

Cortical Substrates of Perceptual Stability during Eye Movements

Peter Thier,^{*,1} Thomas Haarmeier,^{*,†} Subhojit Chakraborty,^{*,2} Axel Lindner,^{*} and Alexander Tikhonov^{*}

^{*}Department of Cognitive Neurology, University of Tübingen, Germany; and [†]Department of General Neurology, University of Tübingen, Germany

Received March 7, 2001

We are usually unaware of retinal image motion resulting from our own movement. For instance, during slow-tracking eye movements, the world around us remains perceptually stable despite the retinal image slip induced by the eye movement. This example of perceptual invariance is achieved by subtracting an internal reference signal, reflecting the eye movement, from the retinal motion signal. If the two cancel each other, visual structures, which do not move, will also be perceived as nonmoving. If, however, the reference signal is too small or too large, a false eye-movement-induced motion of the external world will be perceived. We have exploited our ability to manipulate the size of the reference signal in an attempt to reveal the structures in visual cortex, encoding the perception of self-induced visual motion rather than the retinal motion signal. Using EEG and lately also MEG recordings in human subjects and single-unit recordings in monkeys, we have been able to show that our ability to perceive the world as stationary despite eye-movement-induced retinal image slip is based on “late” parts of the cortical hierarchy of motion processing, sparing the early stages up to cortical area MT and, among others, involving cortex at the junction between the parietal and temporal lobes close to the parieto-insular-vestibular cortex. Lesions of this network in humans render the visual system unable to compensate for the visual consequences of eye movements, giving rise to severe dizziness, whenever the eyes move smoothly. © 2001 Academic Press

Retinal image motion may result from motion of objects in the external world or, alternatively, from movement of our eyes relative to the external world. Of course, as long as we are moving around, most of the time, both sources of image slip will be available, pro-

viding the formidable challenge to the visual system to distinguish the component of retinal image motion resulting from motion in the external world from the one arising from ego motion. The reason is that only the former should be perceived as visual motion, but not the latter. If this were not the case, our concept of a reassuringly stable world, through which we move, in which we act, would inevitably be lost. Smooth-pursuit eye movements may serve as a case in point. They allow us to stabilize the image of a selected object on or close to the fovea in order to make use of the advantages of foveal vision (Haarmeier and Thier, 1999). The inevitable consequence is that the images of all other objects (the “visual background”) will, if stationary in the world, move at a speed corresponding to the speed of the pursuit eye movement carried out, however, without being perceived as moving.

Building on early suggestions by Hermann von Helmholtz (1910), the *inferential theory of motion perception* holds that our ability to distinguish between external stimuli and self-induced sensory stimuli is achieved by subtracting an internal reference signal, termed the *Willensanstrengung* (the effort of will) by von Helmholtz and the *Efferenzkopie* (the efference copy) by von Holst and Mittelstaedt (1950) from the sensory signal (Fig. 1A). Specifically, perceptual stability in the case of smooth-pursuit eye movements is accomplished by subtracting a copy of the eye movement motor command from the retinal motion signal (Wertheim, 1994). If the two cancel each other, visual structures, which do not move, will be perceived as nonmoving, even though their images move on the retina as a consequence of the eye movement. Our ability to compensate for the visual consequences of eye movements is a typical example of an invariance operation, releasing perception from the influence of confounding variables. Other examples of perceptual invariance such as color constancy, i.e., our ability to perceive a constant object color despite changes in the wavelength composition of the light, have been shown to have a cerebrocortical basis. It was therefore close at hand to consider the possibility that also eye movement invariance might have a cerebrocortical basis.

¹To whom correspondence and reprint requests should be addressed at Neurologische Universitätsklinik, Abt. Kognitive Neurologie, Hoppe-Seyler-Strasse 3, 72076 Tübingen, Germany. Fax: +49 7071 29 5326. E-mail: thier@uni-tuebingen.de.

²Present address: Department of Psychology, University College London, England.

