

# Attentive visual motion processing: computations in the log-polar plane

Kostas Daniilidis  
Computer Science Institute  
Christian-Albrechts University Kiel  
Preusserstr. 1-9, 24105 Kiel, Germany

## Abstract

Attentive vision is characterized by selective sensing in space and time as well as selective processing with respect to a specific task. Selection in space involves the splitting of the visual field in a high resolution area - the fovea - and a space-variant resolution area - the periphery. Both in neurobiology and in robot vision, models of the resolution decrease towards the image boundaries have been established. The most convincing model is the theory of log-polar mapping where very high data compression rates are achieved.

In combination with the complexity reduction we believe that the log-polar mapping has further computational advantages which we elaborate in this study. Based on the optical flow we study the computation of 3D motion and structure globally and locally. We present a global method to compute the Focus of Expansion in the case of pure translation. By fixating on an object we show how to estimate ego motion in the presence of translation and rotation of the observer from the flow in the log-polar periphery. Then, we turn to local differential computations and we establish both approximate and exact expressions for the time to collision.

## 1 Introduction

It is well known that when a human observer has to accomplish a navigational task like car driving the retinal information of the entire field of view is used for the estimation of 3D-motion and the control of the vehicle. It is also well known and easy to show experimentally that there is an immense decrease of the resolution towards the periphery of the retina. However, the coarse resolution on the major part of the field of view does not hinder the observer to detect alarming events and to exploit the image periphery for the estimation of ego-motion. In the same scenario, if the car-driver approaches a crossing he tries to estimate the relative motion of the cross-moving cars in order to react appropriately. Then the driver directs his gaze towards the moving object and pursues it for a short time interval. Such a gaze change and holding are carried out with a navigational and not a recognition purpose: the driver needs the information how the cross-moving car is moving rather than how it looks like.

In this paper, we study the properties of the motion field using a space variant resolution scheme, and we propose algorithms for the recovery of 3D-motion which exploit the advantages of space variant sensing and fixation. Space variant sensing is one of the main features of an attentive vision system. Attention arises as a necessity in building robot systems that are able to react according to the visual stimuli in a dynamically changing environment. Even with the fastest image processing architectures it is not possible to process a uniformly sampled image with a wide field of view in a process time enabling immediate reaction. To overcome the reaction time constraints we have to focus on a region of interest. However, even if we select a subset of the retinal information we have still to process the image data outside the focus of attention to detect new events. We, thus, need an image data reduction scheme which conserves high resolution in the center of the image - the fovea- and a gradually decreasing resolution area - the periphery. Since most visual motion tasks necessitate a wide field of view we have to elaborate algorithms

for computing the optical flow and recovering the 3D-motion and structure using space-variant resolution. In this paper, we use the complex logarithmic mapping (or log-polar transformation) to model the non-uniform periphery of an image, and we keep the cartesian model for the uniform fovea. However, we do not only elaborate motion algorithms regarding the logarithmic polar plane as a practical necessity. We show that the log-polar mapping has also computational advantages with respect to motion recovery tasks:

- We propose an algorithm for the computation of the translation direction in case of pure translation using only 4% of the original data.
- We show that fixation enables the computation of the translation direction in case of general motion. Pursuit and optokinetic eye movements have long been considered by the psychophysicists [6] as a means to overcome the limited field of view, to bound the binocular disparity, and to eliminate the motion blur. Here, we prove that eye movements are not only an implication of the inhomogeneous retina, but they support the recovery of 3D-motion.
- In analogy to the approach in [23] we prove that the polar motion field enables the local computation of bounds to the time to collision which are independent of rotation and the slant of the viewed surface.

We proceed with a survey of approaches to space variant sensing and fixation. The first experimental findings on the density of the photoreceptors in the human retina are due to Schulze [18] - for a comparison of further biological findings the reader is referred to [6, 25]. Space variant arrangements concern not only the density of the photoreceptors but the geometry of the receptive fields in the cortex, too. Such retino-cortical mappings have been explicitly formulated by Schwartz [19] for cortical areas of the monkey and by Mallot [13] for cortical areas of the cat. Schwartz [19] proposed the complex logarithmic mapping and showed its invariance properties. The amount of literature on biological work is very large and we will not attempt to give a further review.

Weiman and Chaikin [28] studied the properties of the complex logarithmic transformation as a conformal mapping and they proposed a logarithmic spiral grid as a digitization scheme for both image synthesis and analysis. Sandini et al. [14] studied the recognition of 2D-patterns in the log-polar plane and they proposed optimality criteria for the subsampling of the non-foveal part of the image. The description of the first CCD-sensor with a cartesian fovea and a log-polar periphery can be found in [24]. Wallace et al. [26] implemented the log-polar mapping on a DSP-architecture by averaging sets of CCD pixels. They integrated the log-polar mapping into a miniaturized pan-tilt camera enabling a visual communication using voice-bandwidth channels.

Jain et al. [10] were the first who applied the complex logarithmic mapping in image sequence analysis. Given the motion information from other sensors and assuming pure translation they transformed the images into the log-polar plane using the focus of expansion as the center of the transformation. The depths of the viewed point were easily obtained by inspecting only the shifts along one coordinate of the log-polar plane. Tistarelli and Sandini [22, 23] derived the motion equations for the log-polar plane and proposed a method for the computation of the time to collision. Our work on local computations (section 4) is inspired by their approach. The combination of fixation and space variant sensing appears only in [22]. The following references concern fixation in the cartesian plane. Aloimonos et al. [1] and Bandopadhyay and Ballard [2] showed how the structure from motion problem is simplified in case of active tracking. Fermüller and Aloimonos [8] proved that normal flow measurements are sufficient for bounding the locus of the focus of expansion and the time to collision. Raviv and Herman [17] studied the geometric properties of fixation and derived the locus of 3D-points whose projections yield a zero-optical flow in case of motion with three degrees of freedom (road following). Taalebinezhaad [21] proposed a method that simulates tracking by a pixel shifting process without actively controlling the mechanical degrees of freedom of the camera. The translation direction is then obtained by minimizing the depth deviation from the fixation point using only normal flow measurements. A similar method was proposed in [9] for the derotation of the cartesian motion field.

The organization of the paper is as follows. First, we formally describe the complex logarithmic mapping. We derive the equations of the motion-field in the log-polar plane in section 3. The rest of the study is divided to global motion computations (subsections of 3) and local parallax computations (section 4).

## 2 The complex logarithmic mapping

We use  $(x, y)$  for the cartesian coordinates and  $(\rho, \eta)$  for the polar coordinates in the plane. By denoting with  $z = x + jy = \rho e^{j\eta}$  a point in the complex plane the complex logarithmic (or log-polar) mapping is defined as

$$w = \ln(z) \quad (1)$$

for every  $z \neq 0$  where  $\text{Re}(w) = \ln\rho$  and  $\text{Im}(w) = \eta + 2k\pi$ . To exclude the periodicity of the imaginary part we constrain the range of  $\text{Im}(w)$  to  $[0, 2\pi)$ . The complex logarithmic mapping is a well-known conformal mapping preserving the angle of the intersection of two curves. It is illustrated in Fig. 1 where  $\xi$  denotes  $\ln\rho$ . It is trivial to show that every scaling and rotation about the origin in the  $z$ -plane is represented in the  $w$ -plane by a shift parallel to the real and imaginary axis, respectively.

