Active, Foveated, Uncalibrated Stereovision

**James P. Monaco · Alan C. Bovik ·
Lawrence K. Cormack**

Abstract Biological vision systems have inspired and will
continue to inspire the development of computer vision sys-
tems. One biological tendency that has never been exploited
is the symbiotic relationship between foveation and *uncal-
ibrated* active, binocular vision systems. The primary goal
of any binocular vision system is the correspondence of
the two retinal images. For calibrated binocular rigs the
search for corresponding points can be restricted to epipo-
lar lines. In an uncalibrated system the precise geometry is
unknown. However, the set of possible geometries can be
restricted to some reasonable range; and consequently, the
search for matching points can be confined to regions de-
lineated by the union of all possible epipolar lines over all
possible geometries. We call these regions *epipolar spaces*.
The accuracy and complexity of *any* correspondence algo-
rithm is directly proportional to the size of these epipolar
spaces. Consequently, the introduction of a spatially variant
foveation strategy that reduces the average area per epipo-
lar space is highly desirable. This paper provides a set of
sampling theorems that offer a path for designing foveation
strategies that are optimal with respect to average epipolar
area.

J.P. Monaco (✉) · A.C. Bovik
Department of Electrical and Computer Engineering,
University of Texas at Austin, Austin, TX 78712-1084, USA
e-mail: monaco@ece.rutgers.com

A.C. Bovik
e-mail: bovik@ece.rutgers.com

L.K. Cormack
Department of Psychology, University of Texas at Austin, Austin,
TX 78712-0187, USA
e-mail: cormack@psy.utexas.edu

1 Introduction

Stereoscopic vision systems image the environment from
two distinct vantage points and then correspond the retinal
images in a process called stereopsis. Once matching points
are determined, depth information can be recovered via tri-
angulation. Nature has, thus far, produced the most success-
ful binocular vision systems. Consequently, an understand-
ing and emulation of their advantageous properties are con-
stant goals of many scientists and engineers. The biologi-
cal visual system most often modeled by machines is the
human visual system (HVS). Three important, interrelated
attributes of the HVS which have been adopted to varying
degrees, are the following:

1. *Active vision*—The paradigm of active vision, i.e. the
   ability to dynamically pose the eyes, is of paramount
   importance in the HVS. Active vision systems possess
   a virtually unlimited field-of-view, while confining their
   resources to a specific region-of-interest. From a math-
   ematical standpoint, active systems are able to address
   in a well-posed manner tasks, such as the correspon-
   dence problem (Marr and Poggio 1976), that are ill-
   posed to the passive observer (Aloimonos et al. 1987).
   In general, the ability to dynamically interact with the
   environment has been shown to simplify many prob-
   lems in computer vision (Sandini and Tistarelli 1990;
   Ballard 1991).

2. *Foveation*—In perfect harmony with the active vision paradigm is the adoption of foveated retinas. Foveation allows the high-resolution analysis of a spatially concentrated region-of-interest, while simultaneously providing a low-resolution, wide field-of-view. The use of such foveation strategies as log-polar (Weiman 1995), reciprocal wedge transform (Li and Zhou 1999), foveate wavelet transform (Wei and Li 1998), and fish-eye (Shah and Aggarwal 1994) have been successfully applied to both monocular and binocular vision problems such as foveated compression (Lee et al. 2001; Wang and Bovik 2001; Geisler and Perry 1998), vergence (Manzotti et al. 2001), time-to-impact analysis (Tistarelli and Sandini 1991), and depth recovery of a scene (Chen and Deconinck 1994; Schindler and Bischof 2004). In these instances the sampling strategies were first investigated elsewhere and later applied to areas of stereo vision. In rare cases the foveation strategy has been specifically tailored for a stereo vision task. Both, Basu (1992) and Elnagar (1998) derived optimal sampling schemes with respect to the error discretization of depth measurements. Klarquist and Bovik (1998) designed a real-time foveated stereo technique that adapts to the specific geometry, always producing horizontal epipolar lines.

3. *Lack of calibration*—The HVS is known to be uncalibrated[1] (Stevenson and Schor 1997; Schreiber et al. 2001; Schreiber and Tweed 2003), not possessing precise knowledge of the ocular geometry. The emulation of this trait is far less prevalent in computer vision than the previous two. The advantages of epipolar geometry have driven the majority of the research toward calibrated systems. The disadvantages of such calibrated systems are numerous. Obtaining the calibration information necessary for the accurate estimation of epipolar geometry can be highly complex, involving motorized lens calibration, kinematic calibration, and head/eye calibration (Shih et al. 1998). Such calibrated systems require highly accurate, often expensive equipment. Furthermore, the calibration of active vision systems can be constantly perturbed due to motion, thermal changes, and numerous other unforeseen environmental factors, and thus require constant recalibration. Unarguably, some tasks, such as absolute depth recovery, demand exact calibration. Yet, for a substantial number of visual problems calibration is irrelevant (Hespanha et al. 1999). Evolution of the HVS has found that explicit knowledge of epipolar geometry is unnecessary and/or untenable; and yet, the HVS is able to perform a wider variety of depth related visual tasks with a greater degree of success than any mechanized system.

It is easy to see that the first two traits (active vision and foveation) are interrelated. Both facilitate the concept of a small, highly detailed region-of-interest intertwined with a large, less detailed periphery. The HVS is not the only biological vision system that has evolved toward this interdependence. In fact, all forward-looking animals with active binocular vision systems that have been shown to employ stereopsis have foveated retinas. Thus far this group includes the monkey, cat, falcon, and barn owl (Fite and Rosenfield-Wessels 1975; Rapaport and Stone 1984; Ptito et al. 1991; Fox et al. 1977; Bough 1970; van der Willigen et al. 1998; Wathey and Pettigrew 1989). To underscore the symbiotic nature of foveation and stereopsis, consider the following *afoveate* animals: goldfish (Easter 1972), rabbit (Davis 1929), hamster (Tiao and Blakemore 1976) and rat (Dräger and Olsen 1980). All possess laterally placed eyes with minimal binocular overlap and are not known to have stereoscopic vision. In place of a centralized foveal region, the retinas of rabbits, hamsters, and rats are known to have high-density visual streaks useful for detecting predators along the horizon.

The importance and emulation of the uncalibrated nature of the HVS have been largely absent in the computer vision literature. This is likely due to a couple of reasons. Most prominent, as mentioned above, is the desire to use epipolar constraints to simplify the correspondence problem. Secondly, the discovery that the HVS is uncalibrated is rather recent (and is still unknown for animals). In fact, some previous studies tended to suggest that it was calibrated (Nielsen and Poggio 1984).

In this paper we demonstrate that a symbiosis exists not only between foveation and binocular active vision systems, but more specifically between foveation and *uncalibrated* binocular active vision systems. Although this work is biologically inspired and may even offer biologically relevant conclusions, it is presented solely in the context of computer vision systems for two reasons: (1) the geometrical complexities of biological vision systems makes mathematical analysis exceedingly difficult, (2) we wish to demonstrate from an engineering perspective the inherent advantages of adopting foveated, uncalibrated binocular active vision systems. An outline of the arguments to be presented is as follows: the primary goal of any binocular vision system is the correspondence of the two retinal images. For calibrated binocular rigs the search for corresponding points can be restricted to epipolar lines. In an uncalibrated system the precise geometry is unknown. However, the set of possible geometries can be restricted to some reasonable range; and consequently, the search for matching points can be confined to regions delineated by the union of all possible epipolar lines over all possible geometries. We call these regions *epipolar spaces* (Monaco et al. 2007b). The accuracy and complexity of *any* correspondence algorithm is directly proportional to the size of these epipolar spaces. Consequently,

---

[1] Note that the term uncalibrated is meant to reflect a *complete* lack of calibration information and is not to be construed as lacking only metric calibration.

the introduction of a spatially variant foveation strategy that reduces the average area per epipolar space is highly desirable (Monaco et al. 2007a). This paper provides a set of sampling theorems that offer a path for constructing foveation strategies that are optimal with respect to average epipolar area.

The remainder of this paper proceeds with the following format: First we mathematically quantify the concept of an epipolar space. We then present a sequence of sampling theorems that form the foundation for the design of optimal foveation strategies. Next, we derive the optimal sampling scheme for a simple set of epipolar spaces and quantify the attendant reduction in average epipolar area. Finally, we discuss the impact of this work and how it provides insight into both biological and mechanical vision systems.

2 Epipolar Spaces

Knowledge of the camera geometry can be valuable in the registration process, reducing the search for matching points to epipolar lines (Horn 1986; Bovik 2000). Consider the stereo geometry in Fig. 1. Here two pinhole cameras whose optical centers are located at C_l and C_r converge at the fixation point V. Both cameras have planar imaging surfaces situated at identical focal lengths f from their respective pinholes. The left and right camera rotation angles are θ_l and θ_r. Each camera has an associated right-hand coordinate system originating at its optical center. If $\tilde{m}_l = [x_l, y_l, z_l, 1]^t$ and $\tilde{m}_r = [x_r, y_r, z_r, 1]^t$ are projective world coordinates in the left and right camera frames, respectively, then their relation is $\tilde{m}_r = [R \,|\, T] \tilde{m}_l$, where

$$R = \begin{bmatrix} -\cos(\theta_l + \theta_r) & 0 & -\sin(\theta_l + \theta_r) \\ 0 & 1 & 0 \\ \sin(\theta_l + \theta_r) & 0 & -\cos(\theta_l + \theta_r) \end{bmatrix} \quad (1)$$
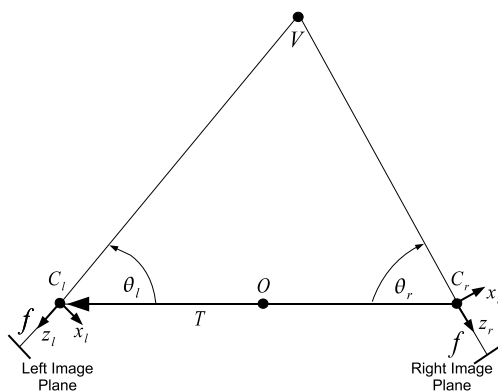

Fig. 1 Stereo geometry. *Left* and *right* cameras with optical centers C_l and C_r converge at fixation point V. Both cameras have focal lengths f. The *left* and *right* camera rotation angles are θ_l and θ_r

is the rotation matrix and

$$T = |T| \begin{bmatrix} -\sin(\theta_r) \\ 0 \\ -\cos(\theta_r) \end{bmatrix} \quad (2)$$

is the translation vector of length $|T|$. Let $m_l = [u_l, v_l, 1]^t$ and $m_r = [u_r, v_r, 1]^t$ be projective coordinates in the left and right image planes. It is well known that the equation relating corresponding epipolar lines is

$$m_r^t A_r^{-t} T_\times R A_l^{-1} m_l = 0, \quad (3)$$

where, in our simplified geometry,

$$A_l = A_r = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

are the identical intrinsic matrices and

$$T_\times = |T| \begin{bmatrix} 0 & \cos(\theta_r) & 0 \\ -\cos(\theta_r) & 1 & \sin(\theta_r) \\ 0 & -\sin(\theta_r) & 0 \end{bmatrix} \quad (5)$$

implements the crossproduct as a matrix. The matrix $E = T_\times R$ is the essential matrix and relates the coordinate frames. The matrix $F = A_r^{-t} E A_l^{-1}$ is the fundamental matrix (Faugeras et al. 1992) and includes the intrinsic parameters of the cameras. The expression in (3) can be simplified to the following:

$$v_l = v_r \frac{f \sin(\theta_l) + u_l \cos \theta_l}{f \sin(\theta_r) - u_r \cos \theta_r}. \quad (6)$$

That is, for a given point (u_r, v_r) in the right image, the matching point in the left image (if not obscured) lies on the line given by (6). It is worth mentioning that the length $|T|$ of the translation vector has no effect on the epipolar equations.

