Mixed visual reference frames: perceiving non-retino-centric visual quantities in a retino-centric frame

A.V. van den Berg[1] R. van Ee[2] and A.J.Noest[1]

1 Functional Neurobiology, Helmholtz Institute, Faculty Biology, University Utrecht, Hugo R. Kruijtgebouw, Padualaan 8, 3584 CH Utrecht, The Netherlands

2 Physics of Man, Buys Ballot Laboratorium, Postbus 80000, 3508 TA Utrecht

Correspondence should be addressed to A.V. van den Berg . E-mail: a.v.vandenberg@bio.uu.nl

Tel: +3130 2534230/4218, Fax: +3130 2542219

Running title:

Mixed visual reference frames

Publication:

Proc. SPIE Vol. 5666, p. 449-461, Human Vision and Electronic Imaging X; Bernice E. Rogowitz, Thrasyvoulos N. Pappas, Scott J. Daly; Eds.

Publication Date:

Mar 2005

Abstract

It is a useful competence to see motion relative to the head or to the external world, although those quantities are not directly given on the retina. The same holds for judgement of the shape of an object. We argue that the required transformations can be and are done independent of the associated direction transformations. This creates perceptual channels with retinal apertures but non-retinocentric motion- or shape-sensitivity.

In order to arrive at units that perform such a mixed transformation, the substructure of the retinotopic receptive field needs to be dynamically adjusted, using extra-retinal

signals (or equivalent measures like vertical disparity). Here we show that detectors tuned to disparity X retinal-direction can extract (metric) object curvature from the retinal disparity field in one step. We point out the correspondence to a previously proposed model of heading detection, which contains detectors that become tuned to head-centric flow by dynamically changing their preferred structure of the flowfield in a retinal aperture, depending on the eye movement.

Introduction

A recurrent question in visual neurophysiology deals with the reference frame. For example a neuron with auditory and visual sensitivity receives inputs that by their nature refer to different reference frames. To what reference frame does the combined activity refer to? This question is also interesting because it forms a natural point of contact to *perceptual* studies that deal with a similar question: what reference frame does perception of visual direction or depth refer to? We begin to establish that the question of THE reference frame is ambiguous, and needs further specification. We proceed to discuss two examples of visual perception based on mixed reference frames, i.e. perception of visual quantities that refer to one reference frame are established through viewing apertures that refer to another reference frame. Using these perceptual results, we discuss a model how neurons with mixed reference frames as defined above could be constructed.

Visual receptive fields and visual trigger properties

Typically when one attempts to establish a neuron's reference frame one asks whether the neuron's viewing aperture or 'receptive field' maintains a fixed direction with respect to some landmark (for example the fovea) when the eye looks in different directions. When it does, the neuron has a retinal receptive field. In contrast, one concludes that the receptive field of the neuron is head-centric when under various manipulations, like rotation of the eye, the head or the body, the neuron is activated only when the visual stimulus is placed at the same location relative to some head-centric landmark like the nose.

In perceptual studies a similar logic is employed using the instrument of adaptation. One deals not with the activity of a neuron, but with a 'perceptual channel' that likely corresponds to some mass activity of the brain. Prolonged stimulation of that perceptual 'channel' that is tuned to some visual quantity (e.g. motion) leads to a reduction of its sensitivity to a test stimulus: adaptation. The reference frame of the perceptual channel can now be established by asking whether the test stimulus must be presented at the same retino- (or head- or body-) centric *location* as the adapting stimulus to observe the reduced sensitivity.

The question whether the *visual stimulus* that triggers the neuron (or the perceptual channel for that matter) refers to the retina or the head is often ignored in such studies. For simple flashed point targets this causes little harm. After all, the flash is too limited spatially and temporally to have any such bearing independent of its location. Natural vision however, deals with much more complicated visual structure than flashed points. When for example visual motion triggers

the neuron or perceptual channel the question to what reference frame the visual trigger belongs becomes highly relevant. Visual motion can reflect the motion relative to the retina, motion relative to the head, motion relative to the hand etc. Likewise, binocular depth sensitivity can refer to distance relative to the fixation point, distance relative to the head or relative to some external object. The question for the reference frame of a perceptual channel or neuron thus entails two parts: what is the reference frame of the viewing aperture and what reference frame does the preferred visual stimulus within that aperture refer to? If these reference frames are the same we speak of a 'natural' reference frame representation otherwise the representation is 'mixed'.

Perception in a Mixed reference frame

Heading perception from visual flow uses a mixed reference frame

A display of camera motion through a real or virtual environment can produce a vivid impression of self-motion. This self-motion percept results from the visual flow, that radiates outward from the heading direction when the camera makes a pure translation through the environment (Gibson, 1966). Importantly, the apparent self-motion direction is based on the pattern of motion relative to the head, because perceived heading does not change when the subject chooses to pursue an object in the environment to the side of the path. This pursuit eye movement causes a shift of the centre of the radial pattern on the retina relative to the heading direction. The absence of a shift in the percept (Royden, Crowell & Banks, 1994, van den Berg, 1992, Warren & Hannon, 1990), is evidence that eye signals compensate for the change of the retinal flow.

