



A review of log-polar imaging for visual perception in robotics

V. Javier Traver^{a,b,*}, Alexandre Bernardino^c

^a Computer Languages and Systems Dept., Universitat Jaume I, Castellón, Spain

^b Institute of New Imaging Technologies (INIT), Universitat Jaume I, Castellón, Spain

^c Instituto de Sistemas e Robótica, Instituto Superior Técnico, Lisbon, Portugal

ARTICLE INFO

Article history:

Received 31 July 2008

Received in revised form

6 October 2009

Accepted 13 October 2009

Available online 30 October 2009

Keywords:

Foveal imaging

Log-polar mapping

Real-time robotics

Active vision

ABSTRACT

Log-polar imaging consists of a type of methods that represent visual information with a space-variant resolution inspired by the visual system of mammals. It has been studied for about three decades and has surpassed conventional approaches in robotics applications, mainly the ones where real-time constraints make it necessary to utilize resource-economic image representations and processing methodologies. This paper surveys the application of log-polar imaging in robotic vision, particularly in visual attention, target tracking, egomotion estimation, and 3D perception. The concise yet comprehensive review offered in this paper is intended to provide novel and experienced roboticists with a quick and gentle overview of log-polar vision and to motivate vision researchers to investigate the many open problems that still need solving. To help readers identify promising research directions, a possible research agenda is outlined. Finally, since log-polar vision is not restricted to robotics, a couple of other areas of application are discussed.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Both natural and artificial visual systems have to deal with large amounts of information coming from the surrounding environment. When real-time operation is required, as happens with animals or robots in dynamic and unstructured environments, image acquisition and processing must be performed in a very short time (a few milliseconds) in order to provide a sufficiently fast response to external stimuli. Appropriate sensor geometries and image representations are essential for the efficiency of the full visual processing stream. To address this problem it is wise to look for the solutions present in biological systems, which have been optimized by millions of years of evolution. For instance, the visual system of many animals exhibits a non-uniform structure, where the *receptive fields*¹ represent certain parts of the visual field more densely and acutely. In the case of mammals, whose eyes are able to move, retinas present a unique high resolution area in the center of the visual field, called the *fovea*. The distribution of receptive fields within the retina is fixed and the fovea can be redirected

to other targets by ocular movements. The same structure is also commonly used in robot systems with moving cameras [2–6].

In the late 70s computer vision researchers broke new ground by considering the foveal nature of the visual systems of primates as an alternative to conventional uniform resolution sensors for artificial perception in computers and robots. Earlier on, biological findings in the visual cortex of monkeys [7] had shown that the displacement of a light stimulus in the retina produces displacements in the cortex that are inversely proportional to the distance to the fovea. This effect, also known as *cortical magnification*, indicates a general scaling behavior by which both receptive field spacing and size increase linearly with eccentricity, i.e. the distance from the fovea [8]. It was found that responses to linear stimuli originating in the fovea lie roughly along lines in the cortex, and circular stimuli centered on the fovea produce linear responses in the cortex at approximately orthogonal orientations [9]. Thus, the information transmitted between the retina and the visual cortex is organized in an approximate logarithmic-polar law [10].

The foveal structure of the retina of some animals is, together with their ability to move the eyes, a fundamental mechanism in the control of visual perception. In the late 80s and early 90s, researchers started exploiting eye movements to achieve complex visual tasks. The paradigm of active vision emerged as a powerful concept to endow an active observer with the ability to find more efficient solutions to problems that, from a passive vision perspective, were ill-posed and non-linear [11]. The idea can be generalized beyond pure perception by including manipulation. The smart usage of robot arms and hands, for instance, opens up many more

* Corresponding author at: Computer Languages and Systems Dept., Universitat Jaume I, Castellón, Spain.

E-mail addresses: vtraver@uji.es (V. Javier Traver), alex@isr.ist.utl.pt (A. Bernardino).

¹ Receptive fields are spatially organized biological computational elements and, according to [1], are probably the most prominent and ubiquitous computational mechanism employed by biological information systems. The concept is revisited in Section 2.

possibilities for better visual perception [12]. A great deal of excitement was aroused at that time with regard to the present and future possibilities of active vision [13]. For instance, by purposefully moving the eyes, an observer with a foveal low resolution sensor can acquire a “virtual” high resolution image of its entire field of view [14]. Therefore, following on as the next natural step forward, a new concept appeared —*space-variant active vision* [15].

Since then, efforts have been made to explore the advantages that foveal-like log-polar imaging can bring to robotic applications. However, after almost three decades since those initial studies, no systematic and comprehensive work has been published that attempts to review past research on the topic. While a careful review of log-polar models was conducted ten years ago in [16], its focus was a detailed study and comparison of log-polar mapping templates and models with overlapping receptive fields. More recently, the motivations for retina-like sensors and the properties of the log-polar mapping were nicely considered in [17,18]. Another paper [19] surveyed foveated sensors with a particular emphasis on image processing issues.

Thus, while these few review-like papers did not consider the applications and usages of log-polar images in depth, we feel that such an analysis is needed, to reflect on past achievements, discuss current challenges, and predict future developments. Additionally, this literature overview would be a valuable aid to any researcher interested in approaching the field, particularly to beginners. Finally, another important benefit of such a survey is that of helping to promote further work, both on the theoretical and practical sides of log-polar vision.

Therefore, the present survey aims to complement these previous reviews, by looking further into the variety of applications of log-polar sensing that have been proposed. Furthermore, our analysis pays particular attention to robotic applications. Hence, the usage of the log-polar transform for pattern recognition issues, though important, is not considered here. Studies and reviews on this other perspective can be found elsewhere [20–23].

An overview of the log-polar mapping (Section 2) allows the readers to become familiar with the basics of this transform. There are different ways to obtain log-polar images either from conventional images or directly from a scene, using software-and/or hardware-based solutions (Section 3). In Section 3 we also address issues regarding how the mapping parameters may influence the visual task and whether the selection of optimal parameters can be automated. The area of visual attention and salience computation under foveal vision (Section 4) has not been explored very much, even though it plays a key role in active object search and recognition, in exploratory gaze strategies, and in the proper integration of different visual tasks in practical scenarios. One of the visual processes where log-polar imaging is most suitable is probably active target tracking (Section 5), and substantial research has been devoted to this topic. Some advantages have also been found in estimating the observer's motion using log-polar images (Section 6), basically due to its polar geometric nature which fits particularly well with time-to-collision computation and other navigation tasks in mobile robots. Binocular depth estimation has been considered with a joint usage of log-polar imaging and active vergence movements (Section 7). There are also a number of less conventional sensor arrangements and less known properties of log-polar imaging that deserve some consideration. It is our prediction that many of these issues will open up the door to fascinating new research challenges in automatic foveal vision not only within robotics but also in other fields of application (Section 8).

2. Log-polar mapping

Log-polar mapping is a geometrical image transformation that attempts to emulate the topological reorganization of visual

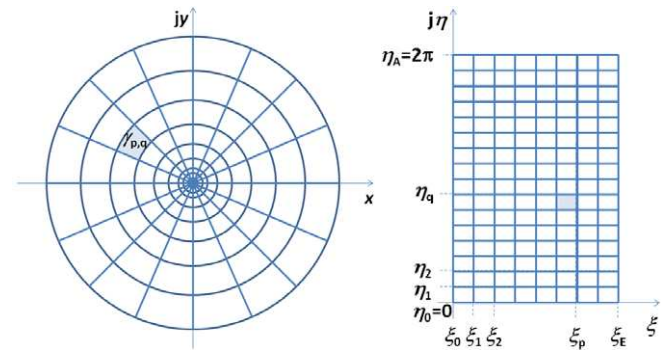


Fig. 1. The $\log(z)$ model for retino-cortical mapping, where a central circle of small radius has been left out of the mapping in order to deal with the singularity problem. The retinal plane (left) is mapped onto the cortical plane (right) via $w = \log(z)$. Concentric circumferences and radial lines in the retinal plane become straight lines in the cortical plane. Rectangular cells in the transform domain correspond to sections of concentric annuli in the original domain.

information from the retina to the visual cortex of primates. It can also be found in the literature under different names, such as *log-polar transformation* or the *log(z) model*. The reason for this last denomination comes from the fact that the mapping can be mathematically modeled by the complex logarithmic function $\log(z)$, where z is the complex variable representing points on the image plane.

2.1. Definition

Let us consider the complex retinal and cortical (log-polar) planes, represented by the variables $z = x + jy$ and $w = \xi + j\eta$, respectively (j is the complex imaginary unit). The complex log-polar mapping is:

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

and the log-polar coordinates ξ (eccentricity) and η (angle) are given by:

$$\xi = \log(|z|) = \log \sqrt{x^2 + y^2}$$

$$\eta = \arg(z) = \text{atan2}(y, x)$$

where $\text{atan2}(y, x)$ denotes the two-argument arctangent function that considers the sign of x and y in order to determine the quadrant of the resulting angle.

This mapping transforms concentric circumferences and radial lines in the retinal plane into straight lines along the ξ and η directions in the cortical plane, respectively (Fig. 1).

2.2. Properties

The main properties of the mapping and some of their practical implications are as follows:

Conformal mapping: The cortical image (also called *log-polar image*) preserves oriented angles between curves and neighborhood relationships, almost everywhere, with respect to the retinal image. In theory, this property predicts that image processing operations developed for Cartesian images can be applied directly to log-polar images. In practical terms, however, specific algorithms are required in many applications.

Elegant trade-off solution between these three mutually opposing criteria: wide field of view, high visual resolution and little data to process. For robotics applications, there are two significant benefits of this particular sampling. On the one hand, the reduced size of log-polar images

(as much as 30 times smaller than uniformly-sampled Cartesian images have been reported) hugely facilitates real-time visual data processing. On the other hand, the radially logarithmic sampling entails that a higher resolution is devoted to the center of the scene (fovea area) which, in turn, means that foveal information is represented by a big number of pixels in the log-polar image. One of the most interesting practical implications of this *foveal predominance*, is that foveated targets can be tracked without being explicitly segmented from the background. Additionally, the segmentation of a verged target in a binocular system becomes easier.

Biological plausibility: It approximates the receptive field distribution and retino-cortical mapping in the visual system of mammals. Depending on the application, this mimicry of biological solutions can be seen as an experimental support tool to neurophysiology, either to validate its findings or to propose new hypotheses. On the other hand, from an engineering point of view, it is a rich source of inspiration for sound strategies developed in the natural animal world.

Rotation and scaling “invariance”: When the original image is rotated or scaled with respect to its center, patterns in the log-polar image only undergo translations, thus preserving their shape. This geometric property, also known as *edge* or *shape* invariance, is particularly helpful for rotation- and scale-invariant pattern recognition, but it has also been exploited for motion estimation in active tracking scenarios, as described in Section 5.

The first two properties (conformality and data selection) are common to other space-variant imaging models, while the last two (biological inspiration and edge invariance) are specific advantages of the log-polar mapping.

2.3. Singularity at the origin

One weakness of the $\log(z)$ model is the existence of a singularity in the center of the image, since it is not possible to evaluate $\log(0)$. This means that receptive fields would become infinitely small toward the origin. Therefore, points in the fovea cannot be represented with this mapping. Two solutions are commonly used to overcome this problem: either using a different mapping for the fovea (e.g. the identity mapping or a pure polar structure) or applying the $\log(z + a)$ model. This other model was proposed in [24] as a better approximation to the retino-topic mapping of monkeys and cats. The $\log(z + a)$ model transforms points in the first and fourth quadrants ($x \geq 0$) of the retinal plane via

$$w = \log(z + a), \quad (2)$$

where a is a positive real number. In coordinates, we have:

$$\xi' = \log \sqrt{(x + a)^2 + y^2}$$

$$\eta' = \arctan \left(\frac{y}{x + a} \right).$$

Because x is positive in the first quadrant and a is a positive constant, the minimum value of ξ' is finite ($\min \xi' = \log(a)$), therefore avoiding the singularity present in the $\log(z)$ model (Fig. 2(a,b)).

The mapping for the other quadrants is obtained by symmetry, which saves some computation time. Additionally, pixels in the log-polar plane are rearranged to match the quadrants of the Cartesian plane and to be connected at the origin (Fig. 2(c,d)). The resulting mapping is:

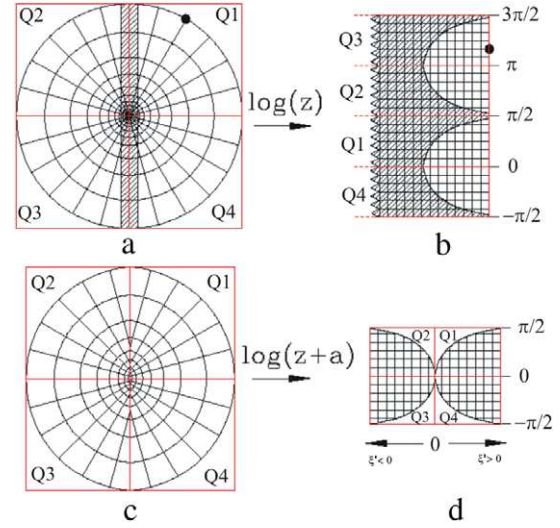


Fig. 2. The $\log(z + a)$ model [25] can be seen as removing the shaded region in (a) from the $\log(z)$ model. The new log-polar coordinates are designed to have connectivity at the origin and regions organized in the usual quadrants (Q1, Q2, Q3, Q4). Figure adapted from [25].

Source: With kind permission from Springer Science + Business Media: International Journal of Computer Vision, Volume 13, 1995, pages 75, R. Wallace, P. Ong, B. Bederson, and E. Schwartz, Fig. 3.

- first and fourth quadrants ($x \geq 0$):

$$\xi' = \log \sqrt{(x + a)^2 + y^2} - \log(a)$$

$$\eta' = \arctan \left(\frac{y}{x + a} \right)$$

- second and third quadrants ($x < 0$):

$$\xi' = -\log \sqrt{(-x + a)^2 + y^2} + \log(a)$$

$$\eta' = \arctan \left(\frac{y}{-x + a} \right).$$

The additional term $\log(a)$ produces a shift in the radial coordinates such that points at the origin result in $\xi = 0$, thus forcing the connectivity of the different quadrants. The use of the $\arctan(\cdot)$ function instead of $\text{atan2}(\cdot, \cdot)$ is chosen so that the angular range is $[-\frac{\pi}{2}, \frac{\pi}{2}]$, which promotes the arrangement of the quadrants in the usual sequence.

The $\log(z + a)$ model lacks the property of exact scale invariance. However, this is often tolerated in practical applications. Other solutions to the singularity problem have also been proposed [16,26–28].

2.4. Discretization

The $\log(z)$ and $\log(z + a)$ are conceptual models defined in continuous coordinates. They tell us how retinal and cortical coordinates are related, but, in practice, the mapping must be discretized. The conventional approach considers the cortical plane to be uniformly discretized as if it was an ordinary Cartesian image, i.e. covered with a dense grid of rectangular regions. Thus, let us consider a grid of $E \times A$ rectangular pixels (also known as *cortical cells*) whose corners are at coordinates $w_{p,q} = \xi_p + j\eta_q$, $p \in \{1, \dots, E\}$, $q \in \{1, \dots, A\}$, with E as the number of eccentricities (radial rings) and A the number of angular sectors. In the retinal plane, the corresponding regions, $\gamma_{p,q}$, are shaped like angular sections of concentric annuli. Such regions are called *retinal cells* or *receptive fields* (RFs). Fig. 1 illustrates the partition of the cortical plane with a uniform grid and the corresponding retinal cells.

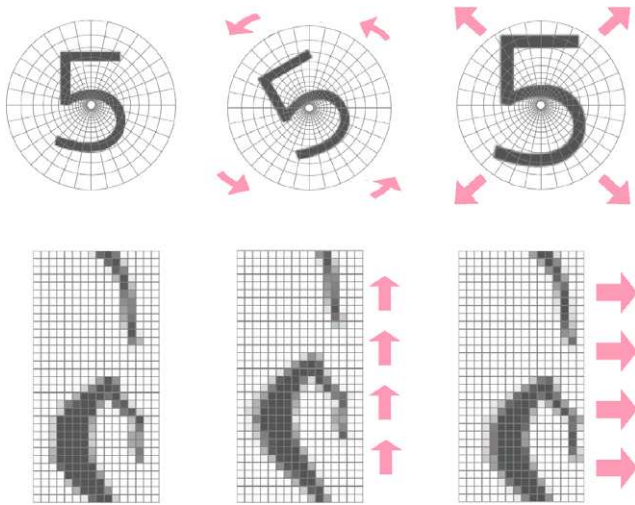


Fig. 3. Top: Retinal images on the log-polar mapping template. Bottom: The corresponding log-polar images. To illustrate the edge invariance property, the image on the left is rotated (middle) and scaled (right), which correspond to approximate translations in the log-polar domain, in the angular (η) and radial (ξ) directions, respectively, as shown with the arrows.

Each retinal cell has a weighing function $\phi_{p,q}(z)$ associated to it that represents the way the value of each cortical pixel is obtained from the information in the retinal array. This can be modeled by

$$c_{p,q} = \langle f, \phi_{p,q} \rangle, \quad (3)$$

where f is the function representing the Cartesian image. With uniform weighting functions, this operation represents the simple averaging of the retinal information within the spatial support of each retinal cell. Neighbor cells in the retinal domain are also neighbors in the cortical domain, except along the angular discontinuity and the radial singularity. Shape invariance to centered rotations and scalings no longer holds perfectly for the discretized $\log(z)$ model. However the approximation is good enough for practical applications, if discretization is not too coarse (Fig. 3).

2.5. Other implementation details

Models following biological data more closely have overlapping cells. They are computationally more expensive than non-overlapping ones but gain in smoothness of the log-polar image pixels. The models presented in [29] and [30] have cells with a log-polar distribution but with circular shapes and a moderate amount of overlap with the neighbors. In [29] a tessellation is proposed with a linear relation between receptive field size and eccentricity, and a cell overlap of 50%, as shown in Fig. 4(a). The proposal in [30] also uses circular receptive fields but attempts to minimize the amount of overlap between them by using a slightly different organization of receptive fields where direct neighbors are not in the same ring (Fig. 4(b)). Besides those seen above, there are other implications of log-polar sampling, as well as less obvious properties. For instance, discrete log-polar images and their internal representation introduce some practical difficulties, as the shapes and motions of the objects are distorted in a complex non-linear and inhomogeneous way. Appropriate algorithms and strategies have to be devised carefully to deal with these issues. Another example is the difference in resolution between the fovea and the periphery which calls for active strategies from the observers so that the gaze can be redirected according to the scene, the goals of the task and the ongoing visual events.