We now consider the situation where the camera configuration actively changes and we no longer know the specific geometry. In such a situation we will not be able to restrict our search for corresponding points to epipolar lines. However, even though we may not know the precise values of parameters such as focal length, baseline distance, and camera rotation angles, we can establish acceptable ranges for these values. Consequently, we can still restrict the location of corresponding points across images. For a given point in one image, the matching point in the other is confined to a region defined by the union of all corresponding epipolar lines produced over all possible camera configurations. We call these continuous regions *epipolar spaces*.

The goal of the remainder of this section is to quantify these epipolar spaces for a stereo rig with a fixed baseline and fixed focal length as shown in Fig. 1. In this configuration the only variable parameters that effect the epipolar

geometry are the rotation angles θ_l and θ_r. Translation and rotation of the entire stereo rig about O, while allowed, do not influence the epipolar geometry. We establish the range of rotation angles by confining them to the interval

$$\theta_l, \theta_r \in [\theta_M, \pi - \theta_M], \tag{7}$$

where θ_M is the minimum angle relative to the baseline.

Although theoretically a matching point can lie anywhere on the corresponding epipolar lines, the search is usually restricted to a maximum horizontal disparity. Bounding the horizontal disparity d has the effect of limiting the depth around the horopter at which objects can be fused. For our purposes, we assume a maximum horizontal disparity defined by

$$|d| = |u_l - u_r| \le D. \tag{8}$$

For the right image point I_r Fig. 2(a) illustrates several corresponding epipolar lines in the left image. Each separate epipolar line results from a unique geometric configuration, i.e. a unique combination of θ_l and θ_r. The dashed vertical lines delimit the maximum allowable horizontal disparities defined by (8).

The restriction imposed by (8) determines the leftmost and rightmost bounds of the epipolar spaces. The upper and lower bounds are determined by maximizing and minimizing (6) with respect to both θ_l and θ_r, respectively. The maximum can be found by separately maximizing the numerator and minimizing the denominator. Taking the derivative of the numerator with respect to θ_l, setting it to zero, and solving for θ_l yields the maximizing value $\theta_l = \operatorname{atan}(f/u_l)$. The denominator (which is positive so long as the cameras never image each other, i.e. the images do not contain the epipoles) is minimized when $\theta_r = \theta_M$. Inserting these results into (6), the upper bound becomes

$$v_{l,\max} = v_r \frac{\sqrt{f^2 + u_l^2}}{f \sin(\theta_M) - u_r \cos(\theta_M)}. \tag{9}$$

The minimization of (6) is performed similarly with $\theta_l = \pi - \theta_M$ and $\theta_r = \operatorname{atan}(-f/u_r)$, producing

$$v_{l,\min} = v_r \frac{f \sin(\theta_M) - u_l \cos(\theta_M)}{\sqrt{f^2 + u_r^2}}. \tag{10}$$

Since the horizontal extents are limited by (8), u_l in the preceding equations can be rewritten as $u_l = u_r + \epsilon$, where $\epsilon \in [-D, D]$. In practice $f \gg \epsilon$, allowing the following approximation: $u_l \approx u_r$. Incorporating this result into (9) and (10) produces

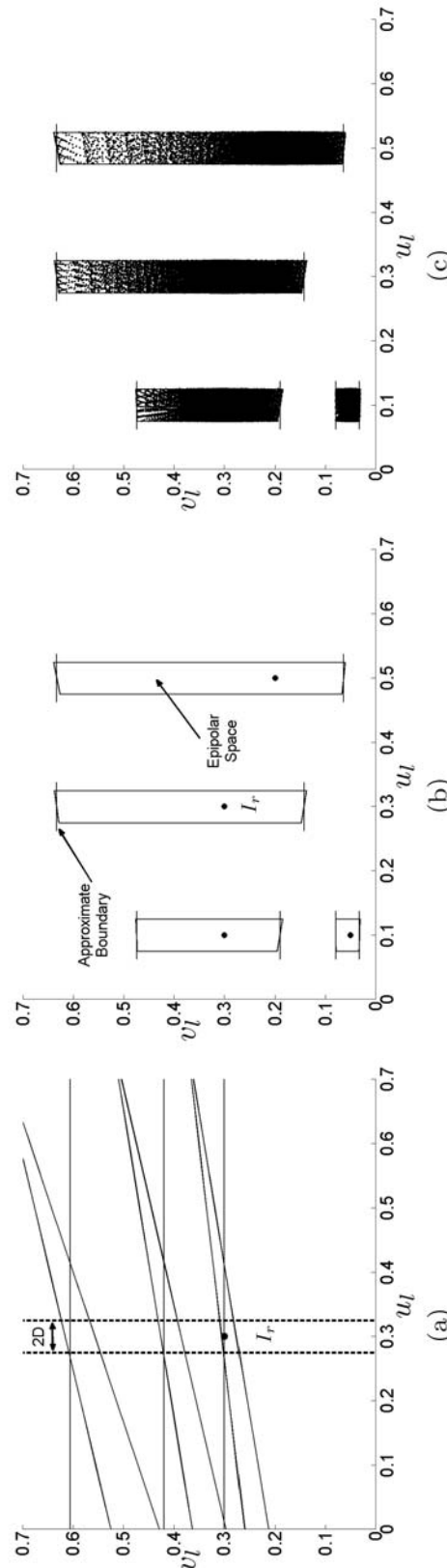$$v_{l,\max} \approx v_r \frac{\sqrt{f^2 + u_r^2}}{f \sin(\theta_M) - u_r \cos(\theta_M)} = v_r c(u_r), \tag{11}$$

Fig. 2 Epipolar Spaces. (**a**) *Left image* epipolar lines that correspond to *the right image* point I_r for several unique geometric configurations ($\theta_r, \theta_l \in \{\pi/4, \pi/3, \pi/2\}$ and $f = 1$). (**b**) *The large dots* represent points from *the right image*. Each point is enclosed in its attendant epipolar space in *the left image* ($f = 1$ and $\theta_M = \pi/4$). *The thick lines* represent the precise boundaries described by (8), (9), and (10). *The thin lines* denote the approximate upper and lower bounds determined from (11) and (12). (**c**) Each point I_r shown in (**b**) represents the projection of one of many possible 3D points whose locations depend upon the camera configuration. For all possible configurations we project a sampling of these 3D points onto *the left image* plane. Only those projections satisfying (8) are shown. As expected, they lie within the epipolar spaces

$$v_{l,\min} \approx v_r \frac{f \sin(\theta_M) - u_r \cos(\theta_M)}{\sqrt{f^2 + u_r^2}} = \frac{v_r}{c(u_r)}, \qquad (12)$$

where

$$c(u_r) = \frac{\sqrt{f^2 + u_r^2}}{f \sin(\theta_M) - u_r \cos(\theta_M)}. \qquad (13)$$

Remarkably, an epipolar space is well modeled by a rectangle. This fact is illustrated in Fig. 2(b). Each large dot represents a coordinate in the right image plane; they are enclosed in their corresponding epipolar spaces in the left image plane. The thick lines represent the precise boundaries of the regions described by (8), (9), and (10). The thin lines denote the approximate upper and lower bounds determined from (11) and (12). Epipolar spaces are nonuniform in area, increasing in size with increasing values of u and v.

For further clarification we recapitulate the concept of epipolar spaces. Consider a point I_r in the right image. Though many possible 3D points can project to this coordinate, they are constrained to the 3D line connecting I_r and the camera pinhole C_r. This line changes as the right camera rotates. The projection of the 3D line onto the left image plane defines the coordinates in the left image that may correspond with I_r. This projection depends upon the rotation angle of the left camera. We provide the following simulation: for each possible geometric camera configuration (i.e. left and right camera rotation angles) we uniformly sample the 3D line containing I_r and C_r and then project the samples onto the left image plane. Those projected points whose horizontal disparities satisfy (8) are plotted in Fig. 2(c). This simulation is performed for the four right image points I_r shown in Fig. 2(b). As expected the projected points fall within their associated epipolar spaces.

3 Optimal Sampling

The complexity and accuracy of any stereopsis algorithm, whether feature-based (Dhond and Aggarwal 1989) or phase-based (Sanger 1988; Jenkin et al. 1991; Monaco et al. 2008), are directly proportional to the average size of the regions that must be searched to locate matching points. For uncalibrated active vision systems these regions are epipolar spaces. The goal of this section is to discuss a sequence of theorems (presented in detail in the Appendix) that will form the foundation for creating optimal sampling schemes. An optimal sampling scheme is defined as the foveation strategy that minimizes the average number of points per epipolar space. Though sampling is inherently a discrete process, it can be modeled continuously. Working in a continuous domain simplifies the analysis, allowing the use of powerful mathematical tools that are either unavailable or extremely cumbersome in a discrete framework.

