# Vergence Control for Robotic Heads using Log-polar Images *

Alexandre Bernardino José Santos-Victor
Instituto de Sistemas e Robótica
Instituto Superior Técnico
1096 Lisboa Codex, Portugal
{ajmb,jasv}@isr.ist.utl.pt

**Abstract**

*This paper describes a real-time vergence control mechanism based on log-polar images, developed for a robot head.*

*The real-time control of active vision systems imposes strong constraints on the computational complexity of the vision algorithms. In this paper, we illustrate that vergence of a stereo head can be achieved at reduced computational cost using log-polar images. These images have higher resolution at the center, where the attention is focused on, and the rest of the visual field is covered at a coarser resolution, still enabling the detection of events at the image periphery.*

*The main advantages of using a non-uniform image sampling mechanism, such as the log-polar images, are related both to perceptual and algorithm complexity issues. We show that, when using correlation measures to control vergence, log-polar images give better results than cartesian images. Additionally, as log-polar images are smaller, the computation time is reduced.*

*Two algorithms for closed loop vergence control, using correlation measures over log-polar images, are proposed. In the test examples described in the paper, we compare the two algorithms and address the problem of designing an adequate log-polar sensor, by introducing a performance analysis to evaluate the various log-polar sensor layouts.*

## 1 Introduction

The research interests in Active Vision have increased in the past few years providing efficient ways of combining the control of robotic systems and advanced visual sensing at a moderate computational cost [1, 2].

Various stereo heads have been designed and built in many research laboratories [5] and were often inspired after their biological analogues [4, 11]. Several ocular movements, like saccades, smooth pursuit and vergence, found in biological vision systems have been implemented and tested.

In nature however, one of the main reasons for the existence of active head-eye control is the space-variant density of photo-receptors in the eye, distinguishing between a central high resolution area - the fovea, and the low resolution peripheral zone.

The non-uniform resolution of the human retina has been found to follow a logarithmic-polar law, taking into account the linear increment of the photoreceptors size with the distance to the center [14]. Combining this feature with the ability to move the eyes over the visual field, it is possible to obtain the information needed to perform the desired tasks, without considering all the data contained in uniformly sampled images. Hence, the processing effort can be concentrated in the foveal area.

In practice, log-polar images are claimed to have advantages over the cartesian images, like shape invariance to scaling and rotation [17], improved binocular fusion [18] and adequate correlation properties to vergence control [3].

The main contributions of this paper are threefold and deal with the imaging geometry, the vergence control approaches and a performance evaluation of the log-polar sensor layouts.

First we exploit the geometric properties of the log-polar images to implement a vergence control system for an active robot head. The vergence control is based on correlation measurements computed in the log-polar images, hence taking advantage of the higher resolution at the image center as opposed to the image periphery. Compared to the cartesian images we show that both computational and performance advantages arise from the use of the space variant images.

A second aspect worth mentioning is the use of two distinct vergence control strategies that use correlation measurements directly to control the vergence angle of the cameras without requiring any intermediate step to reconstruct the environment or calibrate the cameras.

Finally, we address the problem of characterizing the log polar sensor layout and study the influence of the various parameters in the overall system performance.

In general, we believe that these problems : sensor geometry, performance analysis and a control strategy which integrates directly the visual information in the control system, may have a significant impact on various applications of computer vision and robotics [12].

*Alexandre Bernardino was supported by a grant from PRAXIS-XXI research program. This work was partially funded by projects PRAXIS/3/3.1/TPR/23/94 and JNICT-PBIC/TPR/2550/95.

## 2 Log-Polar Mapping

The main reason for ocular movements in many biological systems consists on the existence of a space variant resolution in the retina. In the human visual system, the receptors at the retina are distributed in space with increasing density toward the center of the visual field (the fovea). Therefore, eye movements are used by the human oculomotor system to point the central high resolution fovea in the direction of the scene areas to observe more carefully.

Apart from few exceptions [9], most active vision systems use cartesian image representations. Instead, the vergence system described in this paper will be based on a log-polar space variant image sampling. The main advantage of this kind of sampling is an important data reduction obtained at the cost of reducing the resolution in the periphery of the images, but the high resolution in the fovea is sufficient to accomplish most of the visual tasks [8].

The log-polar transformation is a conformal mapping from the points on the cartesian plane $(x, y)$ to points in the log-polar plane $(\xi, \eta)$, [14, 17, 18]. The mapping is described by :

$$\begin{cases} \xi & = \log_k(\rho) \\ \eta & = \theta \end{cases} \tag{1}$$

where $k$ represents a decaying resolution factor and $\rho$ and $\theta$ denote the commonly used polar coordinates.

