A)*G';
Goods=goods(1:s);

Cn=Cn+1;

  s=s+1;
 else
     flag1=0;
 end

end
```

```
s=s-1;
A=A_aux(1:s,1:s)+eye(s,s);
param=inv(A)*G';
Goods=goods(1:s);
Explain=Expl(1:s)'*100;
sem=0;
val_fac=P(:,Goods);
```

% Delete collinear elements
% Eliminate redundant elements on the design matrix

```
function [P2,A2,GOODS2,M2,alpha2]=del_colinear(expl_i,P1,A1,GOODS1,M1,alpha1)


[nrows ncols]=size(A1);
index=1:ncols;
ind=find(expl_i ~= -1000);
ind2=find(expl_i == -1000);

P2=P1(:,ind);

A2=A1(ind,ind);

GOODS2=GOODS1(ind);

alpha2=alpha1(ind);

n=length(ind2);

M2=M1-n;
```

Bibliography

Acosta-Mesa, H., G., J. Mayhew, et al. (2001). "Functional anatomy of stereoscopic visual process." Third Mexican International Conference on Computer Science, Aguascalientes, Mexico, Sep. 15-19, 2001.

Aertsen, A. and H. Preissl (1991). "Dynamics of activity and connectivity in physiological neuronal networks." VCH Publishers Inc.: 281-302.

Aguirre, G. K., E. Zarahn, et al. (1998). "The variability of human, BOLD hemodynamic responses." Neuroimage 8(4): 360-9.

Aloimonos, J. (1988). "Shape from texture." Biol Cybern 58(5): 345-60.

Amunts, K., A. Malikovic, et al. (2000). "Brodmann's areas 17 and 18 brought into stereotaxic space-where and how variable?" Neuroimage 11(1): 66-84.

Arbuckle, J. L. and W. Wothke (1995). "AMOS 4.0 User's Guide." SmallWaters Corporation.

Backus, B. T., D. J. Fleet, et al. (2001). "Human cortical activity correlates with stereoscopic depth perception." J Neurophysiol 86(4): 2054-68.

Bakin, J. S., K. Nakayama, et al. (2000). "Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations." J Neurosci 20(21): 8188-98.

Bavelier, D., A. Tomann, et al. (2000). "Visual attention to the periphery Is enhanced in congenitally deaf Individuals." J Neurosci 20(17): RC93.

Bendat, J. S. (1990). "Nonlinear system analysis and identification from random data." John Wiley & Sons.

Bentler, P. M. (1985). "Theory and implementation of EQS, a structural equations program." BMDP Statistical software.

Berry, W. D. (1984). "Nonrecursive causal models." SAGE.

Billings, S. A., S. Chen, et al. (1989). "Identification of MIMO non-linear systems using a forward-regression orthogonal estimator." Int. J. Control. 49(6): 2157-2189.

Billings, S. A. and W. S. F. Voon (1986b). "A prediction-error and stepwise-regression estimation algorithm for non-linear systems." International journal of control 44: 803-822.

Blasdel, G. G. (1992). "Orientation selectivity, preference, and continuity in monkey striate cortex." J Neurosci 12(8): 3139-61.

Bollen, K. A. (1989). "Structural equations with latent variables." Wiley-/Interscience publication.

Braddick, O. J., J. M. O'Brien, et al. (2001). "Brain areas sensitive to coherent visual motion." Perception 30(1): 61-72.

Brainard, D. H. (1997). "The psychophysics toolbox." Spatial vision 10: 433-436.

Buchel, C. and K. J. Friston (1997). "Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI." Cereb Cortex 7(8): 768-78.

Buchel, C. and K. J. Friston (2000). "Assesing interactions among neuronal systems using functional neuroimaging." Neural Networks. 13: 871-882.

Buchel, C., O. Josephs, et al. (1998). "The functional anatomy of attention to visual motion. A functional MRI study." Brain 121(Pt 7): 1281-94.

Bullmore, E., B. Horwitz, et al. (2000). "How good is good enough in path analysis of fMRI data?" Neuroimage 11(4): 289-301.

Burkhalter, A., D. J. Felleman, et al. (1986). "Anatomical and physiological asymmetries related to visual areas V3 and VP in macaque extrastriate cortex." Vision Res 26(1): 63-80.

Burkhalter, A. and D. C. Van Essen (1986). "Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey." J Neurosci 6(8): 2327-51.

Burt, P. and J. B. (1980). "A disparity gradient limit for binocularfusion." Science 208: 615-617.

Buxton, R. B., E. C. Wong, et al. (1998). "Dynamics of blood flow and oxygenation changes during brain activation: the balloon model." Magn Reson Med 39(6): 855-64.