This remainder of this section proceeds as follows: in Sect. 3.1 we recapitulate the concept of an optimal sampling strategy and provide intuitive examples. Section 3.2 considers the optimal sampling strategies for a one-dimensional image interval. Analysis in one-dimension simplifies the mathematics and provides a sound framework for the two-dimensional adaptation. Section 3.3 extends the optimal one-dimensional sampling strategies to two dimensions. Note that the following derivations are not restricted to the specific epipolar spaces described by (8), (11), and (12), but instead, are generalized to epipolar spaces of a variety of shapes and sizes.

3.1 Preliminaries

It is important to mention two points that are crucial in understanding this paper: (1) both image planes are assumed to have identical foveation patterns and (2) the pattern of epipolar spaces is symmetric about both axes in the image plane. The inherent symmetry of the stereo vision problem seems to require that these properties be satisfied. The practical effect of these conditions is to allow us to constrain our analysis and illustrations to a single image plane. For example, consider again the plots presented Fig. 2. In all three illustrations the points to be matched are from the right image plane, and the possible matching points are from the left. Even if the roles of the two image planes were reversed the plots would remain the same. Though we will continually return to the matching problem between image planes, the design of an optimal foveation strategy can be conceptually restricted to a single image. This will become clear with the following example.

Fig. 3(a) represents a uniform sampling scheme applied to both the left and right image planes. To recover depth information each sample in one image must be paired with its matching sample in the other. This is called the correspondence problem (Marr and Poggio 1979). If we assume an uncalibrated stereovision system as defined in Sect. 2 then for each point in the right image we can confine the search in the left to the appropriate epipolar space. For example, consider the three white samples in Fig. 3(a) which represent points in the right image for which we want to locate matches in the left (or equivalently, points in the left image for which want to locate matches in the right). Each point is enclosed by its associated epipolar space. Notice how the center epipolar space only contains a single point. Obviously, the process of identifying its corresponding point is both fast and accurate. By contrast, the upper-left epipolar space is much larger, containing 15 candidate matches (i.e. possible matching samples). We need to consider on average $(15+1+12)/3 \approx 9.3$ samples when attempting to match the three points.

Next consider the alternative sampling strategy in Fig. 3(b). Notice that the number of samples stays fixed.

Fig. 3 (**a**) Uniform sampling of *left and right image* planes. (**b**) Nonuniform sampling of *left and right image* planes. *White dots* and accompanying *rectangles in both figures* indicate sample points in *the right/left image* and their associated epipolar spaces in *the left/right*
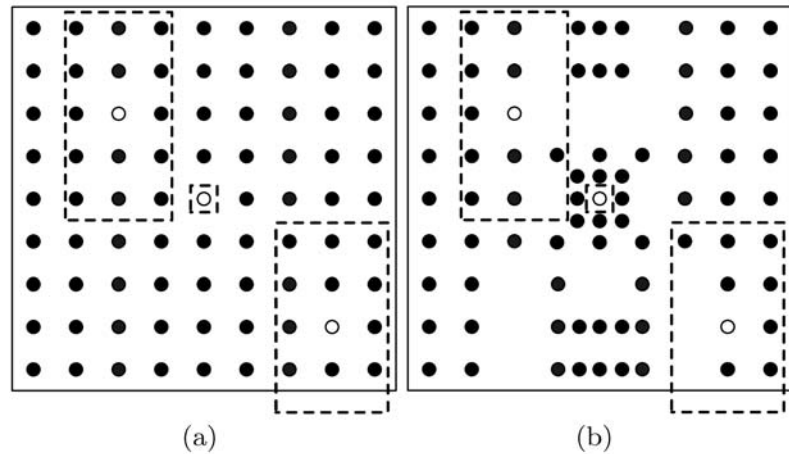
This foveation pattern reduces the average number of candidate points to $(11 + 1 + 9)/3 = 7$. Consequently, by varying the sampling scheme we have reduced the average number of points that must be searched in order to correspond the samples. Though this average considers only three points and should be extended to the remainder, it nevertheless becomes clear that we can simplify the correspondence problem by adapting the sampling strategy.

We define the optimal sampling strategy as the arrangement that minimizes the average number of samples per epipolar space with the caveat that the number of total samples remains constant. Requiring a fixed number of samples prevents the trivial solution realized by removing all the samples. Unfortunately, determining the optimal sampling scheme in a discrete framework is difficult. We can, however, arrive at a solution when modeling the problem continuously. Though we will subsequently discuss the specifics of the continuous reformulation, it is helpful to now provide continuous counterparts for several discrete concepts: average number of samples per epipolar equates to average area per epipolar space, a discrete sampling strategy becomes a continuous function, and the condition requiring a fixed number of samples translates to a condition requiring the preservation of total area.

Before proceeding to the optimal sampling functions we consider the occurrence of clipping. Invariably, portions of some epipolar spaces, such as the bottom-right epipolar space in Figs. 3(a) and 3(b), will extend outside any finite sized imaging surface. Obviously, we would never search for corresponding points in these regions; and consequently, they are clipped from the epipolar spaces. Unfortunately, clipping is a nonlinear operation and is difficult to model mathematically. Therefore, we disregard its effects when deriving the optimal sampling functions. This results in slightly suboptimal sampling strategies in localized regions near the border of the imaging surface. The effects of clipping will be illustrated in later sections.
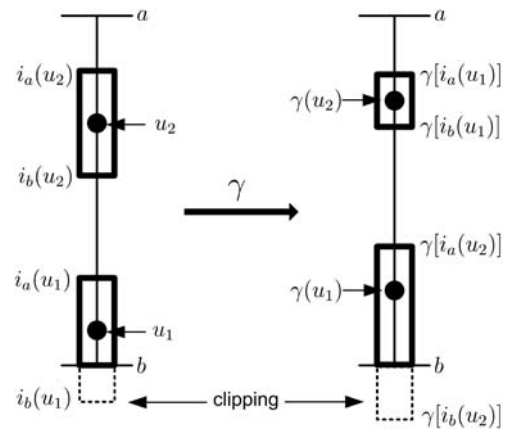
Fig. 4 *This figure* shows two intervals that range from *a* to *b*. *The dots and rectangles in the leftmost interval* illustrate two samples and their associated epipolar intervals. *The dots and rectangles in the rightmost interval* illustrate these samples and epipolar intervals after the application of the sampling function γ. *The dotted areas* indicate the portions of the epipolar spaces that would extend outside the interval if they were not clipped at the boundaries

3.2 One-Dimensional Foveation

Let the interval $I = [a, b]$ represent a one-dimensional image plane (interval). Please refer to the leftmost interval in Fig. 4 for an illustration. For each point u in I the matching point must lie within the corresponding epipolar interval. The upper and lower extents of these epipolar intervals can be expressed as functions of u. Let $i(u)$ represent the epipolar interval associated with the point u, and let $i_b(u)$ and $i_a(u)$ indicate the upper and lower bounds of $i(u)$, respectively. The goal of an optimal sampling strategy is to place N points in interval I in such a fashion as to minimize the average number of candidate matches for each of the N points. A candidate match for point u_j is defined as any point u_k that lies within the epipolar interval $i(u_j)$.

A discrete sampling strategy can be modeled as a continuous sampling function (Zeevi and Shlomot 1993).

A sampling function is defined as any increasing, invertible function $\gamma(u)$. For a given sampling strategy $\gamma(u)$ and set of epipolar intervals $i(u)$, the average epipolar length $E(\gamma; i, I)$ over I can be expressed as

$$E(\gamma; i, I) = \frac{1}{b-a} \int_a^b e[\gamma; i(u)]\, \gamma'(u)\, du, \qquad (14)$$

where

$$e[\gamma; i(u)] = \int_{i_a(u)}^{i_b(u)} \gamma'(\tilde{u})\, d\tilde{u} \qquad (15)$$

is the length of the epipolar interval associated with the point u after transformation. The optimal sampling function is the function $\gamma(u)$ that minimizes (14) subject to the constraint of length preservation:

$$\Lambda(\gamma; I) = \int_a^b \gamma'(u)\, du = b - a. \qquad (16)$$

This length preservation constraint is the continuous analog to the discrete requirement that the number of samples remain fixed. Without this stipulation the solution devolves to the trivial case of $\gamma \equiv 0$.

Consider the leftmost interval in Fig. 4. The dots and their associated rectangles represent the points u_1 and u_2 and their corresponding epipolar intervals. The transformation γ maps this interval onto the rightmost interval, warping the points and epipolar intervals accordingly. Notice that the epipolar intervals of u_1 and $\gamma(u_1)$ were clipped (denoted by the line) so that they would not extent outside $[a, b]$.

In Appendix A we show that the optimal transformation is the function that warps the epipolar intervals into intervals of identical length. This implies that a discrete sampling strategy should arranged the samples so that each epipolar interval contains the same number. That is, the sampling density should be proportional to the epipolar length: smaller epipolar spaces are sampled more finely than larger ones.

We conclude this subsection with an analogy to lend intuition to the previous mathematical abstractions. Imagine we stretch a rubber band (one-dimensional image I) taut and secure its ends with two nails. We then draw along it lengthwise many overlapping intervals (epipolar intervals $i(u)$) of different size. We sum the interval lengths and determine the average (average epipolar length). Next we take the middle of the rubber band, pull (warping function γ) it—along the line of the rubber band—closer to one of the end points, and use a third nail to hold this deformation in place. Assume that the rubber band remains taut and collinear. Though the total length of the stretched rubber band remains the same, the intervals will have expanded or contracted depending on which side of the third nail they lie. If we were to again determine the average interval length it would almost assuredly

be different. By effecting many such deformations in combination we can find the total deformation that minimizes the average length of the intervals (optimal sampling strategy).

3.3 Two-Dimensional Foveation

Consider any two-dimensional, connected region R. For reference, please consult Fig. 5. For each point $\boldsymbol{u}$ in R the matching point must lie within the corresponding epipolar space $r(\boldsymbol{u})$. The goal of a two-dimensional optimal sampling strategy is to place N points in the region R in such a fashion as to minimize the average number of candidate matches for each of the N points. A candidate match for point $\boldsymbol{u}_j$ is defined as any point $\boldsymbol{u}_k$ that lies within the epipolar space $r(\boldsymbol{u}_j)$.