In terms of data structures in computer implementations, a log-polar image may not differ from a conventional image, except for

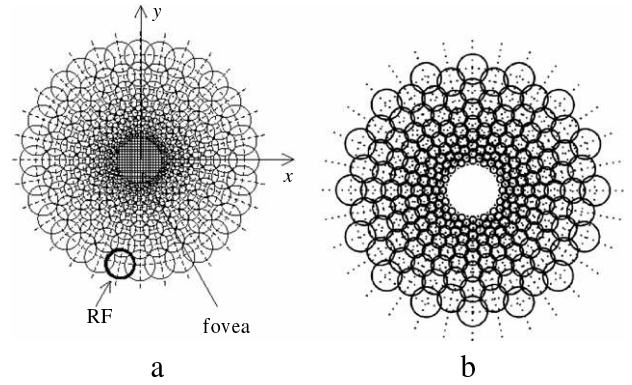


Fig. 4. Overlapping models: (a) the model of [29] implemented in [31], and (b) the model in [30] (figure taken from [16]).²

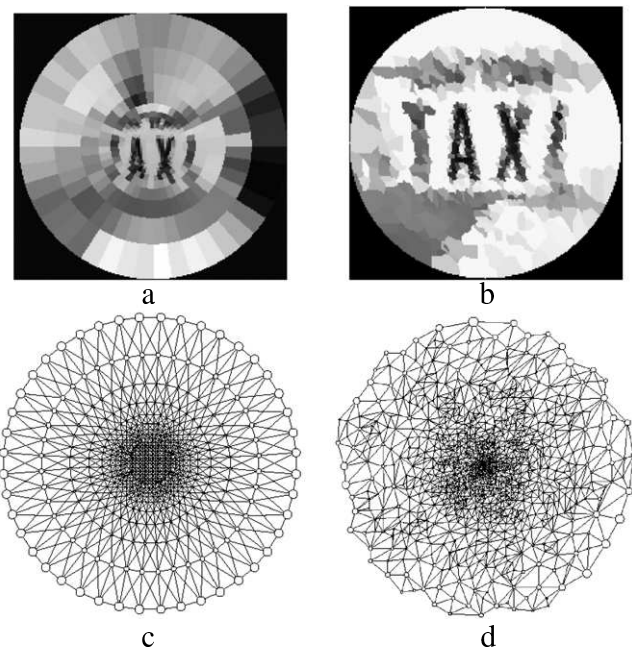


Fig. 5. The connectivity graph (CG) [33,34]: (a, b) sensors with (a) a log-polar and (b) an arbitrary pixel arrangement; (c, d) their corresponding CGs.

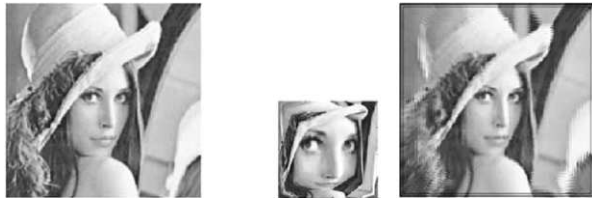
Source: With kind permission from Springer Science + Business Media: International Journal of Computer Vision, Volume 13, 1995, pages 77, 79, R. Wallace, P. Ong, B. Bederson, and E. Schwartz, Fig. 4 and Fig. 6.

the different meaning of the axes (row and column indices). In this case, one common difficulty to be tackled refers to the circularity of the angular axis, since the two ends of the rectangular log-polar image have to be processed as if they were really connected. To solve these kinds of technical complexities, an alternative representation, the connectivity graph, was proposed [32]. It has the advantage of being very general and able to accommodate arbitrary sensor lattices, as illustrated in Fig. 5. Local image operations, such as edge detection, can therefore be applied without special cases (e.g. image boundaries) in mind. As a disadvantage, graphs have to be defined, built, and processed, which may result in a loss of efficiency in particular circumstances.

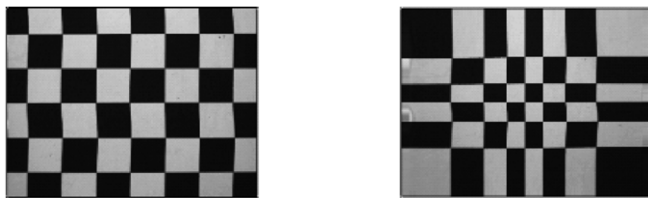
² Source: Reprinted from Computer Vision and Image Understanding, Vol. 69, No. 2, M. Bolduc and M. D. Levine, A review of biologically motivated space-variant data reduction models for robotic vision, Pages 170–184, Copyright (1998), with permission from Elsevier.



(a) Reciprocal Wedge Transform (RWT) [46].
© 1995 IEEE.



(b) Cartesian Foveal Geometry (CFG) [52].
© 2006 IEEE.



(c) Dimensionally-independent exponential mapping (DIEM) [59].
© 1998 IEEE.

Fig. 6. Some alternative foveal imaging methods. Left column: original Cartesian image; Middle column: image in the transform space; Right column: Cartesian image after reconstruction from transformed image.

2.6. Noise issues

Three aspects that relate log-polar images and noise issues can be identified. One of them is the occurrence of noise problems that need to be dealt with during the design or implementation of a log-polar sensor. A second situation may arise in image processing operations or computer vision applications that are performed favorably in terms of noise tolerance thanks to the nature of the log-polar image. Finally, procedures can be designed to reduce the noise present in an existing log-polar image. The scarce literature existing on these three aspects is commented briefly below.

With respect to sensor design, the implementation of hardware-based log-polar sensors (more on this in Section 3) brings noise-related problems, such as those related with the different size of the sensing elements used at different eccentricities, and that have to be addressed [35–38].

Regarding image noise, the advantage of log-polar images has been considered for scale-invariant feature recognition [39]. Since the noise in different Cartesian pixels is uncorrelated, the signal-to-noise ratio (SNR) improves in log-polar pixels when the log-polar transform is computed from Cartesian images [40]. The averaging procedure associated to the each receptive field to compute the gray-level value of a log-polar pixel is essentially a low-pass filter, and that is why an enhanced noise tolerance is also observed when performing image diffusion in the log-polar domain [41]. Among linear filters, a simple moving average is shown to be optimal in reducing random white noise and preserving sharpness in the step response, with noise being reduced by a factor of $\sqrt{n}$ for n pixels being averaged [42, Ch. 5]. Since the size of the RFs increases with eccentricity, resolution decays, but the SNR increases with distance to the center. Furthermore, for a given layout of RFs, their overlapping allows each of the RFs to be larger, which improves its SNR if the averaged signals are locally correlated, as it usually

happens in natural images [43]. An optimal amount of overlap can be found as a trade-off between a greater SNR and a smaller data redundancy [43]. Recently, projection-based motion estimation has been shown to exhibit higher noise robustness in log-polar images than in Cartesian images [44]. But log-polar imaging is not only good at tackling sensor noise in the periphery, it is also highly insensitive to *perceptual noise* (such as unmodeled issues, image clutter and distractors), since these undesirable artifacts are blurred or de-emphasized with respect to the dominant foveal visual data. This advantage has been demonstrated in a number of visual tasks, including tracking, depth computation and vergence control, as commented in Sections 5 and 7 below. Certainly, this very fact becomes a disadvantage if some occlusion happens at the foveal area.

Concerning noise reduction, in [45] a discrete wavelet transform is applied to reduce white Gaussian noise in log-polar images. Interestingly, as a consequence of the implicit low-pass filtering discussed above, if low-pass filtering is required in log-polar images, this process can be limited to the central foveal area since the peripheral pixels, if large enough, already implement this. As a result, this can save further computational cost.

2.7. Other foveal methods

While log-polar imaging is the most common foveal method, many alternative foveating approaches have been proposed. These are generally aimed at (a) preserving linearities; (b) having a shiftable fovea, adaptive-size fovea, or multiple foveae, or (c) the flexibility of a family of parameterized transformations. All the techniques share as their common goal that of reducing the amount of visual information and they differ in the different geometries and data structures chosen to suit certain purposes.

The Reciprocal Wedge Transform (RWT) [46,47] (Fig. 6(a)) preserves linear features, which is not possible in the log-polar domain. The RWT has been shown to be suitable in road following and depth recovery tasks. Multiresolution, pyramid-based foveal-like mechanisms, such as the Cartesian exponential topology (CET), have been used in active vision [48], video transmission [49], and motion and image segmentation [50,51]. The Cartesian Foveal Geometry (CFG) [52] (Fig. 6(b)) is similar to the CET, but the receptive fields are of constant size regardless of their eccentricity and no actual multiresolution is used. While CET and CFG have been shown to work in some tasks, no particular benefit over log-polar imaging has been demonstrated besides the claimed advantage that conventional image processing procedures require less modification than in the log-polar case. In general, foveated systems have been proposed to reduce the required communication bandwidth by exploiting the limitations (space-variant nature) of the human visual system [53–55].

Multiple foveae systems are described in [56] and [57] for stereo and video-conference applications, respectively. The potential benefit of having a relocatable fovea or several foveae in active vision contexts is that no hardware is required for camera movements, and their costly and/or inaccurate movements due to their mechanical nature are avoided [58]. In practical terms, multi-foveae systems have virtually not been applied in robotic systems.

The dimensionally-independent exponential mapping (DIEM) [59] (Fig. 6(c)) can sample along the vertical and horizontal dimensions differently. Possibly, the main advantage of DIEM is its flexibility, so that its parameters can easily be chosen to fit the problem at hand by granting more importance to one dimension or the other, and by sampling more densely at the center or at the periphery.

In short, no foveation mechanism is suitable for all tasks, and all of them exhibit advantages and disadvantages. With respect to other techniques, the log-polar method is biologically motivated,

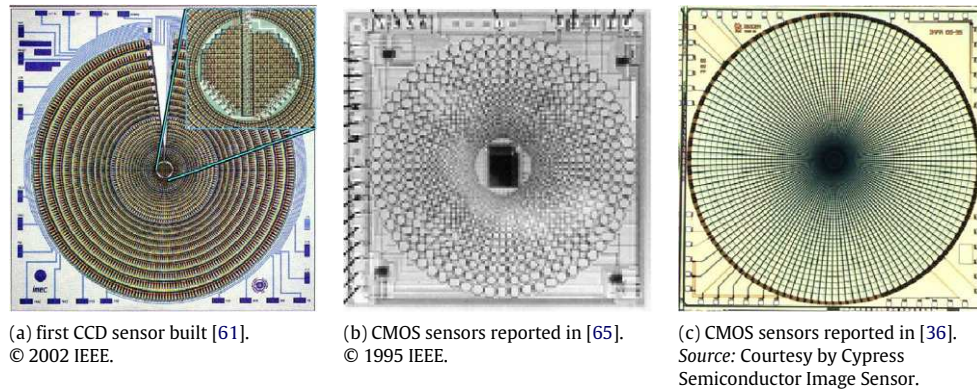


Fig. 7. Retina-like sensors.

it is one of the best known transforms, and it has been largely studied and shown to be useful in a wide range of applications. While linearities and translation invariance are certainly lost in log-polar images, the limitations associated with this lack can be compensated by proper algorithm design and/or appropriate active vision mechanisms, as reported in a significant part of the literature surveyed in this article. Real-time active target tracking and time-to-impact estimation are possibly the tasks that log-polar imaging is clearly best suited to, due to its favorable geometric arrangement.

3. Sensor design

While the market of commercial visual cameras is dominated by sensors with conventional Cartesian lattices, research and development on foveal imaging during the 90s led to the design and construction of several prototypes of the so-called *retina-like* cameras with sensitive elements arranged following a log-polar pattern. The first versions of these sensors and cameras were based on the Charged Coupled Device (CCD) technology [60–64], while later models used the Complementary Metal Oxide Semiconductor (CMOS) technology [65,66,36]. Technical difficulties were gradually overcome and the sensor was improved by increasing the number of pixels and introducing color sensing [67]. Three of these sensors are depicted in Fig. 7. A detailed report on vision chips, including foveal ones, is provided in [68,69].

The main advantage of these hardware-based log-polar imaging devices is that the captured image is already in the log-polar format, which allows or promises fast real-time processing in time-critical applications like active vision tasks. One main drawback, however, is represented by the technical obstacles faced either during their design and construction or even during their usage, which limits the range of applications. Furthermore, the hard-coded geometry layout makes it very difficult, if not impossible, to experimenting with different sensor parameters. An alternative (not log-polar) biologically inspired kind of visual sensor is offered by curved sensors [70], which mimic the hemispherical geometry of the human eye and bring forward additional challenges and possibilities.

Because flexibility in setting the log-polar parameters is such an important issue for experimental work, software that obtains log-polar images from Cartesian ones has been the most frequently used method. In this case, however, care should be taken to perform the log-polar transformation properly. The literature contains several works that provide hints for log-polar mapping computation [71–74], design or resolution considerations [75–77], and alternative fovea designs [78].

Unlike hardware-based sensors, software-based ones have the unavoidable cost of obtaining the log-polar image from

the Cartesian image. While this represents no (significant) disadvantage for proof-of-concept studies, it may be a drawback in real applications. To overcome this problem and still count on the flexibility of software, an intermediate solution is represented by virtual log-polar sensors [79,80]. This alternative still needs Cartesian images as input, but the implementation of the log-polar mapping is performed in special-purpose hardware cards, thus making the conversion faster while providing the possibility of setting the sensor parameters. Other possibilities, not considered to date, include parallel implementations of the log-polar transform, or exploring the speed-up provided by the now popular Graphics Processing Units (GPUs) [81,82].

In the case of hardware-based retina-like sensors, the maximum resolution achievable is dictated by the limits of the smallest sensing element that current imaging technology has to offer. In contrast, the maximum resolution of log-polar images formed by remapping the contents of Cartesian images depends on the size of the sensing element used in conventional Cartesian images. Therefore, while it is possible to oversample a Cartesian pixel to set the brightness of a set of corresponding log-polar pixels, this strategy can only lead to redundant visual information in the foveal area. For this reason, and because of the singularity of the logarithmic function, a central region is usually kept out of the mapping, thus becoming a blind area, as can be appreciated in Fig. 3 (top). The reverse phenomenon, under-sampling, occurs when many Cartesian pixels contribute to the same log-polar pixel. At the borders between receptive fields, a Cartesian pixel can be recursively subdivided, up to a desired precision, so that its contribution to a given log-polar pixel is weighted according to the percentage of its area lying in the associated receptive field [26,83].

Receptive fields can be defined to have different shapes, locations, amounts of overlapping and averaging functions [16]. Depending on the requirements, different solutions can be adopted. While for higher precision, sub-pixel methods [26] are possible, sampling without receptive fields averaging is sometimes used to speed-up computations [73]. Unfortunately, as far as we know, no systematic study has been performed to date on the actual practical impact that these different sampling strategies have.

There has been some interest in analyzing the influence of sensor parameters on performance in some visual tasks, such as vergence control [84] or depth recovery³ [85]. Formulating this problem in a more general way, one can be interested in finding the sensor parameters that are best suited to a particular task automatically. As a step further towards this automation, a

³ The concepts of *vergence control* and *depth recovery* are covered later, in Section 7.

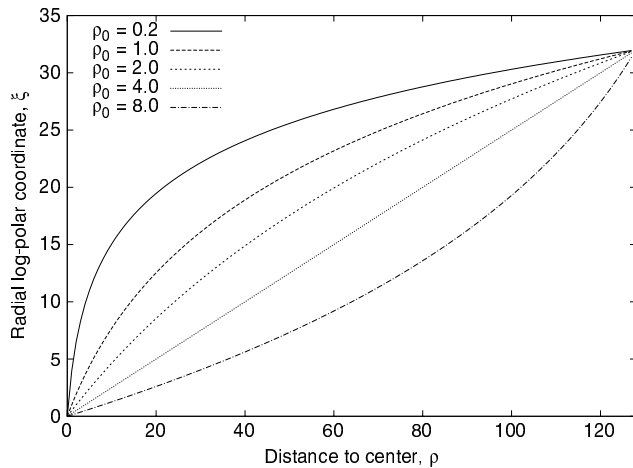


Fig. 8. Customizable mappings. Under Jurie's model [26], different profiles for the radial log-polar coordinate (ξ) can be obtained with different values of a parameter (ρ_0), as a function of the eccentricity (ρ): $\xi = \log_a \left(\frac{\rho + \rho_0}{\rho_0} \right)$.

generic framework is proposed in [86], where a genetic algorithm⁴ finds the geometric parameters of a log-polar sensor that fulfill a set of given design criteria. A further challenging step would be to automatically find the best sensor configuration adaptively, depending on the particular visual task currently being carried out.

While log-polar imaging is adequate for many tasks, it turns out to be limited in certain scenarios. For instance, the log-polar mapping is used in [88] in a context of human-robot interaction to extract depth information through vergence movements of the cameras of a binocular head mounted on a robotic arm, and to center and track a target (in this case, a user's mouth) [89]. This work illustrates that having a central region with too much information becomes a disadvantage when a target is too close to the camera, because of the amount of redundancy. Therefore, in such a situation they use a modified log-polar transform [90], moving the location of maximum resolution towards the periphery. It can be shown [91] that Jurie's log-polar mapping [26] can accomplish this resolution-reversing effect just by properly setting one parameter, even though this was not a planned prerequisite of its design, and neither did the author explicitly mention this property. The logarithmic profiles for different values of this parameter of Jurie's model are shown in Fig. 8. However, as can be easily verified [92], this log-polar model lacks the scaling invariance property.

This is a nice example giving rise to the need for flexible general visual transforms so that switching between them can be performed quickly and easily as a function of the requirements of the task at hand. In particular, a generalized (log-)polar transform, with a reduced and meaningful set of changeable parameters, could be a useful tool in a number of computer vision problems. Viewing the mapping-related computations as matrix operations can be a conceptually elegant way to get insight into this desired generalization [93].

To perform experiments under the philosophy of active vision, some researchers propose the use of foveal sensors in conjunction with hardware-based mobile platforms. One of the first of such, conceived for the specifications required by an active vision context, is the spherical pointing motor, a miniature pan-tilt system [94] that had a CCD sensor mounted on it, and the log-polar transform was applied to obtain a complete space-variant active vision system, Cortex-I [95]. Binocular systems with log-polar

imaging have also been studied, either as single heads [96,97,78] or as part of bigger humanoid-like robots [98]. Very often, off-the-shelf, commercial (pan-tilt) cameras are used, rather than more sophisticated, high performance, expensive platforms. For some proof-of-concept research these simpler solutions are enough, however. It is important to notice that although all these works use log-polar images in their robotic mechanisms, all of them simulate log-polar images by software conversion. Indeed, little published material exists reporting the direct usage of hardware-based, retina-like sensors, one example being the work by Yeung and Barnes [99,100]. Finally, a few authors also simulate the camera and the environment by using simple graphics-based setups [101,77,102–104].

