

# Joint encoding of binocular disparity and direction of motion

Martijn Barendregt<sup>a</sup>

<sup>a</sup>*Graduate School of Life Sciences, Utrecht University, The Netherlands*

---

## Abstract

There are competing views regarding the existence of joint encoding of binocular disparity and direction of motion within the human visual system. It is known that cells within the sensory cortices can have preferential tuning to more than one dimension of a stimulus. For example, most cells in the human middle temporal area (hMT+), which is considered a critical part of the cortical motion processing pathway, are sensitive to both binocular disparity and direction of motion. An important question is how these stimulus dimensions are encoded in such cells and in what way this affects the manner in which information can be extracted from them. Here, data from both electrophysiology and computational studies is reviewed and compared in order to evaluate the current arguments for and against joint encoding of these stimulus dimensions. Physiological data shows that this type of joint encoding does exist in cats and monkeys, but there is not current evidence proving its existence in human visual cortex. The computational models that are discussed attempt to either prove or disprove the necessity of joint encoding of these dimensions, but are found to lack the generalizability needed to provide a definitive answer. To gain new insights into the possible role of joint encoding of binocular disparity and direction of motion, a novel neuroimaging experiment is proposed to investigate the existence of joint encoding in human visual cortex.

*Keywords:* joint encoding, binocular disparity, motion, Pulfrich effect

---

## 1. Introduction

Perception of depth and perception of motion are both vital for survival and some of the most critical tasks the visual system has to perform. There is a close relation between the way percepts of depth and percepts of motion are computed by the brain. Both make use of corresponding points in the visual field perceived by both eyes and either spatial (depth) or temporal (motion) interactions between those points. The similarity between these stimulus dimensions could be an indication that both depth and motion are processed using the same computations or the same underlying neural circuitry.

Extensive research has identified the neural basis for both depth perception and motion perception, as well as the computational mechanisms that are involved. From this research it is also becoming increasingly evident that there may be a close relation between these functions. This may also mean that both binocular disparity (depth) and direction of motion are processed together in the brain, possibly even within the same neural mechanisms.

An example of a phenomenon that appears to point in the direction of joint encoding for binocular disparity and direction of motion is the disparity-contingent motion aftereffect (Anstis, 1974). This effect is generally created by first viewing two transparent planes placed in front of each other in depth (i.e. at different disparities). Each plane consists of coherently moving dots, with the direction of motion opposite in each plane. After adapting to this stimulus, two planes at the same disparities but with

randomly moving (i.e. incoherent) dots are shown. Despite the lack of coherent motion, the dots in each plane seem to move coherently in the direction opposite to that of the previously adapted plane. This phenomenon shows a direct interaction between the encoding of binocular disparity and direction of motion within the visual system.

In recent literature, there is evidence from both electrophysiological and psychophysical studies that binocular disparity and direction of motion are not only processed together but that they may also be jointly encoded. For various stimulus dimensions, cells have been found that are tuned to combinations of these dimensions (e.g. spatial frequency and orientation in primary visual cortex) in many areas of the brain. However, the fact that a single neuron is tuned to more than one dimension of a stimulus does not automatically entail that these dimensions are jointly encoded within this neuron. For example, it is known that human primary visual cortex contains many cells that are tuned to combinations of multiple stimulus dimensions, yet the majority of these encodings are separable (Grunewald & Skoumbourdis, 2004). Although joint encoding is generally recognised as a viable way to integrate multiple stimulus dimensions in visual processing, it is important to consider why the visual system should prefer joint over separable encoding.

From a psychophysical point of view, joint encoding of binocular disparity and direction of motion is an appealing concept because there is an inherent ambiguity in the signals that arrive from both eyes with respect to space and

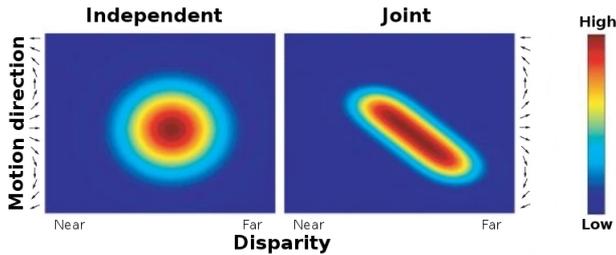


Figure 1: *Hypothetical neuronal response profiles for a neuron implementing either independent (left) or joint encoding (right). Adapted from Grunewald & Skoumbourdis (2004)*

time. A binocular signal arriving in visual cortex might contain a spatial disparity because of a temporal offset between the eyes (as will be further explained in relation to the Pulfrich effect) or because there is a real spatial disparity in the stimulus. This inability to differentiate between spatial and temporal aspects of the stimulus would cause neurons in this area to signal a spatial disparity regardless of whether the binocular difference is caused by a temporal or spatial offset between the eyes.