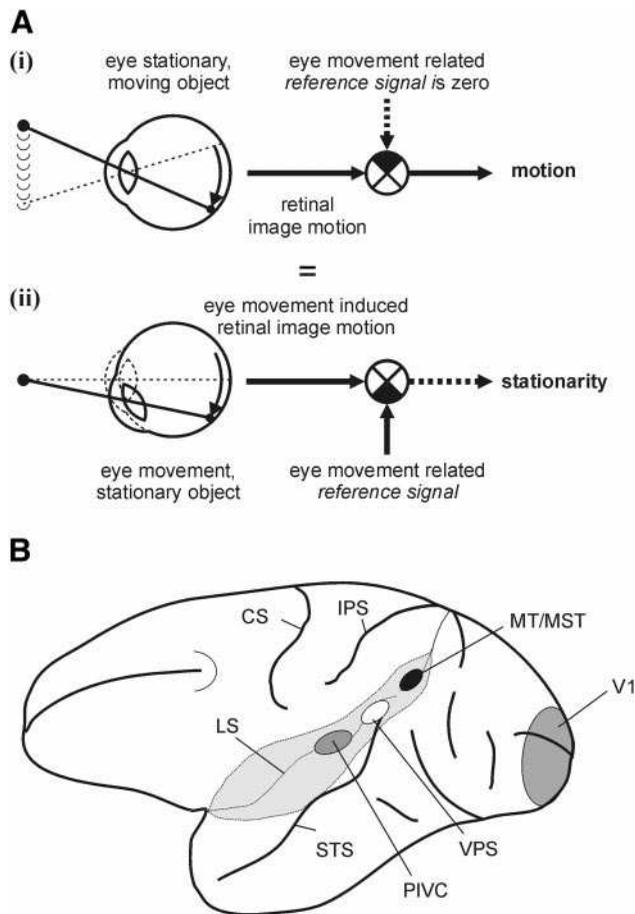


FIG. 1. (A) Cartoon of the inferential mechanism underlying vision during eye movements, capturing the essentials of the von Holst and Mittelstaedt (1950) reafference principle. (i) The eyes are stationary and the object moves in the external world, resulting in externally induced retinal image motion. In this situation the eye movement related reference signal is zero and the output of the element subtracting the reference signal from the retinal signal corresponds to the retinal image motion input. The interpretation of this result is object motion in the world. (ii) The eyes move while the object is stationary. In this case retinal image motion is exclusively due to the eye movement. This *reafference* in the terminology of von Holst and Mittelstaedt is cancelled by the nonzero reference signal. The interpretation of the zero output of the comparator is stationarity of the object. (B) Schematic view of the left hemisphere of the rhesus monkey with the sylvian (lateral) sulcus and parts of the superior temporal sulcus unfolded (gray, surrounded by dashed lines) to show the visual areas explored in the experiments of Erickson and Thier (1991), Ilg and Thier (1996), and Chakraborty and Thier (2000) as well as PIVC (parietoinsular vestibular cortex), a vestibular representation in the lateral sulcus discovered by Grüsser and co-workers (see Guldin and Grüsser, 1998). LS, lateral sulcus; CS, central sulcus; STS, superior temporal sulcus; IPS, intraparietal sulcus.

Our early work on the visual cortex of nonhuman primates (Erickson and Thier, 1991; Ilg and Thier, 1996) had indicated that eye movement invariance must be an achievement of later stages of the visual hierarchy of cerebrocortical areas involved in the processing of visual motion (Fig. 1B). This was suggested

by the fact that neurons in area 17, the primary visual cortex, and those in area MT, the prototypical motion processing area in the superior temporal sulcus (STS) of the monkey brain, typically discharged invariably in response to smooth, slow retinal image slip, irrespective of its source being object movement in the world or, alternatively, a smooth eye movement across a stationary object. On the other hand, a significant number of neurons in posterior parietal area MST on the anterior bank of the STS seemed to differentiate between the two forms of retinal image stimulation, often responding selectively or at least much more strongly to externally induced image slip compared to self-induced image slip. Very similar conclusions were suggested by a later study of visually evoked potentials in humans (Haarmeier and Thier, 1998), whose presentation requires a preceding discussion of the psychophysical measurement of motion perception during eye movements and several key findings on human motion perception based on these techniques.

