

# Neural correlates of monocular and binocular depth cues based on natural images: A LORETA analysis

Florian Ph.S. Fischmeister\*, Herbert Bauer

*Department of Clinical Psychology, Biological Psychology and Differential Psychology, Faculty of Psychology,  
University of Vienna, Liebiggasse 5, A-1010 Vienna, Austria*

Received 1 December 2005; received in revised form 3 March 2006

## Abstract

Functional imaging studies investigating perception of depth rely solely on one type of depth cue based on non-natural stimulus material. To overcome these limitations and to provide a more realistic and complete set of depth cues natural stereoscopic images were used in this study. Using slow cortical potentials and source localization we aimed to identify the neural correlates of monocular and binocular depth cues. This study confirms and extends functional imaging studies, showing that natural images provide a good, reliable, and more realistic alternative to artificial stimuli, and demonstrates the possibility to separate the processing of different depth cues.  
© 2006 Elsevier Ltd. All rights reserved.

*Keywords:* Natural images; Slow cortical potentials (SCP); Depth perception; Stereopsis; Monocular and binocular depth cues

## 1. Introduction

Sense of depth and the associated perception of a three-dimensional (3D) world is something natural to us. For this seemingly simple task our visual system has to re-construct this 3D-world from its two-dimensional (2D) representations on the two retinas. Pointers or “depth cues” for the reconstruction of depth are embedded in the 2D retinal images. Wheatstone (1938) was the first to demonstrate that binocular disparity, based on the interocular difference between the left and the right eye, can elicit perception of depth. Monocular depth cues also elicit depth perception based on the input of one eye, e.g., textures, shading, geometrical perspective, or occlusion (Reading, 1983).

Functional imaging techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) suggest the involvement of occipital and parietal cortices in 3D-vision (see, e.g., Gulyás, Roland, Heywood, Poppellwell, & Cowey, 1994; Nishida et al., 2001;

Naganuma et al., 2005). Within these cortices particularly caudal parietal regions respond to different levels of disparity (Rutschmann & Greenlee, 2004), 3D-geometrical features (Nishida et al., 2001; Naganuma et al., 2005), or surface orientation (Shikata et al., 2001). Activations in the temporal cortex near the occipito-temporal junction were reported while perceiving depth from spatial features or shadings (Faillenot, Decety, & Jeannerod, 1999; Taira, Nose, Inoue, & Tsutsui, 2001) and while extracting depth from motion (Iwami et al., 2002). In addition, frontal regions get activated when perceiving depth information (Gulyás et al., 1994; Taira et al., 2001).

However, the above-mentioned studies have several limitations. At first, most of these studies use only one type of depth cue, either binocular disparity or monocular cues. Binocular disparity alone, although producing impressive changes in 3D-perception, cannot explain certain depth phenomena, e.g., what is located at certain positions (Todd, 2004). Monocular cues on the other side are not equally effective but help to understand the perception of 3D-shape and perceptual grouping. They are therefore necessary to complete the reconstruction of depth from our 2D-retinal images (Todd, 2004; Anderson, 1998). Thus, using only one

\* Corresponding author. Fax: +43 0 1 4277 47859.

E-mail address: [florian.fischmeister@univie.ac.at](mailto:florian.fischmeister@univie.ac.at) (F.P.S. Fischmeister).

URL: <http://brl.psy.univie.ac.at> (F.P.S. Fischmeister).

set of depth cues does not provide a complete picture of the neural processing of depth.

A further limitation comes from the visual stimulus material. For instance, random dot stereograms (RDS) (Julesz, 1960), based on spatial differences between corresponding dots, provide an effective way to generate 3D-shapes that are solely defined by binocular disparity. These shapes do not contain any monocular depth cues and can therefore only be seen after fusion of the two retinal images. Yet the fusion of these RDS images has to be learned (Ramachandran, 1976). Using evoked-potentials Skrandies and Jedynak (1999) showed that this perceptual learning influences the spatial distribution of activated neuronal assemblies in the parietal–occipital cortex. Monocular depth perception is constrained by other limitations like view-point dependency (cf. Inui et al., 2000, investigating depth perception of Necker cubes), or attentional biases (Meng & Tong, 2004). Finally, most studies so far applied only artificial stimulus material, e.g., random dot stereograms, textures, or shadings, to investigate the cortical mechanisms underlying the perception of depth.

To overcome some of these limitations it is necessary to use visual material containing both types of depth pointers: monocular and binocular cues. Such a combination of cues can be found in real world viewing conditions. Therefore, in contrast to previous studies, natural stereoscopic images are used to achieve a realistic and complete set of depth cues to investigate the perception of depth as a whole. To further emphasize real world viewing conditions all images were presented by means of a head mounted display presenting to each eye its own visual stimulus.

This study pursues to first answer the question which cortical areas get activated while perceiving natural stereoscopic images. Second, we aimed to separate the processing of monocular and binocular depth cues and attempted to connect these aspects of depth perception to their cortical counterparts. Thus, extending previous work using functional imaging techniques, this is the first study aiming to determine the neural correlates of monocular and binocular cues based on natural stereoscopic images.

## 2. Materials and methods

### 2.1. Subjects

Thirty male subjects aged between 18 and 22 years participated in this study. All subjects were right-handed as tested with the Annett handedness inventory (Annett, 1985) and had normal or corrected-to-normal vision. None of the participants reported any difficulties perceiving depth. All subjects gave their written informed consent to participate in this study.

