EPIPOLAR SPACES AND OPTIMAL SAMPLING STRATEGIES

James Monaco, Alan C. Bovik

University of Texas at Austin
Department of Electrical
and Computer Engineering
Austin, TX 78712-1084 USA
{monaco,bovik}@ece.utexas.edu

Lawrence K. Cormack

University of Texas at Austin
Department of
Psychology
Austin, TX 78712-0187 USA
cormack@mail.utexas.edu

ABSTRACT

If precise calibration information is unavailable, as is often the case for active binocular vision systems, the determination of epipolar lines becomes untenable. Yet, even without instantaneous knowledge of the geometry, the search for corresponding points can be restricted to areas called *epipolar spaces*. For each point in one image, we define the corresponding epipolar space in the other image as the union of all associated epipolar lines over all possible system geometries. Epipolar spaces eliminate the need for calibration at the cost of an increased search region. One approach to mitigate this increase is the application of a space variant sampling or foveation strategy. While the application of such strategies to stereo vision tasks is not new, only rarely has a foveation scheme been specifically tailored for a stereo vision task. In this paper we derive a foundation of theorems that provide a means for obtaining optimal sampling schemes for a given set of epipolar spaces. An optimal sampling scheme is defined as a strategy that minimizes the average area per epipolar space.

Index Terms— Stereo vision, Image sampling, Active vision, Image registration

1. INTRODUCTION

Active binocular vision systems (ABVS) are especially well suited for the recovery of depth information [1]. This recovery process, accomplished through registration, can be greatly simplified if the geometric configuration is known, allowing the search for corresponding points to be restricted to epipolar lines. Unfortunately, ascertaining the actively changing stereo geometry requires accurate calibration, a complex procedure. Though it is possible to register images without geometric information, such unconstrained algorithms are usually more time consuming and prone to error. In this paper we propose a compromise. Even without knowledge of the exact stereo geometry we can restrict the region of correspondence by imposing limits on the possible range of configurations. That is, by restricting the range of vergence angles, baseline distances,

and focal lengths etc., we can confine our search for matching points to what we refer to as *epipolar spaces*. For each point in one image, we define the corresponding epipolar space in the other image as the union of all associated epipolar lines over all possible system geometries.

Epipolar spaces eliminate the need for calibration at the cost of an increased search region. One approach to mitigate this increase is the application of a space variant sampling or foveation strategy. The application of such a strategy to stereo vision tasks is not new. Space variant transformations such as log-polar [2], reciprocal wedge transform [3], and fish-eye [4] have been successfully applied to binocular vision problems such as vergence [5] and depth recovery of a scene [6]. Yet, only rarely has a foveation strategy been specifically tailored for a stereo vision task. Both Basu [7] and Elnagar [8] derived optimal sampling schemes with respect to the error discretization of depth measurements. Klarquist and Bovik [9] designed a real-time foveated stereo technique that adapts to the specific geometry, always producing horizontal epipolar lines.

In this paper we first formalize the concept of epipolar spaces. We then discuss the general mathematics of applying nonuniform sampling strategies to epipolar spaces. Finally, we present an optimal sampling scheme specifically designed to minimize the average area per epipolar space.

2. EPIPOLAR SPACES

Given the specific stereo geometry shown in Fig. 1, the expression for corresponding epipolar lines is

$$v_l = v_r \frac{f \sin(\theta_l) + u_l \cos(\theta_l)}{f \sin(\theta_r) - u_r \cos(\theta_r)}. \quad (1)$$

Consider the situation where the camera configuration actively changes and we no longer know the specific geometry. Computing epipolar lines becomes untenable. 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.
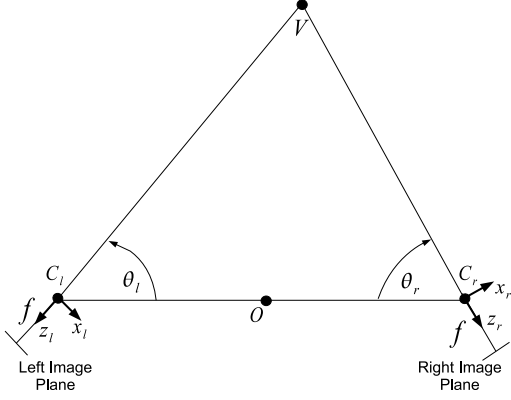
Fig. 1. Stereo geometry.

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 camera 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

$$\theta_l, \theta_r \in [\theta_M, \pi - \theta_M], \qquad (2)$$

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. For our purposes, we assume a maximum horizontal disparity defined by

$$|d| = |u_l - u_r| \leq D. \qquad (3)$$

The restriction imposed by (3) determines the leftmost and rightmost bounds of the epipolar spaces. The upper and lower bounds are determined by maximizing and minimizing (1) with respect to both θ_l and θ_r. For reasonable image sizes and focal lengths these bounds can be approximated as follows:

$$v_{l,max} \approx v_r \frac{\sqrt{f^2 + u_r^2}}{f \sin(\theta_M) - u_r \cos(\theta_M)} = v_r c(u_r) \qquad (4)$$

$$v_{l,min} \approx v_r \frac{f \sin(\theta_M) - u_r \cos(\theta_M)}{\sqrt{f^2 + u_r^2}} = \frac{v_r}{c(u_r)}, \qquad (5)$$

where

$$c(u_r) = \frac{\sqrt{f^2 + u_r^2}}{f \sin(\theta_M) - u_r \cos(\theta_M)}. \qquad (6)$$