The goal of the current study is to review and evaluate the currently available evidence from both electrophysiological data and computational modeling in order to answer the question of why joint encoding of binocular disparity and direction of motion would be a feasible or even necessary arrangement within the human visual system.

## 2. Joint vs independent encoding

In literature the use of the terms joint and independent encoding can lead to confusion, because they can refer to different actual arrangements. In this thesis, the use of the term joint encoding is intended to refer to the arrangement where one neuron is tuned to two different stimulus dimensions but these are not separable, e.g. it is not possible to 'read-out' the response to any one of these dimensions without interference from the other. The opposite arrangement, independent encoding, is thus a situation where one neuron encodes two stimulus dimensions in a separable implementation. It would for example be possible for such a neuron to have outgoing connections to two different types of cells, one for each stimulus dimension.

To illustrate the difference between the intended meanings of the terms joint and independent, see Figure 1. The graphs depict response characteristics for a hypothetical neuron that responds either independently (left) or jointly (right) to its two encoded stimulus dimensions (represented as the two axes).

## 3. Physiology

The neural basis that supports perception of motion-in-depth, is not well understood. Both binocular disparity

and direction of motion are important cues to compute the velocity of a stimulus moving in depth (Rokers et al., 2008). Motion-in-depth produces both changes in binocular disparity over time and a difference in stimulus velocity between the two eyes which can be used by the visual system to compute three-dimensional motion from two-dimensional retinal images (Harris et al., 2008). However, this does not provide much insight into *how* these potential cues are used or where in the brain these computations are carried out. Much effort has been put into gaining insights into these computations by examining the way in which these cues are processed by the neural circuitry underlying visual perception.

From physiological studies it is known that neurons in visual area MT can be tuned to direction of motion (Zeki, 1974) as well as binocular disparity (Maunsell & Van Essen, 1983). These cells are organized by direction tuning in columns that run perpendicular to the cortical surface (Albright et al., 1984) and in a topographic map of binocular disparity tuning (DeAngelis & Newsome, 1999). Regions of cells with strong binocular disparity tuning are separated by regions of weak binocular disparity tuning. While both weak and strongly tuned regions contain cells sensitive to motion direction, only the regions of strong binocular disparity tuning show a columnar organization for both binocular disparity and motion direction, see Figure 2 (DeAngelis & Newsome, 1999).

Electrophysiological studies in the visual motion processing areas of the cat (area 17/18; Anzai et al. (2001)) and monkey (area MT/MST; Bradley et al. (1995); Pack et al. (2003); DeAngelis & Newsome (2004)) have found that these areas typically show joint encoding of motion direction and binocular disparity. For example, DeAngelis & Newsome (2004) used microstimulation to show that stimulation of sites that are not selective for binocular-disparity produced a large bias in the perceived direction of motion of a visual stimulus, whereas stimulating sites with strong disparity-selectivity produced little or no bias.

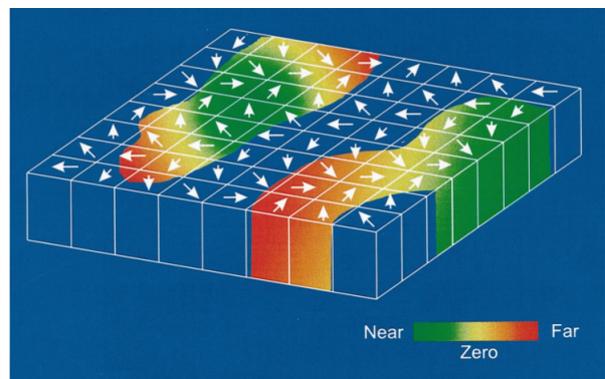


Figure 2: *Schematic overview of the direction columns that are inter-mixed with patches of strong disparity-tuning in macaque area MT. The arrows indicate the direction preference for each column and the colours within the patches indicate disparity preference. Adapted from DeAngelis & Newsome (1999)*