### 2.2. Visual stimuli and presentation

The visual stimuli used consisted of three different scenes (a tree, an arcade, and a building) and were presented in three different modes: stereoscopic, monoscopic, and scrambled. To obtain these images, three digital cameras of the same model (JVC GR-DV J70) were mounted next to each other on a tripod. Stereoscopic images were taken with the two outer cameras; camera-lenses were spaced 8 cm apart. The third camera, placed

exactly in between, was used to take images serving as monoscopic stimuli. This procedure ensures that all scenes are photographed with identical height and distance to the object, level of detail, and focus of the camera. Moreover, it ensures identical photographic quality since resolution, sharpness, and exposure of all three cameras were identical. Using the image of the centre camera a scrambled version of each scene was generated by randomly displacing single pixels of the image. The resulting scrambled image is physically equal to the original monoscopic image but has no meaningful resolvable content.

The images of the different scenes generated this way were presented with a head mounted display (HMD: Sony LDI-D100B 3D SVGA). Via this HMD it was possible to present to each eye a different view of the same scene via a small SVGA-LCD device. For the stereoscopic presentation each eye received one of the photographs taken with the two outer cameras (the left eye from the left side and vice versa) to induce a solid 3D-effect. For the monoscopic and the scrambled presentations, both eyes received exactly the same input.

### 2.3. Experimental conditions

All trials were presented via the HMD and consisted of the three scenes that randomly switched between scrambled, monoscopic, and stereoscopic presentation mode at randomly varied intervals of 8–16 s (mean 12 s). One block consisted of 75 such changes of mode of one scene (trials) and was repeated once. All together 450 mode changes, e.g., from monoscopic to stereoscopic, in 6 blocks, two per scene, were presented. Subjects were instructed to follow these constant changes actively and attentively. To preserve natural viewing conditions, no instruction to fixate a certain point was given and thus no crosshair was presented within the images. Furthermore they were not required to give an overt response. In functional mapping of specific visual cortical areas such a passive viewing procedure has been shown to be as effective as procedures requiring the subjects to actively judge certain stimulus features, e.g., by pressing a button (Kanwisher, Chun, McDermott, & Ledden, 1996).

To determine which cortical areas get activated while perceiving the different presentation modes, the following conditions were extracted: “switch from scrambled to monoscopic” (*ScMo*), “switch from scrambled to stereoscopic” (*ScSt*), “switch from monoscopic to stereoscopic” (*MoSt*), and “switch from stereoscopic to monoscopic” (*StMo*) presentation.

The major advantage of this design is that it allows to target neural responses related to specific aspects of visual perception. Since, certain parts of the visual information are constant across switching presentation modes, any neuronal activity observed after the switch can be attributed to new or missing attributes of the visual image. Therefore, it can be expected that any neuronal activity observed in the *ScMo* condition can be related to monoscopic depth perception and general perception of the image content, since brightness, colour, and other physical properties are not altered. *ScMo* and *ScSt* are distinguishable only by the binocular disparity cues included in the latter condition. In condition *MoSt* only those cortical areas should be active that are sensitive for binocular disparity since monocular cues, scene, and physical properties do not vary across that mode-change. Finally, *StMo* should be inverse to *MoSt*, and only activities not related to disparity should be observable.

### 2.4. Data acquisition

EEG was recorded using a multi-channel DC-amplifier with high baseline stability and an input impedance of 200 G $\Omega$ . Signals were sampled with 125 Hz within a frequency range from DC to 35 Hz. To allow the elimination of artefacts related to EOG and blinks, vertical eye movements, using electrodes above and below the right eye, and horizontal eye movements, using electrodes on the left and right outer canthi, were recorded simultaneously. Electrodes were mounted on small plastic adaptors that were individually fixed to the subjects' scalp and skin using collodium or adhesive tape, respectively. To minimise skin potential artefacts and to ascertain an electrode impedance that is homogeneous, stable, and below 1 k $\Omega$ , the skin was slightly scratched at all recording sites using a

sterile single-use needle. Electrodes and adaptors were previously filled with degassed electrode gel (Electro-Gel, Electrode-Cap International, Inc., Eaton/OH, USA) (Bauer, 1998).

EEG was recorded from 22 locations evenly distributed across the scalp surface against a true non-cephalic, sterno-vertebral reference (Stephenson & Gibbs, 1951). Individual 3D coordinates of these electrodes referenced to nasion, inion, and the two preauricular points of all subjects were measured using a photogrammetric head digitizer (PHD: Bauer et al., 2000). Mean electrode coordinates were close to the coordinates of the 10–20 electrode system by Jasper (1958).

### 2.5. Analyses of EEG

Prior to analyses, eye movement and blink related artefacts were eliminated offline using a linear regression approach with channel specific correction parameters. EOG parameters were determined separately for vertical and horizontal eye movements in an EOG calibration trial. Blink coefficients were calculated using a template matching procedure (see Vitouch, Bauer, Gittler, Leodolter, & Leodolter, 1997; Lamm, Fischmeister, & Bauer, 2005, for a detailed description). Using this parameters weighted eye movement and blink related signals were subtracted from each EEG channel trial by trial. Afterwards all EOG and blink corrected trials were visually inspected to exclude artefact contaminated trials and trials with insufficient EOG or blink correction.

Artefact-free single trials were averaged per condition and subject using the mean amplitude in the interval of 500 ms preceding the presentation switch as baseline. Such a long baseline ensures that only stable activations are subtracted and therefore, changes induced by the presentational switch can be optimally assessed. Furthermore the grand mean for each condition across all subjects was calculated. Subsequently, the mean amplitudes over 50 ms epochs in the range of 1250 ms after stimulus onset (switch of presentation modes) to 1750 ms were determined for all single-subject averages and the grand means. During this time range settled and to a large extent stable slow cortical potentials evoked only by the presented images can be expected.