The log-polar mapping exhibits a singularity in the origin of the coordinate plane, as $\rho$ tends to zero. This means that the mapping will not be defined in a neighborhood of the origin :

$$\xi \in [\log_k(\rho_{min}), \log_k(\rho_{max})],$$

where $\rho_{max}$ depends on the cartesian plane dimension.

Figure 1 shows the discretized cartesian and log-polar planes.

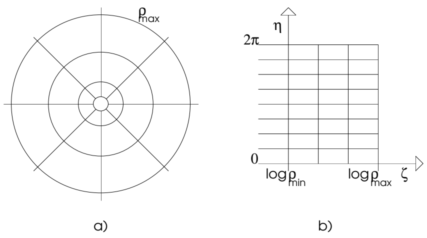


Figure 1: Space variant mapping from the cartesian plane, (a), to the log-polar plane, (b).

## 3 Vergence using correlation

When following a target, the motion of the human eyes can be decomposed into saccadic, smooth pursuit and vergence movements [11]. Saccadic and smooth pursuit movements use image position and velocity errors, in order to compensate lateral motion of the target. Vergence movements, instead, are based on disparity cues to compensate motion in depth.

Vergence is one of the basic eye movements to focus visual attention on a given scene point. When observing a target, moving away along the same gaze direction, our eyes rotate by equal amounts in opposite directions. The purpose of the vergence movements is to control the angle between the optic axes, in such a way that a fixation point along some specified gaze direction is kept at the center of the visual field. When this is the case, we say that the target is “foveated”. Figure 2 illustrates the vergence angle $\nu$, in a binocular system.

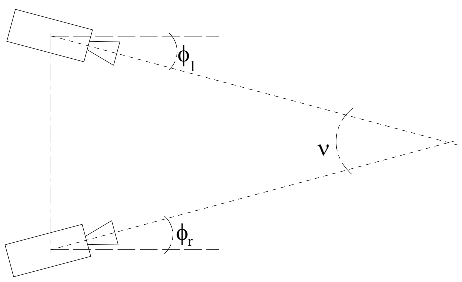


Figure 2: Stereo cameras verging at a point. The vergence angle, $\nu = \phi_l + \phi_r$, is shown.

The vergence process enables foveated animals to register objects of interest on the fovea of each eye. As a consequence, it allows the extraction of the greatest possible amount of information. Moreover, when verging on a target, it is possible to find its retinotopic location through disparity filtering, thus allowing to isolate the target from the background [6].

As the vergence angle is directly related to the target distance, depth cues can be used to help the vergence system in compensating depth changes. Depth cues such as binocular disparity and accommodation are important inputs to the vergence control system, and humans rely also on other cues such as motion and shading [4].

Traditional algorithms to calculate explicit disparity measures include cepstral filtering and phase correlation, but these techniques are very time consuming, and therefore, not very adequate for real-time applications. By relying on more active visual processes, it is possible to avoid the explicit computation of disparity and still detect correct vergence configurations.

The basic strategies to control vergence eye movements are still under discussion in the biological literature. Hypotheses include continuous feedback control [11], sequential preprogrammed motor responses [15], and channel-like processing of disparity [10].

In the approach used here we test and compare two algorithms for vergence control. Both of them rely on closed loop control and on simple correlation measures of log-polar images, avoiding explicit disparity calculation.

The first and simpler algorithm consists basically in estimating, over time, a similarity measure between two stereo images and controlling the vergence angle in order to increase their overlap, in a fashion similar to [9].

The second algorithm consists in a set of channels tunned on different disparities. A different channel is activated depending on the instantaneous level of disparity and, as disparity changes, control switches to other channel, resulting in a sequence of pre-programmed responses. This approach follows, in some sense, the ideas of [15] and [10].

As both algorithms use correlation measures, we start by analyzing the differences between the correlation of images of a stereo pair in the vicinity of vergence situations, on the cartesian and log-polar domains.

Suppose that a stereo pair is composed by one image $I(x,y)$ and a translated version $I(x+\delta_x, y)$. For small disparity $\delta_x$ we can write:

$$I(x+\delta_x, y) = I(x,y) + \frac{\partial I}{\partial x}\cdot\delta_x \tag{2}$$

As a measure of similarity we use, in this analysis, the integral of the squared difference between the two images. Considering correlation in cartesian coordinates, we have:

$$c_c = \iint_{x,y} \left[I(x+\delta_x, y) - I(x,y)\right]^2 dxdy \tag{3}$$

Substituting (2) in (3), we have:

$$c_c = \delta_x^2 \cdot \iint_{x,y} \left(\frac{\partial I}{\partial x}\right)^2 dxdy \tag{4}$$