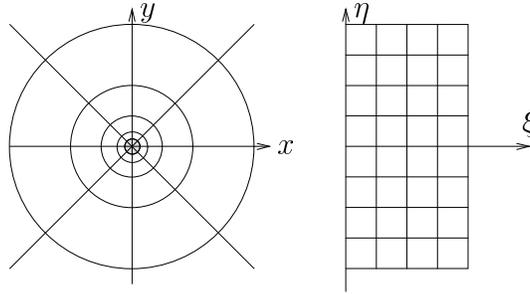


Figure 1: The complex logarithmic mapping maps radial lines and concentric circles into lines parallel to the coordinate axes. The log-polar variables  $\xi$  and  $\eta$  are defined in (2).

We apply the log-polar mapping on the non-foveal part of a retinal image. Therefore, we define as the domain of the mapping the ring-shaped area  $\rho_0 < \rho < \rho_{max}$  where  $\rho_0$  and  $\rho_{max}$  are the radius of the fovea and the half-size of the retinal image, respectively. Furthermore, a hardware CCD-sensor with the log-polar property or a software implementation of the mapping needs a discretization of the  $w$ -plane -which we will call log-polar plane in contrast to the cartesian plane. By assuming that  $N_r$  is the number of cells in the radial direction and  $N_a$  is the number of cells in the angular direction the mapping from the polar coordinates  $(\rho, \eta)$  to the log-polar coordinates  $(\xi, \gamma)$  reads as follows (see also [22])

$$\begin{aligned} \xi &= \log_a\left(\frac{\rho}{\rho_0}\right) \\ \gamma &= \frac{N_a}{2\pi}\eta \end{aligned} \quad (2)$$

where the logarithmic basis  $a$  is obtained from the foveal radius  $\rho_0$ , the image radius  $\rho_{max}$  and the radial resolution  $N_r$

$$a^{N_r} = \frac{\rho_{max}}{\rho_0} \quad \text{or} \quad a = e^{\frac{1}{N_r} \ln\left(\frac{\rho_{max}}{\rho_0}\right)}. \quad (3)$$

From now on we will use only  $\eta$  ranging from 0 to  $2\pi$  for the angular component of the motion field vector.

The mapping of the gray-value function  $I(x, y)$  in the cartesian plane to the gray-value function  $J(\xi, \eta)$  in the log-plane is by no means trivial. This issue concerns the software implementation

of the log-polar mapping given a cartesian image. Every log-polar cell corresponds to a receptive field in the cartesian plane. The image  $J(\xi, \eta)$  is the result of a space-variant filtering that affects all subsequent computations on the log-polar plane like spatiotemporal filtering appearing later in this paper. We will not delve in this issue here. It has been extensively studied in [3] but it still remains an open problem. In our implementation we used non-overlapping averaging receptive fields as implemented in the emulation of the space-variant sensor in [22].

### 3 The motion field in the log-polar plane

We begin this section with a brief summary of the cartesian motion equations. Let an object be moving with translational velocity  $\mathbf{v} = (v_x, v_y, v_z)^T$  and angular velocity  $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)^T$  relative to the camera. We denote by  $\mathbf{X}$  the position of a point on the object with respect to the camera coordinate system, by  $\hat{\mathbf{z}}$  the unit-vector in the  $z$ -axis taken as the optical axis, and by  $\mathbf{x} = (x, y, 1)^T$  the projection of  $\mathbf{X}$  on the image plane  $Z = 1$ . The motion field vector reads [15]

$$\dot{\mathbf{x}} = \frac{1}{\hat{\mathbf{z}}^T \mathbf{X}} \hat{\mathbf{z}} \times (\mathbf{v} \times \mathbf{x}) + \hat{\mathbf{z}} \times (\mathbf{x} \times (\mathbf{x} \times \boldsymbol{\omega})). \quad (4)$$

In case of ego-motion of the camera with the above velocities and a stationary environment, the above equation as well as all following equations have to be read with the opposite sign for  $\mathbf{v}$  and  $\boldsymbol{\omega}$ . The 3D-motion estimation problem is known as the recovery of all the depths, the direction of translation  $\mathbf{v}$ , and the angular velocity from the motion field  $\dot{\mathbf{x}}$ . The magnitude of translation cannot be recovered due to the scale ambiguity that couples translation and depth.

We first compute the motion field vectors in the polar plane. The definition of the polar coordinates yields

$$\begin{aligned} \dot{\rho} &= \dot{x} \cos \eta + \dot{y} \sin \eta \\ \dot{\eta} &= \frac{1}{\rho} (-\dot{x} \sin \eta + \dot{y} \cos \eta). \end{aligned} \quad (5)$$

The radial component of the log-polar motion field can be easily obtained:

$$\dot{\xi} = \frac{1}{\ln a} \frac{\dot{\rho}}{\rho}. \quad (6)$$

In order to make the equations for the log-polar plane more readable we introduce the polar unit-vectors  $\hat{\boldsymbol{\rho}} = (\cos \eta, \sin \eta)^T$  and  $\hat{\boldsymbol{\eta}} = (-\sin \eta, \cos \eta)^T$ . Furthermore we introduce the vectors  $\mathbf{v}_{xy} = (v_x, v_y)^T$  and  $\boldsymbol{\omega}_{xy} = (\omega_x, \omega_y)^T$  to describe motion parallel to the image plane. We will carry out the computations for the polar motion field  $(\dot{\rho}, \dot{\eta})$ . The log-polar motion field is different only in the radial component  $\dot{\xi}$  which can be computed from  $\dot{\rho}$  straightforward (6). This will also enable us to find out that most of the advantages are due to the polar nature of log-polar plane.

Using the cartesian motion field (4) and the transformation rules (5) we obtain the following expressions which relate the polar motion field to the 3D geometry (depths  $Z$ ) and motion  $(\mathbf{v}, \boldsymbol{\omega})$  of the scene:

$$\begin{aligned} \dot{\rho} &= -\frac{\rho v_z}{Z} + \frac{\mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}}}{Z} + (1 + \rho^2) \boldsymbol{\omega}_{xy}^T \hat{\boldsymbol{\eta}} \\ \dot{\eta} &= \frac{\mathbf{v}_{xy}^T \hat{\boldsymbol{\eta}}}{\rho Z} - \frac{\boldsymbol{\omega}_{xy}^T \hat{\boldsymbol{\rho}}}{\rho} + \omega_z. \end{aligned} \quad (7)$$

Hence, the log-component of the motion field reads

$$\ln a \dot{\xi} = -\frac{v_z}{Z} + \frac{\mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}}}{\rho Z} + \left(\frac{1}{\rho} + \rho\right) \boldsymbol{\omega}_{xy}^T \hat{\boldsymbol{\eta}}. \quad (8)$$

Equivalent relations have been derived in [23].

We see that the motion field is the sum of the translational part including the depth information and a rotational part like in the cartesian formulation. However, we note that the influence of the motion-components parallel to the optical axis are decoupled. The translation  $v_z$  appears only in  $\dot{\rho}$  and the rotation  $\omega_z$  only in  $\dot{\eta}$ . This was expected due to the properties of the complex logarithmic mapping regarding two-dimensional expansions and rotations, respectively. However, as it is already known the main problem in recovery of 3D-motion is to totally decouple the rotational from the translational effects. We will show later how to attack this problem globally by an active technique, and locally by a purposive technique.