Interestingly, two recent studies in the primary visual cortex (V1) of the macaque (Pack et al., 2003; Read & Cumming, 2003) showed that joint encoding of binocular disparity and motion direction is quite rare in cells within this area. Another recent study by Grunewald & Skounbourdis (2004) found that although most cells in monkey V1 (around 80%) encode multiple stimulus dimensions, nearly all of these cells show separate encoding for their multiple dimensions. Since a large portion of the cells in this area are sensitive to disparity but very few are sensitive to motion direction, it would appear that most cells that are selective for motion direction will also be sensitive to binocular disparity (MT+ cells), while many cells that are sensitive to binocular disparities are not selective for motion direction (V1 cells). In other words, there seems to be a one-way relationship in which cells that are tuned for motion direction will also likely be tuned to binocular disparity but not the other way around.

If joint encoding is a necessary condition to support the perception of motion-in-depth in the human visual system then, as Read & Cumming (2005a) suggest, this would lead to the conclusion that the visual system processes motion in depth solely by spatiotemporal filters in the MT complex and ignores the pure spatial disparity filters in primary visual cortex. This would seem a highly inefficient organization because the disparity information present in area V1 could just be used in combination with the motion information present in other areas in order to perform the necessary computations. Therefore, any theoretical or computational account arguing for the necessity of joint encoding of motion direction and binocular disparity should also provide a clear argument explaining the advantage of such an organization over the (possibly) more simple organization involving the disparity information in area V1.

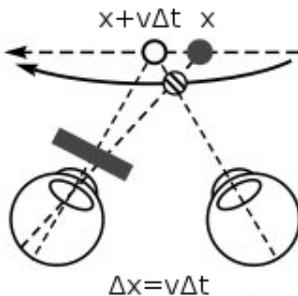


Figure 3: Schematic explanation of the classical Pulfrich effect. The stimulus (filled dot, speed  $v$ ) is seen by the right eye at position  $x$  and time  $t$ . The left eye sees the stimulus at position  $x$  after an interocular delay  $\Delta t$  (created by a neutral density filter). At time  $t + \Delta t$  the stimulus has moved to position  $x + v\Delta t$  in the right eye, which creates a disparity of  $v\Delta t$  between the right and left eye and the perception that the stimulus is following an elliptical path (striped dot). Adapted from Anzai et al. (2001)

#### 4. Pulfrich-like phenomena

The existence of joint encoding within the visual system is often proposed as an explanation for the Pulfrich phenomenon (Pulfrich, 1922) and its more recent derivatives (Tyler, 1974; Burr & Ross, 1979). In these phenomena a perceptual illusion of depth is created by introducing an interocular delay, which can be either artificial using a neutral density filter over one eye to cause a relative delay in processing (Wilson & Anstis, 1969; Carney et al., 1989) or clinical for patients suffering from degeneration of optic fibres. A common explanation for this illusion is that a temporal delay in processing between the two eyes is interpreted by the visual system as a spatial disparity signal (Pulfrich, 1922). Consider an interocular delay  $\Delta t$  between the right and left eye. A stimulus moving with speed  $v$  that appears in the right eye at position  $x$  will appear in the left eye at time  $t + \Delta t$ . However, at this point the stimulus will have shifted position in the right eye to  $x + v\Delta t$  thus producing a real spatial disparity of  $v\Delta t$ . (Figure 3)

This explanation holds for the original Pulfrich effect but not for its more recent variations such as the stroboscopic Pulfrich effect (Burr & Ross, 1979). In this case the stimulus is intermittently seen by each of the two eyes (Figure 4). Since at any point in time, the stimulus is only seen by one of the two eyes the visual system needs to determine the position of the stimulus in the other eye either by extrapolation or by retaining it in perceptual memory in order to binocularly match them.

In recent years the generally accepted explanation for all Pulfrich-like phenomena is that the visual system uses spatio-temporal inseparable filters to process depth information (Anzai et al., 2001; Qian & Andersen, 1997). These neuronal mechanisms use joint encoding of motion direction and binocular disparity and this makes them unable to distinguish an interocular temporal delay from a binocular disparity (Qian & Andersen, 1997).