To validate the inverse solution algorithm used in this study, the P1 component of the visual evoked potential was used. The P1 component is a positive going waveform with a peak latency of around 100 and 130 ms whose generators can be reliably localized to the lateral extrastriate cortex (Brodmann area 18) (Di Russo, Martinez, Sereno, Pitzalis, & Hillyard, 2001, 2005). Base on visual inspection of the grand mean waveforms and topographies of all conditions per subject, the P1 was defined as the maximum within the occipital electrodes O3, Oz, and O4 in the interval of 80–130 ms after mode-change.

The extracted ten amplitude values plus the P1 component were transformed by means of low resolution electromagnetic tomography (LORETA: Pascual-Marqui, Michel, & Lehmann, 1994; Pascual-Marqui, 1999) using individual electrode coordinates acquired via PHD and cross-registered to the standard Talairach atlas (Talairach & Tournoux, 1988). According to Pascual-Marqui et al. (1994) current density distributions found by LORETA are the smoothest of all possible solutions that optimally describes the observed surface potentials. The LORETA implementation for this study uses a three-shell spherical head model registered to a standardized stereotactic space (Talairach & Tournoux, 1988). The solution space is restricted to cortical grey matter and hippocampus, defined via a reference brain from the Montreal Neurological Institute (MNI), and is subdivided into 2394 voxels with a spatial resolution of  $7 \times 7 \times 7$  mm. For the transformation the smoothest possible inverse solution, where the regularisation parameter is zero, was used.

Like other distributed source imaging techniques (for a review, see e.g., Michel et al., 2004), LORETA does not require any assumptions about the number, localization, configuration, or extent of neuronal sources. Although there is an ongoing debate about the constraints used by this algorithm, a growing number of studies support the usefulness (e.g., Park et al., 2002; Esslen, Pascual-Marqui, Hell, Kochi, & Lehmann, 2004; Lamm et al., 2005) and validity of LORETA for the localization of surface recorded potentials (e.g., Vitacco, Brandeis, Pascual-Marqui, & Martin, 2002; Mulert et al., 2004). These studies in particular (Mulert et al., 2004),

comparing LORETA with fMRI results, show that LORETA is capable to correctly localize single and multiple active brain sources obtained from small electrode-settings.

For a descriptive analysis of the different viewing conditions, grand mean LORETA averages were calculated and thresholded at  $3 \times 10^{-3} \mu\text{V}/\text{mm}^2$ . Coordinates, anatomical structures, and Brodmann areas (BA) for cortical regions activated above this threshold throughout the whole time range of 1250–1750 ms after stimulus onset were determined using the Talairach atlas implemented in the LORETA software.

### 2.6. Statistical analyses

To target aspects of the perception of depth the following pairwise contrasts were calculated: Contrasting “*StMo* versus *MoSt*” should reveal those neuronal areas that get activated by binocular disparity only. This can be expected since the only difference between these two condition is the presence (*MoSt*), respectively, absence (*StMo*) of disparity information. Differences between the conditions *ScSt* and *MoSt* should yield cortical areas responsive to monoscopic depth cues and general image perception. These two aspects are present since binocular disparity processing is common to both conditions. Finally, activities induced by monoscopic perception compared to stereoscopic perception will be assessed by contrasting conditions *ScMo* and *ScSt*.

All statistical analyses were performed using the approach of Statistical non-parametric Mapping (Holmes, Blair, Watson, & Ford, 1996, SnPM; Nichols & Holmes, 2002) implemented in the LORETA software. Differences between conditions were assessed by calculating dependent-sample *t*-values for every voxel using log-transformed, but not subject-wise normalised—attenuation of possible differences in the amplitudes should be avoided—LORETA-values. The resulting  $T_{\max}$  statistic was based on 5000 permutations—randomly drawn condition configurations tested against the original configuration. The significance threshold was set at  $p \leq 0.05$ ; no correction for multiple comparisons is needed since this is already inherently accounted for (Nichols & Holmes, 2002, see). Coordinates, anatomical structures, and Brodmann areas for all significant activity differences exceeding this threshold will be reported.

## 3. Results

According to post-experimental questionnaire all subjects were motivated and reported a vivid perception of depth during the stereoscopic conditions. Differences between the monoscopic and the stereoscopic conditions were easily recognisable for all participants. Some subjects reported initial problems perceiving a 3D-image that they could overcome with a single blink immediately after the mode-change. These initial problems are presumably only related to the presentational switching from monoscopic to stereoscopic.

The EEG data were generally of good quality and no subject had to be excluded from the study. In the grand mean (i.e., average across all conditions and subjects), the maximum current density value for the P1 could be localized to the Cuneus (BA 18:  $X=4$ ,  $Y=-74$ ,  $Z=8$ ). Evaluating the localization of the P1 component separately for all single conditions revealed consistent current density maxima within the striate area (BA 17) and in directly adjoining visual areas (BA's 18, 19, and 7).

Descriptive analysis of the different viewing conditions revealed large and stable activity in temporal and parieto-occipital regions. Current density maxima were found bilaterally within the Cuneus (BA 7, 31, and 18) and the Precuneus (BA 7 and 31) for all conditions across all time-

frames. Maxima within temporal cortices were localized in middle and inferior temporal regions (BA 37 and 39), again within all conditions. Additional smaller activity was located in frontal regions—middle and superior frontal gyri—when switching from the monoscopic to the stereoscopic presentation condition and vice versa. Tables 1 and 2 give an overview of the current density maxima found within the different switching conditions.