### 3.1 Pure translation

If there is only translation parallel to the optical axis (Fig. 2a) the polar (and the log-polar) motion field has only one component as can be seen in Fig. 2b. The magnitude variation of the log-component  $\dot{\xi}$  in this case depends only on the depths of the projected 3D structures. If the motion is pure translation in an arbitrary direction (Fig. 2c) the motion field looks like in Fig. 2d.

To better understand this field structure and the method presented next we show a pencil of lines in the cartesian plane and its complex logarithmic mapping in Fig. 3. The lines are the orbits of the motion field.

Closer inspection of the equations of the polar translational motion field

$$\dot{\rho} = -\frac{\rho v_z}{Z} \quad (9)$$

$$\dot{\eta} = \frac{\mathbf{v}_{xy}^T \hat{\boldsymbol{\eta}}}{\rho Z} \quad (10)$$

yields following facts:

1. The angular component reads  $\dot{\eta} = \frac{\mathbf{v}_{xy}^T \hat{\boldsymbol{\eta}}}{\rho Z}$ . There are two lines in the motion field  $\eta = \phi$  and  $\eta = \phi + \pi$  (or  $\phi - \pi$  depending on which of both is in  $[0, 2\pi)$ ) where  $\dot{\eta}$  vanishes for every  $\rho$ . The angle  $\phi$  gives the direction of  $(v_x, v_y)$  or the line where the focus of expansion lies. Assuming that there is no direction  $\eta$  where the integral

$$\int_{\rho_0}^{\rho_{max}} \frac{1}{\rho^2 Z^2} d\rho$$

vanishes the desired directions  $\phi$  and  $\phi \pm \pi$  are given by the global minima of  $\int \dot{\eta}^2 d\rho$  in the presence of noisy measurements. The above sufficient condition is met if the camera does not gaze on a point of infinite depth. For example, the condition is not satisfied if the origin of the cartesian image lies above the horizon's projection. If the integral  $\int \dot{\eta}^2 d\rho$  vanishes everywhere we imply the existence of a pure  $v_z$ -translation.

2. Along the lines  $\eta = \phi + \pi/2$  and  $\eta = \phi - \pi/2$  or  $(\phi + 3\pi/2)$  the flow reads

$$\dot{\rho}_{\phi+\pi/2} = -\rho \frac{v_z}{Z} \quad \dot{\eta}_{\phi+\pi/2} = \pm \frac{\sqrt{v_x^2 + v_y^2}}{\rho Z}.$$

By dividing we obtain

$$\frac{\dot{\rho}_{\phi+\pi/2}}{\rho^2 \dot{\eta}_{\phi+\pi/2}} = \pm \frac{v_z}{\sqrt{v_x^2 + v_y^2}} \quad (11)$$

which is the tangent of the polar angle of the translation direction.

Hence, it is possible by an explicit search for the global minima along the angle coordinate - which is feasible in the polar plane due to the low resolution - to obtain the full translation direction.

As an alternative method, we could proceed as in the cartesian case where we first eliminate the depth from the motion field equation (4) - recall that eq. (4) consists of two equations since  $\dot{\mathbf{x}}$  has two components. If we eliminate depth from (7) we obtain

$$\rho \dot{\eta} (\mathbf{v}_{xy}^T \hat{\rho} - \rho v_z) = \dot{\rho} \mathbf{v}_{xy}^T \hat{\eta}$$

that can be rewritten as

$$\begin{pmatrix} v_x & v_y & v_z \end{pmatrix} \begin{pmatrix} \rho \dot{\eta} \cos \eta + \dot{\rho} \sin \eta \\ \rho \dot{\eta} \sin \eta - \dot{\rho} \cos \eta \\ -\rho^2 \dot{\eta} \end{pmatrix} = 0. \quad (12)$$

This is nothing else than the equivalent of the well known epipolar constraint in the continuous case [5]:

$$\mathbf{v}^T (\mathbf{x} \times \dot{\mathbf{x}}) = 0.$$

### 3.2 Fixation in the log-polar plane

We assume that the camera mount has two controllable degrees of freedom which enable a rotation about an axis parallel to the image plane through the optical center. We denote by  $(\Omega_x, \Omega_y, 0)$  the resulting additional angular velocity. The motion field arising from the relative motion of the camera to the environment (ego and/or object motion) follows from (4)

$$\begin{aligned} \dot{x} &= \frac{v_x - xv_z}{Z} - xy(\omega_x - \Omega_x) + (1 + x^2)(\omega_y - \Omega_y) - y\omega_z \\ \dot{y} &= \frac{v_y - yv_z}{Z} - (1 + y^2)(\omega_x - \Omega_x) + xy(\omega_y - \Omega_y) + x\omega_z. \end{aligned} \quad (13)$$

We define fixation as keeping the same point in the image center along time. The fixated point may be a point on a moving object in case of a stationary or moving observer or a stationary point in the case of a moving observer. Fixation is achieved by closed loop control where controllables are the pan and the tilt angle of the camera. The formally defined fixation criterion we use is the vanishing of the motion field vector at the central point of the image:

$$\dot{\mathbf{x}}|_{\mathbf{x}=0} = \mathbf{0}. \quad (14)$$

This is equivalent to

$$\frac{v_x}{Z_o} + \omega_y - \Omega_y = 0 \quad \text{and} \quad \frac{v_y}{Z_o} - (\omega_x - \Omega_x) = 0, \quad (15)$$

where  $Z_o$  is the depth of the point projected on the image center.

However, this criterion relies on the existence of a well defined mathematical point in the image. In reality, this is only the case when the camera fixates on an object like a light spot. When fixating on a larger object the average flow over an area should be minimized. Assuming that this area is the cartesian fovea we seek control angular velocities  $(\Omega_x, \Omega_y)$  such that

$$\int_{r=0.. \rho_0} \int_{\theta=0..2\pi} \dot{\mathbf{x}} r dr d\theta = \mathbf{0}$$

where  $\dot{\mathbf{x}} = (\dot{x}, \dot{y})$  is given in (13) and  $\rho_0$  is the foveal radius. We note that we do not minimize the integral of the squares of the flow so that we allow foveal motions like rolling around the optical axis. After integration we obtain

$$\Omega_y = \omega_y + v_x \frac{\zeta}{1 + \frac{\rho_0^2}{4}} \quad \text{and} \quad \Omega_x = -\omega_x + v_y \frac{\zeta}{1 + \frac{\rho_0^2}{4}} \quad (16)$$

where  $\zeta$  is the average of the inverse depth. For a small foveal area we can still use the approximation (15).

What kind of control strategy is applied to achieve the fixation criterion is not the subject of this paper. We will assume that the system is in steady state mode and we will study the effects of the satisfied fixation criterion in the peripheral motion-field of the log-polar plane.