To compare the correlation of cartesian images with the correlation of log-polar images, we need to change the coordinates of equation (4) from cartesian to the log-polar domain. We have:

$$\frac{\partial I}{\partial x} = \frac{\partial I}{\partial \xi}\cdot\frac{\partial \xi}{\partial x} + \frac{\partial I}{\partial \eta}\cdot\frac{\partial \eta}{\partial x} \tag{5}$$

where the partial derivatives of the log-polar mapping are:

$$\begin{cases} \frac{\partial \xi}{\partial x} = \frac{\cos\theta}{\rho\cdot\log k} \\ \frac{\partial \eta}{\partial x} = -\frac{\sin\theta}{\rho} \end{cases} \tag{6}$$

The determinant of the coordinate change jacobian matrix $J$ is given by:

$$|J| = \log k \cdot \rho^2 \tag{7}$$

and therefore equation (4) can be rewritten in log-polar coordinates as:

$$c_c = \delta_x^2 \log k \iint_{\xi,\eta} \left(\frac{\partial I}{\partial \xi}\frac{\cos\theta}{\log k} - \frac{\partial I}{\partial \eta}\sin\theta\right)^2 d\xi d\eta \tag{8}$$

Analogous calculations can be done for the correlation in the log-polar domain. For small deviations from vergence we have:

$$I(\xi+\delta_\xi, \eta+\delta_\eta) = I(\xi,\eta) + \frac{\partial I}{\partial \xi}\cdot\delta_\xi + \frac{\partial I}{\partial \eta}\cdot\delta_\eta$$

and the correlation measure can be expressed as:

$$c_l = \iint_{\xi,\eta} \left(\frac{\partial I}{\partial \xi}\cdot\delta_\xi + \frac{\partial I}{\partial \eta}\cdot\delta_\eta\right)^2 d\xi d\eta \tag{9}$$

Using (6), the increments $\delta_\xi$ and $\delta_\eta$ are:

$$\begin{cases} \delta_\xi = \frac{\cos\theta}{\rho\cdot\log k}\cdot\delta_x \\ \delta_\eta = -\frac{\sin\theta}{\rho}\cdot\delta_x \end{cases} \tag{10}$$

Hence the correlation function of the log-polar images is given by:

$$c_l = \delta_x^2 \iint_{\xi,\eta} \frac{1}{\rho^2}\left(\frac{\partial I}{\partial \xi}\frac{\cos\theta}{\log k} - \frac{\partial I}{\partial \eta}\sin\theta\right)^2 d\xi d\eta \tag{11}$$

Observing carefully the expressions for the correlation measures in the cartesian and log-polar domains (equations (8) and (11)), we can see that the main difference is the additional factor of $1/\rho^2$ in the integrand of $c_l$. Thus, in the log-polar images, the importance given to the disparity of peripheral image patches is smaller when compared to more central parts of the image.

For vergence control, this is an important advantage over the cartesian geometry. Suppose that the cameras are verging on an object in the center of the visual field so that the area under vergence has low disparities, while the area in the periphery belongs to the background and exhibits large disparities. In this case we would like to have low values for the correlation function. In cartesian coordinates, however, as all disparities are equally considered, the disparities of the background may significantly increase the values of the correlation, as opposed to the log-polar case, where the contribution of the background disparities is reduced.

For log-polar images, the central area which represents the focus of attention can be reduced by increasing the design parameter $k$, thus allowing to keep effective vergence on smaller targets.

The discrete version of the correlation function we used in the previous analysis is known as the Sum of Squared Differences (SSD):

$$SSD = \sum_{x,y}(I_l(x,y) - I_r(x,y))^2,$$

where $I_l(x,y)$, $I_r(x,y)$ stand for the left and right images.

However, the SSD norm is very sensitive to illumination changes between left and right images, which often occur if the two cameras have different apertures.

For the remaining of this paper, the normalized cross-correlation is used. This function behaves robustly [7] relative to changes in the illumination conditions :

$$NCC = 1 - \frac{\sum(I_l - \mu_l)(I_r - \mu_r)}{\sqrt{\sum(I_l - \mu_l)^2 \sum(I_r - \mu_r)^2}}$$

where $\mu_l$ and $\mu_r$ are the average gray levels for the left and right images, respectively.

We have performed experiments to illustrate the advantages of the log-polar geometry over the cartesian images, for the vergence behavior. An object was placed in front of the camera rig and the cameras moved to sweep the full vergence space. The cartesian images size is 128x128 pixels whereas the log-polar images size is 64x32 pixels, thus achieving a reduction factor of 8 and a better computational efficiency. Figure 3 shows a plot of $\mathcal{C}(\nu)$, the NCC index, as a function of the vergence angle, both for the cartesian and log-polar images.