Carman, G. J., H. A. Drury, et al. (1995). "Computational methods for reconstructing and unfolding the cerebral cortex." Cereb Cortex 5(6): 506-17.

Chen, S., S. A. Billings, et al. (1989). "Orthogonal least squares methods and their application to non-linear system identification." Int. J. Control. 50(5): 1873-1896.

Chklovskii, D. B. (2000). "Binocular disparity can explain the orientation of ocular dominance stripes in primate primary visual area (V1)." Vision Res 40(13): 1765-73.

Chow, G. C. (1983). "Econometrics." McGraw Hills.

Coghill, R. C., C. N. Sang, et al. (1999). "Pain intensity processing within the human brain: a bilateral, distributed mechanism." J Neurophysiol 82(4): 1934-43.

Corbetta, M., E. Akbudak, et al. (1998). "A common network of functional areas for attention and eye movements." Neuron 21(4): 761-73.

Corbetta, M., F. M. Miezin, et al. (1991). "Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography." J Neurosci 11(8): 2383-402.

Cowey, A. and J. Porter (1979). "Brain damage and global stereopsis." Proc R Soc Lond B Biol Sci 204(1157): 399-407.

Cowey, A. and F. Wilkinson (1991). "The role of the corpus callosum and extra striate visual areas in stereoacuity in macaque monkeys." Neuropsychologia 29(6): 465-79.

Creem, S. H. and D. R. Proffitt (2001). "Defining the cortical visual systems: "what", "where", and "how"." Acta Psychol (Amst) 107(1-3): 43-68.

Cumming, B. G. and G. C. DeAngelis (2001). "The physiology of stereopsis." Annu Rev Neurosci 24: 203-38.

Cumming, B. G. and A. J. Parker (1994). "Binocular mechanisms for detecting motion-in-depth." Vision Res 34(4): 483-95.

Cumming, B. G. and A. J. Parker (1999). "Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity." J Neurosci 19(13): 5602-18.

Darby, D. G., A. C. Nobre, et al. (1996). "Cortical activation in the human brain during lateral saccades using EPISTAR functional magnetic resonance imaging." Neuroimage 3(1): 53-62.

DeAngelis, G. C. and W. T. Newsome (1999). "Organization of disparity-selective neurons in macaque area MT." J Neurosci 19(4): 1398-415.

DeJong, B. M., S. Shipp, et al. (1994). "The cerebral activity related to the visual perception of forward motion in depth." Brain 117(Pt 5): 1039-54.

DeValois, R., C. J. Smith, et al. (1958). "Responses of single cells in monkey lateral geniculate nucleus to monochromatic light." Science 127: 238-239.

DeYoe, E. A., G. J. Carman, et al. (1996). "Mapping striate and extrastriate visual areas in human cerebral cortex." Proc Natl Acad Sci U S A 93(6): 2382-6.

Dupont, P., B. De Bruyn, et al. (1997). "The kinetic occipital region in human visual cortex." Cereb Cortex 7(3): 283-92.

Faillenot, I., J. Decety, et al. (1999). "Human brain activity related to the perception of spatial features of objects." Neuroimage 10(2): 114-24.

Felleman, D. J., A. Burkhalter, et al. (1997). "Cortical connections of areas V3 and VP of macaque monkey extrastriate visual cortex." J Comp Neurol 379(1): 21-47.

Felleman, D. J. and D. C. Van Essen (1987). "Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex." J Neurophysiol 57(4): 889-920.

Felleman, D. J. and D. C. Van Essen (1991). "Distributed hierarchical processing in the primate cerebral cortex." Cereb Cortex 1(1): 1-47.

Fortin, A., M. Ptito, et al. (2000). "Extrastriate visual cortical areas in the processing of stereoscopic depth perception." Neuroimage (poster) 11(5).

Frackowiak, R. S. and K. J. Friston (1994). "Functional neuroanatomy of the human brain: positron emission tomography--a new neuroanatomical technique." J Anat 184(Pt 2): 211-25.

Friston, K. J. (1994). "Functional and effective connectivity in neuroimaging: A synthesis." Human Brain Mapping. 2: 56-78.

Friston, K. J. (2002). "Bayesian estimation of dynamical systems: an application to fMRI." Neuroimage 16(2): 513-30.

Friston, K. J. and C. Buchel (2000). "Attentional modulation of effective connectivity from V2 to V5/MT in humans." Proc Natl Acad Sci U S A 97(13): 7591-6.