The basic task in all these later experiments was to carry out smooth-pursuit eye movements in one horizontal direction, usually to the right, across a large, visual background (Fig. 2A). This background likewise moved along the horizontal, at a velocity, whose size and direction was varied. At the end of each trial, subjects were required to indicate whether they had seen background movement to the right or to the left. Of course, this decision is easy, if the background moves at high speed to the left or to the right. On the other hand, as exemplified in Fig. 2B, if the speed is reduced, subjects' decisions become increasingly inconsistent, the more, the more the velocity approaches the velocity of subjective stationarity (VSS). At this VSS, titrated by varying background velocity according to a staircase procedure, subjects guess, indicated by 50% left and 50% right choices. They guess because the background has become *subjectively* stationary. Healthy subjects have a VSS which does not deviate much from physical stationarity (Fig. 2C, dot symbols). In other words, they perceive the background as being stationary, when it is physically stationary despite the fact that the eye movement shifts the background image across the retina. According to the inferential theory of perception, at this VSS the reference signal encoding the eye movement is equal to the afferent signal, reflecting the velocity of the retinal image motion (Mack and Herman, 1973; Wertheim, 1994). Correspondingly, a VSS of zero reflects an ideal reference signal, having a size which allows it to fully compensate for the effects of the eye movement.

Surprisingly, our early psychophysical studies also showed that a very simple variation of the basic psychophysical paradigm allowed us to evoke predictable changes of the percept of eye-movement-induced visual motion (Haarmeier and Thier, 1996, 1998). According to the inferential theory, an illusory movement of the

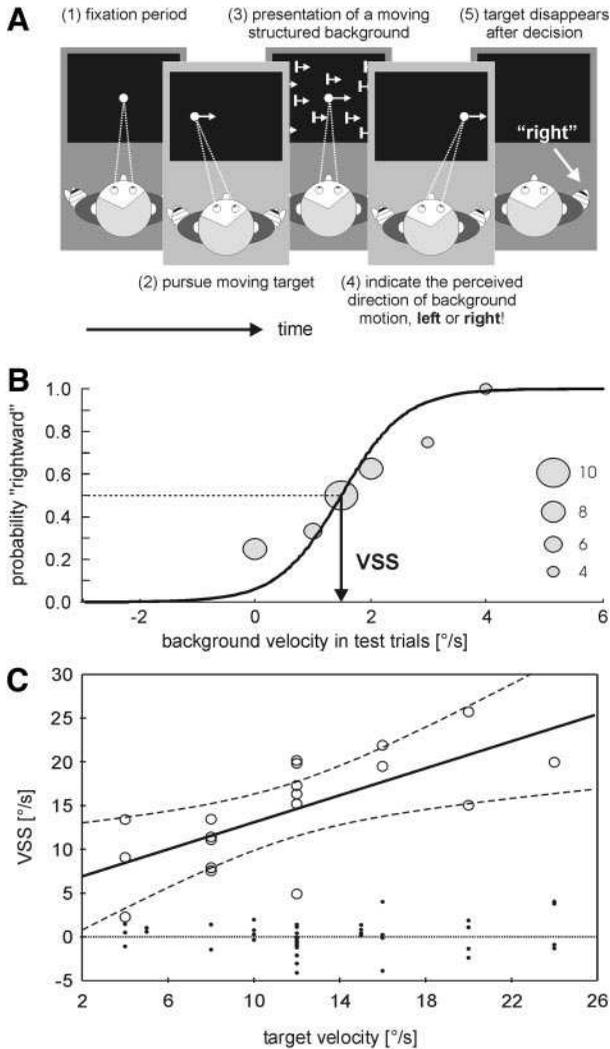


FIG. 2. (A) Paradigm used in order to quantify motion perception during smooth-pursuit eye movements. The subject tracks a small dot which is initially stationary for a few hundred milliseconds, then jumps to the left and from there moves to the right at a constant speed (on the order of 10°/s in most experiments). While the eyes are close to straight ahead, a large-field random dot background comes on for several hundred ms, whose elements move coherently to the left or to the right. At the end of the pursuit trial, subjects have to indicate, whether they perceived the background as moving to the left or to the right (two alternative forced choice). Eye movements are monitored with infrared optics in order to guarantee high precision smooth-pursuit eye movement. (B) The percentage of "rightward" decision of subjects is plotted as a function of the background velocity. Background stimuli moving at comparably high velocities (4°/s) to the right are always perceived as moving to the right. Physically stationary backgrounds (0°/s) were more often perceived as moving leftward than rightward. The turning point of the Probit function fitted gives the velocity of subjective stationarity (VSS), the velocity leading to as many right as left decisions of the subject. (C) Perception of smooth-pursuit-induced visual motion in healthy controls (18 subjects, 50 individual measurements represented by dots, partially lying on top of each other) and patient R.W. (open circles). Reprinted by permission from *Nature* (Haarmeier *et al.*) copyright (1997) Macmillan Magazines Ltd. The plot shows the VSS as function of pursuit target velocity varied from 4 to 24°/s. In this and all other experiments discussed, pursuit target and background were presented on a 19-in. computer monitor at a viewing distance of 57 cm in a dark