We turn now to the motion field on the polar transform of the periphery which reads in case of fixation:

$$\begin{aligned}\dot{\rho} &= -\frac{\rho v_z}{Z} + \frac{\mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}}}{Z} + (1 + \rho^2)(\boldsymbol{\omega}_{xy} - \boldsymbol{\Omega}_{xy})^T \hat{\boldsymbol{\eta}} \\ \dot{\eta} &= \frac{\mathbf{v}_{xy}^T \hat{\boldsymbol{\eta}}}{\rho Z} - \frac{(\boldsymbol{\omega}_{xy} - \boldsymbol{\Omega}_{xy})^T \hat{\boldsymbol{\rho}}}{\rho} + \omega_z.\end{aligned}\quad (17)$$

We rewrite the polar motion field by means of (15)

$$\begin{aligned}\dot{\rho} &= -\frac{\rho v_z}{Z} + \frac{\mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}}}{Z} + (1 + \rho^2) \left( \frac{v_y}{Z_o} - \frac{v_x}{Z_o} \right) \hat{\boldsymbol{\eta}} \\ \dot{\eta} &= \frac{\mathbf{v}_{xy}^T \hat{\boldsymbol{\eta}}}{\rho Z} - \frac{1}{\rho} \left( \frac{v_y}{Z_o} - \frac{v_x}{Z_o} \right) \hat{\boldsymbol{\rho}} + \omega_z.\end{aligned}\quad (18)$$

that leads to the following polar motion field in case of fixation:

$$\begin{aligned}\dot{\rho} &= -\frac{\rho v_z}{Z} + \left( \frac{1}{Z} - \frac{1 + \rho^2}{Z_o} \right) \mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}} \\ \dot{\eta} &= \frac{1}{\rho} \left( \frac{1}{Z} - \frac{1}{Z_o} \right) \mathbf{v}_{xy}^T \hat{\boldsymbol{\eta}} + \omega_z.\end{aligned}\quad (19)$$

We rewrite the equation by introducing  $v'_z = v_z/Z_o$  and  $\mathbf{v}'_{xy} = \mathbf{v}_{xy}/Z_o$ :

$$\begin{aligned}\dot{\rho} &= -\frac{Z_o}{Z} \rho v'_z + \left( \frac{Z_o}{Z} - (1 + \rho^2) \right) \mathbf{v}'_{xy}{}^T \hat{\boldsymbol{\rho}} \\ \dot{\eta} &= \frac{1}{\rho} \left( \frac{Z_o}{Z} - 1 \right) \mathbf{v}'_{xy}{}^T \hat{\boldsymbol{\eta}} + \omega_z.\end{aligned}\quad (20)$$

To reduce the number of symbols we will use again non-primed symbols for the translation. We first note that we obtain as independent unknowns the depths  $\frac{Z_o}{Z}$  relative to the depth of the fixation point which results in an object centered scene representation. Second, the other unknowns are reduced from initially five in the non-fixation case - three for rotation and two for the translation direction - to four: the three scaled translation components, and one angular velocity component. This dimension reduction was already proved in the cartesian plane in [1] and [2].

We next propose an algorithm to recover the motion parameters in case of fixation from the motion field in the log-polar plane. Equivalent to the case of pure translation without fixation our method is based on the observation that the angular component  $\dot{\eta}$  equals  $\omega_z$  everywhere along the line  $\eta = \phi$  where  $\tan \phi = v_y/v_x$ . However, this is not a necessary condition. The angular component  $\dot{\eta}$  equals  $\omega_z$  in the additional case that  $Z = Z_o$  along a radial line, or more general  $\dot{\eta}$  is constant if the relative depth is linearly varying:

$$\frac{Z_o}{Z} - 1 = k\rho. \quad (21)$$

This may happen if the environment is planar. For a particular  $\eta_0$  the radial component for  $\eta = \eta_0$  reads in this case

$$\dot{\rho} = -(1 + k\rho)v_z + (k\rho - \rho^2)\mathbf{v}'_{xy}{}^T \hat{\boldsymbol{\rho}}_{\eta_0}.$$

We can exclude this case if we test  $\dot{\rho}$  subject to quadratic variation with respect to  $\rho$ . The algorithm we propose comprises following steps:

1. We build the average  $\bar{\eta}$  over  $\rho$  for every  $\eta$  and the variance

$$\int_{\rho} (\dot{\eta} - \bar{\eta})^2 d\rho.$$

The variance vanishes at  $\eta = \phi$  and  $\eta = \phi \pm \pi$  and at the angles where  $Z = Z_0$  on an entire radial line which is excluded as described above.

We carry out an one-dimensional search for the global minima of the above integral. We carry out the above described linearity test to exclude the case  $Z = Z_0$ . The global minima give the direction of the line containing the FOE in the cartesian plane as well as  $\omega_z$  (equal the average  $\bar{\eta}$  at  $\eta = \phi$ ). We exclude additional global minima due to linear depth variation as described above.

2. At  $\eta = \phi + \frac{\pi}{2}$  and  $\eta = \phi - \frac{\pi}{2}$  (or  $\phi + \frac{3\pi}{2}$ ) the motion field reads:

$$\begin{aligned}\dot{\rho} &= -\frac{Z_o}{Z}\rho v_z \\ \dot{\eta} &= \frac{1}{\rho}\left(\frac{Z_o}{Z} - 1\right)\sqrt{v_x^2 + v_y^2} + \omega_z.\end{aligned}$$

We build the following averages

$$\begin{aligned}\overline{\rho\dot{\eta}} &= \bar{\rho}\omega_z + \overline{(Z_o/Z - 1)}\sqrt{v_x^2 + v_y^2} \\ \overline{\dot{\rho}/\rho} &= -\overline{Z_o/Z}v_z.\end{aligned}$$

3. Building for every point the deviation from the average we obtain

$$\begin{aligned}\frac{\dot{\rho}}{\rho} - \overline{\dot{\rho}/\rho} &= -v_z\left(\frac{Z_o}{Z} - \overline{Z_o/Z}\right) \\ \rho\dot{\eta} - \overline{\rho\dot{\eta}} - (\rho - \bar{\rho})\omega_z &= \sqrt{v_x^2 + v_y^2}\left(\frac{Z_o}{Z} - \overline{Z_o/Z}\right).\end{aligned}$$

After testing the vanishing of the lhs of the latter equation indicating a frontal translation we obtain

$$\frac{v_z}{\sqrt{v_x^2 + v_y^2}} = -\frac{\frac{\dot{\rho}}{\rho} - \overline{\dot{\rho}/\rho}}{\rho\dot{\eta} - \overline{\rho\dot{\eta}} - (\rho - \bar{\rho})\omega_z} \quad (22)$$

for every point along the lines  $\eta = \phi + \frac{\pi}{2}$  and  $\eta = \phi - \frac{\pi}{2}$  (or  $\phi + \frac{3\pi}{2}$ ). Since  $\omega_z$  is given by the second step we are able to recover the angle between translation and the optical axis. Its tangent is equal to the inverse of the above expression.

We elaborated a method to recover the direction of translation in the case of fixation from the polar motion field based on a feasible 1D search over the angle range. The equations for the logarithmic field are obtained straightforward by substituting (6) in the above expressions. This substitution does not introduce any computational advantages, hence, the potential of the method relies only on the polar nature of the log-polar mapping.

### 3.3 Experiments on the translation computation

In this subsection, we present results on the computation of the optical flow in the log-polar domain as well as on the computation of the translation direction in the case of pure translation.