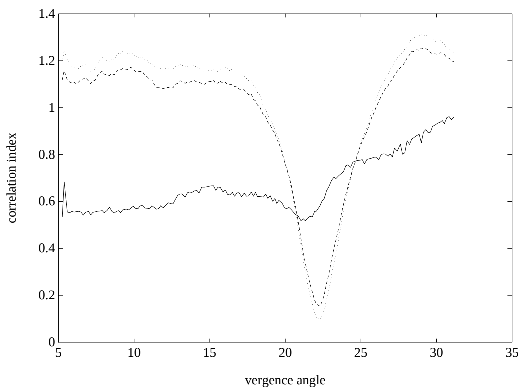


Figure 3: $\mathcal{C}(\nu)$: correlation index as a function of the vergence angle. Solid line: $\mathcal{C}(\nu)$ for a cartesian pair of images. Dashed and dotted lines: $\mathcal{C}(\nu)$ for log-polar images with different layouts. The minimum values represent the correct vergence angle.

This experiment shows that for the cartesian images, the correlation index exhibits various local minima and a nonuniform behavior. On the other hand, for the log-polar images, with 8 times less data, the curve is much smoother and presents a sharp minimum in the vicinity of the correct vergence angle.

Except for very cluttered environments, in the experiments made with NCC of log-polar images, all correct vergence configurations were associated to local minima in $\mathcal{C}(\nu)$.

# 4 Control algorithm

In this section we propose and compare two strategies for vergence control. Both of them use correlation measures from log-polar images to drive the cameras motion, using a standard feedback control architecture, as represented in Figure 4.

Both vergence control mechanisms are based only on correlation cues acquired continuously over time.

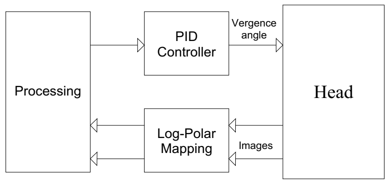


Figure 4: Vergence loop control diagram.

No calibration information is needed in order to accomplish vergence.

The difference among the two strategies remains on the processing step, as will be described in the next sub-sections.

## 4.1 Iterative vergence control

The first algorithm is based on the minimization of the correlation of the log-polar stereo pair, calculated as the NCC between the two images. As we have seen in the former section, for log-polar images the correlation function exhibits a sharp minimum in correct vergence configurations.

The NCC calculation provides a non-negative correlation index, $\mathcal{C}_i$, which is not suitable as direct input for a PID controller and must be further processed.

Usual gradient descent techniques require calculation of the gradient of the function $\mathcal{C}(\nu)$ where $\nu$ is the vergence angle. The control reference for the relative vergence angle value ($\delta\nu_i$) would be given by:

$$\delta\nu_i = -\alpha \frac{\Delta\mathcal{C}_i}{\Delta\nu_i}$$

where $\frac{\Delta\mathcal{C}_i}{\Delta\nu_i}$ is a discrete estimate of $\nabla\mathcal{C}(\nu_i)$.

The determination of the gradient is in practice very sensitive to noise. Furthermore, this algorithm is well suited to parabolic shaped functions, where the reference signal tends to zero smoothly as we approach the minimum, whereas, in the present case, the slope of $\mathcal{C}(\nu)$ is very high.

In [9], an algorithm was proposed to calculate the reference value to control, using the inverse gradient of the correlation, and the correlation value itself, as a normalization factor to smooth the noise coming out from the gradient :

$$\delta\nu_i = -\frac{\mathcal{C}_i}{\nabla\mathcal{C}(\nu_i)}$$

Similarly to the gradient descent, we have found this algorithm very unstable and potentially presents a singularity in the minimum ($\nabla\mathcal{C}(\nu)$ becomes zero), which poses problems to the control.

Since most of the instability of these algorithms is due to the noise in the calculation of the gradient of the correlation index, our approach avoids relying on values of any differential measure to determine an appropriate signal to control. Therefore, for the experiments made, the designed algorithm relies only on the value of the correlation index, the sign of its first difference which indicates the increase or decrease of $\mathcal{C}(\nu)$ and the sign of the previous reference value:

$$s_i = \mathcal{C}_i \cdot [-sgn(\Delta\mathcal{C}_i)] \cdot sgn(s_{i-1}) \qquad (12)$$

As $\mathcal{C}_i$ is always non negative, if we are moving in the right direction, then $\Delta\mathcal{C}_i$ is a negative value and the direction of vergence is maintained. In the other way, if we are moving in the wrong direction, the control reference inverts the direction of motion. The factor $\mathcal{C}_i$ respects the intuitive fact that when far from the minimum, the reference value should be high and when $\mathcal{C}_i$ approaches the minimum, the reference should be low.