A two-dimensional continuous sampling function is defined as any invertible function $\boldsymbol{\gamma}(\boldsymbol{u}) \doteq [\gamma_u(\boldsymbol{u}), \gamma_v(\boldsymbol{u})]^t$ that maps $\mathbb{R}^2 \to \mathbb{R}^2$ and has a Jacobian matrix whose determinant $J_{\boldsymbol{\gamma}}(\boldsymbol{u}) = |\boldsymbol{\gamma}'(\boldsymbol{u})|$ is positive everywhere. For a given sampling function $\boldsymbol{\gamma}(\boldsymbol{u})$ the average area of the epipolar spaces $r(\boldsymbol{u})$ over the region R is defined as

$$E(\boldsymbol{\gamma}; r, R) = \frac{1}{A_R} \iint_R e[\boldsymbol{\gamma}; r(\boldsymbol{u})]\, J_{\boldsymbol{\gamma}}(\boldsymbol{u})\, d\boldsymbol{u}, \qquad (17)$$

where

$$e[\boldsymbol{\gamma}; r(\boldsymbol{u})] = \iint_{r(u)} J_{\boldsymbol{\gamma}}(\tilde{\boldsymbol{u}})\, d\tilde{\boldsymbol{u}} \qquad (18)$$

is the area of the epipolar space associated with the point $\boldsymbol{u}$ after transformation and $A_R = \iint_R d\boldsymbol{u}$. An optimal sampling scheme is a function $\boldsymbol{\gamma}(\boldsymbol{u})$ that minimizes $E(\boldsymbol{\gamma}; r, R)$ subject to the constraint of area preservation:

$$\Lambda(\boldsymbol{\gamma}; R) = \iint_R J_{\boldsymbol{\gamma}}(\boldsymbol{u})\, d\boldsymbol{u} = \iint_R d\boldsymbol{u} = A_R. \qquad (19)$$

In Appendix B we show that the optimal transformation maps the epipolar spaces into regions of uniform area
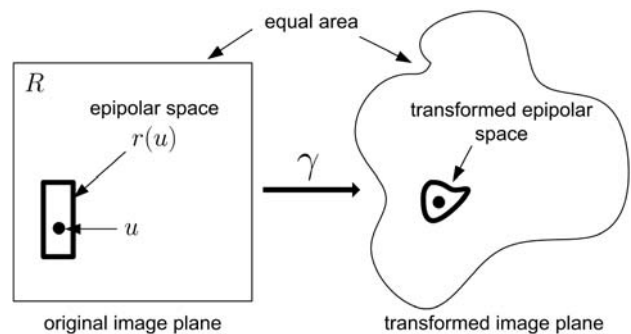

Fig. 5 *Leftmost drawing* illustrates the imaging surface R, sample u, and its epipolar space $r(u)$. *The rightmost image* illustrates the imaging surface, sample, and epipolar space after the application of the sampling function $\boldsymbol{\gamma}$

but with *arbitrary* shape. Consequently, there are an infinite number of optimal transformations. This result is expected given the form of (17), which does not depend on the specific transformation γ, but only the determinant of its Jacobian. This determinant reflects the instantaneous rate of change in area much as the derivative in one dimension indicates the instantaneous rate of change in length (Strang 1988, p. 212). These results imply that a discrete sampling strategy should arrange the samples so that an equal number appear in every epipolar space. Obviously, there are numerous arraignments that can accomplish this. For example, consider the case where the epipolar spaces are already uniform in area. Accordingly, the optimal strategy is uniform sampling. However, this uniform sampling can come in many forms. For example, rectangular uniform sampling whose sampling function is $\gamma(u) = [u\ v]^t$ has a Jacobian with a determinant of one. Yet, so does hexagonal uniform sampling:

$$\gamma(u) = \sqrt{2/3} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}. \tag{20}$$

We were unable to contrive a two-dimensional analogy similar to that of the rubber band. The closest approximation is an elastic sheet littered with overlapping rectangles. The size and shapes of the rectangles can be deformed by internal stretching. However, it is not clear that the possible deformations are as extensive as those realizable by γ.

4 Optimal Sampling Applications

This section uses the results from previous sections to formulate the optimal sampling schemes for the epipolar spaces defined in (8), (11), and (12). First, the optimal scheme is derived independently for each dimension. Then these solutions are integrated to create the optimal two-dimensional sampling scheme.

4.1 One-Dimensional

The epipolar spaces delineated in (8), (11), and (12) are identical in both images and symmetric across image quadrants, therefore we need only consider the positive coordinates of a single image plane. For the point (u, v), the corresponding bounds for a one-dimensional epipolar interval $i_1(u)$ along the horizontal u dimension are given by (8). These boundaries form intervals of uniform length $2D$ (ignoring clipping). As shown in Appendix A, the optimal sampling scheme over the interval $[a, b]$ for epipolar intervals of uniform length is uniform sampling:

$$\gamma_1(u) = u. \tag{21}$$

The average epipolar length for these uniform epipolar intervals under uniform sampling is

$$E(\gamma_1; i_1, I) = 2D - \frac{D^2}{b-a}. \tag{22}$$

This converges to $2D$ as $b - a$ becomes large with respect to D, i.e. clipping becomes negligible.

In the vertical v dimension the boundaries of the epipolar interval $i_2(v)$ are given by (11) and (12). The resulting intervals $[v/c(u), vc(u)]$ are nonuniform in length. The function $\gamma(v)$ that minimizes their average length on I must warp them into intervals of uniform length and satisfy (16). This optimal function is

$$\gamma_2(v) = \beta \ln v, \tag{23}$$

where $\beta = (b-a)/(\ln b - \ln a)$. The average epipolar length under this logarithmic transformation is

$$E(\gamma_2; i_2, I) = 2\beta \ln c - \frac{[\beta \ln c]^2}{b-a}. \tag{24}$$

Since (23) has a singularity at zero, the interval $[a, b]$ must not contain zero.

If instead the interval were sampled uniformly, the average epipolar length would be

$$E(\gamma_1; i_2, I) = \frac{a^2(1-c) + b^2(1-1/c)}{b-a}. \tag{25}$$

Fig. 6 plots the ratio $E(\gamma_1; i_2, I)/E(\gamma_2; i_2, I)$. Since the value of the variable c given in (13) is a function of u, this ratio is plotted over a range of positive u values. For small values of u, the logarithmic sampling function $\gamma_2(v)$ produces epipolar lengths that are, on average, less than half the size of those produced by the uniform sampling function $\gamma_1(v)$.

For additional insight, consider a third sampling method that, like logarithmic sampling, also tends to concentrate more samples near a:

$$\gamma_3(v) = -abv^{-1}, \tag{26}$$

The commensurate average epipolar length is

$$E(\gamma_3; i_2, I) = \frac{a^2(1-c) + b^2(1-1/c)}{b-a}. \tag{27}$$

Interestingly, this is identical to the average length for uniform sampling.

4.2 Two-Dimensional

In the previous subsection each dimension was optimized independently. Unfortunately, the vertical bounds of the epipolar spaces described in (11) and (12) depend on both
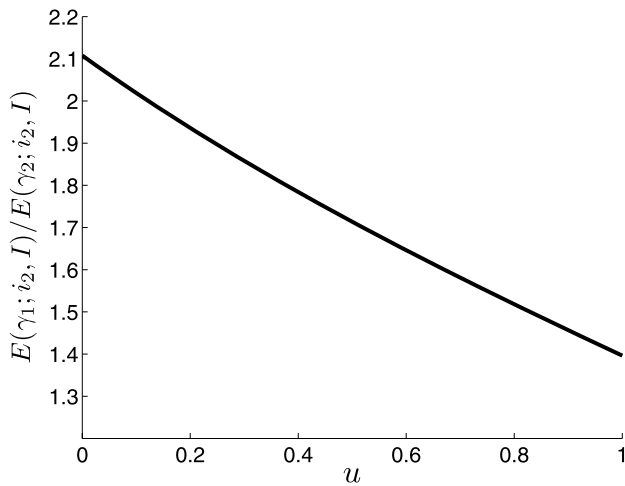
Fig. 6 Ratio of the average lengths $E(\gamma_1; i_2, I)/E(\gamma_2; i_2, I)$ of the epipolar intervals delineated in (11) and (12) under uniform and logarithmic sampling. The vertical interval extends from $a = 0.017$ to $b = 1$. Since the value of the variable $c(u)$ described in (13) is a function of u, the ratio is plotted over a range of positive u values. The determination of $c(u)$ uses the following parameters: $\theta_M = \pi/3$ and $f = 1$

Fig. 7 Area per epipolar space. Each area is normalized by the total area of the image plane. The normalized area remains constant over the majority of the imaging surface. The degradation seen in the upper right corner is caused by clipping

Fig. 8 Epipolar spaces from *Fig. 2(b)* after optimal sampling. The epipolar spaces in *Fig. 2(b)* were warped using the transformation in (28) with $\beta_u = 1$. They are almost perfectly uniform in area

u and v: they are nonseparable. To produce regions of equal area (and, therefore, minimize the average epipolar area) we must compensate for this dependence by normalizing (23). The following function produces spaces of approximately equal area:

$$\boldsymbol{\gamma}(u, v) = \left[\beta_u u, \frac{\beta_v \ln v}{\ln c(u)}\right]^t, \tag{28}$$

where β_u and β_v are constants chosen such that their product satisfies the constraint in (19). The following calculation of the area of an arbitrary epipolar space $r(\boldsymbol{u})$ (that is not clipped) after transformation confirms they have constant area:

$$e\left[\boldsymbol{\gamma}; r(\boldsymbol{u})\right] = \int_{u-D}^{u+D} \int_{v/c(u)}^{vc(u)} \frac{\beta_u \beta_v}{v \ln[c(\tilde{u})]} d\tilde{u} \, d\tilde{v}$$
$$= \int_{u-D}^{u+D} 2\beta_u \beta_v \frac{\ln c(u)}{\ln c(\tilde{u})} d\tilde{u}$$
$$\approx 4D\beta_u \beta_v. \tag{29}$$