Although most authors agree that joint encoding of binocular disparity and direction of motion is the best explanation for the Pulfrich effect, it should be noted that nobody has provided any argument to explain *why* the visual system should work this way. It remains unclear what would be the advantage of jointly encoding binocular disparity and motion direction compared to separately encoding them. Anzai et al. (2001) provides the clearest argumentation as to why the joint encoding of binocular disparity and motion direction would be an optimal solution. It is worthwhile to examine their argument here, because most of the authors of the computational models on the Pulfrich phenomena refer to this argumentation.

Anzai et al. (2001) consider three separate hypotheses regarding the neural mechanism responsible for the observed motion-in-depth in the Pulfrich phenomena. The perception of motion-in-depth could be supported by either a purely spatial offset (interocular spatial offset hypothesis), a purely temporal offset (interocular temporal

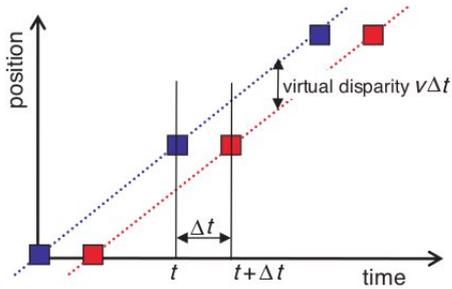


Figure 4: Schematic representation of the stimulus in the stroboscopic Pulfrich effect. The blue and red squares mark the appearance of the stimulus in the right and left eye, respectively. The stimulus appears at the same position in each eye, but there is a temporal delay  $\Delta t$  between the right and left eye. The dotted lines indicate the trajectories of the apparent motion defined by the stimulus. The separation between these lines defines a virtual disparity  $v\Delta t$  (Burr & Ross, 1979). Adapted from Read & Cumming (2005a)

offset hypothesis) or a combination of both spatial and temporal offsets (joint encoding hypothesis).

Anzai et al. argue that the first hypothesis (spatial offset) is incorrect because a purely spatial offset would not be able to explain the stroboscopic Pulfrich effect. This argument is essentially the same argument that is used against the classical, geometric explanation of the Pulfrich effect. However, as Read & Cumming (2005b) show, a set of disparity-tuned neurons with space/time-separable receptive fields would still respond to the disparity present in the stroboscopic Pulfrich illusion. Therefore, disregarding the spatial offset hypothesis based on this argument is not valid because evidence shows that a pure spatial offset can in fact explain the perception of motion-in-depth in the Pulfrich effect.

The second hypothesis, the temporal offset hypothesis, is disregarded because to produce a depth percept based solely on an interocular temporal offset ‘the visual system requires mechanisms tuned to various interocular latency differences at corresponding points’ in the visual field (Anzai et al., 2001). Although it is not entirely clear what is meant by the authors, it would appear that a set of basic motion detectors would fit this description. In this case the brain would meet this requirement set by Anzai et al. since cells that act as motion detectors are obviously present. Besides detection, the visual system would need ‘a mechanism that converts the latency differences into depth’ consistent with the stimulus (Anzai et al., 2001). Again, this description is not very specific, and it would appear that the way in which the visual system computes motion-in-depth from 2d motion cues (Shioiri et al., 2000) is sufficient to meet this requirement.

The reasoning put forward by Anzai et al. (2001) as well as that of some authors of the quantitative accounts of the Pulfrich effect (Qian & Andersen, 1997; Read & Cumming, 2005a) seems to be predicated on the assumption that the computations that are responsible for the illusion of depth

are carried out very early in the visual system. While this assumption is not in itself a problem, it does become problematic when the authors also argue that joint encoding is a requirement to explain the motion-in-depth perception in the Pulfrich effect. Since electrophysiological evidence (as discussed above) shows that joint encoding of binocular disparity and direction of motion is rare within primary visual cortex, it is most likely that these two assumptions are not compatible. Joint encoding of binocular disparity and motion direction does exist in area MT/MST. If these computations are assumed to be located in the motion processing area MT, the requirement of joint encoding of these two stimulus dimensions could be met. However, this would also raise the same question as discussed in the previous section: why would the pure disparity signal present in primary visual cortex be ignored, especially given the fact that this signal alone could possibly suffice to explain the Pulfrich phenomena (Read & Cumming, 2005a; ?).

## 5. Computational models

Over the past decade several computational accounts have attempted to capture the phenomenology observed in the Pulfrich effect. An interesting aspect of these modeling efforts is that the authors have differing points of view with regard to the necessity of joint encoding of binocular disparity and direction of motion to explain the Pulfrich phenomena.