### 4.2 Preprogrammed vergence control

There are several theories to explain the vergence eye movement control. The dual mode theory, proposed in [15], considers that the movement is composed by an initial predetermined motor force program, followed by a slower movement to attain accurate bifixation. Their experiments with human individuals have shown that the first part consists in a step-like behavior, as a result of sequential preprogrammed responses, which are triggered when the retinal disparity is sufficiently large.

A different theory is proposed in [10], based on a number of discrete "disparity" channels that are activated depending on the disparity level. When a channel is activated, it will trigger the control process to produce a preprogrammed movement.

Our second control strategy uses a similar channel based approach. We consider a set of $n$ channels, associated to a set of $n$ horizontal disparities $\mathcal{D} = \{d_1, \cdots, d_n\}$, defined on the cartesian plane. Large disparities are sampled at coarser resolution, thus inducing step-like responses of the control system, corresponding to the first mode of the dual-mode theory. Small disparities are sampled at fine resolutions, hence generating smoother responses (second mode).

For each channel with disparity $d_i$, we determine the corresponding disparity map in the log-polar plane. This procedure can be done off-line. Then, we compute the correlation of the left image with the right image, after warping by the log-polar disparity map corresponding to each channel. The correlation value indicates the likelihood of the images having an horizontal cartesian disparity $d_i$. Therefore, the channel yielding the highest correlation corresponds to the dominant image disparity.

In [10], it is argued that each channel has a distinct low pass filter characteristic with different gains and time constants. In our approach, instead, the selected disparity is directly used by the vergence dynamic controller.

## 5 Results

The vergence system described in the previous sections was implemented on the **Medusa** Stereo Head [13], shown in Figure 5. The head has four mechanical degrees of freedom, including independent vergence for both cameras, and a common tilt and pan axes. All

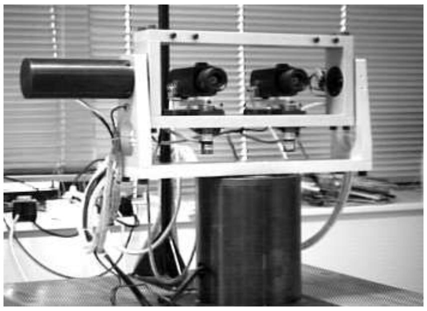

Figure 5: Medusa stereo head used for the experimental tests.

the computations, including image mapping and processing, are done at a PENTIUM PC computer running at 133MHz. In the current implementation, video rate (25Hz) is achieved without the need for special image processing hardware.

Several tests were made in order to evaluate the two main aspects addressed in this paper. First we performed various experimental tests in order to assess the influence of the main design parameters of the log-polar sensor. The second set of experiments illustrates the performance of the proposed vergence control strategies.

### 5.1 Log-polar layout performance analysis

Various authors have used log-polar image representation and have pointed out many of its advantages [16]. However, it is seldom clear what parameters of the log-polar representation to use, and what is their influence on the expected performance of the algorithms.

Here, we will discuss how parameter changes in the log-polar sensor affect the correlation index $\mathcal{C}(\nu)$. The parameters considered are : the angular resolution, which is determined by the number of cells per eccentricity; the radial resolution, determined by the logarithmic factor, and the minimum radius, set by the number of eccentricities, for a given logarithmic factor.

The tests were made with a static scene with several depth planes. The scene is composed by a flat object roughly located at one meter from the cameras, against a second plane of objects at about 3 meters and the background at 5 meters.

Each experiment consists in sweeping the vergence angle space and computing the correlation index for several camera positions, and for the various log-polar geometries. Through the analysis of the shape of the correlation index curve, one can evaluate the effects of changing a given design parameter.

Changes in the angular resolution, minimum radius and radial resolution are considered. The corresponding log-polar layouts are shown in Figures 6, 7 and 8 together with the results of the experiments.

As can be observed in the plots, the major differences in the shape of the correlation function happen when the minimum radius is changed. A large minimum radius reduces the capability to detect correct