Friston, K. J., C. D. Frith, et al. (1993). "Time-Dependent changes on effective connectivity measured with PET." Human Brain Mapping. 1: 69-79.

Friston, K. J., C. D. Frith, et al. (1995). "Characterizing dynamic brain responses with fMRI: a multivariate approach." Neuroimage 2(2): 166-72.

Friston, K. J., C. D. Frith, et al. (1993). "Functional connectivity: the principal-component analysis of large (PET) data sets." J Cereb Blood Flow Metab 13(1): 5-14.

Friston, K. J., L. Harrison, et al. (2002). "Models of effective connectivity." Welcome Department of Cognitive Neurology Internal report.

Friston, K. J., A. P. Holmes, et al. (1999). "How many subjects constitute a study?" Neuroimage 10(1): 1-5.

Friston, K. J., A. P. Holmes, et al. (1994). "Analysis of functional MRI time-series." Hum. Brain Mapp. 1: 153-171.

Friston, K. J., O. Josephs, et al. (1998). "Nonlinear event-related responses in fMRI." Magn Reson Med 39(1): 41-52.

Friston, K. J., A. Mechelli, et al. (2000). "Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics." Neuroimage 12(4): 466-77.

Friston, K. J., P. Ungerleider, et al. (1995). "Characterizing modulatory interaction between areas V1 and V2 in human cortex: A new treatment of functional MRI data." Human Brain Mapping. 2: 211-224.

Gerstein, G. L. and D. H. Perkel (1969). "Simultaneously recorded trains of action potentials: Analysis and functional interpretations." Science 164(828-830).

Ghose, G. M. and D. Y. Ts'o (1997). "Form processing modules in primate area V4." J Neurophysiol 77(4): 2191-6.

Glymour, C., P. Spirtes, et al. (1994). "TETRAD II User's manual." lawrence Earlbaum associates.

Goncalves, M. S., D. A. Hall, et al. (2001). "Can meaningful effective connectivities be obtained between auditory cortical regions?" Neuroimage 14: 1353-1360.

Gonzalez, F., F. Krause, et al. (1993a). "Binocular matching in monkey visual cortex: single cell responses to correlated and uncorrelated dynamic random dot stereograms." Neuroscience 52(4): 933-9.

Gonzalez, F. and R. Perez (1998). "Neural mechanisms underlying stereoscopic vision." Progress in neurobiology 55: 191-224.

Gonzalez, F., J. L. Relova, et al. (1993b). "Cell responses to vertical and horizontal retinal disparities in the monkey visual cortex." Neurosci Lett 160(2): 167-70.

Gossl, C., L. Fahrmeir, et al. (2001). "Bayesian modeling of the hemodynamic response function in BOLD fMRI." Neuroimage 14(1 Pt 1): 140-8.

Gulyas, B. and P. E. Roland (1993). "Binocular disparity discrimination in human cerebral cortex: Functional anatomy by positron emission tomography." Proc. Natl. Acad. Sci. USA (Nuerobiology) 91: 1239-1243.

Hanazawa, A., R. Kawashima, et al. (2000). "The human posterior parietal cortex participates in stereoscopic depth perception. A fMRI study." NeuroImage (poster) 11(5).

Hasebe, H., H. Oyamada, et al. (1999). "Human cortical areas activated in relation to vergence eye movements-a PET study." Neuroimage 10(2): 200-8.

Hasnain, M. K., P. T. Fox, et al. (1998). "Intersubject variability of functional areas in the human visual cortex." Hum Brain Mapp 6(4): 301-15.

Howard, I. P. and B. Rogers (1995). "Binocular vision and stereopsis." Oxford University press.

Howard, R. J., M. Brammer, et al. (1996). "A direct demonstration of functional specialization within motion-related visual and auditory cortex of the human brain." Curr Biol 6(8): 1015-9.

Hubel, D. H. (1995). "Eye, brain and vision." W. H. Freeman.

Hubel, D. H. and M. S. Livingstone (1987). "Segregation of form, color, and stereopsis in primate area 18." J Neurosci 7(11): 3378-415.

Hubel, D. H. and T. N. Wiesel (1970). "Stereoscopic vision in macaque monkey. Cells sensitive to binocular depth in area 18 of the macaque monkey cortex." Nature 225(227): 41-2.

Jennings, J. M., A. R. McIntosh, et al. (1998). "Mapping neural interactivity onto regional activity: an analysis of semantic processing and response mode interactions." Neuroimage 7(3): 244-54.