The first computational account of joint encoding of motion direction and binocular disparity to explain the Pulfrich phenomena was proposed by Qian & Andersen (1997). This integrated model combined an earlier stereo vision model (Qian, 1994) with a general motion energy model (Adelson & Bergen, 1985; Watson & Ahumada, 1985; van Santen & Sperling, 1985) to provide a unified account for both binocular disparity and motion selectivity. Because these joint spatiotemporal complex cells cannot distinguish between the spatial and temporal features of the stimulus, the computations underlying the depth perception are not based on any real disparity that would be present in the stimulus and thus provide a unified account of the Pulfrich-like phenomena.

In an attempt to show that joint encoding of binocular disparity and motion direction is not necessary to explain the Pulfrich phenomena, Read & Cumming (2005a) introduce a model that is based on separate encoding of these stimulus dimensions. Their model encodes disparity and motion in separate neural populations and introduces a “read-out” rule that models how information is extracted from these populations. To explain the stroboscopic Pulfrich effect, their model in fact only needs a population of disparity-tuned neurons that, combined with the read-out rule, reproduces the observed perceptual results.

Neri & Levi (2008) provide an interesting method to investigate the question of joint encoding of binocular disparity. Instead of using the Pulfrich phenomena, they make use of the interaction between binocular disparity

and motion direction in the disparity-contingent motion aftereffect (Anstis, 1974; Verstraten et al., 1994). Using visual adaptation and reverse correlation they derive a set of perceptual filters that allow them to account for the psychophysical data. In contrast to the other two models, which take a bottom-up approach, this top-down approach provides a indication of the type of computations that the visual system would need to perform in order to produce the observed effects in both disparity-contingent motion aftereffects and the Pulfrich phenomena. Since these perceptual filters describe the computations at a higher, more abstract level it would be worth to devise a low-level computational model that implements these filters. Neri & Levi (2008) describe such a model themselves, however this model is still high level and does not model computations at the neuronal level.

A recent review of the currently available computational models on the Pulfrich effect (Qian & Freeman, 2009) emphasizes that the only model that does not implement joint encoding (Read & Cumming, 2005a,b) is not able to accurately explain all Pulfrich phenomena. However, since the author of the review is also one of the authors of the original model that does promote joint encoding of binocular disparity and motion direction to explain the Pulfrich effect (Qian & Andersen, 1997) this review appears to be somewhat biased. Furthermore, although the review claims to present an overview of current models they only compare their own model (Qian & Andersen, 1997) against the separate encoding model by Read & Cumming (2005b) and do not include the other recent model by Neri & Levi (2008) (this paper is cited but only in relation to the motion contingent aftereffect). This, combined with the extensive use of *reductio ad absurdum* to improve their arguments, leads to the conclusion that this paper, although presented as a review, is actually heavily biased towards the authors own interpretation of the evidence on the Pulfrich phenomena.

The proposed quantitative models to explain the Pulfrich phenomena differ in several aspects. For example, the model proposed by Qian & Andersen (1997) is highly detailed and incorporates a large amount of physiological information to produce a model that is as biologically accurate as possible. In contrast, Neri & Levi (2008) have modeled purely psychophysical data at a much more abstract level (perceptual filters) without consideration for the physiological implementation that would be needed, although they state that their results (filters) do reflect possible neural arrangements. As pointed out by Qian & Freeman (2009) the model of Read & Cumming (2005b), although still physiologically motivated, is based on a number of simplifications from physiological data to facilitate the use of the stereo energy model (Read & Cumming, 2005b). Most important they model the response of V1 neurons with a mono-phasic kernel, while physiological evidence shows that these responses are biphasic in nature.

These comments reflect some serious considerations on the one side and the practical limitations of modeling on

the other side. A model is never a complete description of reality and always contains simplifications for either conceptual or computational (e.g. too complex calculations can take too much time) reasons. This should never be a reason to immediately dismiss a model, especially if the model can still provide an accurate account of the phenomenon it aims to describe.

Comparing the discussed modeling efforts of joint/separate encoding of binocular disparity and direction of motion it would seem that an interesting future direction of the computational efforts on this subject would be a combination of the high level model of Neri & Levi (2008) and either of the low-level models (Qian & Andersen, 1997; Read & Cumming, 2005b). This is likely a more fruitful approach to finding a complete computational account instead of the current diversity of competing models.