4. Visual attention

In animals, attention is a cognitive process through which only a reduced subset of all sensory stimuli in the environment is selected. Paying attention to all the information at all times would clearly overwhelm the animal's cognitive capabilities. In robots, this selection is also required since the computational resources are limited, and it is also beneficial to automatically select and process only the most interesting parts. Regarding visual perception, models of primate visual attention [105], as well as computational models [106–108] have been proposed. Important distinctions in all these models are bottom-up vs. top-down approaches, covert vs. overt attention, and object-based vs. space-based attention [109].

Bottom-up attention is driven by low-level visual cues (contrast, orientation, color, motion, etc.) which are different within a neighborhood and thus become visually salient. On the other hand, top-down attention is guided by the requirements of the current task. Computationally, bottom-up attention is typically modeled through salience maps encoding the degree of local conspicuity within an image, while top-down attention is usually represented by providing these maps with task-related biases.

In relation with other visual tasks, visual attention can be understood as a core task orchestrating the functioning of the whole system. For instance, before a target is tracked, it should be detected by drawing the attention of the system over other parts of the scene or even over other competing targets. Most tracking works, however, obviate this attentional part, and assume the target position to be known initially, typically by having the target centered in the visual field.

While a considerable amount of work has been carried out on visual attention in the computer vision and robotics communities, little published material exists that explores the topic within the context of log-polar vision. However, addressing the problem of visual attention combined with foveal vision is relevant and important for several reasons [110]. First, both log-polar vision and visual attention have sound underlying biological principles and motivations; second, a deep interplay between both issues can be expected and this deserves an in-depth study; and third, these related problems are of great interest for robot vision systems.

In his thesis, 15 years ago, Milanese considered a model of visual attention with several elements of biological plausibility [111]. He was possibly a pioneer in considering log-polar images within an attention model. However, this was only lightly dealt with on the surface, as a possible extension of his work. Rao et al. have also explored attention-like problems [112,113] and provided ideas to extend their approaches to space-variant sensing. A simplified model of salience computation with log-polar images was proposed in [114]. Some authors [115,116] have implemented a version of the Itti and Koch's well-known model [117] for log-polar images (Fig. 9). The *FeatureGate* bottom-up attention model [108] is adopted in [118] through banks of oriented filters on log-polar

⁴ Genetic algorithms are a kind of stochastic optimization methods that are particularly useful in searching in large and non-convex search spaces [87].

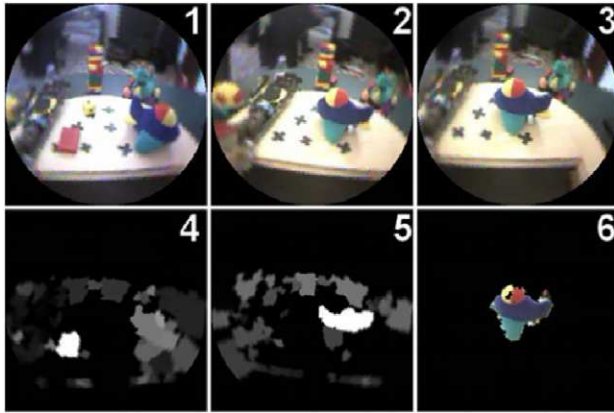


Fig. 9. Saliency maps. Results of a model of object-based visual attention being tested on Babybot, a humanoid robot endowed with log-polar vision [116]: picture 4 is the bottom-up saliency of picture 1; picture 5 is the top-down saliency map of picture 2; and picture 6 shows a segmented object after its recognition.
© 2005 IEEE.

images. Unlike many other approaches, they include the optic flow as one of the feature maps.

The work by Colombo et al. [119,120] consists in computing “cortical pyramids” (image pyramids in the cortical domain) of some given features, an interesting innovative approach in the log-polar literature. As an example, a 3-level cortical pyramid is depicted in Fig. 10. They describe a system that computes a saliency map and models task priorities through scalar weights. From this, they derive a mechanism of gaze selection for attention control. Their simulations with motion features (optical flow magnitude and upper bound on immediacy of collision) allows objects to be recognized after a reduced number of foveations. This setup, while simple, demonstrates the effectiveness of modeling biological theories in computer vision applications. Fig. 11 illustrates how this system works in a toy scenario.

An interest operator, which can be regarded as a very simple version of a saliency-computation model, was defined in [121] on foveated sensors. By using this operator, a sequence of fixation points are selected, and a panoramic-like representation is built from successive snapshots by combining high and low resolution image areas. A measure of convergence of the “scan path” (the sequence of gaze fixations following attentional cues) can be used to evaluate different attention-based strategies [122]. An example of such a scan path is given in Fig. 12. Another possible way to assess the performance of a computational model of visual attention is by (subjectively) comparing its results with those of actual persons obtained through psychophysical experiments, as illustrated in Fig. 13. Sela and Levine [123] propose an interest-point selection method, also based on a symmetry measure,

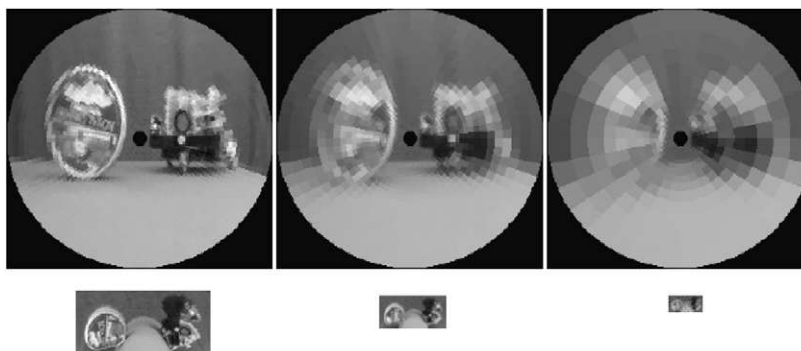


Fig. 10. A 3-level cortical pyramid [120]. Cortical (log-polar) images are on the bottom row and their reconstructions to Cartesian (retinal) space are in the top row. Resolution (and the size of the log-polar images) decreases from left to right.
Source: With kind permission of Springer Science + Business Media.

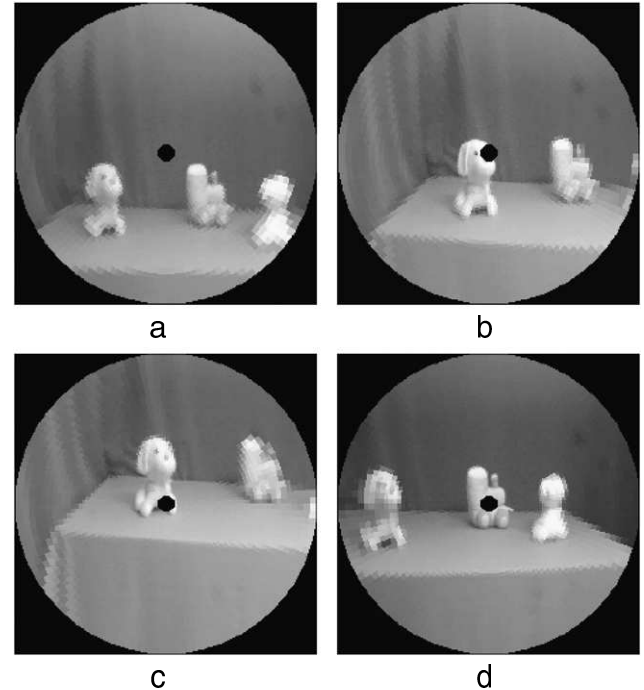


Fig. 11. Illustration of an attentional system [120]. From a first “neutral” foveation in (a), the dog draws the system’s attention in (b) because its immediacy of collision is higher than the speed of the cat, which is moving away. After this, the recognition task is triggered, and a sequence of four foveations are executed to complete the identification of the dog at (c). While the system is recognizing the dog, the train starts moving parallel to the camera, but the system does not interrupt the recognition process because speed detection has a lower priority than the recognition task. Therefore, only after recognition is the system’s attention drawn by the moving train at (d).

Source: With kind permission of Springer Science + Business Media.

which is then adapted to log-polar images. It can be seen that symmetry-based measures have been repeatedly proposed for visual attention since it is believed that humans are particularly good at detecting symmetric stimuli [124]. An interest map, combining features such as blobs, bars, corners and opponent color bands, is also used in [125], where images resulting from narrow-span saccades are reconstructed (fused) into a single representation.

While interest or key points detection is usually adopted in matching and recognition problems, they can also be used to compute measures of saliency, even though this particular meaning of saliency might be somehow different to that implied by researchers who use the term either to denote perceptual saliency or to seek more biologically faithful attentional models. For

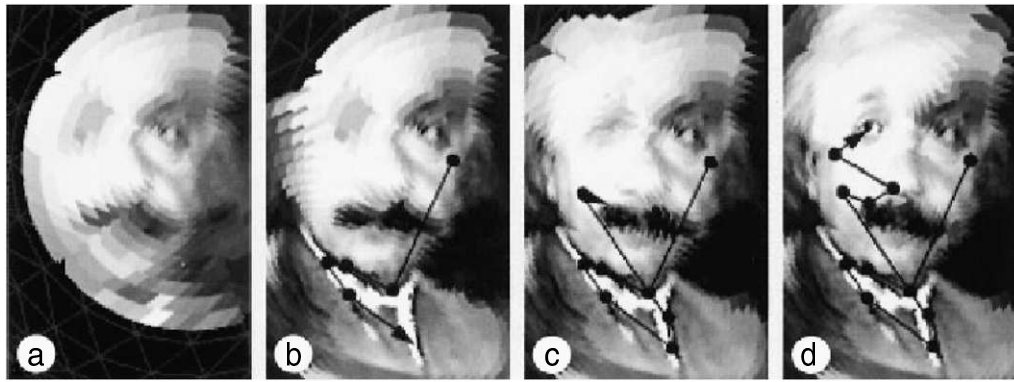


Fig. 12. Example of scan path [121] at (a) first fixation and after (b) 6, (c) 10 and (d) 13 fixations.⁵

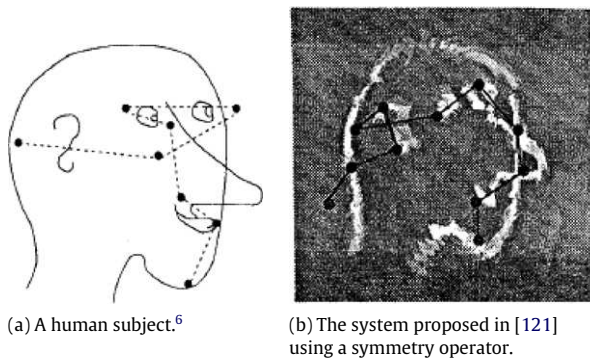


Fig. 13. Similar scan paths.⁵

instance, the popular Kadir–Brady Scale Saliency algorithm [126] (which, despite its name, is also a key-point detector), has recently been tested and compared using log-polar images [110]. Although they do not explicitly demonstrated how it could be achieved, these authors suggest that the same principles could potentially be applied both for interest-point detection and saliency computation, since both problems aim at a different, but very similar goal, namely, finding local distinctiveness. In any case, key-point detection and local image descriptors, which have a richer literature in the context of conventional images [127,128], remain open problems for log-polar images, and in space-variant imaging in general. Achieving general geometric invariance in local descriptors for these unconventional image formats represents an interesting but challenging issue.

Fovea and periphery areas in the image sensor may serve different purposes. Features can be extracted from the periphery to address the “where-to-look-next” problem, while central vision is better suited for a detailed analysis of the visual data [129]. Scene exploration can therefore be performed by the suitable integration of both behaviors. Object recognition can also benefit from active and purposive movements [130,131]. Baba et al. [132] propose a fish-eye system where the log-polar images obtained at each eye fixation are integrated into a global map. A similar idea is provided in [133], where the merging of old and new visual data into the panoramic image considers the resolution at which the data is acquired. Baron et al. [134] compute a saliency map based on a radial symmetry operator and use this map, along with a set of simple heuristics, to perform a sequence of exploratory fixations,

while a long-term memory keeps track of the scene areas that have already been visited. In [135], depth maps computed in log-polar images are input to dynamic neural fields to get distinctive blobs of activation which are used to trigger saccades.

There is not yet full evidence on how saliency influences eye movements and these, in turn, influence further visual saliency and attention shifts. However, Sun et al. [136] propose a computational model to integrate (covert) attentional shifts and (overt) gaze movements, based on object-based visual attention and tested on natural scenes, where eye movements were simulated by selecting image windows (no robotic platform is actually used). Log-polar images of these windows are taken and processed.

Because detailed high-frequency information is lost in peripheral areas of log-polar images, a *filtering before sampling* (FbS) mechanism is suggested in [137, Ch. 6]. The idea is to perform local image operations on Cartesian images (filtering) to get feature maps, and only after that is the log-polar transformation (sampling) of these feature maps performed. While there is some biological motivation underlying this strategy, the alternative approach, *filtering after sampling* (FaS), is not only possible, but also makes sense for several reasons [138], including the probable computational advantage and its agreement with the principles of active vision. Both perspectives, FbS and FaS, have advantages and disadvantages, and further studies can be carried out to explore and compare them more deeply.

Exploring visual saliency and attention under a space-variant sampling is of interest not only in computer and robot vision, but also in animal vision studies, as shown in recent research [139]. The relationship between spatial resolution and visual attention [140] is also of relevance in the context of foveal-based robot vision.

We feel that many important topics on visual attention in log-polar images can still be studied, and a few of them are mentioned briefly here. The role of the periphery for attention is of great importance and clearly still under-explored. The integration of overt attention in robotic heads, possibly inspired by existing models of eye movements [141,142] is a suggestive research avenue for using foveal vision in its natural context of active vision.⁷ While computational models for emotion exist for artificial agents, no attempt has yet been performed, as far as we know, to combine attention and emotion. However, biologically, emotions and attention emerge from the same area in the brain [143], which suggests their connection, in particular, for top-down attention modeling. Research showing the usefulness of attention for

⁵ Source: Reprinted from Vision Research, Vol. 11, D. Noton and L. Stark, Scanpaths in saccadic eye movements while viewing and recognizing patterns, Pages 929–942, Copyright (1971), with permission from Elsevier.

⁷ Journals such as Vision Research (www.sciencedirect.com/science/journal/00426989), Perception (www.perceptionweb.com), or Journal of Vision (www.journalofvision.org) are also excellent sources for robot vision researchers to get inspiration and learn from findings and advances in the field of Vision Sciences.

⁵ Source: Reprinted from Computer Vision and Image Understanding, Vol. 63, No. 1, H. Yamamoto, Y. Yeshurun, and M. D. Levine, An active foveated vision system: Attentional mechanisms and scan path convergence measures, Pages 50–65, Copyright (1996), with permission from Elsevier.

faster, non-brute-force, recognition [144] has still room of further development, particularly for foveal imaging. Since attention provides a link between perception and action [145], it might play a significant role in modeling and implementing affordances [146]. Demonstrating the usefulness of attention in practical real-world, useful applications, and its seamless integration with the rest of the components of a robotic system, is also an appropriate and exciting direction to pursue.

5. Visual tracking

Visual tracking consists in keeping track over time of the movements of one or more moving objects within the visual field of a given observer. When the camera dynamically changes its parameters to track a moving object [147], this can be referred to as *active tracking*. On the contrary, *passive tracking* just keeps record of the target position without moving the camera, and therefore entails a smaller field of view. A simple taxonomy of visual tracking may consider two dimensions: passive/active tracking and uniform/space-variant sampling [148]. Foveal vision makes most sense when used in active vision scenarios and this is thus our focus in this section. In this context, several issues can be raised regarding the impact on performance of a number of factors such as the dynamics of the mechanical visual system, the delays due to computations, the hardware constraints, and the mobility of the target. These relationships were analyzed in [149], using linear optimal control theory.

There are two main issues to consider in active tracking approaches. One has to do with the estimation of the relative motion between the target to be tracked and the camera. This “retinal slip” is one of the main visual features used by humans in controlling the oculomotor behavior known as *smooth pursuit* [150]. Most of the work conducted within the context of visual tracking with log-polar sensors addresses this problem and this will be the main interest in this section.

The second relevant issue to consider is how to control the active tracking device in order to keep the target stabilized in the fovea most of the time. Tracking can only be perfect when the target trajectories are regular. If the system is able to predict these regularities, it can issue the motor commands anticipatively and cancel the delays existing in the control system. Such predictive control schemes are found in infants at early ages and have been implemented recently in humanoid robot heads [151–153]. Constant velocity, constant acceleration and sinusoidal motions are typical motions taken into account in these problems. When the trajectories of the target are not regular, tracking errors will arise and the target may move out of the fovea. In these cases the perception system must be able to identify the target in the periphery of the visual system.

An interesting study by Vincze and Weiman [154] considers the trade-off existing between sampling time and tracking errors. For a certain velocity error, the longer the sampling time is, the farther away the target will be from the predicted position. This would suggest that short sampling times are better, but with short sampling times one can only process small image windows that may not reach the target. Vincze and Weiman [154] show that some space-variant sampling approaches, such as log-polar and pyramids, always improve their performance when their visual search range increases: the rate of growth of computation time is smaller than the rate of growth in the window size. The same does not happen for Cartesian geometries, where the window sizes that maximize performance are limited to prevent long sampling times that would allow the target to escape from the field of view. This is a significant argument for the use of space-variant sampling strategies in active tracking scenarios.

The remainder of this section will review motion estimation methods in log-polar images, either for passive tracking (to update

the tracking window) or in active tracking (to update the camera’s gaze directions). Although the focus is on tracking, the techniques described here can also be helpful in other problems. We will distinguish between approaches which estimate the target motion more *explicitly* from those which estimate the change of target position and, therefore, motion is estimated *indirectly*.

5.1. Explicit motion estimation

Motion estimation is a fundamental and challenging problem in computer vision [155], and has many applications in robotics, including visual tracking. In terms of the locality of the visual information being used and the usage of parametric motion models, two broad groups of motion estimation methods can be distinguished: optical flow-based ones (which tend to be local and non-parametric) and region-based ones (which are generally more global and assume parametric motion models). Both approaches are discussed in the following paragraphs.