We tested the method proposed in section 3 in a log-polar sequence obtained from the cartesian real world sequence ‘‘Marbled Block’’<sup>1</sup> [16]. One image of the sequence and its log-polar transform

<sup>1</sup>Created by Michael Otte at University of Karlsruhe and FhG-IITB, Germany.

are shown in Fig. 6 top and middle, respectively. The log-polar image is drawn such that the  $\eta$ -axis is the horizontal axis and the  $\xi$ -axis is the vertical axis pointing downwards. To interpret the log-polar images we note that the angle  $\eta$  is measured beginning counterclockwise from the  $y$ -axis which is pointing downwards. So moving horizontally in the log-polar plane we first see the transformed lower right quadrant, then the transformed upper right quadrant and so on. The compression rate obtained by the log-polar transformation is 1:25.

The optical flow is computed using the spatiotemporal derivatives and the assumption that the flow is locally affine in the cartesian domain. Based on the computed log-polar flow of the ‘‘Marbled Block’’ sequence we apply the method in section 3 to estimate the direction of translation. We show in Fig. 4 the computed average angular component  $\dot{\eta}$  as a function of the angle. We obtain two global minima - differing by 180 degrees as expected - giving as result an angle of 253 degrees for the direction of the line containing the FOE. By averaging the expression

$$\frac{\dot{\rho}_{\phi+\pi/2}}{\rho^2 \dot{\eta}_{\phi+\pi/2}}$$

over the diameter line corresponding to  $253 - 90 = 163$  degrees we obtain as estimated cartesian FOE-position  $(-247, 102)$ . Using the intrinsic calibration parameters of the sequence we find out that the angle error between the veridical and the estimated translation direction is 5 degrees which is an acceptable estimation error given the fact that the computational effort decreased to its 4%.

#### 4 Local motion computations

In the prior sections we presented methods which exploited the log-polar motion field over the entire peripheral field of view - the domain of log-polar mapping. Here, we delve into the variation of the motion field - also called motion parallax - in the immediate neighborhood of every point. It was repeatedly proved [11, 27, 20, 7] that the differential invariants of the optical flow divergence, curl, and the two shear components contain information on motion and depth.

In this section, we will study the local motion parallax in the log-polar plane. We start with a brief description of the parallax in the cartesian plane. After differentiating the motion field  $\dot{x}$  in (4) with respect to  $x$  and  $y$  we obtain the field divergence

$$\text{div} = \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} = \frac{-2v_z + (pv_x + qv_y) - 3v_z(px + qy)}{Z(1 - px - qy)} - 3y\omega_x + 3x\omega_y, \quad (23)$$

where

$$p = \frac{\partial Z}{\partial X} \quad q = \frac{\partial Z}{\partial Y},$$

and  $(-p, -q, -1)$  is the normal to the tangential plane of the viewed surface. The shear components are similarly obtained

$$\begin{aligned} \text{shear}_1 &= \frac{\partial \dot{x}}{\partial x} - \frac{\partial \dot{y}}{\partial y} = \frac{-p(v_x - xv_z) + q(v_y - yv_z)}{Z(1 - px - qy)} + y\omega_x + x\omega_y \\ \text{shear}_2 &= \frac{\partial \dot{x}}{\partial y} + \frac{\partial \dot{y}}{\partial x} = \frac{\partial \dot{x}}{\partial x} - \frac{\partial \dot{y}}{\partial y} = \frac{-q(v_x - xv_z) - p(v_y - yv_z)}{Z(1 - px - qy)} - x\omega_x - y\omega_y. \end{aligned} \quad (24)$$

We do not consider the curl of the motion field because it contains information about the  $\omega_z$  component which is not of interest to the behavioral tasks we study. The above expressions are simplified at the center of the image ( $x = 0, y = 0$ ) where they contain only translational and depth information. Furthermore if the translation is parallel to the optical axis ( $v_x = 0, v_y = 0$ ) the divergence equals  $v_z/Z$  which is the inverse of the time to collision (TTC) with the object viewed in the image center. If the translation is arbitrary Subbarao [20] proved that the inverse of the TTC can be bounded by the two values  $\text{div} \pm \|shear\|$ . Unfortunately, if points outside the fovea are considered - as is the case in the log-polar motion field- the above parallax

expressions contain rotational information that cannot be eliminated. Parallax methods have either considered the rotational terms as negligible [4] or have used a different image surface. Indeed, if the scene is projected on a spherical surface, the divergence and shear components of the resulting motion field - which is tangential to the image sphere - get rid of the rotation information [11]. This is based on the fact that the rotational *spherical* motion field is linear in  $(x, y)$  whereas the rotational *planar* motion field contains terms quadratic in the image coordinates.

Although there are many practical situations where the cartesian invariants could yield acceptable results for the time to collision and the surface slant - the examples from the literature are very rare - we will search here for exact relations. The inspiration of this study was the result in [23] we will begin with.

We denote the derivatives of depth with respect to  $\rho$  and  $\eta$  with  $Z_\rho$  and  $Z_\eta$ . We first compute the following spatial derivatives:

$$\frac{\partial \dot{\eta}}{\partial \eta} = -\frac{Z_\eta}{\rho Z^2} \mathbf{v}_{xy}^T \hat{\boldsymbol{\eta}} - \frac{1}{\rho Z} \mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}} - \frac{1}{\rho} \boldsymbol{\omega}_{xy}^T \hat{\boldsymbol{\eta}} \quad (25)$$

$$\frac{\partial \dot{\rho}}{\partial \rho} = \frac{\rho Z_\rho - Z}{Z^2} v_z - \frac{Z_\rho}{Z^2} \mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}} + 2\rho \boldsymbol{\omega}_{xy}^T \hat{\boldsymbol{\eta}}. \quad (26)$$

In combination with the scaled field component

$$\frac{\dot{\rho}}{\rho} = -\frac{v_z}{Z} + \frac{1}{\rho Z} \mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}} + \left(\frac{1}{\rho} + \rho\right) \boldsymbol{\omega}_{xy}^T \hat{\boldsymbol{\eta}} \quad (27)$$

We see that the only linear combination of the above three expressions that is independent of rotation is

$$2\frac{\partial \dot{\eta}}{\partial \eta} - \frac{\partial \dot{\rho}}{\partial \rho} + 2\frac{\dot{\rho}}{\rho} = -\frac{v_z}{Z} + \frac{Z_\rho}{Z^2} (\mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}} - \rho v_z) - 2\frac{Z_\eta}{\rho Z^2} \mathbf{v}_{xy}^T \hat{\boldsymbol{\eta}}. \quad (28)$$

If we assume that the surface is locally frontal then the depth derivatives vanish [23] and we obtain the inverse of the time to collision  $\frac{v_z}{Z}$ . However, this assumption is not always valid. In high eccentricities the pixels of the log-polar plane span a considerable viewing angle in space that may contain a non negligible depth variation. Therefore, we compute the depth derivatives  $Z_\rho$  and  $Z_\eta$  as a function of the normal of the surface. Suppose that the surface normal is  $(-p, -q, -1)$  with  $p = \frac{\partial Z}{\partial X}$  and  $q = \frac{\partial Z}{\partial Y}$ . It can be easily proved that

$$\frac{\partial Z}{\partial x} = \frac{pZ}{1 - px - qy} \quad \frac{\partial Z}{\partial y} = \frac{qZ}{1 - px - qy}$$