This follows from the fact that for $\tilde{u} \in [u - D, u + D]$ the function $\ln c_v(u)/\ln c_v(\tilde{u})$ is approximately linear with a mean value of one. Figure 7 illustrates the average area of each epipolar space following transformation. As expected this average remains constant over the majority of the imaging surface. The degradation seen in the upper right corner is caused by clipping. The epipolar spaces in this region (before they are clipped) have relatively large areas extending outside the image boundary.
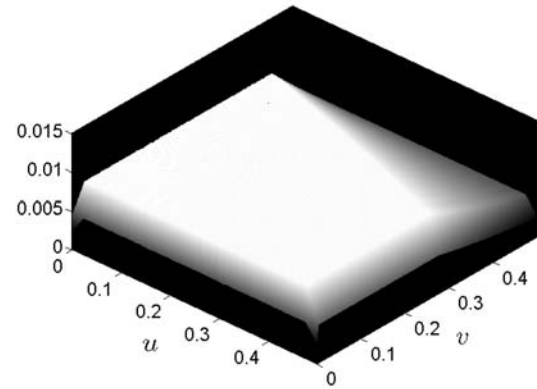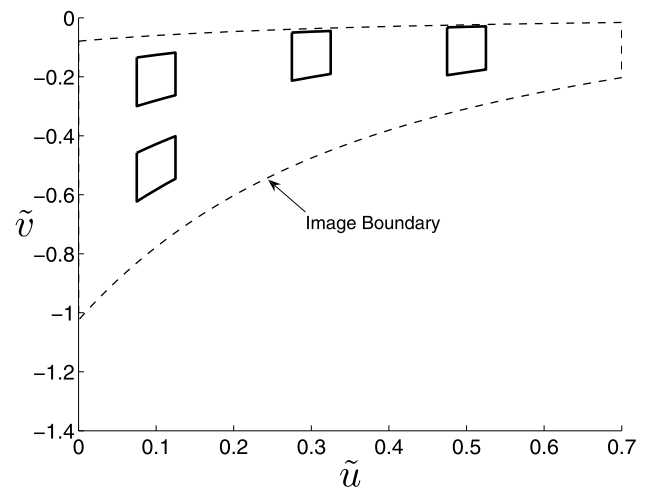
In Sect. 2 the epipolar spaces described in (8), (11), and (12) were depicted as Fig. 2(b). The transformations of these spaces using (28) are shown in Fig. 8. Though the warped boundaries are not identically shaped, they are almost perfectly uniform in area.

As mentioned previously, the constants β_u and β_v in (28) are only restricted in the sense that their product must satisfy (19). This is a consequence of the fact that the average epipolar area is a function of the determinant of the Jacobian of $\boldsymbol{\gamma}(\boldsymbol{u})$, and not $\boldsymbol{\gamma}(\boldsymbol{u})$ itself. That is, an infinite number of functions can have Jacobian matrices with the identical determinant. Figure 9 illustrates the optimal sampling strategies when β_u is set to two different, arbitrarily chosen values: $\beta_u = 1$ and $\beta_u = 4$.

Its is also instructive to view the sampling in the form of a tessellation. Consider the uniform grid in Fig. 10(a). This
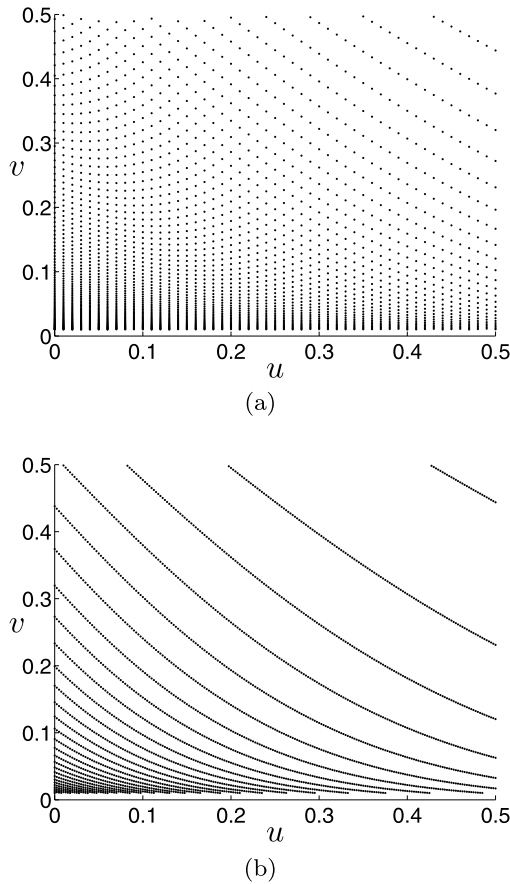
(a)

(b)

Fig. 9 Example of optimal sampling scheme defined in (28) for two different values of β_u. Given β_u, the value of β_v is determined by (19). Additional parameters were also assigned as follows: $f = 1$, $\theta_M = \pi/3$, $a = 0.001$, and $b = 0.5$. (**a**) Optimal sampling scheme with $\beta_u = 1$. (**b**) Optimal sampling scheme with $\beta_u = 4$

(a)

(b)

Fig. 10 Tessellations for optimal epipolar sampling. (**a**) Warped image space resulting from the application of the optimal transformation in (28) to a square image plane. (**b**) Projection of the uniform grid in (**a**) back into the original image plane

grid represents the uniform sampling of the warped image space resulting from the application of the optimal transformation in (28) to a square image plane. Figure 10(b) illustrates the tessellation resulting from projecting the uniform grid in Fig. 10(a) back into the original square image plane.

It may be beneficial to put these results into context by reviewing the meaning of an epipolar space, its importance in stereo registration, and its connection to the optimal sampling strategies. If a point in space projects onto the left image plane, its projection onto the right image plane is restricted to a region called an epipolar space. From a point of view of a stereo registration algorithm, only the pixels in the epipolar space must be examined to identify the matching point. The complexity of this search is proportional to the number of pixels in the epipolar space. Consider the epipolar spaces shown in Fig. 2(b). If pixels were to uniformly cover this space, then the larger epipolar spaces would require a greater search time than the smaller spaces. However, the time to search a particular epipolar space is not
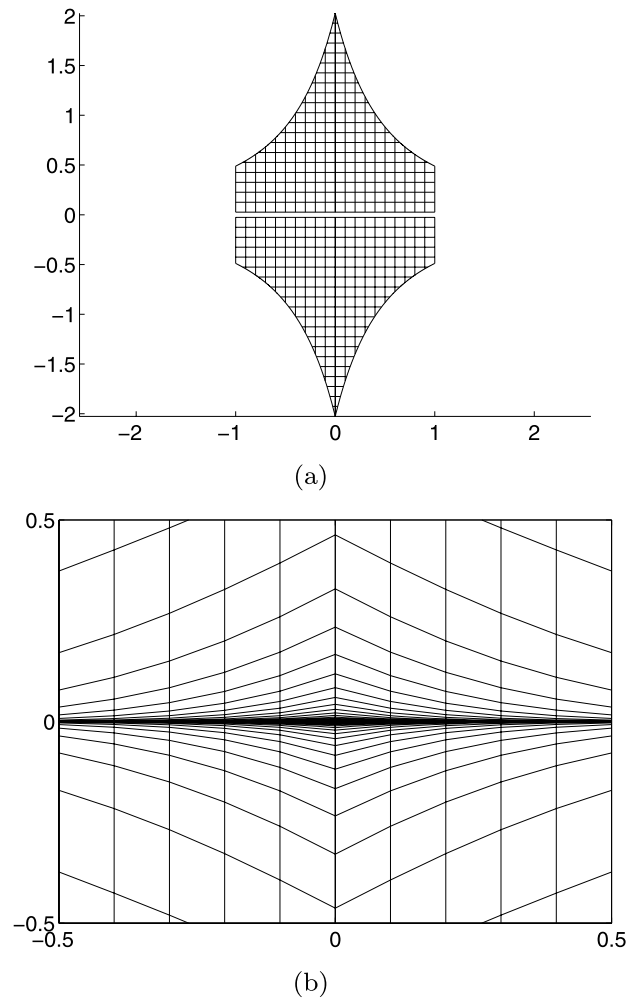
as important as the time it takes to search them all, which is what must be done for dense stereo matching. Now consider if the epipolar spaces in Fig. 2(b) were placed on a sampling grid like those depicted Fig. 9. Were we to count the number of samples appearing in each epipolar space, we would find that they were identical (except in the clipped epipolar spaces); and this number would be the same for both sampling patterns in Fig. 9. More importantly, if a fixed number of samples were placed within the defined image boundary and then the average number of points per epipolar space were computed, no placement strategy would produce an average less than that of those shown in Fig. 9. As an example, Table 1 provides a quantitative comparison of the average number of samples per epipolar space for both uniform sampling and the optimal sampling strategy described in (28). The data were generated for various rectangular image sizes and maximum rotation angles θ_M.

Table 1 Ratio of samples per epipolar space for uniform and optimal sampling. Equivalently, this table presents the ratio of the average area of an epipolar space before and after transformation by the sampling function posed in (28). The leftmost, rightmost, and upper bounds for the rectangular image are $a_u = 0$, $b_u = 0.5$, and $b_v = 0.5$, respectively. The lower bound a_v is specified in the table

a_v	θ_M	$E(\gamma_1; r, R)/E(\gamma_2; r, R)$
0.1	$\pi/4$	1.74
0.1	$\pi/3$	2.06
0.01	$\pi/4$	2.58
0.01	$\pi/3$	3.12
0.001	$\pi/4$	3.47
0.001	$\pi/3$	4.22

Fig. 11 Optimal tessellation using four-degrees-of-freedom model of HVS and log-polar foveation employed by HVS

5 Discussion

Consistent with our own goals, evolution seems to place a premium on uncovering the easiest way to achieve a goal while incurring the least possible expense computationally. Such reasoning may help explain the introduction of foveated retinas into the visual systems of some animals with forward-looking, dynamically configurable eyes. In this work we have demonstrated that the application of such a spatially variant sampling scheme can reduce the average area of a search region by a factor of 3–4 over typical uniform sampling. As we discuss next, were the range of configurations more similar to those of the HVS we would find this reduction factor to be even greater.

To ease mathematical calculations and simplify explanations, the epipolar spaces as delineated in Sect. 2 assumed a constant range of horizontal disparities. As stated previously, the horizontal disparity limits determine the depth about the horopter in which objects can be fused. A constant range corresponds to a constant depth. It can easily be argued that fusible depths should increase with distance from the vergence point. In fact, psychophysical investigations seem to suggest that such a tack is taken by the HVS. In Blakemore (1970) Blakemore measures the expected horizontal disparity limits of qualitative stereopsis (this is slightly different, and more appropriate, than Panum's fusional area (Howard and Rogers 1995; Yeshurun and Schwartz 1999)) as a function of horizontal eccentricity. Recomputing the epipolar spaces and attendant optimal sampling strategy based on these measurements (and setting θ_M in (7) to $\pi/6$—a reasonable value for the HVS) we arrive at the tessellation shown as Fig. 11(a). For these newly constructed epipolar spaces, optimal sampling reduces the average epipolar area to 1/16 of that produced by uniform sampling. For the epipolar spaces in Sect. 2 the spatially variant sampling was really only advantageous in the vertical dimension. Since the new epipolar spaces increase in area with increasing distance in either dimension, the effective-
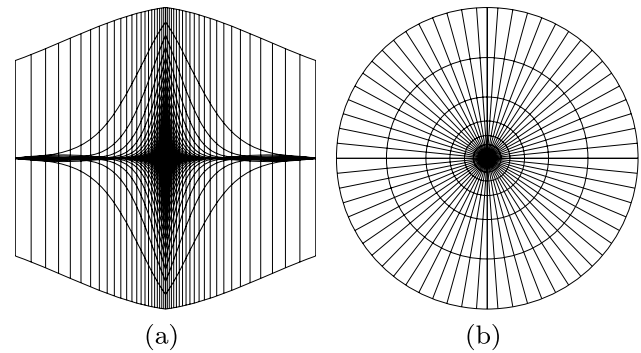
ness of the optimal sampling strategy has essentially been squared.