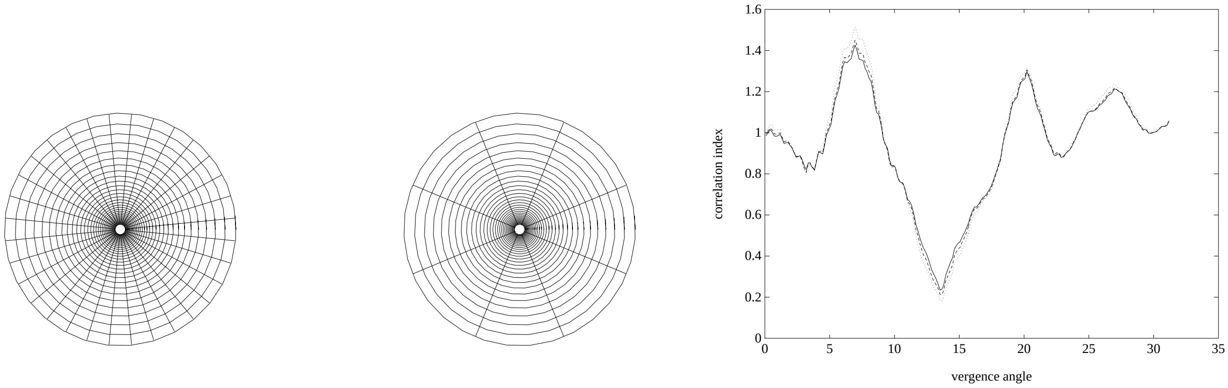


Figure 6: Effect of changing the angular resolution. Plots of $\mathcal{C}(\nu)$ for different angular resolutions. The number of discretization steps are 64 for the solid line, 32 for the dashed line and 16 for the dotted line.

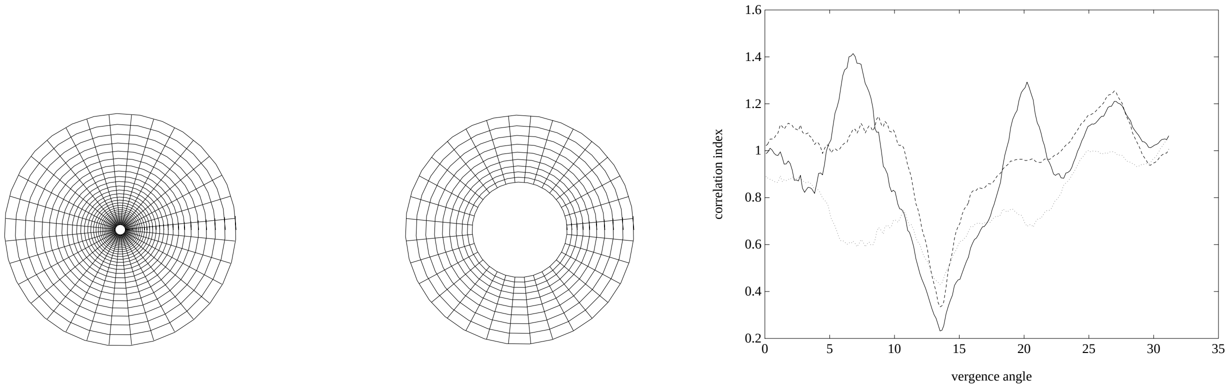


Figure 7: Effect of changing the minimum radius. Plots of $\mathcal{C}(\nu)$ for different minimum radius. The minimum radius is 2.82 pixels for the solid line, 13.43 pixels for the dashed line and 29.32 pixels for the dotted line.

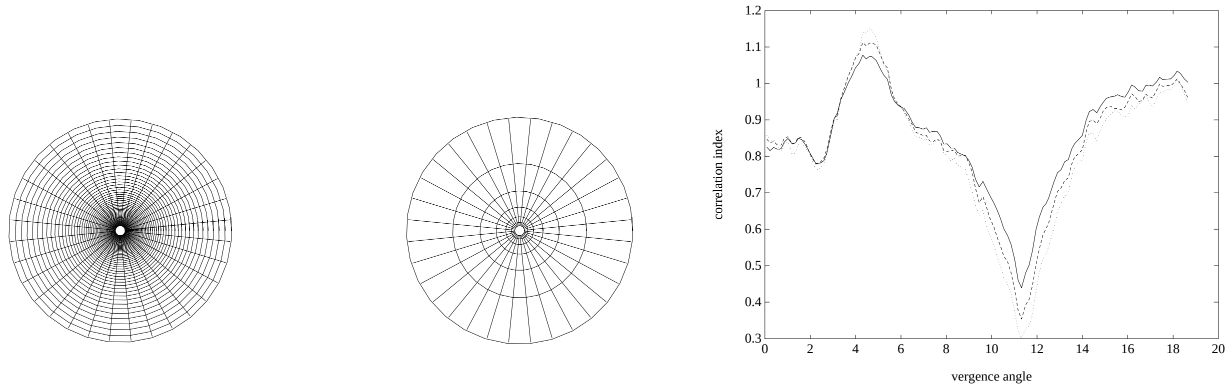


Figure 8: Effect of decreasing the radial resolution. Plots of $\mathcal{C}(\nu)$ for different radial resolutions. The logarithmic factor is $k = 1.02$ for the solid line, $k = 1.15$ for the dashed line and $k = 1.30$ for the dotted line.

vergence configurations because the correlation function presents a smaller "peak". As we want to verge on the object roughly located in the center of the retinas, the existence of a large blind area greatly reduces the amount of information available.