5.1.1. Optical flow-based methods

One fundamental technique to estimate motion is that of optical flow [156–159], where the motion is estimated locally either at each image pixel or only at selected image positions. The main advantage of optical flow lies in its flexibility to represent local motion, which is required when the moving object is non-rigid, or articulated, or when different objects move independently in the same scene. However, optical flow estimation also has some disadvantages, such as the influence of the aperture problem⁸, a high computational cost, and noise sensitiveness.

Theoretically, optical flow can also be computed in log-polar images. Araujo and Dias [74,73] compute the *normal* optic flow, which is the projection of the flow vector *perpendicular* to the local image gradient, since the aperture problem prevents the true motion vector from being computed unless more sophisticated means and assumptions are employed. Daniilidis and Krüger have also made contributions to optical flow computation and 3D motion estimates using log-polar images [161,162].

In practice, care must be taken when computing optical flow in log-polar images [71]. On the one hand, some techniques developed bearing the Cartesian geometry in mind may lead to wrong results if they are directly applied to the log-polar geometry; on the other hand, the particular polar nature and logarithmic sampling of discrete log-polar images motivate the need of devising appropriate filters for computing spatio-temporal derivatives. Similarly, the well-known *brightness constancy assumption* [160] underlying many optical flow methods,⁹ is easily violated when illumination changes, and under these conditions motion could be wrongly detected even with a static scene. To deal with this problem, a *generalized dynamic image model*, which also takes into account the changes in brightness, is used in [163] for log-polar images. In that work, the practical issues related to space-variant sampling are addressed through space-variant local windows and specialized versions of the gradient operator, a topic also considered in other works [164,40].

Despite these caveats, other authors have used conventional optical flow in log-polar images, not only without drawing attention to any disadvantages, but in fact with reported benefits with respect to Cartesian images [165]: under fixation conditions, translational motion is nearly zero, and the other affine motion components (rotation, scaling and shear) can be estimated better in polar-logarithmic images. This is based on the fact that the high-acuity fovea can appreciate small motion, while the coarser sampling at periphery is more suitable for larger motion estimation,

⁸ The aperture problem arises because the limited visual data of a local analysis prevents the true direction of motion from being inferred [160].

⁹ Basically, this assumption states that a change in brightness at an image location can only be due to an underlying motion.

which is something other authors have noticed too [166]. Tunley and Young's approach has been employed by some authors using log-polar images in active vision tasks [27,167,168]. This idea of detecting large-range motions while still being able to accurately estimate them seems to be under-explored, despite its theoretical and practical attractiveness. Recently, experiments with 1D space-variant arrays of *elementary motion detectors* (EMD) have been performed to actively track a single target [169]. The tuning and arrangement of the EMDs within the array are customized according to the expected target velocities and desired precision. Even though this work does not really use log-polar imaging, the concept is however applicable and provides inspiration for further developments.

For actual practical purposes, the optical flow field itself is not very useful, and some kind of post-processing is required. In [73], for instance, the magnitude of the flow vectors are computed and summed along the radial and angular directions. The center of mass of the two resulting magnitude profiles is then used to locate the moving target, and to control the cameras in order to fixate it.

5.1.2. Region-based methods

In contrast to optical flow, which is basically a local approach, global and region-based techniques are attractive for their higher robustness. One known and efficient region-based tracking approach is that proposed by Hager and Belhumeur [170] for Cartesian images, and which inspired two alternative solutions in the log-polar domain. On the one hand, the work in [171] brings further efficiency by reformulating the framework and then applying it to log-polar images. On the other hand, the emphasis in [172] is to exploit the log-polar domain by minimizing the explicit usage of Cartesian coordinates.

While translation is the simplest motion component in Cartesian domain, because it can be represented simply by two constants (the shifts along the x and y axes), it becomes a complicated space-variant warping in log-polar imagery. An early attempt to solve this used a 1D correlation-based approach along individual rows and columns in an alternative log-polar representation that was particularly suitable for the estimation of translational motion, but lacked the scale and rotation invariance [173]. More recently, the problem was better addressed in [174] by proposing and comparing two techniques: a gradient descent optimization approach and a projection-based technique, the latter being much more effective and robust. These two techniques deal with the added complication of how translation maps onto log-polar domain while, unlike [173], still preserving the edge invariance property. Bonmassar and Schwartz proposed the *exponential chirp transform* (ECT) [175], by means of which the shift invariance of a Fourier transform still holds in a space-variant domain such as the log-polar, and that was used for motion deblurring [176]. Although this is an outstanding contribution, the conceptual and implementation complexities behind the ECT probably reduce its appeal in practical applications.

A progressive registration technique [177] has been successfully applied to log-polar images in [178] to track planar structures. The idea is to pre-compute a set of sample templates by assuming a known initial reference template as well as the kind and range of expected image deformations. Planar tracking in log-polar images has also been approached with other methods, such as complex wavelets [179] and spatio-temporal gradient-based least squares [180]. A redundant 2D motion parameterization [171] has proven effective in estimating different parametric motion up to projective models. Basically, a linear combination of partial derivatives is improved by a linear combination of discrete derivatives in several directions and scales.

The effectiveness and robustness of log-polar images in estimating large rotations and scale changes has been exploited

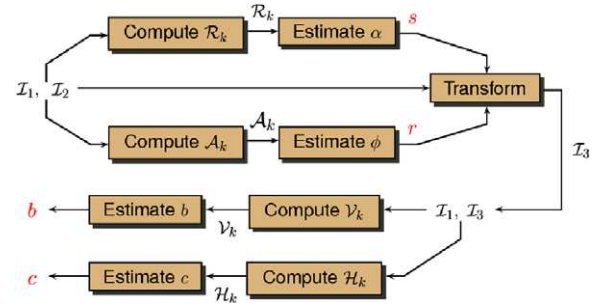


Fig. 14. From projections to motion estimates [148]. Radial and angular projections, $\mathcal{R}_k, \mathcal{A}_k$, of the input images, $\mathcal{I}_k$, $k \in \{1, 2\}$, allow the estimation of shifts s and r along the two axes in the log-polar plane which, through the scale and rotation invariance of this mapping, can be readily converted into the estimated change of scale and rotation angle, respectively. Shifts s and r are then used to rectify one of the input images, so that finally the remaining translational shifts, b and c , are estimated from the vertical and horizontal projections, $\mathcal{V}_k$ and $\mathcal{H}_k$. In turn, b and c are used to control the camera's pan and tilt degrees of freedom.

in image registration problems, allowing the recovery of large affine motions [181]. Motion estimation based on a generalized least-squares (GLS) technique is studied and compared in [182] both on Cartesian and log-polar images as well as on their combination. These two last works are examples of combining Cartesian images with the log-polar transform. While this idea is a powerful means of exploiting the best of both domains, one of its chief disadvantages is that it would not be applicable if a hardware-based log-polar sensor were to be used, since only log-polar images are available in this case.

By using four different kinds of projections (vertical, horizontal, angular and radial), the four parameters of a similarity motion model (horizontal and vertical translation, rotation angle and scaling factor, respectively) can be recovered simultaneously as long as the target is actively tracked, which reduces the amount of relative translation and thus decouples the translational effects from scaling-rotation effects [148]. This is a nice example of exploiting the synergy of active vision, foveal sensing and appropriately devised algorithms: (i) active tracking simplifies motion estimation; (ii) log-polar images allow fast image processing and facilitates rotation and scaling estimation; and (iii) the efficient and effective projection-based motion estimator makes active foveal tracking possible and simpler. A block diagram of this projection-based approach is given in Fig. 14.

A number of approaches have been proposed to take advantage of the fact that changes in scale and rotational motion both map to simple shifts along the log-polar coordinates [172,148]. However, moving beyond similarity motion models or disabling the active tracking scenario, and still exploiting the log-polar domain, seems considerably harder. A generalization of the spatial projections by using the Radon transform for affine motion estimation has been suggested [183,44] with interesting, but still limited, results as regards the number of affine parameters and the range or accuracy with which these parameters can be estimated. Nevertheless, log-polar images have been shown to exhibit higher noise robustness and better results than Cartesian images with the same number of pixels and field of view [44].

Unfortunately, some of the more general approaches [171,184] not only fail to exploit the characteristics of log-polar imaging so smoothly, but also require that the target to be tracked is known in advance, thus posing a significant limit on their practical application. While both ideas are similar, the more recent work [184], based on the idea of the *interaction matrix* proposed for Cartesian images in [185], has some advantages over [171]. This is essentially due to the fact that it does not depend on the motion model (theoretically, any arbitrarily complex motion model can be implicitly adopted) nor on the image sampling scheme (any

sampling other than log-polar can be used). A deeper analysis and experimental comparison of both approaches in both theoretical and practical terms would be interesting. Regarding their common limitations of an assumed initial template, two kinds of interesting, non-trivial solutions can be considered. On the one hand, the computationally heavy off-line procedures now involved in these approaches should be turned into very fast real-time processes, so that they could be applied on-line once the target to be tracked is decided. The other alternative would seek the quite different route of devising some template-independent technique, and still sharing with existing approaches nice properties such as search-free motion estimation.

5.2. Feature-based approaches

The advantage of feature-based tracking is that, once features have been extracted, fewer data have to be processed and matching may be less ambiguous. However, good features need be found and matched, which is not always easy, and the object of interest must have enough texture. Additionally, detecting certain features in log-polar images can become a particularly challenging problem. Corner detection, for instance, was tackled in [99] through a Hough-based straight-line detection. Detecting lines through the Hough transform becomes more natural in log-polar space since the Hough and (log-polar) image spaces are the same, as found almost 20 years ago [186]. Since then, other researchers have been attracted by the problems of line and circle detection in the log-polar domain [187,188]. In general, learning is found to be a better strategy for extracting some features (edges, bars, blobs) than mathematically modeling these features [189,190]. Specialized table-based data structures are proposed [191] to compute image processing operations directly and efficiently onto foveated images. While initializing these tables can be time consuming, this off-line effort is paid off by the real-time on-line application.

As mentioned above, because of their sampling, log-polar images are suitable for detecting large motion in the periphery and smaller motions at the fovea [180,166,73], since these areas have different motion sensitivity. Weiman and Juday proposed three tracking algorithms acting differently, depending on the position of the target blob [192]. An interesting aspect of their approach is that they exploit the nature of log-polar images. For instance, instead of explicitly computing the centroid of a blob, centering it can be achieved actively by camera movements that maximize the number of target pixels. This is a nice example of how properties of log-polar mapping can be considered and exploited. This algorithm is also used in [193], as part of a gaze control mechanism. Similarly, [101] uses the fact that objects increase in size as they approach the fovea.

Tracking simple bright shapes moving on a uniform black background is considered in [194] as a proof-of-concept of the adaptation to the log-polar domain of a number of well-known image processing techniques (erosion, dilation, region labeling, boundary tracking, corner detection and Hough transform of circles, ellipses and polygons).

5.3. Other tracking approaches

Due to their particular geometric nature, log-polar images are particularly well suited for actively tracking one single object at a time. However, the unconventional work by Kang and Lee [195] considers the tracking of several objects passively (i.e. without camera movements) by keeping image templates and updating them over time. In a simulated active vision system, Lim et al. [101] also addressed a simplified version of the problem of tracking multiple targets. They decide which object to foveate and track

next as a function of the size of the objects, how much they have moved between consecutive images, and their distance to the fovea.

Principal Component Analysis is employed in [26] for face detection and tracking in log-polar images. This approach uses a single face model for all possible positions. Under a space-variant imaging representation, however, considering multiple models seems more advantageous, as it takes advantage of the varying resolution, as suggested in [196].

Since color is invariant to a number of geometric and resolution changes, it is an appropriate cue in some visual tasks, allowing distinctively colored objects to be robustly tracked even in log-polar images [197].

Background subtraction is a well-known technique to detect objects moving on a static background [198]. However, these algorithms generally do not account for a moving background, and are therefore not useful with a moving observer. In those situations, a panoramic image can be built, so that a part of the scene can be indexed according to the direction of the observer's gaze. In the case of a pan-tilt unit, spherical panoramas seem to be particularly well suited to a polar-like image representation, as suggested in [133]. An alternative way of segmenting an unknown object from the background can be performed if the object is actively tracked and the background changes over time [172].

Binocular tracking is still another possibility, but it is left to Section 7, where 3D issues (depth recovery and vergence control) will be examined.

To recap on the usage of log-polar vision in tracking, it is worth stressing how the implicit focus-of-attention brought by the foveal predominance (Section 2) has been considered for tracking without explicitly segmenting the target. However, this is only possible if tracking is very accurate and/or the object has a minimum size and certain round-like shapes. For instance, it has been found [148] that centered targets should represent roughly 25% of the field of view to successfully estimate their motion and track them. While this result actually depends on the particular mapping parameters being used and the capabilities of the tracking algorithm, it is nevertheless clear that further research is still required to address the tracking of non-centered, smaller, arbitrarily-shaped targets, and also multiple targets. The actual challenge lies in solving these problems as elegantly as possible so that log-polar imaging, combined with powerful active vision strategies, can still have advantages to offer.

6. Egomotion estimation

Egomotion is the motion of the observer (a pan-tilt camera, a robot arm, or any other dynamic visual agent). Its estimation is important in those situations where the knowledge of actual robot movement is required either for self-localization, or to distinguish it from the motion of objects in the scene. Although egomotion estimation can be tackled with the general tools for motion estimation, some specific approaches have been proposed that exploit the task at hand. Two common characteristics of these systems are their reliance on the computation and analysis of patterns of optic flow, and their hardware implementation (usually, on FPGA-based architectures). Not surprisingly, the polar-like geometric nature of log-polar images makes them very suitable for the estimation of self-motion in the direction of the optical axis. When the focus of expansion is aligned with the center of the log-polar transform, motion flow vectors are only radial and of uniform magnitude in the log-polar plane. As a result, time-to-impact (also called time-to-collision, or its reciprocal, immediacy of collision) computation can be addressed more easily [199–203]. However, some authors find that the main benefit comes from the polar structure and no advantage, besides the computational

speed-up, are gained from the logarithmic radial sampling [204]. Others consider that these radial-like sensors (polar and log-polar) bring no particular advantage for the time-to-collision computation, which is in fact simplified by the assumption of fronto-parallel surfaces [205].

When it is the observer that moves, rather than a visual target, the task of keeping a static target in view is called *fixation*. The problem of fixation is similar to that of tracking, and has also been tackled with log-polar images [180,206,207,99,208]. Ahrens and Neumann [206,207] compare log-polar and Cartesian windows and find that error measures tend to be more robust, with fewer local minima, under a log-polar sampling. This result for the pan angle of a camera, has been repeatedly verified in other contexts, such as in vergence control [209,27,84], or translation estimation [210]. An interesting point in the appearance-based fixation strategy by Ahrens and Neumann is that a confidence-based measure, by comparing the reference and current views, is incorporated to make the procedure more robust against occlusions or other changes of view [207].

Daniilidis proposed two algorithms for the estimation of the focus of expansion, which are based on a 1D global minima search for flow patterns [204]. Similarly, the rotational component of optical flow is used in [211,212] to implement an active docking behavior, which aims to align the camera's optical axis with the robot's direction of motion. Silva and Santos-Victor found that the subset of points whose radial and circular normal flow vectors do not depend on the camera translation are described, respectively, by a special circle and line [213]. These image locations provide relevant geometric information about the observer's motion, and their estimation is reduced to searches along horizontal and vertical lines in the log-polar image, so that the focus of expansion and the rest of the parameters are estimated in a two-step algorithm [213]. In [214], four sets of spatio-temporal Gabor filters are applied: two spatial filters are aligned in the radial and angular directions, while two temporal Gabor filters are tuned to opposite directions of motion, so that expanding and contracting patterns of motion can be discriminated.

Bishay et al. [215] take advantage of the fact that a mobile robot moving along an indoor corridor will perceive the edges arising from the intersections between the walls and the floor and between the walls and the ceiling as horizontal lines in the log-polar plane. Additionally, the varying resolution of log-polar images simplifies a matching procedure, since doors or lamps along the corridor map to very similar patterns even when they are at different depths. This makes the detection of these landmarks simpler and faster for robot navigation. Similar ideas are explored in [216]. Advantages of log-polar and other task-specific image representations are also exploited in [217] for robot navigation in a corridor. In outdoors environments with some structure, such as roads, these kinds of benefits also apply: lane marks in straight road sections become parallel lines in the log-polar space and vehicles are perceived with the same size and shapes. However, this is only true if the vanishing point coincides with the center of the log-polar transform. To fulfill this assumption, the camera has to be fixated on the vanishing point, so that the movements of the vehicle on which the camera is attached can be compensated for [104].

To distinguish ecomotion (the object's motion) from egomotion, a differential algorithm is proposed for log-polar images, and implemented using FPGAs [218–221] based on the fact that the second temporal derivative in the image sequence is non-zero only at the borders of self-moving objects.

Omnidirectional vision, usually consisting of the combination of a curved mirror and a camera pointing towards this mirror, provides a 360-degree view, which can be very effective for robot navigation. Under this arrangement, if conventional cameras are used, the resulting images have to be unwarped for their human or

machine processing. However, the mirror's shape can be designed to match certain design properties and, when combined with a log-polar camera, rectangular panoramic images are obtained directly, with no additional computational effort [222,223]. These ideas are illustrated in Fig. 15.

One problem related to egomotion estimation is image stabilization, where the goal is to compensate for unwanted camera movements so that better images are obtained. Some benefits of using log-polar images in the context of this problem were explored in [224], where optical flow is computed in four selected peripheral regions of the log-polar image.

7. 3D cues and vergence control

One of the most important capabilities of autonomous robots is their ability to perceive the three-dimensional structure of the environment, in order to avoid obstacles, recognize shapes, and manipulate objects. Several cues for depth perception can be used from an image stream. In monocular systems, motion, focus and shading cues (amongst others) have been systematically used to address the problem of depth perception. In binocular systems, depth can be computed via stereo, a conceptually simple and easy methodology. In active vision systems [11], improved perception can be achieved by controlling the camera motion. In this section we will review the main works using log-polar sensors for depth perception, considering both static and active vision systems. In the active case we will also refer to the control of the vision system parameters for improved perception.

7.1. Depth maps

Though a full 3D reconstruction of the environment is not possible in the general case and, indeed, may not even be a desirable goal under a purposive vision philosophy [225], depth maps (i.e. images where each pixel location represents the measured depth of an observed 3D point) are important perceptual structures for robot motion and behavior control.