## 6. Neuroimaging

Despite the numerous studies performed on cats and monkeys, human neurophysiological evidence for joint encoding of binocular disparity and direction of motion has not been found. While some authors seem to disregard such experiments because they would merely serve to ‘only confirm in humans the electrophysiological evidence already obtained in monkeys’ (Neri & Levi, 2008), neuroimaging could provide a useful tool to investigate joint encoding of binocular disparity and direction of motion in the human visual system. With the advent of high-resolution neuroimaging techniques, it is possible to construct an experiment to measure the response characteristics of neurons in the human visual cortex during a task that would produce different results depending on whether or not joint encoding of binocular disparity and direction of motion is involved. From the analysis of the electrophysiological data it follows that the most promising area to look for these effects would be the human motion processing area hMT+. The experiment that is proposed here will focus on area hMT+ but the experiment can equally well be used to investigate the presence or absence of joint encoding of binocular disparity and motion direction within human primary visual cortex.

If binocular disparity and direction of motion are jointly encoded within area hMT+ then a logical consequence is that adaptation to one of those dimensions causes a change in the neural response to the other dimension in this area. This prediction can be tested with a neuroimaging experiment using an adaptation paradigm. The first step of the experiment would be to measure BOLD responses of voxels in area hMT+ to a stimulus containing only binocular disparity, without motion. This would establish a preadapted baseline value for every voxel to compare against the postadapted levels of activity. In order to ensure an optimal response from every voxel, the procedure should consist of displaying a random disparity from a certain range of values (e.g.  $-5^\circ$  to  $+5^\circ$ ) on every TR and

repeating each disparity value at least a certain number of times to increase the signal-to-noise ratio. The next step of the experiment is the adaptation phase. In this phase, the observer is shown a stimulus that only contains incoherent motion for a prolonged period of time. This stimulus will cause adaptation of the neurons that are sensitive to motion, which means their response to motion will be lowered due to neural fatigue. The use of motion without coherence will ensure an optimal response from the entire population of motion-sensitive voxels instead of being limited to the population that is tuned to a specific direction of motion. The final step of the experiment is measuring BOLD responses to the same pure disparity stimulus again, using the same range of values, and comparing the measured neural activity to the preadapted measurements.

Should binocular disparity and direction of motion indeed be jointly encoded within the hMT+ neural populations, then the expected result of the experiment is that the postadapted responses to a certain disparity value will be different from the preadapted response to this value. However, if binocular disparity and direction of motion are encoded independently then the response to each disparity value would be the same as the preadapted response.

This experiment is straightforward in design, yet the technical possibility to perform such an experiment reliably has only become available recently. The disparity-tuned patches identified by DeAngelis & Newsome (1999) are roughly 0.6mm in diameter. Since this is a finding in the macaque monkey the size of the equivalent structure in human cortex will be larger (Tootell & Taylor, 1995). Increases in magnetic field strength coupled with innovations in the scanning protocol have opened up the possibility to measure BOLD responses at voxel sizes of less than 1mm (Cheng et al., 2001; Pfeuffer et al., 2002) which should suffice to measure the disparity-tuned patches within area hMT+ and thus provide reliable measurements for the proposed experiment.

The results of this experiment could provide critical insights into the existence of joint encoding of binocular disparity and motion direction in human visual cortex, specifically area hMT+. If joint encoding of these dimensions is indeed found, then this would be the first direct evidence in humans that the visual system uses joint encoding for these dimensions and support the idea by some authors (Qian & Andersen, 1997; Qian & Freeman, 2009; Neri & Levi, 2008) that this could be involved in the perception of motion-in-depth.

If joint encoding of binocular disparity and direction of motion is not found with this experiment (and a similar experiment investigating primary visual cortex) then the first conclusion would be that the assumption that joint encoding of these stimulus dimensions is necessary to explain the Pulfrich phenomena is invalid. Further, it would also suggest that there might be an important difference between the human visual system and that of the other human primates with regard to how binocular disparity and direction of motion are encoded.

## 7. Discussion

This study reviewed the current evidence in neuroscience literature regarding the existence and necessity of joint encoding of motion direction and binocular disparity within the visual system. Based on both electrophysiological recordings (Anzai et al., 2001; Pack et al., 2003; Read & Cumming, 2003; DeAngelis & Newsome, 2004) and psychophysical data on for example the disparity-contingent motion aftereffect (Anstis, 1974; Verstraten et al., 1994), it is evident that there exists a strong relation between binocular disparity and direction of motion.