physically stationary world in a direction opposite to the direction of smooth-pursuit eye movement would be perceived, if the reference signal were too small. Conversely, false perception of movement of the world in the direction of the eye movement would result from a too large reference signal. If the inferential mechanism providing and using the reference signal were able to sense its insufficiencies, it would use such feedback information to gradually increase the size of the reference signal in the first, the "reference signal too small" condition ("rf_small") and conversely, to decrease it in the second, the "reference signal too large" condition ("rf_large"). This is exactly what we found, when we mimicked such imperfections by asking subjects to execute smooth pursuit across backgrounds moving at a high, constant velocity ("constant background" trials), either in the direction of the eye movement, corresponding to the rf_large condition or opposite to the direction of the eye movement, corresponding to the rf_small situation (Fig. 3A). The sequence of these constant background trials was occasionally interrupted by "test" trials consisting of backgrounds, whose velocity was determined by the staircase procedure needed to pinpoint the VSS. As suggested by the differential development of the staircase procedures shown in Fig. 3A, representing an individual subject in the rf_large and rf_small condition respectively, this manipulation indeed induces changes in motion perception, consistent with an ecologically useful adaptation of the reference signal. This conclusion was first supported by a comparison of the VSSs for a group of 10 subjects tested in the rf_large and rf_small condition and has since then been replicated many times. In summary, these psychophysical experiments show that the perception of eye-movement-induced visual motion can be modified predictably by mimicking insufficiencies of the reference signal, prompting functionally useful changes of the

experimental room. The pursuit target was a red dot (diameter 10 min of arc) and the background subtended $27 \times 27^\circ$ of visual angle and consisted of 350 white dots (diameter 15 min arc, local contrast, 0.01). The background was presented for 300 ms at a time, when the eyes were close to straight ahead. Target motion was always to the right. Positive values of VSS indicate background movement in the same direction as the eyes. The higher the VSS, the less the resulting retinal image slip velocity allowing the percept of stationarity and, thus, the less the ability to cope with self-induced visual motion. In the control subjects the VSS is always close to 0°/s. Hence perceived background motion is tightly related to visual motion in extrapersonal space, that is, stationarity is preserved despite increasing retinal image shifts with growing target velocities. In R.W., however, perceived motion reflects image movement on the retina. The linear regression (with 99% confidence bands) suggests that R.W. is deprived of a sense of stationarity unless the background image is stabilized to some degree on the retina. Accordingly, he will experience an illusory background motion (Filehne illusion) in the order of the eye velocity prevailing.