Another important point to address is the substantial differences between the sampling pattern shown in Fig. 11(a) and the log-polar sampling scheme (Schwartz 1980) adopted by the HVS as illustrated in Fig 11(b). The log-polar sampling scheme is radially symmetric, while our sampling pattern has singularities along both axes, an undesirable property. (Log-polar sampling has a single singularity at the origin.[2]) These singularities result from the fact that for a four-degrees-of-freedom stereo configuration any retinal point falling on either the u or v axis will have a corresponding epipolar space that is one dimensional, i.e. the matching point must lie on the u or v axis in the other image, respectively. The HVS is not well modeled as a four-degrees-of-freedom rig; the human eyes rotate about their optical axes when fixating, a process called cyclovergence. Permitting cyclovergence has both motor and monocular advantages (Kaiser and Hecht 1995; Crowell and Banks 1993). Human eye movements follow Listing's law (Howard and Rogers 1995) when verging on far objects and L2 when fixating on near objects. The movements produce epipolar regions (on spherical image planes) that are far more radially symmetric (Schreiber et al. 2001) than those associated with the four-degrees-of-freedom model. The use of these epipolar spaces by the HVS has been substantiated by psychophysical experiments (Schreiber and Tweed 2003). As to whether or not the optimal sampling strategy for these retinal epipolar spaces is the log-polar transformation is still under investigation.

Finally, we would like to reiterate that though our results may have biological implications with respect to the phy-

---

[2]The actual log-polar transformation presented in Schwartz (1980) is $\log(z + a)$ with $a > 0$. This mapping is applied separately to each cortical hemisphere. With $a > 0$ the mapping does not have a singularity at the origin and is only approximately radially symmetric. A simpler mapping employed prevalently in the engineering literature (and assumed herein) uses $a = 0$, producing the singularity at the origin. For values of a that more precisely model the HVS see (Schwartz et al. 1985; Tootell et al. 1982).

logenetic specialization of retinal topography, we are not proposing a definitive link between foveal evolution and stereo correspondence. In fact, there are many other hypotheses for foveal evolution which consider such factors as predation (Ross 1996), locomotion (Virsu and Hari 1996), optical constraints of the eye (Snyder et al. 1986), behavior (Thompson 1991), and optimal allocation of visual resources (Virsu and Hari 1996). An assiduous comparison amongst evolutionary theories is far beyond the scope of this paper. We merely wish to suggest that there does exist an inherent symbiosis between foveation and uncalibrated binocular active vision systems such as the HVS.

6 Conclusion

Biological vision systems have inspired and will continue to inspire the development of computer vision systems. In this paper we have addressed the symbiotic relation between foveation and uncalibrated active, stereovision systems. We began by introducing the concept of an epipolar space. For a point in one image, the epipolar space was defined as the region in the other image formed from the union of all associated epipolar lines produced over all possible geometric configurations. Epipolar spaces for uncalibrated active vision systems are analogous to the epipolar lines of calibrated vision systems. The sizes of the epipolar spaces are directly related to the accuracy and computational complexity of any correspondence process that registers the two retinal images. Consequently, it is desirable to reduce the average area per epipolar space. Addressing this issue, we introduced a sequence of theorems that provides a path for creating foveation strategies that are optimal with respect to the average epipolar area.

The most important aspect of this work is the identification of a connection between foveation and uncalibrated binocular active vision systems. We are not advocating the immediate use of the sampling strategies introduced in this work. The epipolar spaces presented here were necessarily simple to reduce the mathematics and engender understanding. Additionally, the optimization procedure to select a sampling method focused solely on the metric of epipolar area. In reality, no single optimization criteria should be espoused at the expense of all others. Advantageous properties such as radial symmetry, 3D discretization error, and ease of implementation should also be considered concurrently. Restricting the sampling functions to conformal mappings may also prove highly beneficial. This work introduces a new paradigm, helping us better understand the symbiotic relationship between foveation and uncalibrated binocular active vision. Biological vision systems have exploited this symbiosis for millennia. It is our hope that computer vision systems will begin to do the same.

Appendix A: One-Dimensional Sampling Theorems

In this section we provide proofs for the optimal one-dimensional sampling theorems. Since all the following proofs depend solely on the derivative of $\gamma(u)$, the choice of the constant term in $\gamma(u)$ is arbitrary. Therefore, let any $\gamma(u)$ satisfying (16) map $[a, b]$ onto itself, i.e. $\gamma(a) = a$, $\gamma(b) = b$. Mapping to the same interval removes ambiguity and conceptually simplifies future derivations without reducing the generality of the results.

We begin by considering uniform epipolar intervals $\tilde{i}_c(u)$ defined by $[\max\{u - c/2, a\}, \min\{b, u + c/2\}]$, where $c \geq 0$. The min and max functions clip the epipolar intervals that extend outside of $[a, b]$. In the modeling of practical applications it is reasonable to assume that $b - a \gg c$. This stipulation eliminates the need for clipping, simplifying the interval to $[u - c/2, u + c/2]$. Furthermore, under this assumption (15) can be approximated (for uniform epipolar intervals) as follows:

$$e[\gamma; \tilde{i}_c(u)] \approx c\gamma'(u). \tag{30}$$

Inserting this result into (14) produces

$$E(\gamma; \tilde{i}_c, I) \approx \frac{c}{b - a} \int_a^b [\gamma'(u)]^2 \, du. \tag{31}$$

The next portion of this section provides a sequence of theorems and their proofs that will, in combination, provide a framework for obtaining optimal sampling schemes for a given set of epipolar intervals. A discussion of their significance immediately follows their presentation.

Theorem 1 *If $\gamma(u) \in C^2[a, b]$ then $\gamma(u) = u$ is the unique minimizer of the objective functional given in (31) subject to the constraint posed in (16).*

Proof The Euler–Lagrange equation provides the first order necessary condition for a local minimizer of (31):

$$\frac{\partial L}{\partial \gamma} - \frac{d}{dx}\left(\frac{\partial L}{\partial \gamma'}\right) = 0, \tag{32}$$

where $L(u, \gamma, \gamma')$ is called the Lagrangian (Olver and Shakiban 2009, in preparation). Inserting $L(u, \gamma, \gamma') = [\gamma'(u)]^2$ into (32) produces

$$\gamma''(u) = 0. \tag{33}$$

Integrating this result yields $\gamma(u) = \beta u + \alpha$. The constraint prescribed in (16) forces $\beta = 1$. Since it was previously stipulated that all sampling functions map I onto itself, it follows that $\alpha = 0$. In order to guarantee that this single critical

function is a minimizer, it is necessary to examine the second variation $Q(\gamma; \chi)$:

$$Q(\gamma; \chi) = \int_a^b \left[A\chi^2 + 2B\chi\chi' + C(\chi')^2 \right] du, \qquad (34)$$

where

$$A = \frac{\partial^2 L}{\partial \gamma \partial \gamma}, \qquad B = \frac{\partial^2 L}{\partial \gamma \partial \gamma'}, \qquad C = \frac{\partial^2 L}{\partial \gamma' \partial \gamma'},$$

and $\chi(u) \in C^2[a, b]$ is any arbitrary function with $\chi(a) = 0 = \chi(b)$. If $Q(\gamma; \chi) > 0$ for all χ not identically zero, then $Q(\gamma; \chi)$ is positive definite and $\gamma(u)$ is a strict local minimum (Olver and Shakiban 2009, in preparation). Inserting $L(u, \gamma, \gamma') = [\gamma'(u)]^2$ into the (34) produces

$$Q(\gamma; \chi) = \int_a^b 2(\chi')^2 du. \qquad (35)$$

This quantity vanishes only if χ is a constant. Since $\chi(a) = 0$, it follows that $\chi \equiv 0$. Therefore, (34) is positive definite and $\gamma(u) = u$ is the unique minimizer of (31). $\square$

Theorem 2 *If $\gamma(u) = \Gamma(u)$ is the sampling function that minimizes $E(\gamma; i, I)$ subject to (16) and $\chi(u)$ is another arbitrary sampling function satisfying (16), then $\gamma(u) = \Gamma_\chi(u) = \Gamma[\chi(u)]$ minimizes $E(\gamma; i_\chi, I)$ subject to (16), where $i_\chi(u)$ is the interval with upper bound $i_{b\chi}(u) = \chi^{-1}(i_b[\chi(u)])$ and lower bound $i_{a\chi}(u) = \chi^{-1}(i_a[\chi(u)])$.*

Proof The following equations use the variable substitution $x = \chi^{-1}(u)$.

$$E(\Gamma; i, I) = \frac{1}{b-a} \int_a^b e\left[\Gamma; i(u)\right] \Gamma'(u) \, du$$

$$= \frac{1}{b-a} \int_a^b e(\Gamma; i\left[\chi(x)\right]) \Gamma'\left[\chi(x)\right] \chi'(x) \, dx$$

$$= \frac{1}{b-a} \int_a^b e\left[\Gamma_\chi; i_\chi(x)\right] \Gamma_\chi'(x) \, dx$$

$$= E(\Gamma_\chi; i_\chi, I) \qquad (36)$$

$$e(\Gamma; i\left[\chi(x)\right]) = \int_{i_a(\chi(x))}^{i_b(\chi(x))} \Gamma'(\tilde{u}) \, d\tilde{u}$$

$$= \int_{\chi^{-1}(i_a[\chi(x)])}^{\chi^{-1}(i_b[\chi(x)])} \Gamma'\left[\chi(\tilde{x})\right] \chi'(\tilde{x}) \, d\tilde{x}$$

$$= \int_{i_{a\chi}(x)}^{i_{b\chi}(x)} \Gamma_\chi'(\tilde{x}) \, d\tilde{x}$$

$$= e\left[\Gamma_\chi; i_\chi(x)\right] \qquad (37)$$

$$\Lambda(\Gamma_\chi; I) = \int_a^b \Gamma'\left[\chi(u)\right] \chi'(u) \, du$$

$$= \int_a^b \Gamma'(x) \, dx$$

$$= \Lambda(\Gamma; I) = b - a. \qquad (38)$$

From the above proofs we can draw several important conclusions. First, Theorem 1 demonstrates that uniform sampling becomes optimal for uniform epipolar intervals when the total interval length is large with respect to the lengths of the uniform epipolar intervals. This caveat of relative length is necessary to mitigate the effects of the nonlinear clipping needed to ensure that the epipolar intervals never extend outside I. Second, Theorem 2 provides a means for creating other sets of epipolar intervals and their attendant optimal sampling functions by warping uniform epipolar intervals and their corresponding optimal sampling function $\Gamma(u) = u$. This immediately suggests the following corollary: *if a sampling function $\gamma(u)$ warps intervals of nonuniform length into intervals of uniform length, it is the optimal sampling function with respect to those nonuniform intervals.* This corollary becomes readily apparent by reversing the proof of Theorem 2 with $\Gamma(u) = u$.