Direct comparison of *StMo* to *MoSt* revealed generally higher cortical activities for the *MoSt* condition at the parietal–occipital border (cf. Table 2). Statistically significant activity differences ( $p < 0.05$ , mean  $t$ -threshold = 3.8807) were found in the left Brodmann area 7 (Precuneus) and the

Table 1  
Current density activity maxima separated for the switching from scrambled to monoscopic (*ScMo*) and stereoscopic (*ScSt*) presentation

Anatomical area	Talairach coordinates	Current density
<i>ScMo</i>		
Precuneus (left BA 31)	$X = -3, Y = -74, Z = 29$	5.849
Cuneus (right BA 18)	$X = 4, Y = -74, Z = 15$	5.757
Middle temporal gyrus (right BA 37)	$X = 46, Y = -67, Z = 8$	4.793
Fusiform gyrus (right BA 37)	$X = 46, Y = -60, Z = -13$	4.702
Middle temporal gyrus (left BA 39)	$X = -45, Y = -67, Z = 15$	4.679
Precuneus (left BA 7)	$X = -3, Y = -53, Z = 57$	4.243
<i>ScSt</i>		
Precuneus (left BA 31)	$X = -3, Y = -74, Z = 29$	7.316
Cuneus (right BA 18)	$X = 4, Y = -74, Z = 15$	7.258
Fusiform gyrus (right BA 37)	$X = 46, Y = -60, Z = -13$	5.422
Middle temporal gyrus (right BA 46)	$X = 46, Y = -74, Z = 15$	5.393
Middle occipital gyrus (right BA 37)	$X = 46, Y = -67, Z = 1$	5.393
Middle temporal gyrus (right BA 37)	$X = 46, Y = -67, Z = 8$	5.393
Precuneus (left BA 7)	$X = -3, Y = -53, Z = 57$	5.279
Middle temporal gyrus (left BA 39)	$X = -52, Y = -67, Z = 8$	5.078
Anterior cingulate (left BA 10)	$X = -3, Y = 52, Z = 1$	4.045
Middle temporal gyrus (left BA 22)	$X = -59, Y = -39, Z = 1$	3.701
Middle occipital gyrus (left BA 19)	$X = -31, Y = -88, Z = 8$	3.672
Middle temporal gyrus (right BA 21)	$X = 60, Y = -32, Z = -6$	3.643
Middle temporal gyrus (right BA 22)	$X = 60, Y = -39, Z = 1$	3.012
Uncus (right BA 20)	$X = 39, Y = -11, Z = -34$	3.012
Medial frontal gyrus (left BA 6)	$X = -3, Y = -11, Z = 64$	3.012

Note that only cortical regions activated throughout the whole time range—1250–1750 ms after stimulus onset—exceeding a current density value of  $3 \times 10^{-3}$  are listed.

Note: First column: activated brain area, in parentheses hemisphere and Brodmann areas are given, second column: Talairach coordinates of the maximal activated voxel within this structure in mm for  $X$ -,  $Y$ -, and  $Z$ -axes together with its current density value in  $\mu V/mm^2 \times 10^{-3}$  (third column).

Table 2

Current density activity maxima when switching from monoscopic to stereoscopic (*MoSt*) and vice versa (*StMo*) exceeding a current density value of  $3 \times 10^{-3}$  throughout the whole time range—1250–1750 ms after stimulus onset—are listed

Anatomical area	Talairach coordinates	Current density
<i>MoSt</i>		
Cuneus (right BA 7)	$X = 4, Y = -67, Z = 29$	5.315
Precuneus (left BA 7)	$X = -3, Y = -53, Z = 57$	4.669
Anterior cingulate (left BA 10)	$X = -3, Y = 52, Z = 1$	3.189
Fusiform gyrus (right BA 37)	$X = 46, Y = -60, Z = -13$	3.168
Middle temporal gyrus (left BA 39)	$X = -45, Y = -67, Z = 15$	3.064
Fusiform gyrus (left BA 37)	$X = -45, Y = -60, Z = -13$	3.064
Inferior temporal gyrus (left BA 37)	$X = -45, Y = -67, Z = -6$	3.064
Middle temporal gyrus (left BA 39)	$X = -45, Y = -74, Z = 15$	3.064
Middle temporal gyrus (right BA 21)	$X = 60, Y = -32, Z = -6$	3.001
<i>StMo</i>		
Fusiform gyrus (right BA 37)	$X = 46, Y = -60, Z = -13$	3.973
Cuneus (right BA 18)	$X = 4, Y = -74, Z = 15$	3.568
Middle temporal gyrus (left BA 39)	$X = -52, Y = -67, Z = 8$	3.303
Middle temporal gyrus (left BA 39)	$X = -45, Y = -67, Z = 15$	3.303
Precuneus (left BA 7)	$X = -3, Y = -53, Z = 57$	3.069

Note: See Table 1 for an explanation of the different columns.

Cuneus (BA 19). A further somewhat smaller cluster was found in the right Brodmann area 7 (see Table 3 and Fig. 1A for details).

Examining *ScSt* and *MoSt* large current densities were found in the Precuneus (BA 7 and 31), extending up to the superior parietal lobe and to a smaller extent into the Cuneus (BA 17 and 18). Additional areas of activity were localized at the border between the middle temporal and the occipital lobe of both hemispheres (BA 39 and 37), extending along the middle and inferior temporal (BA 37, 20, and 21) and fusiform gyrus in the scrambled to stereo condition (*ScSt*) (see Tables 1 and 2). Statistically significant differences ( $p < 0.05$ , mean  $t$ -threshold = 3.8050) between these two conditions were found in the middle temporal gyrus; maxima in the right BA 37 and the left BA 39. In the right hemisphere this significant region was larger and extended into the fusiform gyrus of the occipital cortex and the limbic lobe. Two additional significant regions of activity were identified in the inferior parietal lobe (Precuneus, BA 19) and middle frontal cortex (BA 10) exclusively within the left hemisphere. Detailed results are listed in Table 4 and presented in Fig. 1B.