Compensatory shifts in tuning of cells in area MST of the monkey, for the centre of a radial pattern of motion have been reported (Bradley, Maxwell, Andersen, Banks & Shenoy, 1996, Shenoy, Bradley & Andersen, 1999) but have been disputed by others (Page & Duffy, 1999, Upadhyay, Page & Duffy, 2000). All these reports deal with the question whether the neural signals code retinal or head-centric velocity, but the question of the reference frame for the aperture is not dealt with.

Psychophysical studies indicate that motion adaptation occurs within retinocentric locations, suggesting that heading perception from flow may also occur through retinotopic apertures. (van den Berg & Beintema, 2000)did a study that did deal with heading perception (Fig. 2). They investigated precision of heading perception for flow patterns that were shown within a small aperture on the screen. For moving and fixating eyes alike, precision was highest when

the apertured pattern of motion contained a centre on the *retina*. When the *retinal* pattern of flow was tapering towards a centre outside the aperture, precision of the heading percept was reduced. Because the eye movement shifts the center of flow on the retina relative to the centre of flow on the screen, this shows that the *retinal* pattern of flow limits precision not the pattern of flow on the screen. (van den Berg & Beintema, 2000) concluded that a representation of the head-centric flow is built from motion channels that register the retinal *pattern* of flow. The component of the motion pattern caused by eye movement is inaccessible through a head-centric receptive field, because that RF is not moving relative to the head by definition. Thus, the retinal flow limitation on the precision of heading perception points to a *retinal position reference of head-centric flow signals*. This conclusion is supported by a recent study by (Poljac & van den Berg, 2003). They found that pointing towards memorized perceived heading direction showed the same type of retino-centric errors following a saccadic eye movement as pointing towards single LED targets. Because this error is believed to reveal the effect of a retino-centric updating process for eye movements it implicates that the heading direction was coded in a retino-centric format, at least for this pointing task. Hence, we conclude that a mixed reference frame representation for heading perception is present at least in humans.

Shape perception from binocular images

When you observe your friend on the platform while she waves you good by, you would be surprised to see her turn into a card-bord figure while your train leaves the station. Yet, this is precisely what one would predict when the visual system would merely rely on retinal disparity to judge your friend's waist size.

This holds, because retinal disparity scales differently ($\propto 1/R^2$) with distance than the angular extent of an object($\propto 1/R$). Thus, the stereoscopic depth signals shrink much faster than the object's size. If the visual system would rely simply on the ratio of retinal disparity and size to represent 3D shape, an object would be seen to flatten as it recedes in the distance. Apparently, perception relies on more sophisticated methods.

A related problem forms the change of the curvature of the zero-disparity loci in space as a function of the fixation distance. This means that a patch of a fronto-parallel surface that is visible through some aperture of fixed size does not maintain a fixed distribution of disparities for changing distance. Thus by the disparity field alone, one would not even see a true card-bord figure as flat from the departing train.

Several algorithms have been proposed in the past to solve this problem. One line of approach

(e.g. (Garding, Porrill, Mayhew & Frisby, 1995, Mayhew & Longuet-Higgins, 1982)) corrects the disparity with the aid of the eyes' orientations for the distance dependent curvature and slant of the zero-disparity surface. Metric depth -to be used for the final shape computation-is only achieved when the corrected disparities are scaled by a distance dependent factor.
Another line of approach computes the second-order horizontal spatial-derivative of disparity (disparity curvature;(Rogers & Cagenello, 1989)), to arrive at the object curvature directly. However, disparity curvature is proportional to metric object curvature only for a small range of eye version and for highly curved objects. More importantly, this scheme totally ignores the effect of the object's slant on the relation between object curvature and disparity curvature. Thus, Rogers & Cagenello's scheme provides reliable measures of object curvature only for special conditions. Our analysis (presented below) provides a method to relieve these restrictions.
Either of above mentioned two methods are examples of 'mixed' transformation methods, because they arrive at depth estimates (object centred distances) for specific binocular directions.
A third and entirely different proposal uses a 'normal' transformation. In (Erkelens & van Ee, 1998) each eye's retinal signal is combined with that eye's orientation (using oculo-motor signals) to compute the visual direction relative to the head for that eye. In a next stage, the head-centric disparity is computed by subtraction of the head-centric direction signals of the two eyes. Head-centric disparity provides a measure for the head-centric distance, because it relates to the vergence angle required to fixate any visual object in view. Such a field of head-centric distances can then provide the basis for shape perception.