Appendix B: Two-Dimensional

In this section we provide proofs for the optimal two-dimensional sampling theorems. We begin by considering epipolar spaces $\tilde{r}(\boldsymbol{u})$ of equal area c but arbitrary shape. As in the one-dimensional case, it is assumed that $A_R \gg c$, eliminating the need to clip the epipolar spaces that extend outside of R. With these restrictions (18) can be approximated as follows:

$$e\left[\boldsymbol{\gamma}; r(\boldsymbol{u})\right] \approx c J_\gamma(\boldsymbol{u}). \qquad (39)$$

This result states that the average value of the determinant of the Jacobian of the sampling function over the particular epipolar space can be approximated by its value at $\boldsymbol{u}$. Furthermore, since the result in (39) is independent of the shape of the epipolar spaces, it is reasonable to restrict the space of optimal sampling functions to those with Jacobian matrices whose determinants are separable, i.e. $J_\gamma(\boldsymbol{u}) \doteq J_{\gamma_u}(u) J_{\gamma_v}(v)$. Accordingly, the insertion of (39) into (17) produces

$$E(\boldsymbol{\gamma}; \tilde{r}, R) \approx \frac{c}{A_R} \iint_R \left[J_\gamma(\boldsymbol{u})\right]^2 \, du \, dv$$

$$= \frac{c}{A_R} \iint_R \left[J_{\gamma_u}(u) J_{\gamma_v}(v)\right]^2 \, du \, dv$$

$$= \frac{c}{A_R} \iint_R \left[\gamma_u'(u) \gamma_v'(v)\right]^2 \, du \, dv. \qquad (40)$$

The final step results from the fact that a separable sampling function (i.e. $\boldsymbol{\gamma}(\boldsymbol{u}) \doteq [\gamma_u(u), \gamma_v(v)]^t$) is sufficient to produce any separable $J_\gamma(\boldsymbol{u})$. Though this is true, it is important to note that many nonseparable sampling functions also may have separable Jacobians. For example, hexagonal sampling, expressed as

$$\boldsymbol{\gamma}(\boldsymbol{u}) = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}, \tag{41}$$

is not a separable sampling function. However, the determinant of its Jacobian $\sqrt{3}/2$ is also the determinant of the Jacobian of the separable sampling function $\boldsymbol{\gamma}(\boldsymbol{u}) = [u, \sqrt{3}/2\, v]^t$.

The following theorems and proofs form the foundation for constructing optimal two-dimensional sampling strategies.

Theorem 3 *Let R be a rectangular image plane with $u \in I_u$ and $v \in I_v$, where I_u and I_v are the intervals defined by $[a_u, b_u]$ and $[a_v, b_v]$, respectively. If $\gamma_u(u) \in \mathrm{C}^2 [a_u, b_u]$ and $\gamma_v(v) \in \mathrm{C}^2 [a_v, b_v]$ then $\gamma_u(u) = \beta_u u$ and $\gamma_v(v) = \beta_v v$, where $\beta_u \beta_v = 1$, minimize the objective functional given in (40), subject to the constraint posed in (19).*

Proof

$$E(\boldsymbol{\gamma}; \tilde{r}, R)$$
$$= \frac{c}{A_R} \iint_R [\gamma_u'(u) \gamma_v'(v)]^2 \, du \, dv$$
$$= \frac{c_1}{b_u - a_u} \int_{a_u}^{b_u} [\gamma_u'(u)]^2 \, du \frac{c_2}{b_v - a_v} \int_{a_v}^{b_v} [\gamma_v'(v)]^2 \, dv$$
$$= E(\gamma_u; \tilde{i}_{c_1}, I_u) E(\gamma_v; \tilde{i}_{c_2}, I_v), \tag{42}$$

where c_1 and c are arbitrary positive constants satisfying $c_1 c_2 = c$. Equation (42) obtains its minimum when both $E(\gamma_u; \tilde{i}_{c_1}, I_u)$ and $E(\gamma_v; \tilde{i}_{c_2}, I_v)$ obtain their respective minimums. From the proof of Theorem 1 we know this occurs when $\gamma_u = \beta_u u + \alpha_u$ and $\gamma_v = \beta_v v + \alpha_v$. The constraint in (19) forces $\beta_u \beta_v = 1$. Since it was stipulated that any one-dimensional sampling function maps I onto itself, it follows that $\alpha_u = 0$ and $\alpha_v = 0$. $\qquad \square$

Theorem 4 *If $\boldsymbol{\gamma}(\boldsymbol{u}) = \boldsymbol{\Gamma}(\boldsymbol{u})$ is a sampling function that minimizes $E(\boldsymbol{\gamma}; r, R)$ subject to $\Lambda(\boldsymbol{\gamma}; R) = A_R$ and $\boldsymbol{\chi}(\boldsymbol{u})$ is some other sampling function satisfying $\Lambda(\boldsymbol{\chi}^{-1}; R) = A_R$, then $\boldsymbol{\gamma}(\boldsymbol{u}) = \boldsymbol{\Gamma}_\chi(\boldsymbol{u}) = \boldsymbol{\Gamma}(\boldsymbol{\chi}(\boldsymbol{u}))$ minimizes $E(\boldsymbol{\gamma}; r_\chi, R_\chi)$ subject to $\Lambda(\boldsymbol{\gamma}; R_\chi) = A_{R_\chi}$, where $R_\chi = \boldsymbol{\chi}^{-1}(R)$ and $r_\chi(\boldsymbol{u}) = \boldsymbol{\chi}^{-1}(r[\boldsymbol{\chi}(\boldsymbol{u})])$.*

Proof The following equations use the variable substitution $\boldsymbol{x} = \boldsymbol{\chi}^{-1}(\boldsymbol{u})$.

By definition, $\Lambda(\boldsymbol{\chi}^{-1}; R) = A_R$ and $\Lambda(\boldsymbol{\Gamma}; R) = A_R$, consequently

$$A_R = \iint_R |(\boldsymbol{\chi}^{-1})'(\boldsymbol{u})| \, d\boldsymbol{u}$$
$$= \iint_{R_\chi} d\boldsymbol{x} = A_{R_\chi} \tag{43}$$

$$A_{R_\chi} = A_R$$
$$= \iint_R |\boldsymbol{\Gamma}'(\boldsymbol{u})| \, d\boldsymbol{u}$$
$$= \iint_{R_\chi} |\boldsymbol{\Gamma}'[\boldsymbol{\chi}(\boldsymbol{x})]||\boldsymbol{\chi}'(\boldsymbol{x})| \, d\boldsymbol{x}$$
$$= \iint_{R_\chi} |\boldsymbol{\Gamma}_\chi'(\boldsymbol{x})| \, d\boldsymbol{x}$$
$$= \Lambda(\boldsymbol{\Gamma}_\chi; R_\chi). \tag{44}$$

Using these results and the aforementioned variable substitution yields:

$$E(\boldsymbol{\Gamma}; r, R)$$
$$= \frac{1}{A_R} \iint_R e[\boldsymbol{\Gamma}; r(\boldsymbol{u})]|\boldsymbol{\Gamma}'(\boldsymbol{u})| \, d\boldsymbol{u}$$
$$= \frac{1}{A_R} \iint_{\boldsymbol{\chi}^{-1}(R)} e(\boldsymbol{\Gamma}; r[\boldsymbol{\chi}(\boldsymbol{x})])|\boldsymbol{\Gamma}'[\boldsymbol{\chi}(\boldsymbol{x})]||\boldsymbol{\chi}'(\boldsymbol{x})| \, d\boldsymbol{x}$$
$$= \frac{1}{A_{R_\chi}} \iint_{R_\chi} e[\boldsymbol{\Gamma}_\chi; r_\chi(\boldsymbol{x})]|\boldsymbol{\Gamma}_\chi'(\boldsymbol{x})| \, d\boldsymbol{x}$$
$$= E(\boldsymbol{\Gamma}_\chi; r_\chi, R_\chi), \tag{45}$$

$$e(\boldsymbol{\Gamma}; r[\boldsymbol{\chi}(\boldsymbol{x})]) = \iint_{r(\boldsymbol{\chi}(x))} |\boldsymbol{\Gamma}'(\tilde{\boldsymbol{u}})| \, d\tilde{\boldsymbol{u}}$$
$$= \iint_{\boldsymbol{\chi}^{-1}(r[\boldsymbol{\chi}(x)])} |\boldsymbol{\Gamma}'[\boldsymbol{\chi}(\tilde{\boldsymbol{x}})]||\boldsymbol{\chi}'(\tilde{\boldsymbol{x}})| \, d\tilde{\boldsymbol{x}}$$
$$= \iint_{r_\chi(x)} |\boldsymbol{\Gamma}_\chi'(\tilde{\boldsymbol{x}})| \, d\tilde{\boldsymbol{x}}$$
$$= e[\boldsymbol{\Gamma}_\chi; r_\chi(x)]. \qquad \square \tag{46}$$

For practical purposes it is important to note that though Theorem 3 was presented in the context of a rectangular image plane R, the result still holds for any arbitrarily shaped region R with the following caveat: the total area clipped from the epipolar spaces that extend outside of R must be sufficiently small compared to the total area of R. In general, the size and shape of the regions R and R_χ can be arbitrarily chosen to fit a specific circumstance, again, with the caveat that the total clipped epipolar area be relatively small.