Comparisons between monoscopic and stereoscopic perception were done by contrasting the conditions *ScMo* and *ScSt*. In both conditions brain areas of nearly identi-

**Table 3**  
 Statistical significant differences ( $p < 0.05$ , mean  $t$ -threshold = 3.8807) showing higher current density activity when switching from monoscopic to stereoscopic presentation ( $StMo < MoSt$ )

Anatomical area	Talairach coordinates	Max. $t$ -value
Precuneus (left BA 7)	$X = -10, Y = -81, Z = 43$	4.11
Precuneus (left BA 19)	$X = -10, Y = -81, Z = 36$	4.04
Precuneus (right BA 7)	$X = 4, Y = -67, Z = 36$	4.01

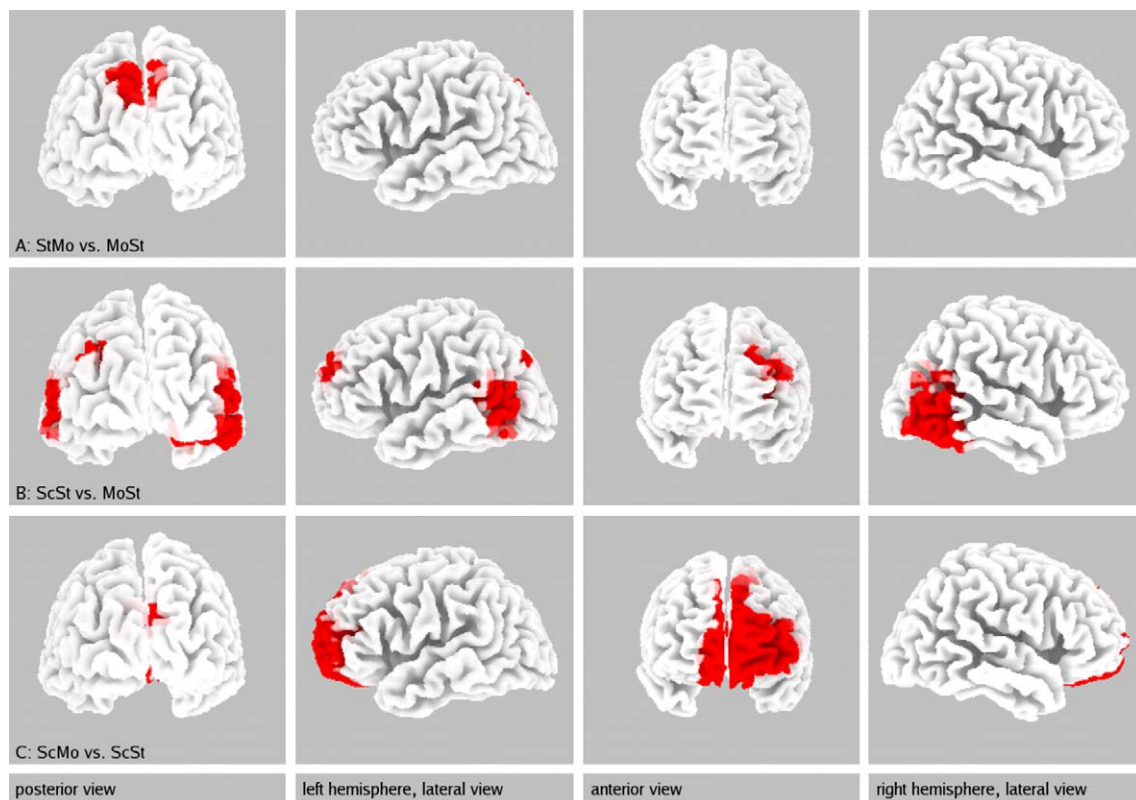
*Note:* First column: activated brain area, in parentheses Brodmann areas and hemisphere are given, second column: Talairach coordinates of the maximal activated voxel within this structure in mm for  $X$ -,  $Y$ -, and  $Z$ -axes, third column  $t$ -value for this voxel. All brain areas are determined via the Talairach demon implemented in LORETA.

cal activation were found in parietal (BA's 7, 31, 37, and 39) and occipital areas (BA's 18 and 19). Descriptive comparisons of the frontal cortex revealed higher current density values when switching from scrambled to stereoscopic presentation  $ScSt$  (see Table 1). Statistically significant differences ( $p < 0.05$ , mean  $t$ -threshold = 3.7601) were found for the superior frontal gyrus (BA 10 and 11) of both hemispheres. On the left hemisphere this cluster extended into the medial frontal gyrus (BA 10) and the anterior cingulate cortex (BA 24 and 32). Another significant cluster, although rather small, was detected within the parietal–occipital junction (BA 31 and 18) (see Table 5 and Fig. 1C).

#### 4. Discussion

The aim of the present study was to assess the neural correlates of monocular and binocular cues in natural stereoscopic images. Natural monoscopic and stereoscopic pictures were used to receive a picture of depth perception as complete as possible. This is in contrast to previous studies that used either random dot stereograms or visual stimuli based on shading, perspective, or surface orientation. The images used were presented via a head mounted display presenting to each eye either its own slightly different view of the same scene (stereocondition), or the same input (monoscopic and scrambled presentation). Thus it was possible to induce a realistic perception of depth as reported by all subjects. Due to the randomised presentation of the different viewing modes and the randomisation of the presentation duration artefacts caused by expectation or fatigue could be excluded. Out of this random sequence of presentational changes we extracted those conditions containing only certain aspects of depth perception.