Beyond the obvious reduction in computational cost, log-polar images have other benefits in computing depth maps. Although the greatest advantages arise in active vision scenarios (see next section), the use of log-polar sensors is advantageous in the following cases:

- *Depth from forward motion:* When the observer is moving forward in natural scenarios, the optical flow amplitude grows, on average, from the center to the periphery of the visual field. Due to the log-polar transformation properties, the magnitude of the optic flow in the periphery is reduced, thus facilitating its computation by local search or gradient methods [226].
- *Depth from convergent stereo:* In converging camera configurations the image planes are rotated with respect to each other. This poses registration difficulties because corresponding points in the two images diverge towards the periphery of the visual field. Again, since log-polar tessellations compress coordinates in the periphery, registration mismatches are reduced, thus facilitating the search for corresponding points [227].

In this section we review some works that compute depth maps in log-polar images. The techniques are similar to their Cartesian counterparts, but must be adapted to cope with the different image topology. Two basic approaches have been used: depth-from-motion and depth-from-stereo.

In depth-from-(ego)motion, a (monocular) moving observer in a static scenario can stabilize the center of the image by a proper gaze control. In these fixation conditions, the depth-from-motion problem becomes well-posed and its dimensionality is reduced. Optical flow corresponding to camera motion can then be used to

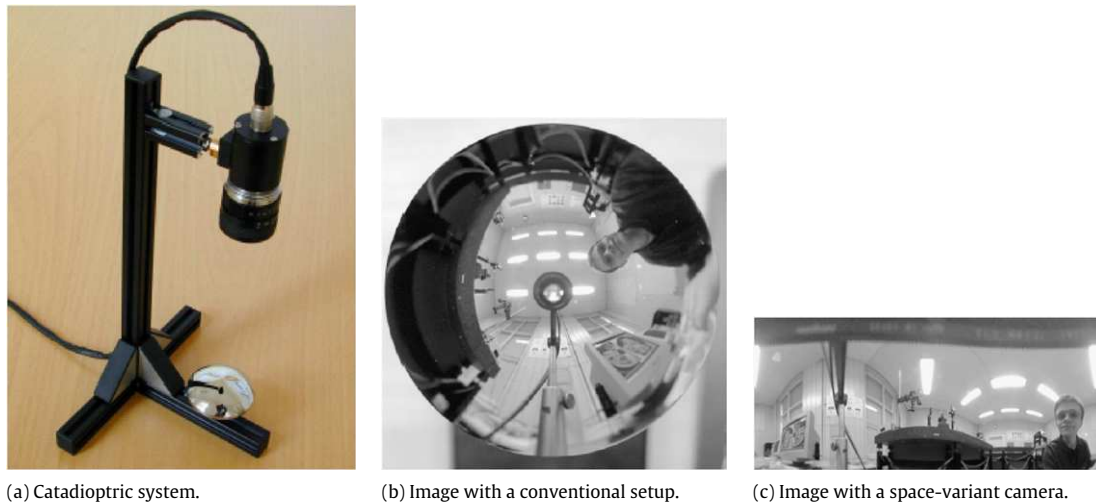


Fig. 15. (a) Catadioptric sensor designed within the OMNIVIEWS EU project [223]; panoramic images obtained with (b) a conventional mirror-camera setup, and (c) a mirror designed to match the geometric features of an existing log-polar camera. © 2002 IEEE.

compute the depth *only* at sufficiently textured pixels (*sparse* depth maps) [226]. Alternatively, a phase-based technique using Gabor filters can be used to compute optical flow vectors and integrate them over time via a Kalman filter [228]. This results in a depth estimation for every pixel (*dense* depth map).

The second approach, depth-from-stereo, uses two cameras and the projection of 3D scene points onto different locations in the retinas. The difference between these corresponding locations, known as *disparity*, depends on the depth of the observed point as well as on the camera setup calibration. Sparse disparity maps can be computed [229] by estimating disparity in log-polar images at pixels with enough texture. Dense disparity maps are computed in real time in [230] by testing several hypotheses between pixels from the left and right cameras and using local pixel interactions to avoid ambiguities in the disparity estimates in non-textured areas.

More recently, epipolar geometry has been proposed to compute dense disparity maps in log-polar images [231], which reduces the general 2D matching problem to a 1D search along the epipolar lines. However, since straight lines in the Cartesian domain become curves in the log-polar space, good line re-sampling algorithms are required for an accurate matching procedure [188].

7.2. Active perception

Log-polar images have their maximum acuity in the central part (the fovea). Furthermore, the rotation and scaling invariance properties only hold when objects are centered in the visual field. Therefore, when analyzing some particular targets of interest, it is natural to consider controlling the cameras motion so as to keep the targets centered in the fovea. This is also the solution adopted by foveated biological vision systems, which are constantly changing the direction of their gaze toward the objects of interest. Fig. 16 shows a couple of active robotic stereo heads that have been used within this paradigm.

The advantages of using this elegant combination of foveated sensors and active vision systems have been demonstrated in a series of works. As far as depth estimation is concerned, Weiman [234] was one of the first to formally analyze the gains obtained from the use of log-polar images (apart from the obvious computation reduction aspect). It was shown that the maximum stereo resolution with a Cartesian binocular setup is obtained inconveniently at the periphery of the visual field, whereas log-polar stereo resolution is maximal in the center of the visual field of both cameras. Therefore, actively controlling the point of



(a) Baltazar's head [232]. (b) Eurohead [233].

Fig. 16. Example of robotic active stereo heads: (a) the head of Baltazar [232], a humanoid robot, and (b) Eurohead [233], a high precision head for quality measurements.

fixation toward the objects of interest allows the visual acuity to be maximized also in the depth dimension.

To be able to fully exploit the stereo resolution properties of log-polar stereo, the fixation point must be controlled to track the object of interest. In active vision systems, oculomotor control is usually inspired in its biological counterparts and is decomposed into vergence and version movements [235]. To track motions in depth, the vergence angle should be adjusted so that the fixation point matches the distance from the object of interest. To track motions in fronto-parallel directions to the observer, the camera's direction is controlled to center the target in the eyes. Interestingly, it has been demonstrated that log-polar images facilitate the tracking problem *per se*, both in depth and in fronto-parallel directions to the cameras [236].

When a target is being tracked, its disparity in the images is near to zero, i.e. there is a close match between the information contained in the same retinal coordinates in both images. Some authors have proposed the maximization of a *binocular fusion index* measured by the correlation between the left and right images [237,238]. The process involves the execution of some exploratory motion on the vergence angle in order to obtain the gradient of the fusion index, and then controlling the vergence angle so as to maximize it. A correct vergence on the object would be attained when the fusion index is at its maximum value. In [27], this method is complemented by estimating the expansion and contraction patterns from the optical flow of the image, velocity feedback measurements being provided for vergence control.

Searching for the maximum fusion index requires active vergence movements, which may lead to some problems in the control loop. Other alternatives have been proposed that attempt to overcome this problem. In [237], it is proposed to

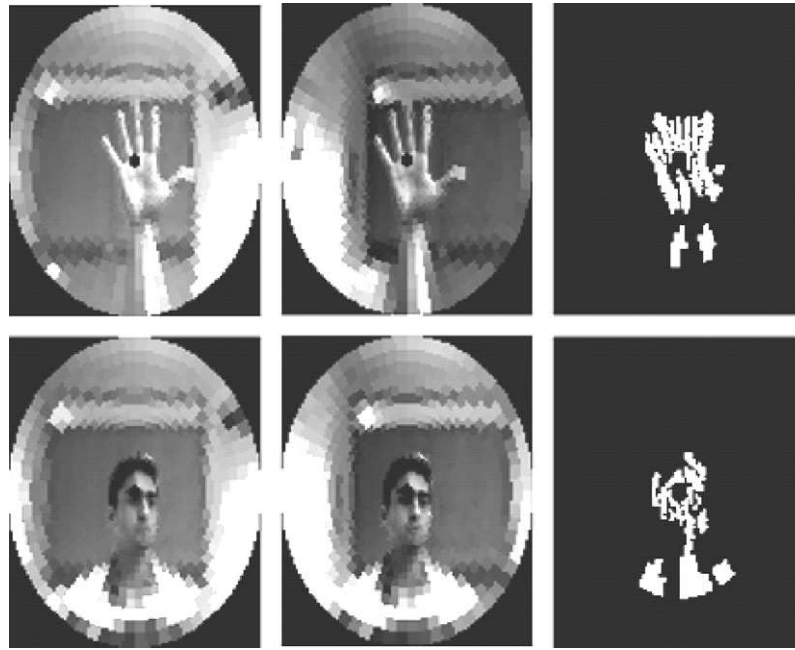


Fig. 17. Zero disparity filtering in log-polar images [236]. When objects are under fixation in both cameras, their disparity is close to zero. A zero disparity filter can thus segment the points belonging to the object (right column). This process is illustrated for a hand (top row) and a face (bottom row). The input stereo images are in the left and middle columns.

compute the dominant disparity in the images explicitly, by evaluating the correlation between the images, where one of the images is transformed according to certain disparity hypotheses. The hypothesis with maximum correlation is taken as the true disparity. This value is then input to a feedback PID filter in order to control the vergence angle.

In [239], a similar approach is taken but using the concept of *virtual horopter* [240] and the output of *zero disparity filters* [241]. The *horopter* is the locus of all points with zero disparity. Points in the horopter can be segmented in the images by matching left and right image patches at the same image coordinates, in a process known as *zero disparity filtering* (ZDF). Such points are then assigned to the target under vergence. Fig. 17 illustrates the ZDF process by which points belonging to the target under vergence are segregated from the background. A virtual horopter, tuned for disparities other than zero, can be obtained by transforming one of the images according to the required disparity. Oshiro et al. [239] create several virtual horopters for a discrete set of disparity hypotheses, and compute the output of zero disparity filtering for each one. The disparity hypothesis leading to the greatest number of segmented points is then fed back to control the vergence angle.

The segmentations provided by zero disparity filtering have proven useful for rejecting points belonging to background elements. Although the log-polar mapping attenuates background elements in the periphery of the visual field in a hard-wired manner, the output of zero disparity filtering has been exploited to provide added robustness to tasks. In [27], affine optical flow parameters are computed on the target region and used for vergence control. In [236] the centroid of the segmented region and the translational parameters of the optical flow are used for tracking in the fronto-parallel directions, controlling the robot head pan and tilt degrees of freedom.

In the above-mentioned works, the set of disparity hypotheses is often distributed in a non-uniform fashion; it samples the small disparity range densely and has a coarse representation of large disparities. The purpose is to have simultaneously, and with limited computational resources, good precision in keeping the target disparity close to zero and the ability to detect and compensate for large disparities. In [242], an explicit rule to determine the set of

disparity hypotheses is proposed, and sub-pixel disparity estimation is achieved by employing quadratic interpolation methods.

Methods based on testing disparity hypotheses have proven to be fast and robust to environmental changes. In [2], it was shown that vergence control with correlation measures in log-polar images allows smaller objects to be tracked than its Cartesian counterpart. The reason provided empirically is that global disparity computation is dominated by the disparity with the largest support, in terms of numbers of pixels. In Cartesian images this is only the case when the object occupies a very large part of the image (>50%). However, in log-polar images, objects can be much smaller if they are centered in the retina, due to the larger number of pixels assigned to the fovea. Fig. 18 shows a sequence where a small object approaches the center of the image. While in the periphery its influence on the global disparity is small and the system does not change its vergence, as it approaches the fovea its weight increasingly grows and the robot locks onto the target.

The global disparity computation in log-polar images using correlation measures is theoretically equivalent to performing the same computation in Cartesian images that are pre-weighted by a function (the Jacobian of the log-polar transformation) that attenuates the value of background pixels [2]. Although this analysis in the Cartesian domain clarifies the advantage of using log-polar images for centered objects, for computational reasons it is preferable to perform the operations in log-polar images due to the smaller number of image pixels to be processed. However, as in the case of tracking, whether (and how) the scenario of non-centered objects can be appropriately coped with remains an open issue.

8. A look to the future

The research on log-polar imaging, with emphasis on the robot vision community, has been reviewed in this article. In the past few years, fundamental properties of this image representation have been studied and exploited in appropriate algorithms.

One criticism that can naturally be posed is whether, taking all into account, it is really worth adopting log-polar images. Regarding the computational advantages, it could be argued

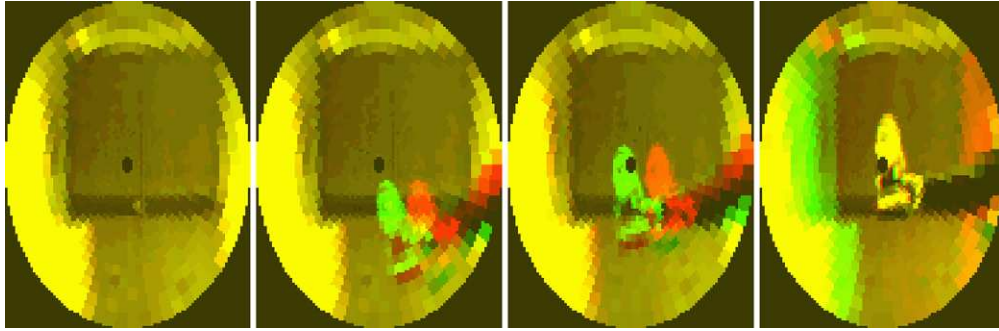


Fig. 18. Active vergence control using log-polar images [237]. Global disparity computation in log-polar images weighs objects in the center more favorably. Four images of an active tracking sequence are shown. Each color image codes both images of the stereo pair: the right image in the red channel and the left image in the green channel. The first image (left) corresponds to the binocular system verging on the background. No other objects are present. In the second image, an object appears in the periphery, but it is still far from the center and the system keeps fixating the background. In the third image, the object approaches the fovea and the system starts moving the cameras in order to gaze at the target. Finally, in the fourth image (right) the object is in the fovea and the system verges perfectly on the target. (Figure to be seen in color).

that, after all, almost any task solved with log-polar images can be (and in many cases has been) approached successfully using conventional images. This is possible due either to the use of special-purpose, high performance equipment or to the advances in hardware technology, which make a reduction in image processing times affordable. While this is certainly true, it should be noted that, in any case, under equal conditions (same computer resources), performance can always be better with log-polar images. We like to compare the computational benefits of log-polar image sensors with those offered by algorithmics. It may be the case that a quadratic cost $O(n^2)$ algorithm exists to solve some known and useful problem. If a novel algorithm is devised that allows this same problem to be solved in logarithmic cost $O(\log n)$, this is seen as an outstanding achievement because the order of magnitude of the speed-up has a more profound and significant impact than those allowed by advances in hardware technology. This is also the scenario with *logarithmic* polar images, where, as a function of the field width ρ , the spatial complexity would be logarithmic, $O(\log \rho)$, in contrast to the quadratic cost, $O(\rho^2)$, exhibited by the uniform resolution images [76].

Like the design of a new algorithm that outperforms existing ones, which may be a rather contrived, intellectually demanding task, the design of retinal sensors faces its own challenges too. Furthermore, having a good sensor does not, by itself, guarantee that any visual task will be performed efficiently and successfully. On the contrary, appropriate algorithms and mechanisms should be conceived to exploit the advantages of such a sensor and to deal with its difficulties.

As described in Section 2, some properties of log-polar imaging such as scale and rotation invariance, the selective data reduction or the implicit focus-of-attention have been widely explored, while others have only been slightly considered (or still remain undiscovered!). One example is the better numerical behavior that log-polar images offer for eye-in-hand visual servoing, as revealed by an analysis of the Jacobian matrix condition number [243,244], and its comparison with polar images.

Finally, as in many algorithmic achievements, the log-polar geometry offers a trade-off solution. Compromises of some kind are at the core of many new discoveries in any field, and particularly in computer science. Benefits are often possible only at the cost of something else. Indeed, progress in scientific research is in many respects determined by well-engineered trade-off solutions. Log-polar sensors are, to our mind, one of such solutions.

The geometric and sampling properties of log-polar images have been emphasized throughout the paper. It has become clear that polar-logarithmic sampling provides interesting benefits in a number of scenarios. But it has also been shown that more general or customizable mappings could be more generally useful.

While log-polar vision is one of the most popular foveal sensing models being used, some works consider the simultaneous use of wide- and narrow-field cameras so that their coordination makes it possible to have both rough visual data from large field of view and detailed information from a narrow part of the scene. Most of these arrangements use conventional Cartesian images [245], but recently log-polar images have been considered in a system inspired by the eyes of birds of prey [103] for easier detection and tracking of objects in a wider visual scene. In their approach, Melnyk and Messner use a driving simulator to illustrate how two images from two cameras of different fields of view can be seamlessly integrated into a single log-polar image where the over-sampling problem (Section 3) is somehow solved.

We would like to point out two potential areas of application where foveal vision is now absent, but could fit in well. In sensor networks, there are four resources that can be limited: power, sensing, communication and computation [246]. In the case of vision networks, all of these resource requirements can benefit from using log-polar sensors. Indeed, we found that our idea about the challenges imposed by the resolution-variant images and the need for proper algorithms addressing these visual limitations finds support in the following quote:

“Finally, tiny inexpensive sensor nodes are often limited in computational capability, so developers may need to implement computationally lightweight algorithms that sacrifice sensing quality but take advantage of the distributed computation resources of the sensor network.” [246, p. 38]

While its authors did not seem to have had foveal vision in mind when they wrote it, we find an appropriate link between this statement and the issues raised by foveal sensing.

A second class of domains where foveal sensing might prove beneficial is in applications involving control loops that require very low latencies. This is the case not only in robot control, but also in perceptual user interfaces (PUIs) [247], and visual-based interfaces in particular [248]. One key problem in PUIs is to guarantee low reaction times to user commands captured through cameras as the input signal. Latencies that are too high may lead to unstable behavior and low usability of the interface and, in the end, low user acceptability of the system. Log-polar images might be an advantageous sensing device for these perceptual interfaces, which may include human–robot interaction [249].

Algorithmic, computational and geometric issues aside, still another insight on why log-polar vision can deserve study and investigation lies in its biological inspiration, for two reasons. One, because nature is wise and we, scientists and engineers, can try to introduce some of this wisdom into the artificial systems we build. And two, because these very artificial agents can provide us with

Table 1
Summary of achievements and open research issues.