The significance of the previous proofs is now addressed. Theorem 3 states that uniform rectangular sampling is optimal for epipolar spaces of uniform area when the area of *R* is large with respect to the individual epipolar areas, i.e. as the amount of clipping becomes negligible. Combining this result with Theorem 4 demonstrates that *any area preserving sampling function $\gamma(u)$ that warps epipolar spaces into regions of uniform area is optimal with respect to their average epipolar area.*

References

Aloimonos, J., Weiss, I., & Bandyopadhyay, A. (1987). Active vision. In *Proceedings of the first international conference on computer vision* (pp. 35–54).

Ballard, D. (1991). Animate vision. *Artificial Intelligence Journal*, *48*, 57–86.

Basu, A. (1992). Optimal discretization for stereo reconstruction. *Pattern Recognition Letters*, *13*(11), 813–820.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *Journal of Physiology*, *211*(3), 599–622.

Bough, E. W. (1970). Stereoscopic vision in the macaque monkey: a behavioural demonstration. *Nature*, *225*(5227), 42–44.

Bovik, A. C. (2000). *The handbook of image & video processing*. San Diego: Academic Press.

Chen, Q. S., & Deconinck, F. (1994). Foveal-view stereopsis using symmetric phase-only matched filtering. *Proceedings of SPIE (The International Society for Optical Engineering)*, *2233*, 46–53.

Crowell, J. A., & Banks, M. S. (1993). Perceiving heading with different retinal regions and types of optic flow. *Perception Psychophysics*, *53*(3), 325–337.

Davis, F. A. (1929). The anatomy and histology of the eye and orbit of the rabbit. *Transactions American Ophthalmological Society*, *27*, 401–441.

Dhond, U. R., & Aggarwal, J. (1989). Structure from stereo—a review. *IEEE Transactions on Systems, Man and Cybernetics*, *19*(6), 1489–1510.

Dräger, U. C., & Olsen, J. F. (1980). Origins of crossed and uncrossed retinal projections in pigmented and albino mice. *Journal of Comparative Neurology*, *191*(3), 383–412.

Easter, S. S. (1972). Pursuit eye movements in goldfish (carassius auratus). *Vision Research*, *12*(4), 673–688.

Elnagar, A. (1998). Optimal error discretization under depth and range constraints. *Pattern Recognition Letters*, *19*(9), 879–888.

Faugeras, O. D., Luong, Q. T., & Maybank, S. J. (1992). Camera self calibration: Theory and experiments. In *Proceedings of the Second European Conference on Computer Vision* (Vol. 588, pp. 321–334).

Fite, K. V., & Rosenfield-Wessels, S. (1975). A comparative study of deep avian foveas. *Brain, Behavior and Evolution*, *12*(1–2), 97–115.

Fox, R., Lehmkuhle, S. W., & Bush, R. C. (1977). Stereopsis in the falcon. *Science*, *197*(4298), 79–81.

Geisler, W. S., & Perry, J. S. (1998). Real-time foveated multiresolution system for low-bandwidth video communication. In B. E. Rogowitz & T. N. Pappas (Eds.) *Society of photo-optical instrumentation engineers (SPIE) conference series* (Vol. 3299, pp. 294–305). Bellingham: SPIE.

Hespanha, J., Dodds, Z., Hager, G., & Morse, A. (1999). What tasks can be performed with an uncalibrated stereo vision system? *International Journal of Computer Vision*, *35*(1), 65–85.

Horn, B. (1986). *Robot vision*. Cambridge: MIT Press.

Howard, I., & Rogers, B. (1995). *Binocular vision and stereopsis*. London: Oxford University Press.

Jenkin, M., Jepson, A., & Tsotsos, J. (1991). Techniques for disparity measurement. *Computer Vision, Graphics, and Image Processing: Image Understanding*, *53*(1), 14–30.

Kaiser, M. K., & Hecht, H. (1995). Time-to-passage judgments in nonconstant optical flow fields. *Perception Psychophysics*, *57*(6), 817–825.

Klarquist, W. N., & Bovik, A. C. (1998). Fovea: A foveated vergent active stereo vision system for dynamic three-dimensional scene recovery. *IEEE Transactions on Robotics and Automation*, *14*(5), 755–770.

Lee, S., Pattichis, M., & Bovik, A. (2001). Foveated video compression with optimal rate control. *IEEE Transactions on Image Processing*, *10*(7), 977–992.

Li, F., & Zhou, Y. (1999). Disparity estimation based on frequency domain. *Shanghai Jiaotong Daxue Xuebao/Journal of Shanghai Jiaotong University*, *33*(5), 516–519.

Manzotti, R., Gasteratos, A., Metta, G., & Sandini, G. (2001). Disparity estimation on log-polar images and vergence control. *Computer Vision and Image Understanding*, *83*(2), 97–117.

Marr, D., & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science*, *194*(4262), 283–287.

Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *204*(1156), 301–328.

Monaco, J., Bovik, A., & Cormack, L. (2007a). Epipolar spaces and optimal sampling strategies. *IEEE International Conference on Image Processing*, *6*, VI-545–VI-548.

Monaco, J., Bovik, A., & Cormack, L. (2007b). Epipolar spaces for active binocular vision systems. *IEEE International Conference on Image Processing*, *6*, VI-549–VI-551.

Monaco, J. P., Bovik, A. C., & Cormack, L. K. (2008). Nonlinearities in stereoscopic phase-differencing. *IEEE Transactions on Image Processing*, *17*(9), 1672–1684.

Nielsen, K. R., & Poggio, T. (1984). Vertical image registration in stereopsis. *Vision Research*, *24*(10), 1133–1140.

Olver, P., & Shakiban, C. (2009, in preparation). *Fundamentals of applied mathematics*. New York: Prentice-Hall.

Ptito, M., Lepore, F., & Guillemot, J. P. (1991). Stereopsis in the cat: behavioral demonstration and underlying mechanisms. *Neuropsychologia*, *29*(6), 443–464.

Rapaport, D. H., & Stone, J. (1984). The area centralis of the retina in the cat and other mammals: focal point for function and development of the visual system. *Neuroscience*, *11*(2), 289–301.

Ross, C. (1996). An adaptive explanation for the origin of the anthropoidea. *American Journal of Primatology*, *40*(2), 205–230.

Sandini, G., & Tistarelli, M. (1990). Active tracking strategy for monocular depth inference over multiple frames. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*(1), 13–27.

Sanger, T. (1988). Stereo disparity computation using Gabor filters. *Biological Cybernetics*, *59*, 405–418.

Schindler, K., & Bischof, H. (2004). The epipolar geometry of the log-polar image plane. *Proceedings of the International Conference on Pattern Recognition*, *4*, 40–43.

Schreiber, K., Crawford, J. D., Fetter, M., & Tweed, D. (2001). The motor side of depth vision. *Nature*, *410*(6830), 819–822.

Schreiber, K. M., & Tweed, D. B. (2003). Influence of eye position on stereo matching. *Strabismus*, *11*(1), 9–16.

Schwartz, E., Tootell, R. B., Silverman, M. S., Switkes, E., & Valois, R. L. D. (1985). On the mathematical structure of the visuotopic mapping of macaque striate cortex. *Science*, *227*(4690), 1065–1066.

Schwartz, E. L. (1980). A quantitative model of the functional architecture of human striate cortex with application to visual illusion and cortical texture analysis. *Biological Cybernetics*, *37*(2), 63–76.

Shah, S., & Aggarwal, J. (1994). Depth estimation using stereo fish-eye lenses. *Proceedings of the IEEE International Conference on Image Processing*, *2*, 740–744.

Shih, S. W., Hung, Y. P., & Lin, W. S. (1998). Calibration of an active binocular head. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, *28*(4), 426–442.

Snyder, A., Bossomaier, T., & Hughes, A. (1986). Optical image quality and the cone mosaic. *Science*, *231*(4737), 499–501.

Stevenson, S. B., & Schor, C. M. (1997). Human stereo matching is not restricted to epipolar lines. *Vision Research*, *37*(19), 2717–2723.

Strang, G. (1988). *Linear algebra and its applications*. San Diego: Harcourt Brace Jovanovich.

Thompson, I. (1991). *Vision and visual dysfunction*, *Vol. 2*. Boca Raton: CRC Press (Chap. 7, pp. 136–151).

Tiao, Y. C., & Blakemore, C. (1976). Regional specialization in the golden hamster's retina. *Journal of Comparative Neurology*, *168*(4), 439–457.

Tistarelli, M., & Sandini, G. (1991). Direct estimation of time-to-impact from optical flow. In *Proceedings of the IEEE workshop on visual motion* (pp. 226–233).

Tootell, R. B., Silverman, M. S., Switkes, E., & Valois, R. L. D. (1982). Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science*, *218*(4575), 902–904.

Virsu, V., & Hari, R. (1996). Cortical magnification, scale invariance and visual ecology. *Vision Research*, *36*(18), 2971–2977.

Wang, Z., & Bovik, A. C. (2001). Embedded foveation image coding. *IEEE Transactions on Image Processing*, *10*(10), 1397–1410.

Wathey, J. C., & Pettigrew, J. D. (1989). Quantitative analysis of the retinal ganglion cell layer and optic nerve of the barn owl tyto alba. *Brain, Behavior Evolution*, *33*(5), 279–292.

Wei, J., & Li, Z. N. (1998). Efficient disparity-based gaze control with foveate wavelet transform. *IEEE International Conference on Intelligent Robots and Systems*, *2*, 866–871.

Weiman, C. F. (1995). Binocular stereo via log-polar retinas. *Proceedings of SPIE—The International Society for Optical Engineering*, *2488*, 309–320.

van der Willigen, R. F., Frost, B. J., & Wagner, H. (1998). Stereoscopic depth perception in the owl. *Neuroreport*, *9*(6), 1233–1237.

Yeshurun, Y., & Schwartz, E. L. (1999). Cortical hypercolumn size determines stereo fusion limits. *Biological Cybernetics*, *80*(2), 117–129.

Zeevi, Y., & Shlomot, E. (1993). Nonuniform sampling and antialiasing in image representation. *IEEE Transactions on Signal Processing*, *41*(3), 1223–1236.