Results of the inverse solution calculated with LORETA provided generally valid results. Maximum current density values for the P1 were correctly localized in cortical areas involved in visual perception. Tomographic analyses of the whole sample revealed widespread cortical networks involved in visual perception. Regardless of the presenta-



**Fig. 1.** “Statistical nonParametric Maps” illustrating significant differences between the three comparisons— $MoSt$  vs.  $StMo$  (top row),  $ScSt$  vs.  $MoSt$  (middle), and  $ScMo$  vs.  $ScSt$  (bottom row)—corresponding to 1500 ms after stimulus onset. Significant different voxels exceeding the critical thresholds at  $p < 0.05$  are in red color. Thresholds and cortical regions are listed in Tables 3–5. Note that each single aspect of depth perception is located at distinct cortical regions.

Table 4  
Significant differences in current density activity for the comparison of *ScSt* and *MoSt* ( $p < 0.05$ , mean  $t$ -threshold = 3.8050)

Anatomical area	Talairach coordinates	Max. $t$ -value
Middle temporal gyrus (right BA 37)	$X = 60, Y = -60, Z = -6$	5.98
Middle temporal gyrus (left BA 39)	$X = -59, Y = -60, Z = 8$	5.04
Fusiform gyrus (right BA 37)	$X = 32, Y = -46, Z = -6$	4.75
Limbic lobe (right BA 30)	$X = 18, Y = -39, Z = -6$	4.71
Precuneus (left BA 19)	$X = -31, Y = -81, Z = 36$	4.09
Superior frontal gyrus (left BA 10)	$X = -31, Y = 52, Z = 29$	4.01
Middle frontal gyrus (left BA 10)	$X = -24, Y = 45, Z = 22$	4.00

Listed cortical regions are higher activated during the *ScSt* condition.

Note: See Table 3 for an explanation of the different columns.

Table 5  
Significant differences for the comparison *ScMo* and *ScSt* ( $p < 0.05$ , mean  $t$ -threshold = 3.7601)

Anatomical area	Talairach coordinates	Max. $t$ -value
Superior frontal gyrus (left BA 10)	$X = -10, Y = 66, Z = 1$	6.97
Medial frontal gyrus (left BA 10)	$X = -10, Y = 52, Z = -6$	6.86
Superior frontal gyrus (left BA 11)	$X = -17, Y = 66, Z = -13$	6.86
Medial frontal gyrus (right BA 11)	$X = 11, Y = 45, Z = -6$	6.70
Anterior cingulate cortex (left BA 32)	$X = -10, Y = 31, Z = 22$	5.55
Anterior cingulate cortex (left BA 24)	$X = -3, Y = 31, Z = 22$	5.55
Cuneus (left BA 18)	$X = -10, Y = -74, Z = 22$	4.23
Precuneus (right BA 31)	$X = 4, Y = -74, Z = 22$	4.18
Precuneus (right BA 31)	$X = 4, Y = -67, Z = 22$	4.17

Higher current density activities were found when switching from scrambled to stereoscopic presentation *ScSt*.

Note: See Table 3 for an explanation of the different columns.

tional mode clusters of activity within striate areas extending into the parietal and the temporal cortex were found. This pattern of results is in line with previous studies associating activities in ventral stream areas (temporal cortex) with “what” information, and dorsal stream areas with “where” information (parietal lobe) (Ungerleider & Mishkin, 1982).

In agreement with functional imaging studies in humans (e.g., Gulyás et al., 1994; Iwami et al., 2002; Naganuma et al., 2005) we observed significantly higher current density values in the parietal cortex (BA 7) extending into occipital regions (BA 19) while processing binocular disparity cues. Additional support for this result can be derived from studies on the sensitivity of dorsal areas of the occipito-parietal cortex to different levels of disparity (Backus, Fleet, Parker, & Heeger, 2001; Rutschmann & Greenlee, 2004). Thus, as already pointed out by (Gulyás et al., 1994) and (Taira et al., 2001), these cortical areas play an important role in the integration of binocular disparity cues.

Nishida et al. (2001); Fortin, Ptito, Faubert, and Ptito (2002), and others reported a dominance of the right hemisphere while processing binocular disparity. In contrast to these studies no systematic asymmetric cortical activation could be observed. This discrepancy might be due to the different stimulus material or the design used in this study.

Neuronal activity related to the processing of monocular cues were found bilaterally in the temporal lobe (BA 37 and 39). In the right hemisphere activations were generally higher and extended into occipital regions (fusiform gyrus). These regions, in particular ventral and lateral occipito-temporal regions, are reported to be involved in representing the shape of objects and in perceiving objects in recognition tasks (Grill-Spector, 2003). This view is also put forward by Blonder et al. (2004) investigating neuronal responses to a variety of stimuli (faces and houses) in the fusiform gyrus. The results, particularly activations in more posterior regions of BA 37, may putatively overlap with regions reported to play a role in the detection of stereo-defined borders (Mendola, Dale, Fischl, Liu, & Tootell, 1999). Right hemisphere activations reported here are also close to brain areas involved in the recognition of 3D-features (Faillenot et al., 1999; Naganuma et al., 2005) and depth from shading (Taira et al., 2001) or motion (Iwami et al., 2002). From these findings we deduce, therefore, that occipito-temporal regions are important for the neural computation of depth based on edges and shapes and the spatial orientation and configuration of objects independent of binocular disparity.