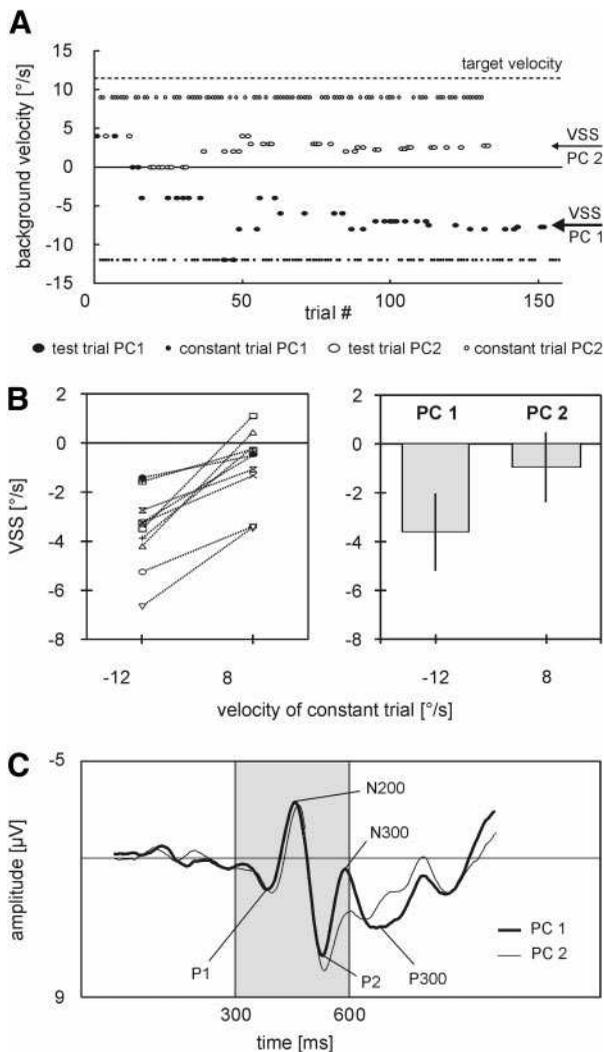
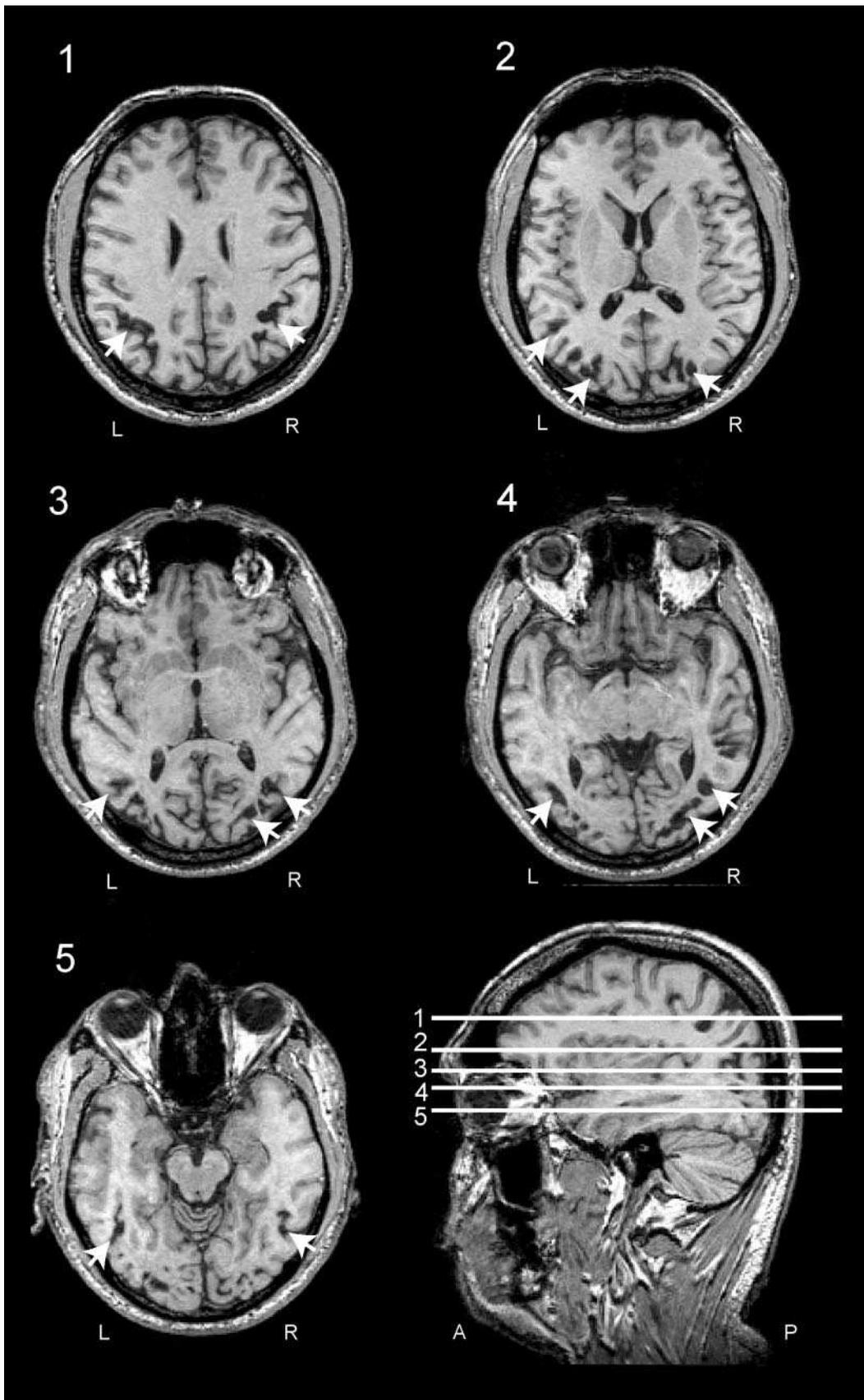