It seems of some interest to establish whether human shape perception uses mixed- or natural reference frames. We already know from two adaptation studies (Domini, Adams & Banks, 2001, Duke & Wilcox, 2003) that perception of a cylindrical surface adapts to the apparent curvature of the surface. Domini et al decoupled disparity curvature and object curvature by presenting curved objects at one adaptation distance and testing at another distance. Duke and Wilcox manipulated perceived curvature of an object by manipulation of the vertical disparity field, producing stimuli with the same apparent curvature for different (horizontal) disparity curvatures. Both studies found adaptation correlated better to the object's apparent curvature than to the presented horizontal disparity curvature.
These studies did not establish whether the adaptation occurred in a retinal or head-centric viewing aperture. Nor did theses studies differentiate between contributions of disparity, and its first or 2nd order derivatives to the adapting channel. To settle some of these issues we

performed an adaptation study.

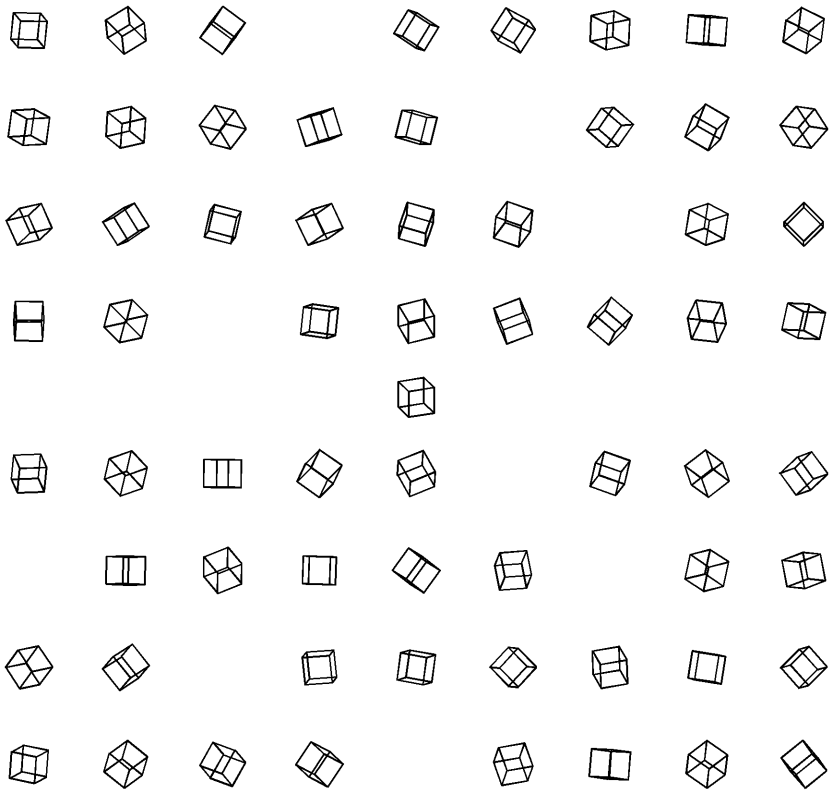

Figure 1. The pattern of randomly oriented cubes. The central row consisted of just one cube that moved horizontally and in depth during the pursuit condition. Subjects were instructed to fixate this cube during adaptation. The adaptation pattern was presented on the left or the right half of the screen, with the remainder of the screen blank. During subsequent testing the pattern was either switched to the other half (transfer condition) or it remained in place. For each switch between test and adaptation pattern a new set of cubes was computed.

Adaptation Experiment

Experimental Procedures

Subjects were seated in a dark room in front of a translucent screen at 1.5 m distance. The head was stabilized by a chin rest. Dichoptic displays (red-green anaglyphs) were generated at a framerate of 75 Hz through an Apple Macintosh G4 computer. Screen dimensions were about 60 deg horizontally x 40 deg vertically. Display resolution was 1024 x 768 pixels. Stimuli were back-projected on the screen by a JVC D-ILA projection monitor. Four subjects participated in the experiment.

Stimuli consisted of a collection of wire-frame cubes (separation: 3 deg; size:1 deg), simulated on a rectangular grid of binocular viewing directions(Fig. 1). Each cube was rotated in 3D about its centre over a random angle about a randomly chosen axis. The binocular distance of the cube corresponded to a location on a cylindrical surface. This arrangement provides a curved cylindrical surface without a density gradient and with the relative disparity of neighbouring cubes' corners unrelated to the global shape.

The adaptation stimulus was presented initially for 100 s followed by 10 s top ups alternated with 2 s test stimuli. The curvature of the successive test stimuli was adjusted in a staircase procedure to null the after-effect

and find the apparent 'flat' stimulus. Two successive staircases paired conditions that were identical in all respects except for the sign of the curvature of the adapting surface.