Like Gulyás et al. (1994) this study revealed significant bilateral activations in the medial and superior gyri of the frontal cortex. Evidence from other functional imaging studies showed either no difference (Paradis et al., 2000) or no coverage of the frontal cortex. Although eye movements may influence frontal activity, they cannot serve as explanation for these activations. First, keeping this possibility in mind, we employed a very strict correction of eye movement related artefacts using a regression approach and subsequent visual inspection of all epoch trails. Second, no activations in the frontal eye fields were found in any of the statistical comparisons. Support against this explanation also comes from imaging studies. Investigating vergent eye movements mainly occipito-parietal activity and no differences within the frontal eye fields were reported (Hasebe et al., 1999). Furthermore, eye movements in response to disparity without depth perception are elicited at ultrashort latencies within 90–160 ms after stimulus presentation (Masson, Busetini, & Miles, 1997). However, the time interval chosen for the analyses in the present study started 1.25 s after the presentational switch. Thus, although we concede that eye movements might play a role here, we think that they cannot serve as an explanation.

One possible interpretation for the observed frontal activity during the perception of depth can be deduced from studies that delve into intercortical connections between frontal and parietal cortical areas. Lesion studies together with various fMRI and PET studies have shown

the importance of the frontal cortex for the encoding, memory and retrieval of spatial layouts (Kessels, Postma, Wijnalda, & de Haan, 2000). Frontal activations in this study were higher when perceiving stereoscopic compared to monoscopic images. Therefore it might be possible that stereoscopic presentation, being a more realistic view of the scene, induced processes of object localization or spatial search. Furthermore, stereoscopic information apparently provides new additional information which is processed within frontal regions. This view is inspired by Frey and Petrides (2000) indicating that orbitofrontal regions are involved in the process of encoding of new information. Future research will have to address the precise role of the frontal lobe for depth perception.

In summary, the results of our study—obtained with slow cortical potentials and LORETA—are in accordance with previous functional imaging studies of depth perception. Therefore, natural stereoscopic images provide a good, reliable, and more realistic alternative to stimuli based solely on one type of depth cues. Furthermore, it has been shown that the presented study approach allows to separate different aspects of depth perception provided by natural images. Therefore, to achieve a more complete picture of depth perception we recommend further studies using images that contain a variety of different binocular and monocular depth cues.

## Acknowledgments

This work was supported by the Austrian Federal Ministry of Defense. Development of the Photogrammetric Head Digitizer used in this study was funded by the Austrian Science Fund (FWF), P12289.

## References

- Anderson, B. L. (1998). Stereovision: beyond disparity computations. *Trends in Cognitive Sciences*, 2(6), 214–222.
- Annett, M. (1985). *Left, right, hand and brain: the right shift theory*. London: Erlbaum.
- Backus, B., Fleet, D., Parker, A., & Heeger, D. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology*, 86(4), 2054–2068.
- Bauer, H. (1998). Slow potential topography. *Behavior Research Methods, Instruments, & Computers*, 30(1), 20–33.
- Bauer, H., Lamm, C., Holzreiter, S., Hollander, I., Leodolter, U., & Leodolter, M. (2000). Measurement of 3D electrode coordinates by means of a 3D photogrammetric head digitizer. *NeuroImage*, 11(5, Suppl. 1), 461.
- Blonder, L. X., Smith, C. D., Davis, C., Kesler/West, M. L., Garrity, T. F., Avison, M. J., & Andersen, A. H. (2004). Regional brain response to faces of humans and dogs. *Cognitive Brain Research*, 20, 384–394.
- Di Russo, F., Matinez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2001). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, 15, 95–111.
- Di Russo, F., Pitzalis, S., Spitoni, G., Aprile, T., Patria, F., Spinelli, D., & Hillyard, S. A. (2005). Identification of the neural sources of the pattern-reversal VEP. *NeuroImage*, 24, 874–886.
- Esslen, M., Pascual-Marqui, R. D., Hell, D., Kochi, K., & Lehmann, D. (2004). Brain areas and time course of emotional processing. *NeuroImage*, 21(4), 1189–1203.
- Faillenot, I., Decety, J., & Jeannerod, M. (1999). Human brain activity related to the perception of spatial features of objects. *NeuroImage*, 10(2), 114–124.
- Fortin, A., Ptito, A., Faubert, J., & Ptito, M. (2002). Cortical areas mediating stereopsis in the human brain: a PET study. *Neuroreport*, 13(6), 895–898.
- Frey, S., & Petrides, M. (2000). Orbitofrontal cortex: a key prefrontal region for encoding information. *Proceedings of the National Academy Sciences of the United States of America*, 97(15), 8723–8727.
- Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, 13, 159–166.
- Gulyás, B., Roland, P., Heywood, C., Popplewell, D., & Cowey, A. (1994). Visual form discrimination from luminance or disparity cues: functional anatomy by PET. *Neuroreport*, 5(17), 2367–2371.
- Hasebe, H., Oyamada, H., Kinomura, S., Kawashima, R., Ouchi, Y., Nobezawa, S., Tsukada, H., Yoshikawa, E., Ukai, K., Takada, R., Takagi, M., Abe, H., Fukuda, H., & Bando, T. (1999). Human cortical areas activated in relation to vergence eye movements—a PET study. *NeuroImage*, 10(2), 200–208.
- Holmes, A., Blair, R., Watson, J., & Ford, I. (1996). Nonparametric analysis of statistic images from functional mapping experiments. *Journal of Cerebral Blood Flow and Metabolism*, 16(1), 7–22.
- Inui, T., Tanaka, S., Okada, T., Nishizawa, S., Katayama, M., & Konishi, J. (2000). Neural substrates for depth perception of the Necker cube; a functional magnetic resonance imaging study in human subjects. *Neuroscience Letters*, 282(2), 145–148.
- Iwami, T., Nishida, Y., Hayashi, O., Kimura, M., Sakai, M., Kani, K., Ito, R., Shiino, A., & Suzuki, M. (2002). Common neural processing regions for dynamic and static stereopsis in human parieto-occipital cortices. *Neuroscience Letters*, 327(1), 29–32.
- Jasper, H. H. (1958). The ten twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 10, 371–375.
- Julesz, B. (1960). Binocular depth perception of computer generated patterns. *Bell System Technical Journal*, 39, 1125–1162.
- Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Cognitive Brain Research*, 5, 55–67.
- Kessels, R., Postma, A., Wijnalda, E., & de Haan, E. (2000). Frontal-lobe involvement in spatial memory: evidence from PET, fMRI and lesion studies. *Neuropsychology Review*, 10(2), 101–113.
- Lamm, C., Fischmeister, F., & Bauer, H. (2005). Individual differences in brain activity during visuo-spatial processing assessed by slow cortical potentials and low resolution electromagnetic tomography. *Brain Research: Cognitive Brain Research*.
- Masson, G., Busetini, C., & Miles, F. A. (1997). Vergence eye movements in response to binocular disparity without depth perception. *Nature*, 389, 283–286.
- Mendola, J.-D., Dale, A., Fischl, B., Liu, A. K., & Tootell, R. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, 19, 8560–8572.
- Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, 4, 539–551.
- Michel, C., Murray, M., Lantz, G., Gonzalez, S., Spinelli, L., & de Peralta, R. G. (2004). EEG source imaging. *Clinical Neurophysiology*, 115, 2195–2222.
- Mulert, C., Jager, L., Schmitt, R., Bussfeld, P., Pogarell, O., Müller, H.-J., Juckel, G., & Hegerl, U. (2004). Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. *NeuroImage*, 22, 83–94.
- Naganuma, T., Nose, I., Inoue, K., Takemoto, A., Katsuyama, N., & Taira, M. (2005). Information processing of geometrical features of a surface based on binocular disparity cues: an fMRI study. *Neuroscience Research*, 51(2), 147–155.
- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping*, 15(1), 1–25.