Physiological data has provided the most direct evidence for the existence of joint encoding of binocular disparity and motion direction within the primary visual cortex of the cat and area MT of monkeys (Anzai et al., 2001; DeAngelis & Newsome, 2004; Pack et al., 2003). Joint encoding of binocular disparity and motion direction within primary visual cortex of the monkey on the other hand is rare (Pack et al., 2003; Read & Cumming, 2003), which would suggest that if these two stimulus dimensions are jointly encoded within the human visual system it is likely that this would only be in the human MT complex rather than human primary visual cortex. However, any neurophysiological or neuroimaging evidence for the existence of joint encoding of binocular disparity and direction of motion in the human visual system is so far lacking. As described above, it is possible to construct an experiment to investigate the existence of joint encoding of these two stimulus dimensions in human visual cortex. The results of this experiment could provide critical evidence to come to a better understanding of how these two stimulus dimensions are processed in the *human* brain.

Based on the physiological data that is available and the psychophysical results showing a close relation between binocular disparity and motion direction, several computational models have been proposed that explain how joint encoding of these two stimulus dimensions could lead to perception of motion-in-depth (Qian & Andersen, 1997; Neri & Levi, 2008; Qian & Freeman, 2009). Most of these models are based on the Pulfrich effect (Pulfrich, 1922) as a test case motion-in-depth processing and the authors reason that if their model can explain the perceptual properties observed with this illusion, the model could also serve as a more general account of how the brain processes motion-in-depth. Although this is a rather strong claim, it can be valid if this line of reasoning would lead to a single, unified model that would explain the Pulfrich phenomena to a sufficient extent and, most important, could be generalized to perception of motion-in-depth in general. Unfortunately, the debate on whether or not joint encoding is a necessary condition to explain the Pulfrich phenomena has shifted attention away from finding this generalized model and towards investigating the Pulfrich phenomena in ever more detail.

So far not one computational account of this illusion has been accepted as a complete explanation because not

one model has been able to explain every single detail on the phenomena. As a result the question of the necessity of joint encoding has not been answered. Although there are promising computational models that incorporate joint encoding (Qian & Andersen, 1997; Neri & Levi, 2008), there is a competing model that performs roughly equal in explaining the Pulfrich phenomena but does not use joint encoding (Read & Cumming, 2005a,b). Despite the fact that there might be several issues with each of the proposed models to explain the Pulfrich phenomena, the fact that all of the models can provide at least a general account of the computations needed to reproduce the perceptual characteristics of the Pulfrich effect shows that the necessity of joint encoding to account for these phenomena has not been confirmed.

Joint encoding of binocular disparity and direction of motion could be an efficient use of the resources available to the visual system. Evidence from both physiology and psychophysical experiments seems to indicate that at the very least there is close relationship between the processing of binocular disparity and direction of motion within the visual system. However, lacking clear human neurophysiological results it is difficult to prove the existence of joint encoding of binocular disparity and direction of motion within the human visual system. In an attempt to provide evidence for the necessity of joint encoding of binocular disparity and motion direction, quantitative accounts of the Pulfrich effect unfortunately fall short of providing this evidence because they are too focussed on the Pulfrich phenomena.

It can thus be concluded that so far there is no compelling evidence to support either the existence of or the necessity for joint encoding of binocular disparity and direction of motion within the human visual system. Although there are strong indications that this type of encoding may exist, the studies that have attempted to prove or disprove this idea have so far not succeeded in providing compelling evidence to support it.

## 8. Acknowledgements

I would like to thank Prof. Bas Rokers for his endless patience, support and helpful suggestions during the writing of this thesis.

## References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A*, *2*, 284–299.

Albright, T. D., Desimone, R., & Gross, C. G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology*, *51*, 16–31.

Anstis, S. M. (1974). Movement aftereffects contingent on binocular disparity. *Perception*, *47*, 153–168.

Anzai, A., Ohzawa, I., & Freeman, R. D. (2001). Joint-encoding of motion and depth by visual cortical neurons: neural basis of the pulfrich effect. *Nature Neuroscience*, *4*, 513–518.

Bradley, D. C., Qian, N., & Andersen, R. A. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature*, *373*, 609–611.