Let us denote  $(p, q)$  with  $\mathbf{p}$ . The required derivatives read

$$Z_\rho = \frac{Z \mathbf{p}^T \hat{\boldsymbol{\rho}}}{1 - \rho \mathbf{p}^T \hat{\boldsymbol{\rho}}} \quad \text{and} \quad Z_\eta = \frac{Z \rho \mathbf{p}^T \hat{\boldsymbol{\eta}}}{1 - \rho \mathbf{p}^T \hat{\boldsymbol{\rho}}}. \quad (29)$$

Hence,

$$2\frac{\partial \dot{\eta}}{\partial \eta} - \frac{\partial \dot{\rho}}{\partial \rho} + 2\frac{\dot{\rho}}{\rho} = \frac{-v_z + \mathbf{v}_{xy}^T \hat{\boldsymbol{\rho}} \mathbf{p}^T \hat{\boldsymbol{\rho}} - 2\mathbf{v}_{xy}^T \hat{\boldsymbol{\eta}} \mathbf{p}^T \hat{\boldsymbol{\eta}}}{Z(1 - \rho \mathbf{p}^T \hat{\boldsymbol{\rho}})}. \quad (30)$$

The exact expression has the advantage of being rotation independent but depends on the unknown normal of the tangential plane of the viewed surface. We know [20] that we can only *bound* the time to collision by using both divergence and shear. However, in a planar image this is only possible in the *rotation-free* case. Using a spherical image the bounds become rotation-independent because the dependence of the motion field on rotation becomes linear  $\boldsymbol{\omega} \times \mathbf{x}$  (cf. the quadratic rotational terms in (4) and (5)). Earlier approaches have assumed that near the center a planar image can be approximated by a spherical image [7]. However, this approximation is not valid in the non-foveal part considered in the log-polar mapping. We will avoid this approximation by transforming exactly onto the polar plane.

Let us assume that the image is the unit half-sphere and that the position on this sphere is given by the polar angle  $\theta$  and the azimuth angle  $\phi$ . Regarding the polar coordinate system above, we have

$$\rho = \tan \theta \quad \text{and} \quad \eta = \phi.$$

The velocity  $\dot{\mathbf{s}}$  of a point  $\mathbf{s}$  on the image sphere is on the tangential plane at that point and is given by

$$\dot{\mathbf{s}} = \dot{\theta} \hat{\theta} + \dot{\phi} \sin \theta \hat{\phi} \quad (31)$$

where  $\hat{\theta}$  and  $\hat{\phi}$  are the unit-vectors of the spherical coordinate system respectively. In analogy to the differential invariants on a spherical image we next propose parallax expressions in the polar motion field that provide bounds for the time to collision which are independent both from the surface slant and the rotation. The divergence of the spherical motion field is [12]

$$\text{div} \dot{\mathbf{s}} = \frac{\partial \dot{\theta}}{\partial \theta} + \frac{\partial \dot{\phi}}{\partial \phi} + \frac{\dot{\theta}}{\tan \theta} \quad (32)$$

and the two shear components read [11]

$$\text{shear}_1 \dot{\mathbf{s}} = \frac{\partial \dot{\theta}}{\partial \theta} - \frac{\partial \dot{\phi}}{\partial \phi} - \frac{\dot{\theta}}{\tan \theta} \quad \text{and} \quad \text{shear}_2 \dot{\mathbf{s}} = \frac{1}{\sin \theta} \frac{\partial \dot{\theta}}{\partial \phi} + \sin \theta \frac{\partial \dot{\phi}}{\partial \theta}. \quad (33)$$

By rewriting the right hand sides of the above equations using a polar coordinate system ( $\rho = \tan \theta, \eta = \phi$ ) on the image plane we obtain

$$\begin{aligned} \text{div} \dot{\mathbf{s}} &= \frac{\partial \dot{\rho}}{\partial \rho} + \frac{\partial \dot{\eta}}{\partial \eta} + \frac{\dot{\rho}}{1 + \rho^2} \left( \frac{1}{\rho} - 2\rho \right) \\ \text{shear}_1 \dot{\mathbf{s}} &= \frac{\partial \dot{\rho}}{\partial \rho} - \frac{\partial \dot{\eta}}{\partial \eta} - \frac{\dot{\rho}}{1 + \rho^2} \left( \frac{1}{\rho} + 2\rho \right) \\ \text{shear}_2 \dot{\mathbf{s}} &= \frac{\sqrt{1 + \rho^2}}{\rho} \left( (1 + \rho^2) \frac{\partial \dot{\eta}}{\partial \rho} + \frac{1}{1 + \rho^2} \frac{\partial \dot{\rho}}{\partial \eta} \right). \end{aligned} \quad (34)$$

Then we build the squared shear magnitude

$$\text{shear}^2 \dot{\mathbf{s}} = \text{shear}_1^2 \dot{\mathbf{s}} + \text{shear}_2^2 \dot{\mathbf{s}}$$

and after tedious calculations we find out that it is factorizable:

$$\text{shear}_1 \dot{\mathbf{s}}^2 + \text{shear}_2 \dot{\mathbf{s}}^2 = \|\mathbf{S}_\alpha\|^2 \|\mathbf{S}_\beta\|^2 \quad (35)$$

with

$$\mathbf{S}_\alpha = \frac{1}{Z} \begin{pmatrix} Z_\rho (1 + \rho^2) + \rho \\ \frac{Z_\eta \sqrt{1 + \rho^2}}{Z_\rho} \end{pmatrix} \quad \mathbf{S}_\beta = \frac{1}{Z} \begin{pmatrix} \frac{\mathbf{v}_{xy}^T \hat{\rho} - \rho v_z}{1 + \rho^2} \\ \frac{\mathbf{v}_{xy}^T \hat{\eta}}{\sqrt{1 + \rho^2}} \end{pmatrix}.$$

On the other hand, trying to express the divergence as a function of  $\mathbf{S}_\alpha$  and  $\mathbf{S}_\beta$  we find out that

$$\text{div} \dot{\mathbf{s}} = -\frac{2(\rho \mathbf{v}_{xy}^T \hat{\rho} + v_z)}{Z(1 + \rho^2)} - \mathbf{S}_\alpha^T \mathbf{S}_\beta. \quad (36)$$

Since  $\mathbf{S}_\alpha^T \mathbf{S}_\beta \leq \|\mathbf{S}_\alpha\| \|\mathbf{S}_\beta\|$  we obtain

$$\text{div} - \|\text{shear}\| < -\frac{2(\rho \mathbf{v}_{xy}^T \hat{\rho} + v_z)}{Z(1 + \rho^2)} \leq \text{div} + \|\text{shear}\|. \quad (37)$$

The expression in the middle can be rewritten as

$$\frac{(\rho \mathbf{v}_{xy}^T \hat{\rho} + v_z)}{Z(1 + \rho^2)} = \frac{\mathbf{v}^T \hat{\mathbf{x}}}{\|\mathbf{X}\|}$$

where  $\hat{\mathbf{x}}$  is the unit-vector in the space direction of the point  $\mathbf{x} = (x, y, 1)$  and  $\|\mathbf{X}\|$  is the distance to the 3D-point  $\mathbf{X}$ . This expression is referred in the literature as the inverse of the time to collision [7] or looming [17]. However, it differs from the more frequently definition  $v_z/Z$ . The former deals with distances  $\|\mathbf{X}\|$  to the points, whereas the latter deals with the depths  $z$  of the points. The physical interpretation is shown in Fig. 5. The time to collision  $Z/v_z$  is the time required for the observer to hit the plane  $\alpha$  at depth  $Z$  or the time required by an object like a frontoparallel plane  $\alpha$  to hit the observer. It is always geometrically plausible in the sense of a scaled depth map but it is physically plausible only in case of frontal translation. On the other hand, the looming expression  $\mathbf{v}^T \hat{\mathbf{x}} / \|\mathbf{X}\|$  is the time required by the observer to hit the plane  $\beta$  perpendicular to the line of sight at the considered point. It is neither a scaled depth nor a scaled range but it can be interpreted as the inverse of the time to collision if the line of sight is in the translation direction. Furthermore, it achieves its maximum at this point - the focus of expansion - if we assume that all points are equidistant  $\|\mathbf{X}\| = \text{constant}$ .