SENSOR DESIGN
Past achievements: • Hardware sensors prototypes in both CCD and CMOS technologies • Commercial-level camera (Giotto) developed • Special-purpose hardware cards built • Influence of sensor parameters on performance of visual tasks studied • Framework to automate sensor design from design criteria Research agenda: • <i>Faster software-based log-polar mapping implementations</i> • <i>Parallel/GPU-based efficient mapping implementations</i> • <i>More flexible and customizable hardware-based silicon retinas</i> • <i>Autonomous task-dependent reconfigurable log-polar sensor</i>
VISUAL ATTENTION
Past achievements: • Bottom-up salience computation • Top-down biasing algorithms and experiments • Multiresolution analysis (cortical pyramids) Research agenda: • <i>Proper combination of salience-based algorithms with eye and head movements</i> • <i>Demonstration of the actual benefits of visual attention in realistic contexts</i> • <i>Integration of visual attention algorithms with practical and useful visual tasks</i> • <i>Devise specialized methods to process peripheral information</i> • <i>Develop space-invariant image descriptors</i>
VISUAL TRACKING
Past achievements: • Implicit focus-of-attention property studied and exploited • Translation-invariant motion estimation and tracking • Off-line (object-dependent) learning-based motion estimation Research agenda: • <i>Estimation of parametric motion (affine or higher models) with better exploitation of the log-polar space, without relying on Cartesian coordinates</i> • <i>Higher robustness when tracking small, occluded and/or non-centered objects</i> • <i>Exploit fovea–periphery dichotomy or geometry arrangement to segment the object or improve tracking efficiency or accuracy</i> • <i>Dealing with the multi-target tracking problem</i>
3D CUES AND VERGENCE CONTROL
Past achievements: • Computation of sparse and dense disparity maps • Demonstration of the advantages of log-polar images in stereo depth resolution • Depth from motion with fixation at a static point in the environment • Vergence control with static and dynamic depth cues, and demonstration of advantages in dealing with small objects under fixation • Zero disparity filtering and object segmentation Research agenda: • <i>Object segmentation in disparity maps</i> • <i>Integration of multiple cues (motion, color, edges, etc.)</i> • <i>High resolution depth maps by active scanning</i> • <i>Detection of occlusions and object boundaries</i>

cues for widening our still limited understanding of human vision and related brain processes.

In conclusion, while some work has been done on log-polar artificial vision in the past, this can be seen only as the germ of future work addressing new challenges. Table 1 is intended to be a brief guide to identify both problems that have already been solved and some open research issues. We would like the review offered in this article to be of some help for those looking back to the past in order to head into the future.

Note: Matlab code for the (direct and inverse) log-polar mapping can be found at <http://www.isr.ist.utl.pt/~alex/resources.html>.

Acknowledgements

The authors acknowledge the Integrated Actions programs funded by the Portuguese Government and Spanish Ministry of Science and Education, through projects E 47-06 and HP2005-0095, respectively, under the topic “FOVEAR: Foveal vision for robotic applications”, and the Spanish research programme Consolider Ingenio-2010 CSD2007-00018. The authors are also grateful to the anonymous reviewers for their comments, the “Servei de Llengües i Terminologia” at Universitat Jaume I for their professional English revision service, Thies Pfeiffer for telling us about Ref. [70],

and publishers and authors for their kind permission to reproduce figures used in this paper.

References

- [1] S. Edelman, Receptive fields for vision: From hyperacuity to object recognition, in: R.J. Watt (Ed.), Vision, MIT Press, 1996.
- [2] A. Bernardino, J. Santos-Victor, Binocular tracking: Integrating perception and control, IEEE Transactions on Robotics and Automation 15 (6) (1999) 1080–1094.
- [3] C.F.R. Weiman, Video compression via log-polar mapping, in: SPIE Symposium on OE/Aerospace Sensing, Orlando, Florida, Apr. 1990.
- [4] F. Panerai, C. Capurro, G. Sandini, Space variant vision for an active camera mount. Technical Report TR 1/95, LIRA, DIST, Univ. Genova, Italy, Feb. 1995.
- [5] C. Breazeal, A. Edsinger, P. Fitzpatrick, B. Scassellati, Active vision for sociable robots, Proc. IEEE Transactions on Systems, Man and Cybernetics Part A 31 (5) (2001) 443–453.
- [6] B. Bederson, A miniature space-variant active vision system, Ph.D. Thesis, New York University, 1992.
- [7] P. Daniel, D. Whitteridge, The representation of the visual field on the cerebral cortex in monkeys, Journal of Physiology 159 (1961) 203–221.
- [8] T. Lindeberg, L. Florack, Foveal scale-space and the linear increase of receptive field size as a function of eccentricity. Tech. Report ISRN KTH/NA/P-94/27-SE, 1994.
- [9] R. Tootell, M. Silverman, E. Swikes, R. DeValois, Deoxyglucose analysis of retinotopic organization in primate striate cortex, Science 218 (1982) 902–904.

- [10] E.L. Schwartz, Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception, *Biological Cybernetics* 25 (1977) 181–194.
- [11] J.Y. Aloimonos, I. Weiss, A. Bandyopadhyay, Active vision, *International Journal of Computer Vision (IJCV)* (1988) 333–356.
- [12] R. Bajcsy, Active perception and exploratory robotics, in: Dario, Sandini, Aebischer (Eds.), *Robots and Biological Systems: Towards a New Bionics?*, Springer-Verlag, 1993, pp. 3–20.
- [13] M. Swain, M. Stricker, Promising directions in active vision, *International Journal of Computer Vision (IJCV)* 11 (2) (1993) 109–126.
- [14] D.H. Ballard, Animate vision, *Artificial Intelligence* 48 (1991) 57–86.
- [15] E.L. Schwartz, D.N. Greve, G. Bonmassar, Space-variant active vision: Definition, overview and examples, *Neural Networks* 8 (7–8) (1995) 1297–1308.
- [16] M. Bolduc, M.D. Levine, A review of biologically motivated space-variant data reduction models for robotic vision, *Computer Vision and Image Understanding (CVIU)* 69 (2) (1998) 170–184.
- [17] G. Sandini, G. Metta, Retina-like sensors: Motivations, technology and applications, in: F.G. Barth, J.A. Humphrey, T.W. Secomb (Eds.), *Sensors and Sensing in Biology and Engineering*, Springer-Verlag, 2002.
- [18] F. Berton, G. Sandini, G. Metta, Anthropomorphic visual sensors, in: C. Grimes, E. Dickey, M.V. Pishko (Eds.), in: *Encyclopedia of Sensors*, vol. X, American Scientific Publishers, 2006, pp. 1–16.
- [19] M. Yeasin, R. Sharma, Foveated vision sensor and image processing—A review, in: B. Apolloni, A. Ghosh, F.N. Alpaslan, L.C. Jain, S. Patnaik (Eds.), *Machine Learning and Robot Perception*, in: *Studies in Computational Intelligence*, vol. 7, Springer, 2005, pp. 57–98.
- [20] L. Massone, G. Sandini, V. Tagliasco, 'Form-Invariant' topological mapping strategy for 2D shape recognition, *Computer Vision, Graphics, and Image Processing* 30 (1985) 169–188.
- [21] R.A. Messner, H.H. Szu, An image processing architecture for real time generation of scale and rotation invariant patterns, *Computer Vision, Graphics, and Image Processing Journal* (1985) 50–66.
- [22] J.C. Wilson, R.M. Hodgson, A pattern recognition system based on models of aspects of the human visual system, in: 4th Intl. Conf. on Image Processing and Applications, 1992, pp. 258–261.
- [23] V.J. Traver, F. Pla, The log-polar image representation in pattern recognition tasks, in: Iberian Conf. on Pattern Recognition and Image Analysis, IbPRIA, Port d'Andratx, Mallorca, Spain, June 2003, in: *Lecture Notes in Computer Science*, vol. 2652, Springer-Verlag, 2003, pp. 1032–1040.
- [24] E. Schwartz, Computational anatomy and functional architecture of the striate cortex, *Vision Research* 20 (1980) 645–669.
- [25] R. Wallace, P. Ong, B. Bederson, E. Schwartz, Space variant image processing, *International Journal of Computer Vision (IJCV)* 13 (1) (1994) 71–90.
- [26] F. Jurie, A new log-polar mapping for space variant imaging. Application to face detection and tracking, *Pattern Recognition* 32 (1999) 865–875.
- [27] C. Capurro, F. Panerai, G. Sandini, Dynamic vergence using log-polar images, *International Journal of Computer Vision (IJCV)* 24 (1) (1997) 79–94.
- [28] L. Florack, Modeling foveal vision, in: *Scale Space and Variational Methods in Computer Vision (SSVM)*, in: *Lecture Notes in Computer Science*, vol. 4485, Springer, 2007, pp. 919–928.
- [29] S. Wilson, On the retino-cortical mapping, *International Journal on Man-Machine studies* 18 (1983) 361–389.
- [30] G. Sandini, V. Tagliasco, An anthropomorphic retina-like structure for scene analysis, *Computer Graphics and Image Processing* 14 (3) (1980) 365–372.
- [31] M. Bolduc, M. Levine, A real-time foveated sensor with overlapping receptive fields, *Real-Time Imaging: Special Issue on Natural and Artificial Imaging and Vision* 3 (3) (1997) 195–212.
- [32] R.S. Wallace, P.-W. Ong, B.B. Bederson, E.L. Schwartz, Space variant image processing, *International Journal of Computer Vision (IJCV)* 13 (1) (1994) 71–90.
- [33] R.S. Wallace, P.-W. Ong, B.B. Bederson, E.L. Schwartz, Space-variant image processing, Technical Report 589-R256, Vision Applications, Inc., Nov. 1991.
- [34] R.S. Wallace, P.-W. Ong, B.B. Bederson, E.L. Schwartz, Space variant image processing, Technical Report 633, Courant Institute of Mathematical Sciences, New York University, NY, Apr. 1993.
- [35] J.A. Boluda, F. Pardo, T. Kayser, J.J. Pérez, J. Pelechano, A new foveated space-variant camera for robotic applications, in: IEEE Intl. Conference on Electronics Circuits and Systems, ICECS'96, Rhodes, Greece, Oct. 1996, pp. 680–683.
- [36] F. Pardo, J.A. Boluda, J.J. Pérez, B. Dierickx, D. Scheffer, Design issues on CMOS space-variant image sensors, in: SPIE Conf. on Advanced Focal Plane Arrays and Electronic Cameras, AFPAEC, vol. 2950, Berlin, Germany, Oct. 1996, pp. 98–107.
- [37] R. Wodnicki, G.W. Roberts, M.D. Levine, A log-polar image sensor fabricated in a standard 1.2- μ m ASIC CMOS process, *Journal of Solid-State Circuits* 32 (8) (1997) 1274–1277.
- [38] F. Pardo, B. Dierickx, D. Scheffer, Space-variant nonorthogonal structure CMOS image sensor design, *Journal of Solid-State Circuits* 33 (6) (1998) 842–849.
- [39] S. Meikle, R. Yates, Building smarter sensors—Lessons learned from computer vision, in: *Intelligent Vehicles Symposium*, 2000, IV 2000, in: *Proceedings of the IEEE*, Dearborn, MI, USA, Oct. 2000, pp. 210–214.
- [40] A.M. Wallace, D.J. McLaren, Gradient detection in discrete log-polar images, *Pattern Recognition Letters* 24 (14) (2003) 2185–2208.
- [41] B. Fischl, M.A. Cohen, E.L. Schwartz, Rapid anisotropic diffusion using space-variant vision, *International Journal of Computer Vision (IJCV)* 28 (3) (1998) 199–212.
- [42] S.W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Pub, 1997, <http://www.dspguide.com>.
- [43] B.G. Borghuis, C.P. Ratliff, R.G. Smith, P. Sterling, V. Balasubramanian, Design of a neuronal array, *The Journal of Neuroscience* 28 (12) (2008) 3178–3189.
- [44] V.J. Traver, F. Pla, Motion analysis with the Radon transform on log-polar images, *Journal of Mathematical Imaging and Vision* 30 (2) (2008) 147–165.
- [45] S. Shimizu, J. Burdick, Eccentricity compensator for log-polar sensor, in: IEEE Intl. Conf. on Robotics and Automation, ICRA, Apr. 2007, pp. 4214–4219.
- [46] F. Tong, Z.-N. Li, Reciprocal-Wedge transform for space-variant sensing, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 17 (5) (1995) 500–511.
- [47] F. Tong, Reciprocal-Wedge Transform: A Space-Variant Image Representation, Ph.D. Thesis, School of Computing Science, Simon Fraser University, May 1995.
- [48] F. Arrebola, C. Urdiales, P. Camacho, F. Sandoval, Vision system based on shifted fovea multiresolution retinotopologies, in: 24th annual conference of the IEEE Industrial Electronics Society, IECON'98, vol. 3, Aachen, Germany, Aug.–Sept. 1998, pp. 1357–1361.
- [49] C. Urdiales, A. Bandera, J. Rodríguez, F. Sandoval, Video flow active control by means of adaptive shifted foveal geometries, in: SPIE International Symposium on Intelligent Systems for Advanced Manufacturing, ISAM'2000: Intelligent Robots and Computer Vision XIX, vol. 4197, Boston, MA, USA, Nov. 2000.
- [50] J.A. Rodríguez, C. Urdiales, A. Bandera, F.S. Hernández, A multiresolution spatiotemporal motion segmentation technique for video sequences based on pyramidal structures, *Pattern Recognition Letters* 23 (14) (2002) 1761–1769.
- [51] R. Marfil, L. Molina-Tanco, A. Bandera, J.A. Rodríguez, F.S. Hernández, Pyramid segmentation algorithms revisited, *Pattern Recognition* 39 (8) (2006) 1430–1451.
- [52] J. Martínez, L. Altamirano, FPGA-based pipeline architecture to transform cartesian images into foveal images by using a new foveation approach, in: Proc. of the IEEE International Conference on Reconfigurable Computing and FPGAs, ReConfig, Sept. 2006.
- [53] E.-C. Chang, C.K. Yap, A wavelet approach to foveating images, in: ACM Symp. on Computational Geometry, 1997, pp. 397–399.
- [54] E.C. Chang, C.K. Yap, T.-J. Yen, Real-time visualization of large images over a thinwire, in: IEEE Visualization (Late Breaking Hot Topics), 1997.
- [55] P.T. Kortum, W.S. Geisler, Implementation of a foveated image-coding system for bandwidth reduction of video images, in: B. Rogowitz, J. Allebach (Eds.), *Human Vision and Electronic Imaging*, in: SPIE Proceedings, vol. 2657, 1996, pp. 350–360.
- [56] P. Camacho, F. Coslado, M. González, F. Sandoval, Multifoveal imager for stereo applications, *International Journal of Imaging Systems and Technology* 12 (4) (2002) 149–165.
- [57] A. Basu, K.J. Wiebe, Enhancing videoconferencing using spatially varying sensing, *IEEE Transactions on Systems, Man and Cybernetics* 28 (2) (1998) 137–148.
- [58] J.A. Rodríguez, C. Urdiales, A. Bandera, F. Sandoval, Nonuniform video coding by means of multifoveal geometries, *International Journal of Imaging Systems and Technology* 12 (1) (2002) 27–34.
- [59] M.W. Peters, S. Arcot, A real-time variable sampling technique: DIEM, in: Intl. Conf. on Pattern Recognition, ICPR, Brisbane, Australia, 1998.
- [60] G. Sandini, P. Dario, M. DeMicheli, M. Tistarelli, Retina-like CCD sensor for active vision, in: *Computer and Systems Sciences (NATO ARW on Robots and Biological Systems)*, Il Ciocco, Tuscany, Italy, 1989, Springer-Verlag, 1989.
- [61] J. Van der Spiegel, R. Etienne-Cummings, M. Nishimura, Biologically inspired vision sensors, in: 23rd International Conference on Microelectronics (MIEL 2002), volume 1, 12–15 May 2002, pp. 125–131.
- [62] G. Kreider, J. Van der Spiegel, I. Born, C. Clays, I. Debusschere, G. Sandini, P. Dario, Intelligent Robots and Computer Vision IX: Algorithms and Techniques, in: Proc. of SPIE, vol. 1381, 1990.
- [63] F. Pardo, E. Martinuzzi, Hardware environment for a retinal CCD visual sensor, in: EU-HCM Smart Workshop: Semi-Autonomous Monitoring and Robotics Technologies, Ispra, Italy, Apr. 1994.
- [64] C.W. Shin, S. Inokuchi, K.I. Kim, Retina-like visual sensor for fast tracking and navigation robots, *Machine Vision and Applications (MVA)* 10 (1997) 1–8.
- [65] R. Wodnicki, G.W. Roberts, M.D. Levine, A foveated image sensor in standard CMOS technology, in: Custom Integrated Circuits Conf., Santa Clara, May 1995.
- [66] F. Pardo, J. Boluda, J. Perez, S. Felici, B. Dierickx, Design of a foveated log-polar image sensor using standard CMOS technology, in: XI Design of Integrated Circuits and Systems, Sitges, Spain, Nov. 1996, pp. 49–54.
- [67] G. Sandini, P. Questa, D. Scheffer, B. Dierickx, A. Mannucci, A retina-like CMOS sensor and its applications, in: Proc. 1st IEEE SAM Workshop, Cambridge, USA, Mar. 2000.
- [68] A. Moini, Vision chips or seeing silicon, Technical Report, The Centre for High Performance Integrated Technologies and Systems. The University of Adelaide, Australia, Mar. 1997.
- [69] A. Moini, Vision Chips, Kluwer Academic Publishers, Norwell, MA, USA, 1999.
- [70] H.C. Ko, M.P. Stoykovich, J. Song, V. Malyarchuk, W.M. Choi, C.-J. Yu, Joseph, J. Xiao, S. Wang, Y. Huang, J.A. Rogers, A hemispherical electronic eye camera based on compressible silicon optoelectronics, *Nature* 454 (7205) (2008) 748–753.