Jöreskog, K. G. and D. Sörbom (1989). "LISREL 7 User's reference guide." Scientific software, Inc.

Julesz, B. (1960). "Binocular depth perception of computer generated patterns." Bell system Tech. 39: 1125-1162.

Kastner, S., P. De Weerd, et al. (2000). "Texture segregation in the human visual cortex: A functional MRI study." J Neurophysiol 83(4): 2453-7.

Kenny, D. A. and M. C. Judd (1984). "Estimating the nonlinear and interactive effects of latent variables." Psychological bulletin 96(1): 201-210.

Korenberg, M., S. A. Billings, et al. (1988). "Orthogonal parameter estimation algorithm for non-linear stochastic systems." Int. J. Control. 48(1): 193-210.

Kourtzi, Z. and N. Kanwisher (2000). "Cortical regions involved in perceiving object shape." J Neurosci 20(9): 3310-8.

Kwee, I. L., Y. Fujii, et al. (1999). "Perceptual processing of stereopsis in humans: high-field (3.0-tesla) functional MRI study." Neurology 53(7): 1599-601.

Lai, S. H. and M. Fang (1999). "A novel local PCA-based method for detecting activation signals in fMRI." Magn Reson Imaging 17(6): 827-36.

Lay, D. C. (2000). "Linear algebra and its applications." Addison Wesley Longman.

Le Bihan, D., J. F. Mangin, et al. (2001). "Diffusion tensor imaging: concepts and applications." J Magn Reson Imaging 13(4): 534-46.

Leontaritis, I. J. and S. A. Billings (1985). "Input-output parametric models for non-linear systems." International journal of control 41: 303-344.

Lueck, C. J., S. Zeki, et al. (1989). "The colour centre in the cerebral cortex of man." Nature 340(6232): 386-9.

Malik, J. and R. Rosenholtz (1994). "A computational model for shape from texture." Ciba Found Symp 184: 272-83.

Marr, D. (1982). "Vision." W. H. Freeman & Company.

Marr, D. and T. Poggio (1979). "A computational theory of human stereo vision." Proc R Soc Lond B Biol Sci 204(1156): 301-28.

Mayhew, J. (1982). "The interpretation of stereo-disparity information: the computation of surface orientation and depth." Perception 11(4): 387-403.

Mayhew, J. E. and H. C. Longuet-Higgins (1982). "A computational model of binocular depth perception." Nature 297(5865): 376-8.

McIntosh, A. R. (2000). "Towards a network theory of cognition." Neural Networks. 13: 861-870.

McIntosh, A. R., R. E. Cabeza, et al. (1998). "Analysis of neural interactions explains the activation of occipital cortex by an auditory stimulus." J Neurophysiol 80(5): 2790-6.

McIntosh, A. R. and F. Gonzalez-Lima (1992). "Structural modeling of functional visual pathways mapped with 2-deoxyglucose: effects of patterned light and footshock." Brain Res 578(1-2): 75-86.

McIntosh, A. R. and F. Gonzalez-Lima (1994a). "Network interactions among limbic cortices, basal forebrain, and cerebellum differentiate a tone conditioned as a Pavlovian excitor or inhibitor: fluorodeoxyglucose mapping and covariance structural modeling." J Neurophysiol 72(4): 1717-33.

McIntosh, A. R. and F. Gonzalez-Lima (1994b). "Structural equation modelling and its application to network analysis in functional brain imaging." Human Brain Mapping. 2: 2-22.

McIntosh, A. R., C. L. Grady, et al. (1994c). "Network analysis of cortical visual pathways mapped with PET." J Neurosci 14(2): 655-66.

McKeefry, D. J., J. D. Watson, et al. (1997). "The activity in human areas V1/V2, V3, and V5 during the perception of coherent and incoherent motion." Neuroimage 5(1): 1-12.

Mendola, J. D., A. M. Dale, et al. (1999). "The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging." J Neurosci 19(19): 8560-72.

Mishkin, M., L. G. Ungerleider, et al. (1983). "Object vision and spatial vision: Two cortical pathways." Trends Neuroscience 6: 414-417.

Nishida, Y., O. Hayashi, et al. (2001). "Stereopsis-processing regions in the human parieto-occipital cortex." Neuroreport 12(10): 2259-63.

Nobre, A. C., D. R. Gitelman, et al. (2000). "Covert visual spatial orienting and saccades: overlapping neural systems." Neuroimage 11(3): 210-6.