In contrast first to the rotation dependent cartesian bounds of  $v_z/Z$ , and second to the slant dependent expression in polar coordinates in (30), we derived both rotation and slant independent bounds on a new expression for the time to collision in the polar plane. All the expressions appearing in (34) may be easily transformed from the polar to the logpolar plane using

$$\begin{aligned} \dot{\rho} &= \ln a \rho_0 a^\xi \dot{\xi} & \frac{\partial \dot{\rho}}{\partial \rho} &= \ln a \dot{\xi} + \frac{\partial \dot{\xi}}{\partial \xi} \\ \frac{\partial \dot{\rho}}{\partial \eta} &= \ln a \rho_0 a^\xi \frac{\partial \dot{\xi}}{\partial \eta} & \frac{\partial \dot{\eta}}{\partial \rho} &= \frac{\rho_0 a^\xi}{\ln a} \frac{\partial \dot{\eta}}{\partial \xi}. \end{aligned} \quad (38)$$

Despite the significant data reduction no further advantages are obvious from the transition from polar to log-polar coordinates.

## 5 Conclusion

An essential feature of an attentive visual system is space-variant sensing. In this study we used the complex logarithmic mapping to model the resolution decrease in the image periphery. The achieved compression ratio of up to 1:25 enables reactive behavior in real time.

We introduced two new methods for computing the focus of expansion by exploiting the structure of the flow patterns in the log-polar motion field. Both are based on a 1D global minimum search which is inexpensive due to the low angular resolution. The second method necessitates the pursuit movement of the camera proving, thus, that a full attentional mechanism based on both fixation and space-variant sensing enables motion estimation in case of general motion. We implemented the first method in a translating real word sequence obtaining an estimation error of  $5^\circ$  in the translation direction. We plan to implement the second method in the near future using the controllable degrees of freedom of an active binocular head. The second part of our study was devoted to derive motion parallax expressions which yield the time to collision. Based on the parallax equations on a spherical image we derived exact equations for the transformation onto the polar plane. Then, we recovered bounds on the time to collision that are independent of rotation and the surface slant at every position of the image periphery. The mathematical advantages of our approach are due to the polar representation in the log-polar image whereas the complexity advantages are due to the logarithmic nature of the log-polar image.

### Acknowledgements

The Vision Group at the DIST Laboratory, University of Genova provided the software emulation of the log-polar sensor. I am grateful to Volker Krüger for the inspiring discussions and his contribution to the software implementation of the flow algorithms. I am indebted to Joannis Papavassiliou for his helpful comments and his hospitality.

### References

- [1] J.(Y.) Aloimonos, I. Weiss, and A. Bandyopadhyay. Active Vision. *International Journal of Computer Vision*, 1:333–356, 1988.

- [2] A. Bandopadhyay and D.H. Ballard. Egomotion perception using visual tracking. *Computational Intelligence*, 7:39–47, 1990.
- [3] M. Bolduc and M.D. Levine. A foveated retina system for robotic vision. In *ECCV-94 Workshop on Natural and Artificial Visual Sensors*, 1994.
- [4] P. Bouthemy and E. Francois. Motion segmentation and qualitative dynamic scene analysis from a moving camera. *International Journal of Computer Vision*, 10:157–182, 1993.
- [5] A. Bruss and B.K.P. Horn. Passive navigation. *Computer Vision, Graphics, and Image Processing*, 21:3–20, 1983.
- [6] R.H.S. Carpenter. *Movements of the Eyes*. Pion Press, London, 1988.
- [7] R. Cipolla and A. Blake. Surface orientation and time to contact from image divergence and deformation. In *Proc. Second European Conference on Computer Vision*, pp. 187–202. Santa Margherita, Italy, May 23-26, 1992.
- [8] C. Fermuller and Y. Aloimonos. The role of fixation in visual motion analysis. *International Journal of Computer Vision*, 11:165–186, 1993.
- [9] R. Guissin and S. Ullman. Direct computation of the focus of expansion from velocity field measurements. In *Proc. IEEE Workshop on Visual Motion*, pp. 146–155, Princeton, NJ, Oct. 7-9, 1991.
- [10] R. Jain, S.L. Bartlett, and N. O'Brien. Motion Stereo Using Ego-Motion Complex Logarithmic Mapping. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 9:356–369, 1987.
- [11] J.J. Koenderink and A.J. van Doorn. Local structure of movement parallax of the plane. *Journal Opt. Soc. Am.*, 66:717–723, 1976.
- [12] G.A. Korn and T.M. Korn. *Mathematical Handbook for Scientists and Engineers*. McGraw-Hill, New York, 1968.
- [13] H.A. Mallot, W. von Seelen, and F. Giannakopoulos. Neural mapping and space-variant image processing. *Neural Networks*, 3:245–263, 1990.
- [14] L. Massone, G. Sandini, and V. Tagliasco. “Form-Invariant” Topological Mapping Strategy for 2D Shape Recognition. *Computer Vision, Graphics, and Image Processing*, 30:169–188, 1985.
- [15] S. Negahdaripour and B.K.P. Horn. Direct passive navigation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-9:168–176, 1987.
- [16] M. Otte and H.-H. Nagel. Optical flow estimation: advances and comparisons. In *Proc. Third European Conference on Computer Vision*, pp. 51–60. Stockholm, Sweden, May 2-6, 1994.
- [17] D. Raviv and M. Herman. Visual servoing from 2-D image cues. In Y. Aloimonos, editor, *Active Perception*, pp. 191–226. Lawrence Erlbaum Associates, Hillsdale, NJ, 1993.
- [18] M. Schultze. Zur Anatomie und Physiologie der Retina. *Archiv Mikrosk. Anatomie*, 2:175–286, 1866.
- [19] E.L. Schwartz. Spatial mapping in the primate sensory projection: analytic structure and relevance to perception. *Biological Cybernetics*, 25:181–194, 1977.
- [20] M. Subbarao. Bounds on translational and angular velocity components from first order derivatives of image flow. In *Proc. Conf. American Association of Artificial Intelligence*, pp. 744–748. Seattle, WA, July 13-17, 1988.
- [21] M.A. Taalebinezhad. Direct recovery of motion and shape in the general case by fixation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-14:847–853, 1992.
- [22] M. Tistarelli and G. Sandini. Dynamic aspects in active vision. *CVGIP: Image Understanding*, 56:108–129, 1992. Special Issue on Purposive, Qualitative, Active Vision, Y. Aloimonos (ed.).
- [23] M. Tistarelli and G. Sandini. On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-15:401–410, 1993.
- [24] J. Van der Spiegel, G. Kreider, C. Claeys, I. Debusschre, G. Sandini, P. Dario, F. Fantini, and P. Belluti G. Soncini. A foveated retina-like sensor using CCD technology. In C. Mead, editor, *Analog VLSI and Neural Systems*. Addison-Wesley, Reading, MA, 1989.
- [25] A.J. van Doorn, J. Koenderink, and M.A. Bouman. The influence of the retinal inhomogeneity on the perception of spatial patterns. *Kybernetik*, 10:223–230, 1972.
- [26] R.S. Wallace, P.-W. Ong, B.B. Bederson, and E.L. Schwartz. Space variant image processing. *International Journal of Computer Vision*, 13:71–90, 1994.
- [27] A.M. Waxman, B. Kamgar-Parsi, and M. Subbarao. Closed-form solutions to image flow equations for 3D structure and motion. *International Journal of Computer Vision*, 1:239–258, 1987.
- [28] C.F.R. Weiman and G. Chaikin. Logarithmic spiral grids for image processing and display. *Computer Graphics and Image Processing*, 11:197–226, 1979.