The difference of each pair of curvature s that nulled the aftereffects was taken and normalized with respect to the curvature difference between the two adapting stimuli (Fig. 2). We call this response measure 'the coefficient of adaptation': CA. This simple subtraction removes the bias of subjects to perceive convex or concave surfaces.

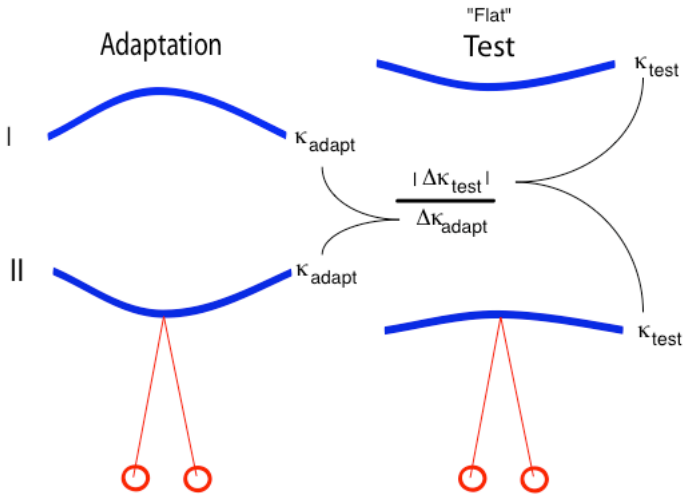

Figure 2.
The curvature of a test pattern (right half of the figure) was adjusted in a staircase procedure until it appeared flat. Two successive staircases paired conditions that were identical in all respects except for the sign of the curvature of the adapting surface (pattern I or pattern II). The pair of settings was subtracted and normalized by the difference in curvature of the adapting patterns

Conditions and motivation

We investigated 2x2x4 conditions:

2 adaptation shapes: convex/concave with curvature ±0.75/m

2 types of fixation during adaptation : pursuit along a Lissajous trajectory with components of motion in depth (Δvergence:1.6 deg) and lateral motion (Δversion: 25 deg) or steady fixation. During the test, fixation was steady. Each test phase contained one cycle of the Lissajous motion.

2x2 types of adaptation->test combinations: (R->R), (L->L) (R->L) or (L->R). (R: 15 deg right, L: 15 deg left of the screen's centre). During adaptation and testing the subject fixated a fixation cube that was located at the centre of the stimulus patch or moved within the boundaries of the stimulus patch.

These conditions allowed us to make the following qualitative predictions:

If adaptation occurs in head-centric apertures then $CA_{R->L}$ and $CA_{L->R}$ should be close to zero because the adaptation does not transfer between different head-centric directions.
If adaptation occurs in retino-centric apertures, the level of adaptation for the 'transfer' ($CA_{R->L}$ and $CA_{L->R}$) and the non-transfer conditions ($CA_{R->R}$ and $CA_{L->L}$) should be equal.
If retinal disparity or its gradient contributes significantly to adaptation, then moving-fixation should reduce CA relative to steady fixation because the lateral eye movement smears out retinal disparity and its gradient (local change 5 times the difference between the convex and the concave stimulus). The change in disparity curvature during the pursuit eye movement was only a fraction (25%) of the difference in disparity curvature between the convex and the concave adapting stimulus.

Results

Data were collected in several sessions with four repeats of each of the 16 conditions. For each subject we present CA values averaged across repeats and pairs of transfer/no-transfer conditions. Figure 3 summarizes the results.
The data for fixation and pursuit cluster about the diagonal, indicating that the adaptation is equally strong for 'transfer' and 'no transfer' conditions. This implies that curvature is perceived through detectors with retino-centric NOT head-centric apertures.
Note that the CA values for pursuit are systematically lower than corresponding CA for fixation. The difference is relatively small however with $Ca_{fixation}$ - $CA_{pursuit}$ ranging between 0.02 and 0.06. Thus, our results indicate a relatively minor contribution of disparity or disparity-gradient related mechanisms to shape adaptation. Our results are consistent with dominance of curvature based adaptation. Combining with earlier results on curvature adaptation (Domini et al., 2001, Duke & Wilcox, 2003) we conclude that perceptual channels tuned to object curvature use retino-centric apertures or receptive fields.