FIG. 3. (A) Dependence of the VSS on the choice of the “constant background” trial. Shown are temporal sequences of background velocities for a healthy subject, carrying out smooth pursuit eye movements to the right at 12°/s. Each dot marks one trial as illustrated in Fig. 2A. The temporal sequence of two experimental sessions are shown (PC1, filled symbols characterizing the constant trial backgrounds and the test trial backgrounds; PC2, open symbols). In the first run, the “constant trial” background moved at $-12^{\circ}/s$ (PC1), in the second run, it moved at $9^{\circ}/s$ (PC2). (B) VSS as a function of the velocity of the “constant trial” background for ten healthy subjects, represented by individual symbols (left panel) and their means/standard deviations (right panel). Negative velocities indicate the direction opposite to the eye movement. Smooth pursuit was to the right at $12^{\circ}/s$. The VSS depends on the choice of the “constant trial” background. A velocity of $-12^{\circ}/s$ mimics a viewing condition, in which the reference signal is too small, causing a compensatory increase of the reference signal and therefore a shift of the VSS toward stronger movement to the left (PC1, perceptual condition 1). A velocity of $+8^{\circ}/s$, conversely, mimics a reference signal which is too large, inducing the opposite changes (PC2, perceptual condition 2). (C) Grand average VEPs for the O1 recording site based on 5 of 10 subjects who showed the strongest difference between the perceptual conditions PC1 (bold line) and PC2 (thin line). The responses are plotted for a 1000-ms period starting 300 ms before onset of the stationary background. The vertical gray column marks the period of background presentation. B and C adapted from Haarmeier and Thier (1998). Reproduced, by permission of MIT Press.

reference signal. Control experiments rule out that the changes of motion perception are a simple consequence of motion adaptation (Lindner *et al.*, 2000). We add that the psychophysical responses of nonhuman primates are very similar to those of humans in experiments mimicking insufficiencies of the reference signal (Chakraborty, 2001), suggesting that the visual system of nonhuman and human primates share a common inferential mechanism.

Our ability to associate qualitatively different percepts of motion with one and the same amount of eye-movement-induced retinal image slip is obviously an ideal tool to reveal those parts of cerebral cortex, reflecting the subjective percept of self-induced visual motion rather than the sensory signal conveyed by the retina. The first study based on this approach, alluded to earlier, was a study of visually evoked potentials (VEPs), picked up from the scalp with 12 surface electrodes (Haarmeier and Thier, 1998). VEPs were evoked by pursuit-induced slip of stationary backgrounds (share 25%), presented randomly interleaved with test backgrounds (share 25%) needed to measure the VSS and a much larger number of constant backgrounds (share 50%) required to modify the percept. The constant backgrounds came in two flavors in two subsequent blocks, giving rise to the different VSS values as presented in Fig. 3B. It is important to understand that the VEPs collected in these two subsequent blocks were based on stationary backgrounds only and that neither the constant nor the test trials contributed to the VEP. Since the eye movements did not differ between blocks, the size of self-induced retinal stimulation was the same. Figure 3C shows the grand average VEP waveforms for the O1 lead for a group of five out of ten subjects whose percept varied the most. The VEP responses are shown for a period of time centered on the presentation of the stationary background. The two curves superimposed are the averages for the two perceptually different conditions. The two waveforms are identical, independent of the percept of motion, for more than 200 ms. The first component which differed significantly between conditions was a negativity peaking at about 300 ms, dubbed the N300. Since its am-