(a)

(b)

Figure 3: The complex logarithmic mapping of a line pencil. The line through the origin is mapped into a vertical line. The horizontal axis is the  $\eta$ -axis and the vertical axis is the  $\xi$ -axis.

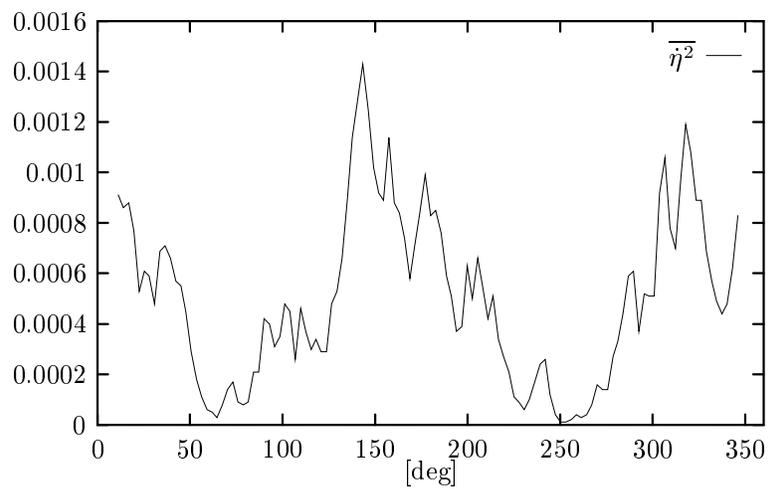


Figure 4: The average of the squared angular components along the radius as a function of the angle  $\eta$  for the Marbled Block sequence.

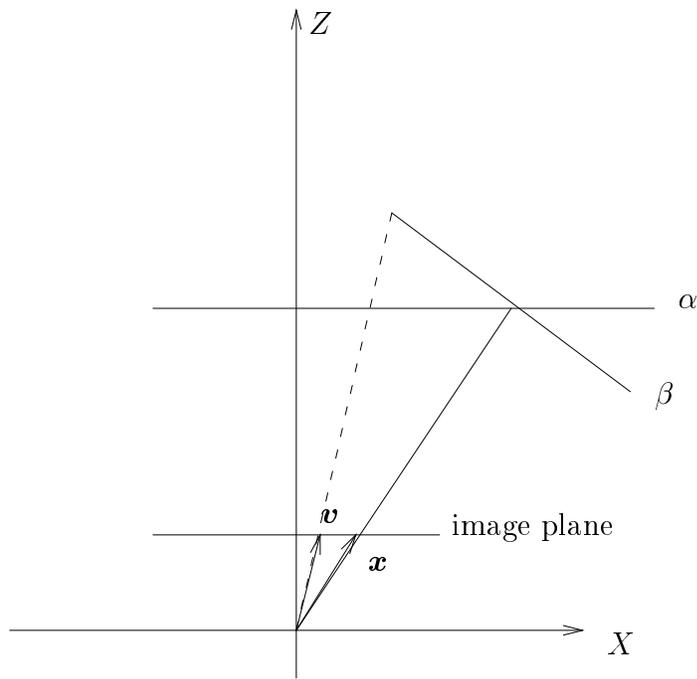


Figure 5: Geometric interpretation of different definitions for the time to collision. The time to collision  $Z/v_z$  is the time required for the observer to hit the plane  $\alpha$  at depth  $Z$  or the time required by an object like a frontoparallel plane  $\alpha$  to hit the observer. The second definition  $\mathbf{v}^T \hat{\mathbf{x}} / \|\mathbf{X}\|$  is the time required by the observer to hit the plane  $\beta$  perpendicular to the line of sight at the considered point.

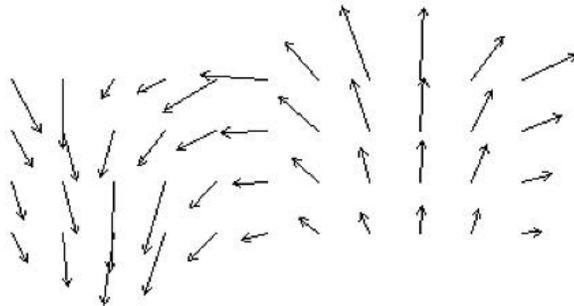
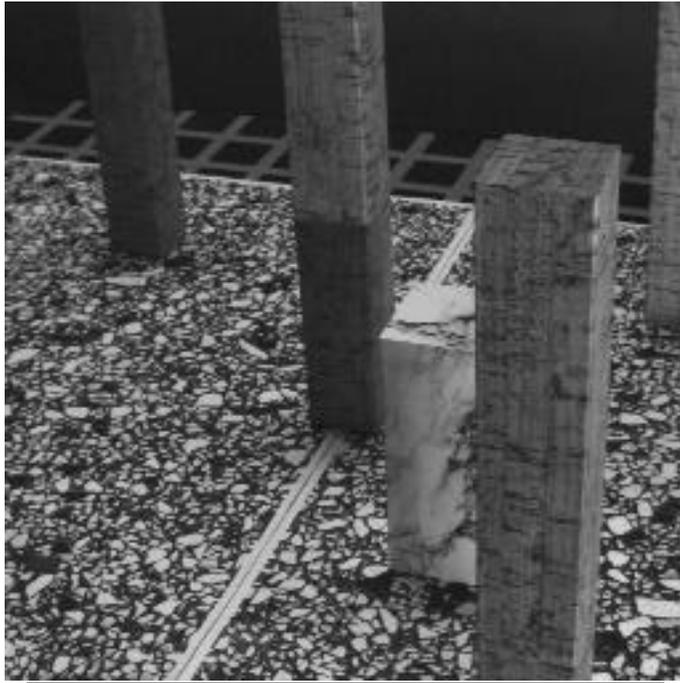


Figure 6: The original cartesian image (above), the log-polar transformed image (middle) and the computed optical flow in the log-polar domain (below) for the Marbled Block sequence.