A second observation can be made in relation to the slight improvement obtained with the change in the radial resolution. For higher values of $k$ the profile of the correlation function around the correct vergence angle is wider and deeper, which should lead to better performance. As discussed in section 3, image patches are weighted according to $1/k^{2\xi}$. Large values of $k$ result in a bigger weight given to the central area of the images, increasing the sensitivity to changes in the vergence angle. For small values of $k$ the sensitivity to overall disparity becomes similar to the cartesian case (compare equations (8) and (11)).

Finally, the experiment corresponding to the analysis of the angular resolution reveals that this issue does not influence the shape of the correlation functions, as was expected from equation (11) since there is no dependency on the angular resolution.

To conclude, we can say that, as a rule of thumb, we should choose a low value for the minimum radius $\rho_{min}$ and a high value for the logarithmic factor $k$. Obviously these parameters should remain within certain limits. Very low values for the minimum radius increase the redundant information contained in the foveal area. This redundancy, measured in the percentage of redundant pixels in the total number of pixels in the log-polar image, should be kept low, for efficiency purposes, depending on the constraints of each particular application.

Very large values for the logarithmic factor tend to reduce excessively the number of eccentricities. Although we have not tested the effect of extremely reduced number of eccentricities, it should be kept under reasonable limits, or otherwise, the loss of resolution in the foveal area can be important.

### 5.2 Controlled Vergence

We have performed various experiments to test the proposed closed loop vergence control systems. A variety of objects, located at different distances from the stereo head, have been used. In the majority of the situations both approaches were able to verge on the targets located in the center of the images, once the initial disparity was kept within appropriate limits.

To evaluate the response to step-like inputs, we placed an object at a given distance from the camera and removed it afterwards. The head moved from its initial position to verge on the target and then verged on the background. A proportional closed loop controller was used for the two proposed algorithms.

Figure 9 shows the time evolution of the vergence angle for this experiment. The time scale units are sampling periods (about 40ms).

The main point to notice is the oscillation exhibited by the iterative vergence control when verging in the target. The reason is that this algorithm relies on the sign of the temporal difference of the correlation index. To obtain this signal, the cameras must move continuously, even when the desired vergence angle is

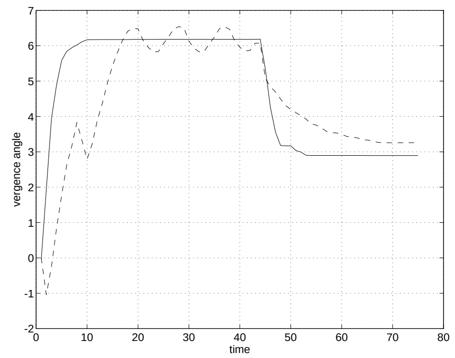


Figure 9: Response of the closed loop control system to step-like stimuli. Solid plot: preprogrammed vergence control. Dashed plot: iterative vergence control.

achieved. This fact constrains the value of the controller gain, thus limiting the raising time of the response. For a higher gain controller, the amplitude of the oscillations would be larger.

The control signal (equation 12), depends on the correlation index, which is a function of the overall mismatch of the images. Therefore, when the system is verging on the background, oscillations are not perceivable because the lower correlation value leads to very small control signals.

The preprogrammed vergence control shows a better performance. The correct vergence angles are reached after a short transient period, and in a fashion similar to a first order dynamic system. Furthermore, we noticed that the range of initial disparities for which the system converges is bigger for the preprogrammed vergence control. This depends, obviously, in the range of the disparity channels used which is easy to define, while in the iterative vergence control the dependency arises on the size of the region of convergence of the correlation function, which can be as small as 5 degrees.

## 6 Conclusions and further work

In this paper we have shown that fast vergence control can be achieved, using simple correlation measures over log-polar images. Vergence can be attained with the close integration of perception and control, without the explicit measurement of disparity and the need to calibrate the cameras.

The work described here is closely related to biological vision systems in at least two ways. On one hand the log-polar image sampling mechanism is inspired after the topology of the human retina. On the other hand the human vergence control system is based on a continuous stream of visual information, similarly to the approaches adopted.

The radial decrease of the resolution in the log-polar images contributes to an important reduction of the amount of data to be processed and was found to be more adequate for vergence control, when compared to cartesian images.