- Nishida, Y., Hayashi, O., Iwami, T., Kimura, M., Kani, K., Ito, R., Shiino, A., & Suzuki, M. (2001). Stereopsis-processing regions in the human parieto-occipital cortex. *Neuroreport*, *12*(10), 2259–2263.
- Paradis, A., CornilleuaPeres, V., Droulez, J., Van de Moortele, P., Lobel, E., Berthoz, A., Le Bihan, D., & Poline, J. (2000). Visual perception of motion and 3-D structure from motion: an fMRI study. *Cerebral Cortex*, *10*, 772–783.
- Park, H., Kwon, J., Youn, T., Pae, J. S., Kim, J. J., Kim, M., & Ha, K. (2002). Statistical parametric mapping of LORETA using high density EEG and individual MRI: application to mismatch negativities in schizophrenia. *Human Brain Mapping*, *17*, 168–178.
- Pascual-Marqui, R., Michel, C., & Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, *18*(1), 49–65.
- Pascual-Marqui, R. D. (1999). Review of methods for solving the EEG inverse problem. *International Journal of Bioelectromagnetism*, *1*(1), 75–86.
- Ramachandran, V. (1976). Learning-like phenomena in stereopsis. *Nature*, *262*, 382–384.
- Reading, R. W. (1983). *Binocular vision: Foundations and applications*. Butterworth: Boston.
- Rutschmann, R. M., & Greenlee, M. W. (2004). BOLD response in dorsal areas varies with relative disparity level. *Neuroreport*, *15*(4), 615–619.
- Shikata, E., Hamzei, F., Glauche, V., Knab, R., Dettmers, C., Weiller, C., et al. (2001). Surface orientation discrimination activates caudal and anterior intraparietal sulcus in humans: an Event-related fMRI study. *Journal of Neurophysiology*, *85*, 1309–1314.
- Skrandies, W., & Jedyak, A. (1999). Learning to see 3-D: psychophysics and brain electrical activity. *Neuroreport*, *10*(2), 249–253.
- Stephenson, W., & Gibbs, F. (1951). A balanced non-cephalic electrode reference. *Electroencephalography and Clinical Neurophysiology*, *3*, 237–240.
- Taira, M., Nose, I., Inoue, K., & Tsutsui, K.-i. (2001). Cortical areas related to attention to 3D surface structures based on shading: an fMRI study. *NeuroImage*, *14*(5), 959–966.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Stuttgart: Thieme.
- Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences*, *8*(3), 115–121.
- Ungerleider, L. G. and Mishkin, M. 1982. Two cortical visual systems. In: D. J. Ingle, M. A. Goodale, and R. J. W., Mansfield (Eds), Analysis of visusal behaviour. Cambridge: MIT Press.
- Vitacco, D., Brandeis, D., Pascual-Marqui, R., & Martin, E. (2002). Correspondence of event-related potential tomography and functional magnetic resonance imaging during language processing. *Human Brain Mapping*, *17*, 4–12.
- Vitouch, O., Bauer, H., Gittler, G., Leodolter, M., & Leodolter, U. (1997). Cortical activity of good and poor spatial test performers during spatial and verbal processing studied with slow potential topography. *International Journal of Psychophysiology*, *27*, 183–199.
- Wheatstone, C. (1938). Some remarkable phenomena of vision. I. *Philosophical Transactions of the Royal Society*, *13*, 371–395.