FIG. 4. MRI of patient R.W. shows bilateral cyst-like widenings (arrows) of the sulci of the parietal and occipital lobes mainly affecting areas 18, 19, and possibly 37 on the lateral aspect of the hemispheres and areas 18 and 19 on the inferior aspect. In addition, cortex in and around the intraparietal sulcus of the parietal lobes is involved. The lesions represent strictly cortical defects and local cortical atrophy with no signs of progression. Subcortical white matter and basal ganglia are intact. Although the lesional pattern does not allow a definitive pathogenetic assignment, posthypoxic cortical laminar necrosis seems to be the most likely explanation. As a possible cause, R.W. had suffered from a severe pertussis infection in early childhood which had required artificial respiration. A, anterior; P, posterior; L, left; R, right.



plitude modulation correlated significantly with the perceptual differences between conditions, we take the N300 as an electrophysiological correlate of the percept of visual motion. The N300 was followed by a P300, also differing between conditions, but probably arising too late to be a direct reflection of the percept of motion. The fact that the N300 followed an earlier N200, known to arise from the human MT/V5 complex (Bach and Ulrich, 1994; Kubová *et al.*, 1995) suggests that the generators of the N300 are located upstream of MT/V5. We have recently obtained independent support for this notion from single-unit recording experiments in monkeys (Chakraborty and Thier, 2000) and MEG-recording experiments in humans (Tikhonov *et al.*, 2001), based on the same approach underlying the VEP-experiments discussed before. This most recent work indicates that the neuronal substrates underlying our percept of motion seem to be concentrated in area VPS (visual posterior sylvian), a region close to the caudal end of the lateral fissure (Fig. 1B). Monkey area VPS has been shown to receive input from motion processing area MST in the superior temporal sulcus (Guldin and Grüsser, 1998) and, moreover, to be reciprocally interconnected with the parieto-insular-vestibular cortex (Guldin and Grüsser, 1998; Grüsser *et al.*, 1990a,b), a multimodal representation of head movement relative to external space. These findings may therefore suggest that the caudal lateral sulcus is a major element in a specialized cortical network underlying primates' *subjective* sense of spatial stability, supplementing the visuomotor contributions of the posterior parietal components of the classical dorsal stream.

Primary visual cortex and probably also the next early stages of the cortical processing of visual motion, up to and including area MT/V5, are ignorant of the percept of motion during eye movements. This is suggested by the lack of any significant percept-related influence on the early parts of the VEP response and the fact that—as discussed before—neurons in monkey area V1 usually do not distinguish between self-induced and externally induced retinal image slip, although it is very likely that the percepts affiliated with these two forms of retinal image slip will differ also in the monkey. We emphasize that this view is by no means invalidated by the observation of a few neurons in the early parts of the cortical visual system which may show subtle differences in their activity evoked by self-induced compared to externally induced visual motion (Galletti *et al.*, 1984; Chakraborty and Thier, 2000).

Further support for the notion of a representation of visual motion awareness in human, excluding area 17, is provided by the case of a patient (R.W.; Haarmeier *et al.*, 1997), suffering from bilateral lesions involving large parts of dorsal extrastriate and posterior parietal cortex, while sparing area 17 (Fig. 4). R.W. has normal

eye movements and normal vision if the eyes do not move. However, he is completely unable to compensate for the visual consequences of his smooth-pursuit eye movements. Instead of perceiving the image shifted across the retina by the eye movement as stationary, he perceives image movement. Lack of compensation for his smooth-pursuit eye movements is indicated by the fact that his velocity of subjective stationarity roughly equals his eye velocity (Fig. 2C, open symbols). He perceives the world as stationary only if the world is moved with the eyes, thereby stabilizing its image physically on the retina. Destabilizing the image of the world on the retina by moving the eyes jeopardizes his sense of visual stability and gives rise to complaints of severe dizziness. In accordance with this interpretation, dizziness immediately subsides if he closes the eyes or stops moving his eyes.