Two vergence control strategies are proposed. The iterative vergence algorithm is based on a hill climbing procedure applied to the correlation index space. Differential measurements are avoided because of noise sensitivity. The control law is solely dependent on the value of the similarity index between two images. The preprogrammed vergence control algorithm follows a channel based approach. Depending on the level of disparity present in the log-polar images, a different channel is selected. The selection is made according to correlation values calculated over all channels defined. The selected channel determines the input to the controller module.

The system was implemented, in real-time, on the stereo head Medusa. A large number of experiments are described in the paper, using realistic scenes and cluttered environments. In a first set of tests, we show how the log-polar sensor parameters affect the performance of the vergence system. These results establish a criterion to choose the log-polar sensor parameters, in order to improve performance.

Finally, results with controlled vergence are described. Both control algorithms confirm the usefulness of correlation measures to control vergence. However, major differences in their dynamic behavior are clear. The preprogrammed vergence control strategy was found to be more stable and robust. Furthermore, in dynamic situations it shows faster reaction and settling times.

Future work should address the problem of integrating the vergence system with the other similar ocular movements, following the same approach. Particularly, we envisage various other visual behaviors specialized in different parts of the visual field (foveal or peripheral), mutually cooperating in the task of following a moving target, or more complex tasks such as docking, grasping, etc.

The joint analysis of a purposive sensor geometry design, adequate control strategies directly integrating perception and action and overall system performance characterization provides, in our opinion, an important framework to cope with a large set of applications on vision and robotics.

## References


[1] Y. Aloimonos, I. Weiss, and A. Banddophaday. Active vision. *Int. Journal of Computer Vision*, 1(4):333–356, January 1988.

[2] R. Bajcsy. Active perception. *Proc. of the IEEE*, 76(8):996–1005, August 1988.

[3] A. Bernardino and J. Santos-Victor. Sensor geometry for dynamic vergence: characterization and performance analysis. In *Proc. of the Performance Characteristics of Vision Algorithms Workshop, ECCV96*, Cambridge, England, April 1996.

[4] R. Carpenter. *Movements of the eyes*. Pion, London, 1988.

[5] H. Christensen, K. Bowyer, and H. Bunke, editors. *Active Robot Vision : camera heads, model based navigation and reactive control*, volume 6. World Scientific, 1994.

[6] D. Coombs and C. Brown. Real-time binocular smooth pursuit. *Int. Journal of Computer Vision*, 11(2):147–164, October 1993.

[7] O. Faugeras, P. Fua, B. Hotz, R. Ma, L. Robert, M. Thonnat, and Z. Zhang. Quantitative and qualitative comparison of some area and feature based stereo algorithms. In W. Förstner and St. Ruwiedel, editors, *Robust Computer Vision*. Wichmann, 1992.

[8] F. Ferrari, J. Nielsen, P.Questa, and G. Sandini. Space variant sensing for personal communication and remote monitoring. In *Proc. of the EU-HCM Smart Workshop*, Lisbon, Portugal, April 1995.

[9] F. Panerai, C. Capurro, and G. Sandini. Space variant vision for an active camera mount. In *Proc. SPIE AeroSense95*, Florida, USA, April 1995.

[10] M. Pobuda and C. Erkelens. The relation between absolute disparity and ocular vergence. *Biological Cybernetics*, 68:221–228, 1993.

[11] D. Robinson. The oculomotor control system: A review. *Proc. of the IEEE*, 56(6), June 1968.

[12] J. Santos-Victor, G. Sandini, F. Curotto, and S. Garibaldi. Divergent stereo in autonomous navigation : From bees to robots. *Int. Journal of Computer Vision*, 14(2):159–177, 1995.

[13] J. Santos-Victor, F. van Trigt, and J. Sentieiro. Medusa - a stereo head for active vision. In *Proc. of the Int. Symposium on Intelligent Robotic Systems*, Grenoble, France, July 1994.

[14] E.L. Schwartz. Spatial mapping in the primate sensory projection : Analytic structure and relevance to perception. *Biological Cybernetics*, 25:181–194, 1977.

[15] J. Semmlow, G. Hung, J. Horng, and K. Ciuffreda. Disparity vergence eye movements exhibit preprogrammed motor control. *Vision Reserach*, 34(10):1335–1343, 1994.

[16] M. Tistarelli and G. Sandini. On the advantages of polar and log-polar mapping for direct estimation of the time-to-impact from optical flow. *IEEE Trans. on PAMI*, 15(8):401–411, April 1993.

[17] R. Wallace, P. Ong, B. Bederson, and E. Schwartz. Space variant image processing. *Int. Journal of Computer Vision*, 13(1):71–90, September 1995.

[18] C. Weiman. Log-polar vision for mobile robot navigation. In *Proc. of Electronic Imaging Conference*, pages 382–385, Boston, USA, November 1990.