Disussion

We have presented two examples of perceptual studies that lead to the conclusion that the visual system uses mixed-reference frame coding, meaning that within retino-topic apertures visual quantities are encoded that refer to non-retinal reference frames.
A 'normal' transformation scheme that would operate through head-centric (or allo-centric) receptive fields would need to deal with the complications that arise from the inhomogeneous retina. If a head-centric aperture is constructed from the retinal input combined with eye position signals the resolution that the visual system can supply within that head-centric receptive field varies with the eye position. Thus, the computations within that receptive field must be able to

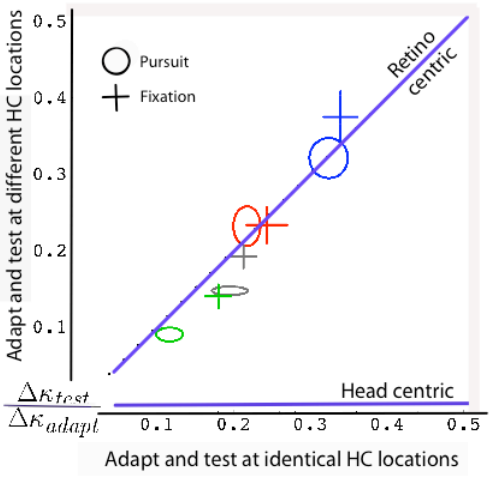

Figure 3. Results of curvature adaptation experiment of four subjects. Each symbol is the pair of coefficients of adaptation for transfer and non-transfer conditions. Ovals indicate pursuit and crosses fixation conditions. Predicted outcomes if apertures of the adapting channels are retino-centric or head-centric are indicated by the continuous lines

support variable resolution of the input signals or if a fixed resolution is used, the fine detail that the visual system offers around the fovea must be blurred to the lower resolution supplied by more eccentric retinal locations which contribute to the head-centric receptive field for certain eye positions. In contrast, mixed coding schemes respect the retinal inhomogeneity, and can profit from the fine detail that the visual system offers in its central view while the same computational structure can operate at a coarser scale in the periphery. Thus, the increased visual resolution at low eccentricity becomes available also for non-retinal visual quantities like the shape and motion of objects. We proceed by briefly describing two transformation schemes that result in mixed-reference frame coding for head-centric flow and object shape.

Transformation to head-centric motion sensitivity

The goal of models for heading perception is to design a scheme that shows invariant activity in some motion template for which the preferred heading corresponds to the actual heading. The retinal motion pattern is not invariant when the eye rotates, causing a shift or degeneration of the radial pattern of flow on the retina relative to the radial motion when the eye is not rotating. One approach assumes a local nature for the compensation, using eye signals to subtract vectorially the effect of the eye rotation from the retinal flow. This local vector-subtraction type of compensation appears to be barred by the observation that the radial pattern of *retinal* flow not head-centric flow limits heading precision (van den Berg & Beintema, 2000). This means that the interaction between eye movement signals and retinal flow occurs rather late, when integration of local flow signals into pattern of motion sensitive templates has occurred. These template signals are then combined into head-centric headcentric flow sensitive units.

The change of the flow pattern depends on the depth structure of the scene as well as the direction and speed of translation and rotation. Thus, attempts to compensate for these changes have involved laborious template schemes that include a wide range of rotation conditions and depth structures (e.g.(Stone & Perrone, 1997)). Essentially such schemes rely on a winner-take-all competition between the various templates.

A more parsimonious type of coding was advocated by (Beintema & van den Berg, 1998)as described below. An array of motion templates covers the set of different heading directions relative to the retina. Each of these expansion templates is backed up by a pair of rotation templates that prefer opposite directions of rotation about the same axis. The rotation templates integrate motion along circles concentric with the preferred retinal direction of the centre of the 'governing' expansion template. The local motion preference along each circle is tangential to the circle. Moreover, the rotation template is unresponsive to the magnitude of the expansion component of the flow, that is perpendicular to the circle. The activity of opponent rotation templates is subtracted and gain modulated by an eye velocity signal. Finally, this signal is added to the response of the retinal expansion template to create a template that is tuned to a particular direction of heading irrespective of the eye rotation. Note, however, that the receptive field of each motion template still refers to the retinal reference frame. Thus the centre of expansion is still coded by a retinal direction line.

The tuning to the pattern of retinal flow becomes dynamic by the eye-velocity modulated rotation templates. When moving to the right(left) the preferred centre-location of the expansion on the retina shifts right(left)ward. (Beintema & van den Berg, 1998)showed that this model could reproduce virtually perfect heading perception for real eye pursuit during simulated forward motion and a reduction of that capacity when eye rotation was simulated through a 3D environment. Large errors occurred both in humans and in the model when depth differences in the scene were very limited. Recent fMRI studies have found modulation of a large part of human area MST by eye movement signals whereas no such modulation is found in area MT (Goossens et al, in prep.).

Transformation to object curvature sensitivity

We focus on object curvature from disparity that is collected within the plane of regard, or any elevation plane that is derived by rotation of the plane of regard about the interocular axis. Thus our solution is limited to the curvature of the intersection of the object's surface and an elevation plane. The curvature of the intersection between the object and a plane that is tilted relative to the plane of regard will be considered elsewhere.