In summary, our work clearly supports the notion that the perception of visual motion is based on an inferential mechanism involving a nonretinal reference signal capturing the visual consequences of the eye movement. It furthermore indicates that this inferential mechanism resides in rather “late” parts of the cortical hierarchy of motion processing sparing the early stages up to cortical area MT and, finally, that disturbances of this mechanism may give rise to severe disturbances of vision and spatial orientation as exemplified by R.W. In general, support for an inferential mechanism underlying motion perception residing in cortex strengthens the view that one important function of cerebral cortex is the extraction of invariances.

REFERENCES

- Bach, M., and Ullrich, D. 1994. Motion adaptation governs the shape of motion-induced cortical potentials. *Vision Res.* **34**: 1541–1547.
- Chakraborty, S., and Thier, P. 2000. A distributed neuronal substrate of perceptual stability during smooth-pursuit eye movements in the monkey. *Soc. Neurosci. Abstr.* **26**: 674.
- Chakraborty, S. 2001. Neuronal mechanisms underlying perceptual stability during smooth-pursuit eye movements. Ph.D. Thesis, Univ. Tübingen.
- Erickson, R. G., and Thier, P. 1991. A neuronal correlate of spatial stability during periods of self-induced visual motion. *Exp. Brain Res.* **86**: 608–616.
- Galletti, C., Squatrito, S., Battaglini, P. P., and Grazia-Maioli, M. 1984. ‘Real-motion’ cells in the primary visual cortex of macaque monkeys. *Brain Res.* **301**: 95–110.
- Grüsser, O.-J., Pause, M., and Schreiter, U. 1990a. Localization and responses of neurons in the parieto-insular vestibular cortex of awake monkeys (*Macaca fascicularis*). *J. Physiol.* **430**: 537–557.
- Grüsser, O.-J., Pause, M., and Schreiter, U. 1990b. Vestibular neurons in the parieto-insular cortex of monkeys (*Macaca fascicularis*): Visual and neck receptor responses. *J. Physiol.* **430**: 559–583.
- Guldin, W. O., and Grüsser, O.-J. 1998. Is there a vestibular cortex? *Trends Neurosci.* **21**: 254–259.

- Haarmeier, T., and Thier, P. 1996. Modification of the Filehne illusion by conditioning visual stimuli. *Vision Res.* **36**: 741–750.
- Haarmeier, T., Thier, P., Repnow, M., and Petersen, D. 1997. False perception of motion in a patient who cannot compensate for eye movements. *Nature* **389**: 849–852.
- Haarmeier, T., and Thier, P. 1998. An electrophysiological correlate of visual motion awareness in man. *J. Cogn. Neurosci.* **10**: 464–471.
- Haarmeier, T., and Thier, P. 1999. Impaired analysis of moving objects due to deficient smooth pursuit eye movements. *Brain* **122**: 1495–1505.
- Ilg, U., and Thier, P. 1996. Inability of visual area V1 to discriminate between self-induced and externally-induced retinal image slip. *Eur. J. Neurosci.* **8**: 1156–1166.
- Kubová, Z., Kuba, M., Spekreijse, H., and Blakemore, C. 1995. Contrast dependence of motion-onset and pattern-reversal evoked potentials. *Vision Res.* **35**: 197–205.
- Lindner, A., Haarmeier, T., and Thier, P. 2000. Psychophysical evidence for an adaptable eye movement reference signal, guaranteeing spatial constancy during smooth pursuit eye movements. *Soc. Neurosci. Abstr.* **26**: 1716.
- Mack, A., and Hermann, E. 1973. Position constancy during pursuit eye movements: An investigation of the Filehne illusion. *Q. J. Exp. Psychol.* **25**: 71–84.
- Tikhonov, A., Haarmeier, T., and Thier, P. 2001. A magnetoencephalographic correlate of perceptual spatial stability in man. In *Tübinger Wahrnehmungskongferenz*, Vol. 4, p. 145. Knirsch-Verlag, Kirchentellinsfurt.
- von Helmholtz, H. 1910. *Handbuch der physiologischen Optik*, Vol. 3. Voss, Leipzig.
- von Holst, E., and Mittelstaedt, H. 1950. Das Reafferenzprinzip. *Naturwissenschaften* **37**: 464–476.
- Wertheim, A. H. 1994. Motion perception during self-motion: The direct versus inferential controversy revisited. *Behav. Brain Sci.* **17**: 293–355.