- [71] V. Krüger, Optical flow estimation in the complex logarithmic plane. Master's Thesis, University of Kiel, Germany, 1995.
- [72] R. Alan Peters II, M. Bishay, T. Rogers, On the computation of the log-polar transform. Technical Report, School of Engineering, Vanderbilt University, Mar. 1996.
- [73] J. Dias, H. Araujo, C. Paredes, J. Batista, Optical normal flow estimation on log-polar images. A solution for real-time binocular vision, *Real Time Imaging* 3 (3) (1997) 213–228.
- [74] H. Araujo, J.M. Dias, An introduction to the log-polar mapping, in: *Proc. of Second Workshop on Cybernetic Vision*, 1996, pp. 139–144.
- [75] C.F.R. Weiman, Exponential sensor array geometry and simulation, in: *Preprint from the Proc. of SPIE, Orlando, Florida*, in: *Digital and Optical Shape Representation and Pattern Recognition*, vol. 938, Apr. 1988.
- [76] A.S. Rojer, E.L. Schwartz, Design considerations for a space-variant visual sensor with complex-logarithmic geometry, in: *Intl. Conf. on Pattern Recognition, ICPR*, 1990, pp. 278–285.
- [77] F.L. Lim, S. Venkatesh, G.A.W. West, Resolution consideration in spatially variant sensors, *Image and Vision Computing (IVC)* 15 (1997) 901–912.
- [78] C.F.R. Weiman, Log-polar binocular vision system. NASA Phase II SBIR Final Report, Dec. 1994.
- [79] T.E. Fisher, R.D. Juday, A programmable video image remapper, in: *SPIE Conf. on Pattern Recognition and Signal Processing*, in: *Digital and Optical Shape Representation and Pattern Recognition*, vol. 938, 1988, pp. 122–128.
- [80] J.R. del Solar, C. Nowack, B. Schneider, VIPOL: A virtual polar-logarithmic sensor, in: *Scandinavian Conf. on Image Analysis, SCIA, Finland*, 1997, pp. 739–744.
- [81] H. Nguyen (Ed.), *GPU Gems 3*, Addison-Wesley Professional, 2007.
- [82] M. Ujaldón, A. Ruiz, N. Guil, On the computation of the circle Hough transform by a GPU rasterizer, *Pattern Recognition Letters* 29 (3) (2008) 309–318.
- [83] C.G. Ho, R.C.D. Young, C.R. Chatwin, Sensor geometry and sampling methods for space-variant image processing, *Pattern Analysis and Application (PAA)* 5 (2002) 369–384.
- [84] A. Bernardino, J. Santos-Victor, Sensor geometry for dynamic vergence: Characterization and performance analysis, in: *Workshop on Performance Characteristics of Vision Algorithms, ECCV*, Cambridge, UK, Apr. 1996.
- [85] S.L. Bartlett, A. Hampapur, M.J. Huber, D. Kortenkamp, S. Moezzi, Vision for mobile robots, in: J.L.C. Sanz (Ed.), *Image Technology: Advances in Image Processing, Multimedia and Machine Vision*, Springer, 1996, pp. 1–37.
- [86] V.J. Traver, F. Pla, Log-polar mapping template design: From task-level requirements to geometry parameters, *Image and Vision Computing (IVC)* 26 (10) (2008) 1354–1370.
- [87] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.
- [88] Z. Bien, W.-K. Song, D.-J. Kim, J.-S. Han, J.-E. Lee, J.-S. Kim, Vision-based control with emergency stop through EMG of the wheelchair-based rehabilitation robotic arm, KARES II, in: *Intl. Conf. on Rehabilitation Robotics*, Evry, France, Apr. 2001.
- [89] W.-K. Song, D.-J. Kim, J.-S. Kim, Z. Bien, Visual servoing for a user's mouth with effective intention reading in a wheelchair-based robotic arm, in: *IEEE Intl. Conf. on Robotics and Automation, ICRA*, Seoul, Korea, May 2001.
- [90] W.-K. Song, J.-S. Kim, Z. Bien, Visual servoing for human-robot interaction in the wheelchair-based rehabilitation robot, in: *IEEE Intl. Conf. on Systems, Man and Cybernetics*, Nashville, Tennessee, USA, Oct. 2000, pp. 1811–1816.
- [91] J. López-Sánchez, V.J. Traver, F. Pla, Estudio comparativo de modelos de conversión de imágenes log-polar, in: *Conf. de la Asoc. Española para la Inteligencia Artificial, CAEPIA*, vol. I, Gijón, Spain, Nov. 2001, pp. 655–664 (in Spanish).
- [92] V.J. Traver, Motion Estimation Algorithms in Log-polar Images and Application to Monocular Active Tracking, Ph.D. Thesis, Dep. Llenguatges i Sistemes Informàtics, Universitat Jaume I, Castellón, Spain, Sept. 2002.
- [93] D. Pamplona, A. Bernardino, Smooth foveal vision with Gaussian receptive fields, in: *9th IEEE-RAS Intl. Conf. on Humanoid Robots*, 2009.
- [94] B.B. Bederson, R.S. Wallace, E.L. Schwartz, A miniature pan-tilt actuator: The spherical pointing motor, *IEEE Transactions on Robotics and Automation* 10 (3) (1994).
- [95] B.B. Bederson, R.S. Wallace, E.L. Schwartz, A miniaturized space-variant active vision system: Cortex-I, *Machine Vision and Applications (MVA)* 8 (2) (1995) 101–109.
- [96] C. Capurro, F. Panerai, E. Grosso, G. Sandini, A binocular active vision system using space variant sensors: Exploiting autonomous behaviors for space applications, in: *3rd. Intl. Conf. on Digital Signal Processing*, Nicosia, July 1993.
- [97] J. Santos-Victor, F. van Trigt, J. Sentieiro, Medusa—A stereo head for active vision, in: *Intl. Workshop on Intelligent Robotic Systems (IRS)*, Grenoble, France, July 1994.
- [98] G. Metta, F. Panerai, R. Manzotti, G. Sandini, Babybot: An artificial developing robotic agent, in: *Intl. Conf. on the simulation of adaptive behavior, SAB*, Paris, France, Sept. 2000.
- [99] A. Yeung, N. Barnes, Towards log-polar fixation for mobile robots—Analysis of corner tracking on the log-polar camera, in: *Intl. Conf. on Pattern Recognition, ICPR*, vol. II, Québec, Canada, 2002, pp. 300–303.
- [100] A. Yeung, N. Barnes, Efficient active monocular fixation using the log-polar sensor, *International Journal of Intelligent Systems Technologies and Applications* 1 (1/2) (2005) 157–173.
- [101] F.L. Lim, G.A.W. West, S. Venkatesh, Tracking in a space variant active vision system, in: *Intl. Conf. on Pattern Recognition, ICPR*, Vienna, Austria, Aug. 1996, pp. 745–749.
- [102] F.L. Lim, G.A.W. West, S. Venkatesh, Use of log polar space for foveation and feature recognition, *IEE Proc. Vis. Image Signal Process.* 144 (6) (1997) 323–331.
- [103] P.B. Melnyk, R.A. Messner, Biologically motivated composite image sensor for deep-field target tracking, in: L.J. Latecki, D.M. Mount, A.Y. Wu (Eds.), in: *Vision Geometry XV*, volume 6499(1), SPIE, 2007, pp. 649905–649905–8.
- [104] P. Melnyk, R.A. Messner, Log-polar based framework for mobile vehicle tracking with road follower, in: *Defense and Security Symposium, SPIE*, Orlando, FL, USA, Apr. 2007.
- [105] L. Itti, Modelling primate visual attention, in: J. Feng (Ed.), *Computational Neuroscience: A Comprehensive Approach*, CRC Press, 2003, pp. 635–655.
- [106] L. Itti, C. Koch, A saliency-based search mechanism for overt and covert shifts of visual attention, *Vision Research* 40 (2000) 1489–1506.
- [107] J.M. Wolfe, Guided search 2.0: A revised model of visual search, *Psychonomic Bulletin & Review* 1 (2) (1994) 202–238.
- [108] K.R. Cave, The FeatureGate model of visual selection, *Psychological Research* 62 (1999) 182–194.
- [109] Y. Sun, R. Fisher, Object-based visual attention for computer vision, *Artificial Intelligence* 146 (2003) 77–123.
- [110] N. Tamayo, V.J. Traver, Entropy-based saliency computation in log-polar images, in: *Intl. Conf. on Computer Vision Theory and Applications*, vol. 1, Funchal, Madeira, Portugal, Jan. 2008, pp. 501–506.
- [111] R. Milanese, Detecting Salient regions in images: From biological evidence to computer implementation, Ph.D. Thesis, Dep. of Computer Science, University of Geneva, Geneva, Switzerland, 1993.
- [112] R.P.N. Rao, Top-down gaze targeting for space-variant active vision, in: *Proc. ARPA Image Understanding Workshop*, Monterey, California, USA, Nov. 1994, pp. 1049–1058.
- [113] R.P.N. Rao, D.H. Ballard, An active vision architecture based on iconic representations, *Artificial Intelligence* 78 (1–2) (1995) 461–505.
- [114] G. Metta, An attentional system for a humanoid robot exploiting space variant vision, in: *Intl. Conf. on Humanoid Robots*, Tokyo, Japan, Nov. 2001, pp. 22–24.
- [115] M. Bellardi, Modello per l'estrazione di caratteristiche salienti ispirato al sistema visivo periferico. Master's Thesis, Laboratorio Integrato di Robotica Avanzata, Dipartimento di Informatica Sistemistica e Telematica, Facoltà di Ingegneria, Università di Genova, 2003.
- [116] F. Orabona, G. Metta, G. Sandini, Object-based visual attention: A model for a behaving robot, in: *Proc. IEEE Computer Vision and Pattern Recognition, CVPR*, San Diego, CA, USA, June 2005.
- [117] L. Itti, C. Koch, E. Niebur, A model of saliency-based visual attention for rapid scene analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 20 (11) (1998).
- [118] O. Stasse, Y. Kuniyoshi, G. Cheng, Development of a biologically inspired real-time visual attention system, in: *BMCV 2000*, Seoul, Korea, in: *LNCS*, vol. 1811, May 2000, pp. 150–159.
- [119] C. Colombo, M. Rucci, P. Dario, Attentive behavior in an anthropomorphic robot vision system, *Robotics and Autonomous Systems* 12 (3–4) (1994) 121–131.
- [120] C. Colombo, M. Rucci, P. Dario, Integrating selective attention and space-variant sensing in machine vision, in: J.L.C. Sanz (Ed.), *Image Technology: Advances in Image Processing, Multimedia and Machine Vision*, Springer, 1996, pp. 109–127.
- [121] H. Yamamoto, Y. Yeshurun, M.D. Levine, An active foveated vision system: Attentional mechanisms and scan path convergence measures, *Computer Vision and Image Understanding (CVIU)* 63 (1) (1996) 50–65.
- [122] Y. Yeshurun, E. Schwartz, Shape description with a space-variant sensor: Algorithms for scan-path, fusion, and convergence over multiple scans, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 11 (11) (1989) 1217–1222.
- [123] G. Sela, M.D. Levine, Real-time attention for robotic vision, *Real Time Imaging* 3 (1997) 173–194.
- [124] G. Kootstra, N. Arco, B. de Boer, Paying attention to symmetry, in: *British Machine Vision Conference*, Leeds, UK, Sept. 2008.
- [125] T.D. Groove, R.B. Fisher, Attention in iconic object matching, in: *Proc. British Machine Vision Conf., BMVC96*, Edinburgh, Sept. 1996, pp. 293–302.
- [126] T. Kadir, M. Brady, Saliency, scale and image description, *International Journal of Computer Vision (IJCV)* 45 (2) (2001) 83–105.
- [127] K. Mikolajczyk, C. Schmid, Scale & affine invariant interest point detectors, *International Journal of Computer Vision (IJCV)* 60 (1) (2004) 63–86.
- [128] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision (IJCV)* 60 (2) (2004) 91–110.
- [129] N. Oshiro, A. Nishikawa, N. Maru, F. Miyazaki, Foveated vision for scene exploration, in: *3rd Asian Conference on Computer Vision, ACCV98*, Hong Kong, China, Jan. 1998, pp. 256–263.
- [130] M. Tistarelli, Active/space-variant object recognition, *Image and Vision Computing (IVC)* 13 (3) (1995) 215–226.
- [131] M. Tistarelli, E. Grosso, Active vision-based face authentication, *Image and Vision Computing (IVC)* 18 (2000) 299–314.
- [132] Y. Baba, T. Kurita, M. Tanaka, S. Akaho, S. Umeyama, T. Mishima, An experimental foveated vision system using fish-eye lens and integration of retinal images, in: *Intl. Technical Conf. on Circuits/Systems, Computers, and Communications, ITC-CSCC*, vol. I, 1999, pp. 1–8.
- [133] F. Pla, V.J. Traver, Spherical panoramas for pan-tilt camera motion compensation in space-variant images, in: M.T. Escrig, F. Toledo, E. Golobardes (Eds.), *Topics in Artificial Intelligence*, in: *Lecture Notes in Artificial Intelligence*, vol. 2504, Springer-Verlag, 2002, pp. 375–387.