Nyberg, L., A. R. McIntosh, et al. (1996). "Network analysis of positron emission tomography regional cerebral blood flow data: ensemble inhibition during episodic memory retrieval." J Neurosci 16(11): 3753-9.

Ogle, K. N. (1962). "The optical space sense." The eye (New York: Academic Press) 4.

Ono, M., S. Kubick, et al. (1990). "Atlas of the cerebral sulci." Thieme medical, New York.

Orban, G. A., P. Dupont, et al. (1995). "A motion area in human visual cortex." Proc Natl Acad Sci U S A 92(4): 993-7.

Orban, G. A., S. Sunaert, et al. (1999). "Human cortical regions involved in extracting depth from motion." Neuron 24(4): 929-40.

Paradis, A. L., V. Cornilleau-Peres, et al. (2000). "Visual perception of motion and 3-D structure from motion: an fMRI study." Cereb Cortex 10(8): 772-83.

Patterson, R., M. Donnelly, et al. (1997). "Speed discrimination of stereoscopic (cyclopean) motion." Vision Res 37(7): 871-8.

Pearl, J. (2000). "Causality." Cambridge university press.

Pelli, D. G. (1997). "The videotoolbox software for visual psychophysics: Transforming numbers into movies." Spatial vision 10: 437-442.

Petit, L. and J. V. Haxby (1999). "Functional anatomy of pursuit eye movements in humans as revealed by fMRI." J Neurophysiol 82(1): 463-71.

Podzebenko, K., G. F. Egan, et al. (2002). "Widespread dorsal stream activation during a parametric mental rotation task, revealed with functional magnetic resonance imaging." Neuroimage 15(3): 547-58.

Poggio, G. F. and B. Fischer (1977). "Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey." J Neurophysiol 40(6): 1392-405.

Poggio, G. F., F. Gonzalez, et al. (1988). "Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity." J Neurosci 8(12): 4531-50.

Poggio, T. and C. Koch (1987). "Synapses that compute motion." Sci Am 256(5): 46-52.

Ptito, A., R. J. Zatorre, et al. (1993). "Localization and lateralization of stereoscopic processing in the human brain." Neuroreport 4(10): 1155-8.

Raichle, M. E., A. M. MacLeod, et al. (2001). "A default mode of brain function." Proc Natl Acad Sci U S A 98(2): 676-82.

Ramachandran, V. S. (1988). "Perception of shape from shading." Nature 331(6152): 163-6.

Rockland and Kaas (1997). "Crebral cortex." Plenum, New York. 12: 694.

Roy, J. P. and R. H. Wurtz (1990). "The role of disparity-sensitive cortical neurons in signalling the direction of self-motion." Nature 348(6297): 160-2.

Saito, H., M. Yukie, et al. (1986). "Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey." J Neurosci 6(1): 145-57.

Salzman, C. D., C. M. Murasugi, et al. (1992). "Microstimulation in visual area MT: effects on direction discrimination performance." J Neurosci 12(6): 2331-55.

Sandell, J. H. and P. H. Schiller (1982). "Effect of cooling area 18 on striate cortex cells in the squirrel monkey." J Neurophysiol 48(1): 38-48.

Schiller, P. H. (1993). "The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey." Vis Neurosci 10(4): 717-46.

Schroeder, C. E., C. E. Tenke, et al. (1990). "Binocularity in the lateral geniculate nucleus of the alert macaque." Brain Res 521(1-2): 303-10.

Schumacker, R. E. and G. A. Marcoulides (1998). "Interaction and nonlinear effects in structural equation modelling." LEA.

Seghier, M., M. Dojat, et al. (2000). "Moving illusory contours activate primary visual cortex: an fMRI study." Cereb Cortex 10(7): 663-70.

Seyama, J. and T. Sato (1998). "Shape from shading: estimation of reflectance map." Vision Res 38(23): 3805-15.

Sharma, S. (1996). "Applied multivariate techniques." John Wiley & Sons: 447-49.

Smith, A. T., M. W. Greenlee, et al. (1998a). "The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI)." J Neurosci 18(10): 3816-30.

Smith, A. T. and N. E. Scott-Samuel (1998b). "Stereoscopic and contrast-defined motion in human vision." Proc R Soc Lond B Biol Sci 265(1405): 1573-81.

Sunaert, S., P. Van Hecke, et al. (1999). "Motion-responsive regions of the human brain." Exp Brain Res 127(4): 355-70.