Burr, D. C., & Ross, J. (1979). How does binocular delay give information about depth? *Vision Research*, *19*, 523–532.

Carney, T., Paradiso, M. A., & Freeman, R. D. (1989). A physiological correlate of the Pulfrich effect in cortical neurons of the cat. *Vision Research*, *29*, 155–165.

Cheng, K., Waggoner, R. A., & Tanaka, K. (2001). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron*, *32*, 359–374.

DeAngelis, G. C., & Newsome, W. T. (1999). Organization of disparity-selective neurons in macaque area MT. *Journal of Neuroscience*, *19*, 1398–1415.

DeAngelis, G. C., & Newsome, W. T. (2004). Perceptual 'read-out' of conjoined direction and disparity maps in extrastriate area MT. *PLoS Biology*, *2*, 394–404.

Grunewald, A., & Skoumbourdis, E. K. (2004). The integration of multiple stimulus features by V1 neurons. *The Journal of Neuroscience*, *24*, 9185–9194.

Harris, J. M., Nefs, H. T., & Grafton, C. E. (2008). Binocular vision and motion-in-depth. *Spatial Vision*, *21*, 531–547.

Maunsell, J. H. R., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, *49*, 1148–1167.

Neri, P., & Levi, D. M. (2008). Evidence for joint encoding of motion and disparity in human visual perception. *Journal of Neurophysiology*, *100*, 3117–3133.

Pack, C. C., Born, R. T., & Livingstone, M. S. (2003). Two-dimensional substructure of stereo and motion interactions in macaque visual cortex. *Neuron*, *37*, 525–535.

Pfeuffer, J., van de Moortele, P.-F., Yacoub, E., Shmuel, A., Adriany, G., Andersen, P., Merkle, H., Garwood, M., Ugurbil, K., & Hu, X. (2002). Zoomed functional imaging in the human brain at 7 Tesla with simultaneous high spatial and high temporal resolution. *NeuroImage*, *17*, 272–286.

Pulfrich, C. (1922). Die stereoskopie im dienste der isochromen und heterochromen photometrie. *Naturwissenschaften*, *10*, 553–564.

Qian, N. (1994). Computing stereo disparity and motion with known binocular cell properties. *Neural Computation*, *6*, 390–404.

Qian, N., & Andersen, R. A. (1997). A physiological model for motion-stereo integration and a unified explanation of the Pulfrich-like phenomena. *Vision Research*, *37*, 1683–1698.

Qian, N., & Freeman, R. D. (2009). Pulfrich phenomena are coded effectively by a joint motion-disparity process. *Journal of Vision*, *9*, 24, 1–16.

Read, J. C., & Cumming, B. G. (2005b). All Pulfrich-like illusions can be explained without joint encoding of motion and disparity. *Journal of Vision*, *5*, 901–927.

Read, J. C. A., & Cumming, B. G. (2003). Testing quantitative models of binocular disparity selectivity in primary visual cortex. *Journal of Neurophysiology*, *90*, 2795–22817.

Read, J. C. A., & Cumming, B. G. (2005a). The stroboscopic Pulfrich effect is not evidence for the joint encoding of motion and depth. *Journal of Vision*, *5*, 417–434.

Rokers, B., Cormack, L. K., & Huk, A. C. (2008). Strong percepts of motion through depth without strong percepts of position in depth. *Journal of Vision*, *8*, 6, 1–10.

van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt detectors. *J Opt Soc Am A*, *2*, 300–321.

Shioiri, S., Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research*, *40*, 2565–2572.

Tootell, R. B. H., & Taylor, J. B. (1995). Anatomical evidence for MT and additional cortical visual area in humans. *Cerebral Cortex*, *1*, 39–55.

Tyler, C. W. (1974). Stereopsis in dynamic visual noise. *Nature*, *250*, 781–782.

Verstraten, F. A., Verlinde, R., Fredericksen, R. E., & van de Grind, W. (1994). A transparent motion aftereffect contingent on binoc-

ular disparity. *Perception*, 23, 1181–1188.

Watson, A. B., & Ahumada, A. J. (1985). Model of human visual-motion sensing. *J Opt Soc Am A*, 2, 322–342.

Wilson, J. A., & Anstis, S. M. (1969). Visual delay as a function of luminance. *American Journal of Psychology*, 82, 350–358.

Zeki, S. M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology*, 236, 549–573.