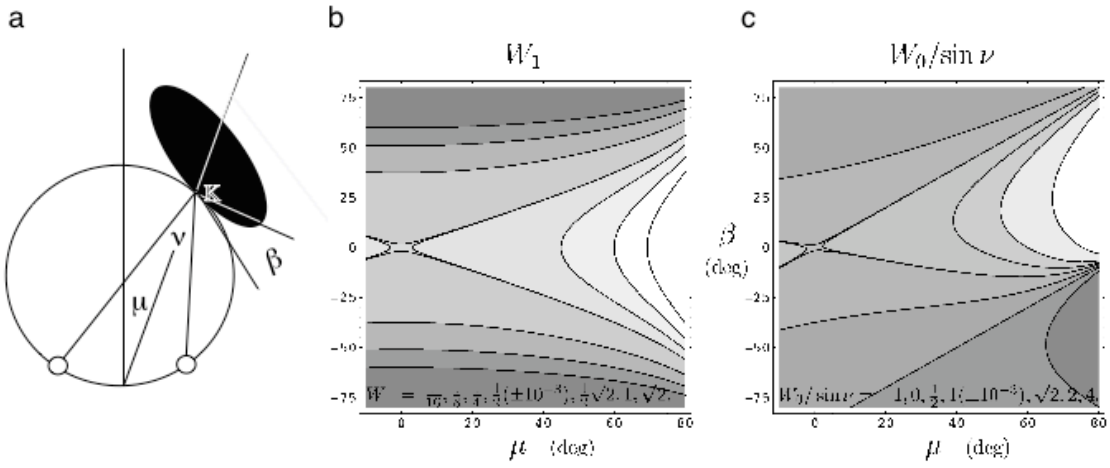

Figure 4. (a)Binocular viewing geometry: version (μ) vergence (ν) and slant (β) angles.
K denotes the object curvature that is computed from retinal disparity ($\nu_{\mu\mu}$)curvature.
(b) The linear gain W_1 (of $\nu_{\mu\mu}$) as a function of version and slant. (c) The correction W_0 (of $\nu_{\mu\mu}$) as a function of version and slant. The effect of eye vergence is an overall scaling by sin(ν).

The object's curvature (K) within the plane of regard scaled by the half of the interocular distance (a) is given by:

$$aK = \frac{b\nu_{\mu\mu} + c}{d}$$

$$b = 0.5\cos^2(\mu) + O(\nu^2)$$

$$c = \sin(\nu)\left(\cos^2(\mu) + 0.5\sin(2\mu)\tan(\beta)\right) + O(\nu^2)$$

$$d = \left(\frac{\cos(\mu)}{\cos(\beta)}\right)^3 + O(\nu^2)$$

where μ denotes the version angle, ν denotes the target vergence (eye vergence + horizontal disparity) and β, the slant angle relative to the cyclopean viewing direction (Fig.4a). This is a far-field approximation, but the error in K resulting from the ($O(\nu^2)$) terms exceeds the curvature discrimination threshold only when $\nu > \pi/2$. Clearly, object curvature is a linear function of the disparity curvature ($\nu_{\mu\mu}$), that is modulated by version, vergence and the slant of the object. Rogers & Cagenello's proposal replaces our expression '$W_1 = b/d$' by a constant and sets '$W_0 = c/d$' to zero, thus they ignore the important contributions of version and vergence. These modulation functions as shown in figure 4b and 4c indicate that significant deviations occur in the normal oculomotor range from W_1 = constant and '$W_0 /\sin(\nu) = 0$. Another important advance with respect to the proposal by(Rogers & Cagenello, 1989) is our quantification of the effect of the object's slant. Object curvature from disparity computations must take into account the object's slant.

The need to take into account the effect of slant is further emphasized when one realizes that the perceived curvature of an object like a shell flips sign when it is rotated about an axis perpendicular to the plane of regard. This sign flip occurs in the region where either eye's viewing direction grazes at the object's surface. Thus, our analysis brakes down in narrow strips near to the contour of the object (at $|\beta| \approx \pi/2$).

Subsequently, we briefly describe a neural-like model that picks up the object's curvature from the retinal disparity field in the neighbourhood of the viewing direction and from signals regarding the viewing parameters (version and vergence from e.g. the vertical disparity field or oculomotor signals). A full derivation and motivation can be found in Noest et al. (under review).