Tabachnik, B. G. and L. S. Fidell (2001). "Using multivariate statistics." Allyn and Bacon Fourth Edition: 764-771.

Takayama, Y. and M. Sugishita (1994). "Astereopsis induced by repetitive magnetic stimulation of occipital cortex." J Neurol 241(9): 522-5.

Talairach, J. and T. P. (1988). "Co-planar stereotaxic atlas of human brain." Thiem medical publishers, New York.

Tanaka, K., H. Saito, et al. (1991). "Coding visual images of objects in the inferotemporal cortex of the macaque monkey." J Neurophysiol 66(1): 170-89.

Taylor, J. G., N. Schmitz, et al. (2000). "The network of brain areas involved in the motion aftereffect." Neuroimage 11(4): 257-70.

Tittle, J. S., J. F. Norman, et al. (1998). "The perception of scale-dependent and scale-independent surface structure from binocular disparity, texture, and shading." Perception 27(2): 147-66.

Tootell, R. B., N. K. Hadjikhani, et al. (1998). "Functional analysis of primary visual cortex (V1) in humans." Proc Natl Acad Sci U S A 95(3): 811-7.

Tootell, R. B., J. D. Mendola, et al. (1997). "Functional analysis of V3A and related areas in human visual cortex." J Neurosci 17(18): 7060-78.

Tyler, C. W. (1973). "Stereoscopic Vision: cortical limitations and disparity scaling effect." Science 181: 276-278.

Uusitalo, M. A., V. Virsu, et al. (1997). "Activation of human V5 complex and rolandic regions in association with moving visual stimuli." Neuroimage 5(4 Pt 1): 241-50.

Vaina, L. M. (1989). "Selective impairment of visual motion interpretation following lesions of the right occipito-parietal area in humans." Biol Cybern 61(5): 347-59.

Vaina, L. M., M. Lemay, et al. (1990). "Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: a case study." Vis Neurosci 5(4): 353-69.

Van Essen, D. C. and H. A. Drury (1997). "Structural and functional analyses of human cerebral cortex using a surface-based atlas." J Neurosci 17(18): 7079-102.

Van Essen, D. C., L. J.W., et al. (2001). "Mapping visual cortex in monkeys and humans using surface-based atlases." Vision reseach.

Van Oostende, S., S. Sunaert, et al. (1997). "The kinetic occipital (KO) region in man: an fMRI study." Cereb Cortex 7(7): 690-701.

Volterra, V. (1959). "Theory of functionals and integral and integro-differenntial equations." Dove, New York.

Watson, J. D., R. Myers, et al. (1993). "Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging." Cereb Cortex 3(2): 79-94.

WDCN (1997). "SPM Course." Welcome Department of Cognitive Neurology, UCL http://www.fil.ion.ucl.ac.uk/spm/.

WDCN (1999). "SPM99 Software." Welcome Department of Cognitive Neurology, UCL http://www.fil.ion.ucl.ac.uk/spm/.

Wright, S. (1918). "On the nature of size factors." Genetics. 3: 367-374.

Wright, S. (1921). "Correlation and causation." Agricultural Research, 20: 557-585.

Wright, S. (1934). "The method of path analysis." Annals of Mathematical Statistics. 5: 161-215.

Xiao, D. K., V. L. Marcar, et al. (1997). "Selectivity of macaque MT/V5 neurons for surface orientation in depth specified by motion." Eur J Neurosci 9(5): 956-64.

Zeki, S. (1993). "A vision of the brain." Blackwell Scientific Publications.

Zeki, S. and S. Shipp (1988). "The functional logic of cortical connections." Nature 335(6188): 311-7.

Zeki, S., J. D. Watson, et al. (1993). "Going beyond the information given: the relation of illusory visual motion to brain activity." Proc R Soc Lond B Biol Sci 252(1335): 215-22.

Zeki, S., J. D. Watson, et al. (1991). "A direct demonstration of functional specialization in human visual cortex." J Neurosci 11(3): 641-9.

Zeki, S. M. (1974). "Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey." J Physiol 236(3): 549-73.

Zeki, S. M. (1978). "Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex." J Physiol 277: 273-90.

Zheng, Y., J. Martindale, et al. (2002). "A model of the hemodynamic response and oxygen delivery to brain." Submited for publication.