- [134] T. Baron, M.D. Levine, Y. Yeshurun, Exploring with foveated robot eye system, in: Intl. Conf. on Pattern Recognition (ICPR), Jerusalem, Israel, Oct. 1994, pp. 377–380.
- [135] I. Ahrens, H. Neumann, Space-variant dynamic neural fields for visual attention, in: Proc. IEEE Computer Vision and Pattern Recognition, CVPR, Fort Collins, Colorado, June 1999.
- [136] Y. Sun, R.B. Fisher, F. Wang, H.M. Gomes, A computer vision model for visual-object-based attention and eye movements, *Computer Vision and Image Understanding (CVIU)* 112 (2) (2008) 126–142.
- [137] A. Bernardino, Binocular Head Control with Foveal Vision: Methods and Applications, Ph.D. Thesis, Vislab, ISR-IST, Lisbon, Portugal, Apr. 2004.
- [138] V.J. Traver, Visual attention under foveated vision in robotic heads, Technical Report, Vislab, ISR-IST, Lisbon, Portugal and DLSI-UJI, Castellón, Spain, 2003.
- [139] B.T. Vincent, T. Troscianko, I.D. Gilchrist, Investigating a space-variant weighted salience account of visual selection, *Vision Research* 47 (13) (2007) 1809–1820.
- [140] J. Intriligator, P. Cavanagh, The spatial resolution of visual attention, *Cognitive Psychology* 43 (2001) 171–216.
- [141] R.H.S. Carpenter, Movements of the eyes, in: London Pion, second edition, 1988.
- [142] J.M. Findlay, I.D.D. Gilchrist, Active Vision: The Psychology of Looking and Seeing, Oxford University Press, 2003.
- [143] R.W. Picard, Affective Computing, The MIT Press, 1997.
- [144] V. Navalpakkam, L. Itti, Sharing resources: Buy attention, get object recognition, in: Intl. Workshop on Attention and Performance in Computer Vision, WAPCV, Graz, Austria, July 2003.
- [145] S.J. Anderson, N. Yamagishi, V. Karavia, Attentional processes link perception and action, *Proceedings of the Royal Society of London B* 269 (May) (2002) 1225–1226.
- [146] L. Montesano, M. Lopes, A. Bernardino, J. Santos-Victor, Learning object affordances: From sensory motor maps to imitation, *IEEE Transactions on Robotics* 24 (1) (2008) 15–26.
- [147] D. Murray, A. Basu, Motion tracking with an active camera, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 16 (5) (1994) 449–459.
- [148] V.J. Traver, F. Pla, Similarity motion estimation and active tracking through spatial-domain projections on log-polar images, *Computer Vision and Image Understanding (CVIU)* 97 (2) (2005) 209–241.
- [149] E. Rivlin, H. Rotstein, Control of a camera for active vision: Foveal vision, smooth tracking and saccade, *International Journal of Computer Vision (IJCV)* 39 (2) (2000) 81–96.
- [150] R. Lencer, P. Trillenber, Neurophysiology and neuroanatomy of smooth pursuit in humans, *Brain and Cognition* 68 (2008) 219–228.
- [151] M. Lopes, A. Bernardino, J. Santos-Victor, C. Hofsten, K. Rosander, Biomimetic eye-neck coordination, in: IEEE 8th International Conference on Development and Learning, ICDL'09, Shanghai, China, 2009.
- [152] E. Falotico, M. Taiana, D. Zambrano, A. Bernardino, J. Santos-Victor, C. Laschi, P. Dario, Predictive tracking across occlusions on the iCub robot, in: 9th IEEE-RAS International Conference on Humanoid Robots, 2009.
- [153] T. Shibata, S. Vijayakumar, J. Conradt, S. Schaal, Biomimetic oculomotor control, *Adaptive Behavior* 9 (3–4) (2001) 187–207.
- [154] M. Vincze, C.F.R. Weiman, A general relationship for optimal tracking performance, in: Proc. of the SPIE'96: Intelligent Robots and Computer Vision XV, Boston, MA, USA, 1996, pp. 402–412.
- [155] C. Stiller, J. Konrad, Estimating motion in image sequences: A tutorial on modeling and computation of 2D motion, *IEEE Signal Processing Magazine* (1999) 70–91.
- [156] B. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, in: Proc. of the International Joint Conference on Artificial Intelligence, 1981, pp. 674–679.
- [157] J.J. Koenderink, Optic flow, *Vision Research* 26 (1) (1986) 161–180.
- [158] S.S. Beauchemin, J.L. Barron, The computation of optical flow, *ACM Computing Surveys* 27 (3) (1995) 433–467.
- [159] S. Baker, I. Matthews, Lucas–Kanade 20 years on: A unifying framework, *International Journal of Computer Vision (IJCV)* 56 (3) (2004) 221–255.
- [160] B.K. Horn, Robot Vision, McGraw-Hill Higher Education, 1986.
- [161] K. Daniilidis, V. Krüger, Optical flow computation in the log-polar plane, in: V. Hlaváč, R. Šára (Eds.), Intl. Conf. on Computer Analysis of Images and Patterns, CAIP, Springer-Verlag, Berlin, Heidelberg, 1995, pp. 65–72.
- [162] K. Daniilidis, Computation of 3D-motion parameters using the log-polar transform, in: V. Hlaváč, R. Šára (Eds.), Intl. Conf. on Computer Analysis of Images and Patterns, CAIP, in: LNCS, vol. 970, Springer-Verlag, Berlin, Heidelberg, 1995, pp. 82–89.
- [163] M. Yeasin, Optical flow in log-mapped image plane: A new approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 24 (1) (2002) 125–131.
- [164] B. Fischl, M.A. Cohen, E.L. Schwartz, The local structure of space-variant images, *Neural Networks* 10 (5) (1997) 815–831.
- [165] H. Tunley, D. Young, First order optic flow from log-polar sampled images, in: J.-O. Eklundh (Ed.), European Conf. on Computer Vision, in: LNCS, vol. 800, Springer Verlag, 1994, pp. 132–137.
- [166] M.R. Blackburn, H.G. Nguyen, Autonomous visual control of a mobile robot, in: ARPA Image Understanding Workshop, Monterey, California, Nov. 1994.
- [167] G. Metta, Baby(rob)ot: A Study on Sensori-Motor Development, Ph.D. Thesis, LIRA-Lab, DIST, University of Genoa, Italy, 1999.
- [168] F. Panerai, Integration of Inertial and Visual Information in Binocular Vision Systems, Ph.D. Thesis, LIRA-Lab (Laboratory for Integrated Advanced Robotics), Department of Communication Computer and Systems Science, Faculty of Engineering, University of Genova, Feb. 1998.
- [169] D. Claveau, C. Wang, Space-variant motion detection for active visual target tracking, *Robotics and Autonomous Systems* 57 (1) (2009).
- [170] G.D. Hager, P.N. Belhumeur, Efficient region tracking with parametric models of geometry and illumination, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 20 (10) (1998) 1025–1039.
- [171] A. Bernardino, J. Santos-Victor, G. Sandini, Foveated active tracking with redundant 2D motion parameters, *Robotics and Autonomous Systems* 39 (3–4) (2002) 205–221.
- [172] V.J. Traver, F. Pla, Motion estimation and figure-ground segmentation in log-polar images, in: Intl. Conf. on Pattern Recognition, ICPR, Québec, Canada, Aug. 2002, pp. 166–169.
- [173] C. Narathong, R. Niño, J. Doner, E. McVey, An algorithm for motion prediction using a biological visual sensor, in: IEEE Intl. Conf. on Robotics and Automation, ICRA, vol. 4, 1987, pp. 1261–1266.
- [174] V.J. Traver, F. Pla, Dealing with 2D translation estimation in log-polar imagery, *Image and Vision Computing (IVC)* 21 (2) (2003) 145–160.
- [175] G. Bonmassar, E.L. Schwartz, Space-variant Fourier analysis: The exponential chirp transform, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 19 (10) (1997) 1080–1089.
- [176] G. Bonmassar, E.L. Schwartz, Real-time restoration of images degraded by uniform motion blur in foveal active vision systems, *IEEE Transactions on Image Processing* 8 (12) (1999) 1838–1842.
- [177] M. Gleicher, Projective registration with difference decomposition, in: Proc. IEEE Computer Vision and Pattern Recognition, CVPR, pp. 331–337, 1997.
- [178] A. Bernardino, J. Santos-Victor, G. Sandini, Tracking planar structures with log-polar images, in: Symp. on Intelligent Robotic Systems, Reading, UK, July 2000.
- [179] N. Okajima, H. Nitta, W. Mitsuhashi, Motion estimation and target tracking in the log-polar geometry, in: 17th Sensor Symposium, Kawasaki, Japan, May 2000.
- [180] H. Tunley, D. Young, Dynamic fixation of a moving surface using log polar sampling, in: British Machine Vision Conference, vol. 2, Univ. York, York, Sept. 1994, BMVA Press, 1994, pp. 579–588.
- [181] S. Zokai, G. Wolberg, Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations, *IEEE Transactions on Image Processing* 14 (10) (2005) 1422–1434.
- [182] R. Montoliu, V.J. Traver, F. Pla, Log-polar mapping in generalized least-squares motion estimation, in: Visualization, Imaging and Image Processing Conf., Benalmádena, Málaga, Spain, Sept. 2002, pp. 656–661.
- [183] V.J. Traver, F. Pla, Radon-like transforms in log-polar images for affine motion estimation, in: Portuguese Conf. on Pattern Recognition, Aveiro, Portugal, June 2002.
- [184] L. Puig-M, V.J. Traver, Efficient tracking of objects with arbitrary 2D motions in space-variant imagery, in: C. Angulo, L. Godo (Eds.), Desé Congrès Català d'Intel·ligència Artificial (CCIA 2007), Sant Julià de Lòria, Principat d'Andorra, in: Frontiers in Artificial Intelligence and Applications, vol. 163, IOS Press, 2007, pp. 171–180.
- [185] F. Jurie, M. Dhome, Real time tracking of 3D objects: An efficient and robust approach, *Pattern Recognition* 35 (2) (2002) 317–328.
- [186] C.F.R. Weiman, Polar exponential sensor arrays unify iconic and Hough space representation, in: SPIE Conf. on Intelligent Robots and Computer Vision VIII: Algorithms and Techniques, vol. 1192, Philadelphia, Nov. 1989.
- [187] D. Young, Straight lines and circles in the log-polar image, in: British Machine Vision Conference, Bristol, UK, Sept. 2000.
- [188] K. Schindler, Geometry and construction of straight lines in log-polar images, *Computer Vision and Image Understanding (CVIU)* 103 (3) (2006) 196–207.
- [189] H.M. Gomes, R. Fisher, Learning-based versus model-based log-polar feature extraction operators: A comparative study, in: XVI Brazilian Symp. on Computer Graphics & Image Processing, SIBGRAPI, San Carlos, Brazil, IEEE Computer Society Press, Oct. 2003.
- [190] H.M. Gomes, R.B. Fisher, Primal sketch feature extraction from a log-polar image, *Pattern Recognition Letters* 24 (7) (2003) 983–992.
- [191] E. Nattel, Y. Yeshurun, Direct feature extraction in a foveated environment, *Pattern Recognition Letters* 23 (2002) 1537–1548.
- [192] C.F. Weiman, R.D. Juday, Tracking algorithms using log-polar mapped image coordinates, in: Intelligent Robots and Computer Vision VIII: Algorithms and Techniques, SPIE, 1989, pp. 843–853.
- [193] R. Young, J. Illingworth, A fixation and viewpoint measure for object-based gaze control, in: British Machine Vision Conference, 1997.
- [194] R.J. Wood, Topography of primary visual cortex: Theory and implementation, Ph.D. Thesis, Graduate School of Arts and Sciences, Boston University, 2001.
- [195] S. Kang, S.-W. Lee, Real-time tracking of multiple objects in space-variant vision based on magnocellular visual pathway, *Pattern Recognition* 35 (10) (2002) 2031–2040.
- [196] V.J. Traver, A. Bernardino, P. Moreno, J. Santos-Victor, Appearance-based object detection in space-variant images: A multi-model approach, in: Intl. Conf. on Image Analysis and Recognition, ICIAR, volume 1, Porto, Portugal, in: Lecture Notes in Computer Science, vol. 3211, Nov. 2004, pp. 538–546.
- [197] G. Metta, A. Gasteratos, G. Sandini, Learning to track colored objects with log-polar vision, *Mechatronics* 14 (9) (2004) 989–1006.
- [198] A. Elgammal, R. Duraiswami, D. Harwood, L.S. Davis, Background and foreground modeling using nonparametric kernel density for visual surveillance, *Proceedings of the IEEE* 90 (7) (2002) 1151–1163.

- [199] C.F.R. Weiman, Log-polar vision for mobile robot navigation, in: Proc. of Electronic Imaging Conference, Boston, USA, Nov. 1990, pp. 382–385.
- [200] M. Tistarelli, G. Sandini, On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 15 (1993) 401–410.
- [201] G. Salgian, D.H. Ballard, Looming detection in log-polar coordinates, in: DARPA Image Understanding Workshop, Monterey, California, 1998, pp. 165–170.
- [202] F. Pardo, J. Boluda, I. Coma, F. Micó, High speed log-polar time to crash calculation for mobile vehicles, *Image Processing & Communications* 8 (2) (2002) 23–32.
- [203] F. Pardo, J.A. Boluda, E. de Ves, Feature extraction and correlation for time-to-impact segmentation using log-polar images, in: A. Laganà, M.L. Gavrilova, V. Kumar, Y. Mun, C.J.K. Tan, O. Gervasi (Eds.), *ICCSA* (4), in: *Lecture Notes in Computer Science*, vol. 3046, Springer, 2004, pp. 887–895.
- [204] K. Daniilidis, Attentive visual motion processing: Computations in the log-polar plane, *Computing* 11 (1995) 1–20.
- [205] C. Colombo, Time to collision from first-order spherical motion, *Robotics and Autonomous Systems* 31 (2000) 5–15.
- [206] I. Ahnrs, H. Neumann, A view-based approach for real-time fixation using log-polar mapping, in: C. Freksa (Ed.), *Proc. in Artificial Intelligence*, June 1998, pp. 89–96.
- [207] I. Ahnrs, H. Neumann, Real-time monocular fixation control using the log-polar transformation and a confidence-based similarity measure, in: A.K. Jain, S. Venkatesh, B.C. Lovell (Eds.), *Intl. Conf. on Pattern Recognition, ICPR*, Brisbane, Australia, Aug. 1998, pp. 310–315.
- [208] A. Bernardino, J. Santos-Victor, G. Sandini, Model-based attention fixation using log-polar images, in: V. Cantoni, A. Petrosino, M. Marinaro (Eds.), *Visual Attention Mechanisms*, Plenum Press, New York, 2002.
- [209] C. Capurro, F. Panerai, G. Sandini, Vergence and tracking fusing log-polar images, in: *Intl. Conf. on Pattern Recognition, ICPR*, IEEE, 1996, pp. 740–744.
- [210] V.J. Traver, F. Pla, An optimization approach for translational motion estimation in log-polar domain, in: W. Skarbek (Ed.), *Intl. Conf. on Computer Analysis of Images and Patterns, CAIP*, Warsaw, Poland, in: *Lecture Notes in Computer Science*, vol. 2124, Springer-Verlag, Sept. 2001, pp. 365–373.
- [211] N. Barnes, G. Sandini, Active docking based on the rotational component of log-polar optic flow, in: W.-H. Tsai and H.-J. Lee (Eds.), *Asian Conf. on Computer Vision*, Taiwan, Jan. 2000, pp. 955–960.
- [212] N. Barnes, G. Sandini, Direction control for an active docking behavior based on the rotational component of log-polar optic flow, in: W.-H. Tsai and H.-J. Lee (Eds.), *European Conf. on Computer Vision*, vol. 2, Dublin, Ireland, June 2000, pp. 167–181.
- [213] C. Silva, J. Santos-Victor, Egomotion estimation using log-polar images, in: *Intl. Conf. on Computer Vision*, Bombay, India, Jan. 1998.
- [214] S. Lam, E. Tsang, Y. Meng, B. Shi, Neuromorphic translational ego-motion estimation using log-polar motion energy, in: *Neural Networks*, 2006. *IJCNN '06. International Joint Conference*, 2006, pp. 5204–5211.
- [215] M. Bishay, R.A.P. II, K. Kawamura, Object detection in indoor scenes using log-polar mapping, in: *IEEE Intl. Conf. on Robotics and Automation, ICRA*, San Diego, CA, 1994, pp. 775–780.
- [216] J. Rett, J. Dias, Autonomous robot navigation—A study using optical flow and log-polar image representation, in: *Proc. of the Colloquium of Automation, Salzhausen* 2003/2004, 2004.
- [217] G. Barattoff, C. Toepfer, H. Neumann, Combined space-variant maps for optical-flow-based navigation, *Biological Cybernetics* 83 (2000) 199–209.
- [218] J.A. Boluda, J. Domingo, F. Pardo, J. Pelechano, Detecting Motion Independent of the Camera Movement Through a Log-Polar Differential Approach, in: *Lecture Notes in Computer Science*, vol. 1296, Springer-Verlag, 1997, pp. 702–710.
- [219] J.A. Boluda-Grau, *Arquitectura de Procesamiento de Imágenes Basada En Lógica Reconfigurable Para Navegación de Vehículos Autónomos con Visión Foveal*, Ph.D. Thesis, Departamento de Informática, Facultad de Física, Universitat de València, Spain, 2000.
- [220] J.A. Boluda, J. Domingo, F. Pardo, J. Pelechano, FPGA implementation of a log-polar motion detection algorithm, in: E. Pagello, et al. (Eds.), in: *Intelligent Autonomous Systems*, 6, IOS Press, Venice, Italy, 2000, pp. 851–858.
- [221] J.A. Boluda, J. Domingo, On the advantages of combining differential algorithms and log-polar vision for detection of self-motion from a mobile robot, *Robotics and Autonomous Systems* 37 (4) (2001) 283–296.
- [222] S. Gächter, T. Pajdla, B. Mičušik, Mirror design for an omnidirectional camera with a space variant imager, in: *Workshop on Omnidirectional Vision Applied to Robotic Orientation and Nondestructive Testing, ICAR* 2001, Aug. 2001.
- [223] G. Sandini, J. Santos-Victor, T. Pajdla, F. Berton, OMNIVIEWS: Direct omnidirectional imaging based on a retina-like sensor, in: *The First IEEE International Conference on Sensors, IEEE Sensors — 2002*, June 2002.
- [224] N. Kyriakoulis, A. Gasteratos, A. Amanatiadis, Image sequence stabilization using fuzzy kalman filtering and log-polar transformation, in: *Intl. Conf. on Computer Vision Theory and Applications*, vol. 1, Funchal, Madeira, Portugal, Jan. 2008.
- [225] B. Mertsching, S. Schmaiz, Systems and Applications, in: *Active Vision Systems*, vol. 3, Academic Press, 1999, pp. 197–219.
- [226] M. Tistarelli, G. Sandini, Estimation of depth from motion using an anthropomorphic visual sensor, *Image and Vision Computing (IVC)* 8 (4) (1990) 271–278.
- [227] N.C. Griswold, J.S. Lee, C.F.R. Weiman, Binocular fusion revisited utilizing a log-polar tessellation, *Computer Vision and Image Processing* (1992) 421–457.
- [228] I. Ahnrs, H. Neumann, Combining log-polar mapping and Gabor filtering for real-time depth estimation, in: C. Freksa (Ed.), *Proc. in Artificial Intelligence, Proc. Workshop Dynamische Perzeption*, June 1998, pp. 119–126.
- [229] E. Grosso, M. Tistarelli, Log-polar stereo for anthropomorphic robots, in: *European Conf. on Computer Vision*, Dublin, Ireland, July 2000, pp. 299–313.
- [230] A. Bernardino, J. Santos-Victor, A binocular stereo algorithm for log-polar foveated systems, in: *Proc. of the 2nd Workshop on Biologically Motivated Computer Vision*, Nov. 2002, pp. 127–136.
- [231] K. Schindler, H. Bischof, The epipolar geometry of the log-polar image plane, in: *Intl. Conf. on Pattern Recognition, ICPR*, Cambridge, UK, Aug. 2004, pp. 40–43.
- [232] M. Lopes, A. Bernardino, J. Santos-Victor, A developmental roadmap for task learning by imitation in humanoid robots: Baltazar's story, in: *AISB 2005 Symposium on Imitation in Animals and Artifacts*, Apr. 2005.
- [233] A. Gasteratos, R. Martinotti, G. Metta, G. Sandini, Precise 3D measurements with a high resolution stereo head, in: *First Intl. Workshop on Image and Signal Processing and Analysis, IWISPA'00*, Pula, Croatia, June 2000, pp. 171–176.
- [234] C. Weiman, Binocular stereo via log-polar retinas, *SPIE* 2488 (1995) 309–320.
- [235] R. Carpenter, *Eye Movements*, The Macmillan Press, 1991.
- [236] A. Bernardino, J. Santos-Victor, Visual behaviors for binocular tracking, *Robotics and Autonomous Systems* 25 (1998) 137–146.
- [237] A. Bernardino, J. Santos-Victor, Vergence control for robotic heads using log-polar images, in: *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems, IROS*, vol. 3, Osaka, Japan, Nov. 1996, pp. 1264–1271.
- [238] E. Grosso, R. Manzotti, R. Tiso, G. Sandini, A space-variant approach to oculomotor control, in: *Intl. Symposium on Computer Vision*, Coral Gables, Florida, Nov. 1995, pp. 509–514.
- [239] N. Oshiro, N. Maru, A. Nishikawa, F. Miyazaki, Binocular tracking using log polar mapping, in: *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems, IROS*, vol. 2, Osaka, Japan, Nov. 1996, pp. 791–798.
- [240] S. Rougeaux, N. Kita, Y. Kuniyoshi, F.C.S. Sakano, Binocular tracking based on virtual horopters, in: *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems, IROS*, vol. 3, 1994, pp. 2052–2057.
- [241] D. Coombs, C. Brown, Real-time binocular smooth pursuit, *International Journal of Computer Vision (IJCV)* 11 (2) (1993) 147–164.
- [242] R. Manzotti, A. Gasteratos, G. Metta, G. Sandini, Disparity estimation on log-polar images and vergence control, *Computer Vision and Image Understanding (CVIU)* 83 (2001) 97–117.
- [243] Z. Bien, W.-K. Song, Novel wheelchair-based robotic arm with visual servoing capability for human-robot interaction, in: *Workshop on Service Automation and Robotics*, Kowloon Tong, Hong Kong, June 2000, pp. 5–17.
- [244] W.K. Song, Intelligent Visual Servoing using Space Variant Vision with Application for Rehabilitation Robots, Ph.D. Thesis, KAIST, Korea, 2001.
- [245] B. Scassellati, A binocular, foveated active vision system. Technical Report MIT AI Memo 1628, MIT Artificial Intelligence Lab., Cambridge, MA, USA, Mar. 1998.
- [246] J. Liu, M. Chu, J.E. Reich, Multitarget tracking in distributed sensor networks, *IEEE Signal Processing Magazine* 24 (3) (2007) 36–46.
- [247] M. Turk, G. Robertson, Perceptual user interfaces (introduction), *Communications of the ACM* 43 (3) (2000) 32–34.
- [248] M. Porta, Vision-based user interfaces: Methods and applications, *International Journal of Human-Computer Studies* 57 (2002) 27–73.
- [249] T. Spexard, M. Hanheide, G. Sagerer, Human-oriented interaction with an anthropomorphic robot, *IEEE Transactions on Robotics* 23 (5) (2007) 852–862.



V. Javier Traver earned a B.Sc. degree in Computer Science from Technical University of Valencia (Valencia, Spain) in 1995, and a Ph.D. degree in Computer Engineering from Universitat Jaume I (Castellón, Spain) in 2002. Since 2000, he has been a full-time lecturer of Computer Engineering at Universitat Jaume I. His research includes foveal imaging, active vision, and image sequence analysis.



Alexandre Bernardino received the Ph.D. degree in Electrical and Computer Engineering in 2004 from Instituto Superior Técnico (IST). He is an Assistant Professor at IST and Researcher at the Institute for Systems and Robotics (ISR-Lisboa) in the Computer Vision Laboratory (VisLab). He participates in several national and international research projects in the fields of robotics, cognitive systems, computer vision and surveillance. He has published several articles in international journals and conferences, and his main research interests focus on the application of computer vision, cognitive science and control theory to advanced robotic and surveillance systems.