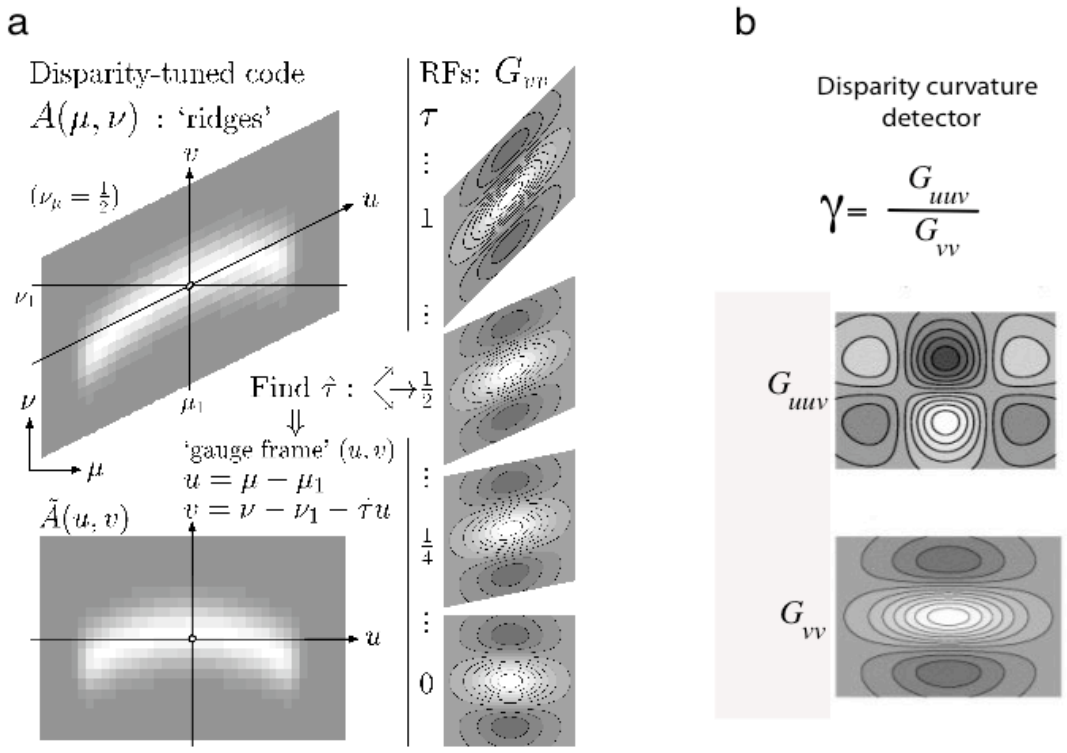

Figure 5. (a) Curved bright ridges in the left panels denote patterns of activity (A(μ,ν)) evoked by a curved object in a patch of retinotopically arranged, disparity tuned detectors. The neighbourhoods of some base disparity ($v = \nu - \nu_1$) and some base retinal direction ($u = \mu - \mu_1$) are depicted. The two panels illustrate the activity ridges for two objects with different slant ($\nu_\mu = \tau = 0$ or 0.5). The receptive fields that are tuned to different slants of the ridge are depicted on the right ($\tau = 0, 0.25, 0.5, 1.0$). Note that these receptive field are related to one another by a shear, NOT a rotation.
(b) The components of a disparity curvature detector. Two receptive fields integrate the activity in the neighbourhood of the base disparity and base retinal direction ($G_{uuv} * A$ and $G_{vv} * A$). The ratio of these integrated activities is a measure of the curvature (γ) of the disparity ridge.

For the visual part of the implementation of above equations one needs detectors tuned to the disparity curvature and the slant plus disparity of the object. The disparity curvature can be found by analogy to the detection of line curvature (Koenderink & Richards, 1988), using two-dimensional operators, constructed by differentiation of the Gaussian kernel. Now, the two-dimensional operators $\delta^2_\mu\delta_\nu G(\mu,\nu)$ and $\delta^2_\nu G(\mu,\nu)$ (shorthand: $G_{\mu\mu\nu}$ and $G_{\nu\nu}$) integrate the activities of detectors tuned to a particular combination of disparity and retinal location. Specifically, curvature of the ridge of activity is proportional to $G_{\mu\mu\nu} / G_{\nu\nu}$.
The slant of the object is not directly given, but slant tuning of a neuron can be achieved through dynamic selection of the appropriate disparity gradient detector, using version and vergence as control signals. This follows from the geometric expression for the slant (β) as a function of the disparity gradient (ν_μ):

$$\nu_\mu = \tan(\nu)\,(\tan(\beta) - \tan(\mu))$$

This formula specifies how for given β, the corresponding ν_μ depends on the viewing parameters (μ,ν).
For each slant a separate 'rotated' curvature operator is built that is optimally tuned to the curvature of the slanted surface. The rotation of the detectors is not Euclidian, but respects the binocular viewing geometry. I.e. it conforms to the requirement that curvature detection is limited to slants that differ from $\pi/2$ and that curvature flips sign when that slant extremum is traversed. Rotated copies of the detectors are related to one another by a shear 'τ' $(\mu,\nu) \rightarrow (\mu, \nu-\tau\mu)$ of the coordinate frame (Fig. 5).
Note that in order to be consistent with the findings that adaptation occurs to object curvature rather than disparity curvature, one is obliged to construct a single detector that directly extracts object curvature from the retinal disparities. Intermediate stages would show up in the adaptation experiments of Domini et al., (2001), and Duke & Wilcox, (2003)as support for disparity curvature adaptation. Direct transformation to object curvatures can be done as follows:

$$aK = W_1 \frac{G_{\mu\mu\nu} * A}{G_{\nu\nu} * A} + W_0 = \frac{(W_1 G_{\mu\mu\nu} + W_0 G_{\nu\nu}) * A}{G_{\nu\nu} * A}$$

where A stands for the activities of the disparity tuned units that provide input to the Gaussian operators.
Note that the denominator is a simple line (in disparity x direction space) detecting operator, while $(W_1 G_{\mu\mu\nu} + W_0 G_{\nu\nu})$ represents a *dynamic* receptive field that adjusts its substructure depending on the ratio of W_1 and W_0. The gain factors W_i (i=0,1) depend smoothly on the viewing parameters and the slant of the object as shown in Fig. 4b,c.

Because the detector extracts object curvature directly from the disparity field, no depth computations and their intrinsic vergence-error sensitivity are involved. The model is essentially blind to depth, because it goes straight for the curvature. Our model does imply slant sensitivity, but the slant is not a precursor to the curvature computation. It only serves to 'select' the disparity curvature detector with the appropriate shear. If this slant selection is crude, little error in the curvature computation is introduced. This property of the model allows it to live happily with the fact that precise curvature discrimination can occur while depth estimates are way off.

Thus, we observe that both the shape from disparity and the heading from visual flow problems can be solved through dynamic tuning by modulating eye signals. The receptive fields of the units remain retino-topic, but the activity of these hypothetical units reflects visual quantities that are lifted out of the retinal reference frame.

Acknowledgement.
This work was supported by a Human Frontier program grant to Albert van den Berg and NWO-VIDI grant to Raymond van Ee.

References

Beintema, J.A., & van den Berg, A.V. (1998). Heading detection using motion templates and eye velocity gain fields. *Vision Res, 38* (14), 2155-2179.
Bradley, D.C., Maxwell, M., Andersen, R.A., Banks, M.S., & Shenoy, K.V. (1996). Mechanisms of heading perception in primate visual cortex. *Science, 273* (5281), 1544-1547.
Domini, F., Adams, W., & Banks, M.S. (2001). 3D after-effects are due to shape and not disparity adaptation. *Vision Res, 41* (21), 2733-2739.
Duke, P.A., & Wilcox, L.M. (2003). Adaptation to vertical disparity induced-depth: implications for disparity processing. *Vision Res, 43* (2), 135-147.
Erkelens, C.J., & van Ee, R. (1998). A computational model of depth perception based on headcentric disparity. *Vision Res, 38* (19), 2999-3018.
Garding, J., Porrill, J., Mayhew, J.E., & Frisby, J.P. (1995). Stereopsis, vertical disparity and relief transformations. *Vision Res, 35* (5), 703-722.
Koenderink, J.J., & Richards, W. (1988). Two-dimensional curvature operators. *J Opt Soc Am, 5*, 1136-1141.
Mayhew, J.E., & Longuet-Higgins, H.C. (1982). A computational model of binocular depth perception. *Nature, 297* (5865), 376-378.

Page, W.K., & Duffy, C.J. (1999). MST neuronal responses to heading direction during pursuit eye movements. *J Neurophysiol, 81* (2), 596-610.
Poljac, E., & van den Berg, A.V. (2003). Representation of heading direction in far and near head space. *Exp Brain Res, 151* (4), 501-513.
Rogers, B., & Cagenello, R. (1989). Disparity curvature and the perception of three-dimensional surfaces. *Nature, 339* (6220), 135-137.
Royden, C.S., Crowell, J.A., & Banks, M.S. (1994). Estimating heading during eye movements. *Vision Res, 34* (23), 3197-3214.
Shenoy, K.V., Bradley, D.C., & Andersen, R.A. (1999). Influence of gaze rotation on the visual response of primate MSTd neurons. *J Neurophysiol, 81* (6), 2764-2786.
Stone, L.S., & Perrone, J.A. (1997). Human heading estimation during visually simulated curvilinear motion. *Vision Res, 37* (5), 573-590.
Upadhyay, U.D., Page, W.K., & Duffy, C.J. (2000). MST responses to pursuit across optic flow with motion parallax. *J Neurophysiol, 84* (2), 818-826.
van den Berg, A.V. (1992). Robustness of perception of heading from optic flow. *Vision Res, 32* (7), 1285-1296.
van den Berg, A.V., & Beintema, J.A. (2000). The mechanism of interaction between visual flow and eye velocity signals for heading perception. *Neuron, 26* (3), 747-752.
Warren, W.H., Jr., & Hannon, D.J. (1990). Eye movements and optical flow. *J Opt Soc Am A, 7* (1), 160-169.