Remarkably, an epipolar space is well modeled by a rectangle. This fact is illustrated in Fig. 2. The thick lines represent the precise boundaries of the regions. The thin lines denote the approximate upper and lower bounds determined from (4) and (5). Epipolar spaces are nonuniform in area, increasing in size with increasing values of u and v.
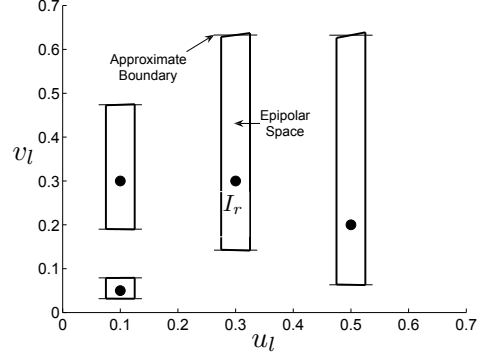

Fig. 2. Various epipolar spaces.

3. OPTIMAL SAMPLING THEOREMS

Epipolar spaces eliminate the need for calibration at the expense of an increased search space. To mitigate this increase, a sampling strategy can be applied. The optimal strategy is defined as the one 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.

Consider any two-dimensional, connected region R. For each point u in R the matching point must lie within the corresponding epipolar space $r(u)$. The goal of an 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 u_j is defined as any point u_k that lies within the epipolar space $r(u_j)$.

A discrete sampling strategy can be represented as a continuous sampling function [10]. Such a function is defined as any invertible function $\gamma(u) \doteq [\gamma_u(u), \gamma_v(u)]^t$ that maps $\mathbb{R}^2 \to \mathbb{R}^2$ and has a Jacobian matrix with a determinant that is positive everywhere. For a given sampling function $\gamma(u)$ the average area of the epipolar spaces $r(u)$ over the region R is defined as

$$E(\gamma; r, R) = \frac{1}{A_R} \iint_R e[\gamma; r(u)] |\gamma'(u)| du, \qquad (7)$$

where

$$e\left[\boldsymbol{\gamma}; r\left(\boldsymbol{u}\right)\right] = \iint_{r(u)} \left|\boldsymbol{\gamma}'\left(\tilde{\boldsymbol{u}}\right)\right| d\tilde{\boldsymbol{u}} \tag{8}$$

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

$$\Lambda\left(\boldsymbol{\gamma}; R\right) = \iint_R \left|\boldsymbol{\gamma}'\left(\boldsymbol{u}\right)\right| d\boldsymbol{u} = \iint_R d\boldsymbol{u} = A_R. \tag{9}$$

We begin by considering hypothetical uniform rectangular epipolar spaces $\tilde{r}\left(\boldsymbol{u}\right)$ defined by $\tilde{u} \in \left[u-c_u, u+c_u\right]$ and $\tilde{v} \in \left[v-c_v, v+c_v\right]$, where $c_u \geq 0$ and $c_v \geq 0$. It is assumed that $A_R \gg c_u c_v$, eliminating the need to clip the epipolar spaces that extend outside of R. Furthermore, since the epipolar spaces are separable and not spatially variant, it is reasonable to restrict the space of optimal sampling functions to the space of separable functions, i.e. $\boldsymbol{\gamma}\left(\boldsymbol{u}\right) \doteq \left[\gamma_u\left(u\right), \gamma_v\left(v\right)\right]^t$. With these restrictions (8) can be well approximated (for uniform epipolar spaces) as follows:

$$e\left(\boldsymbol{\gamma}; \tilde{r}\right) \approx c_u c_v \gamma_u'\left(u\right) \gamma_v'\left(v\right). \tag{10}$$

Assuming separability, the insertion of (10) into (7) produces

$$E\left(\boldsymbol{\gamma}; \tilde{r}, R\right) \approx \frac{4 c_u c_v}{A_R} \iint_R \left[\gamma_u'\left(u\right) \gamma_v'\left(v\right)\right]^2 du dv. \tag{11}$$

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

Theorem 3.1 *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 $\left[a_u, b_u\right]$ and $\left[a_v, b_v\right]$, respectively. If $\gamma_u\left(u\right) \in \mathrm{C}^2\left[a_u, b_u\right]$ and $\gamma_v\left(v\right) \in \mathrm{C}^2\left[a_v, b_v\right]$ then $\gamma_u\left(u\right) = \beta_u u$ and $\gamma_v\left(v\right) = \beta_v v$, where $\beta_u \beta_v = 1$, minimize the objective functional given in (11), subject to the constraint posed in (9).*

Proof This is omitted due to space restrictions.

Theorem 3.2 *If $\boldsymbol{\gamma}\left(\boldsymbol{u}\right) = \boldsymbol{\Gamma}\left(\boldsymbol{u}\right)$ is a sampling function that minimizes $E\left(\boldsymbol{\gamma}; r, R\right)$ subject to $\Lambda\left(\boldsymbol{\gamma}; R\right) = A_R$ and $\boldsymbol{\chi}\left(\boldsymbol{u}\right)$ is some other sampling function satisfying $\Lambda\left(\boldsymbol{\chi}^{-1}; R\right) = A_R$, then $\boldsymbol{\gamma}\left(\boldsymbol{u}\right) = \boldsymbol{\Gamma}_\chi\left(\boldsymbol{u}\right) = \boldsymbol{\Gamma}\left[\boldsymbol{\chi}\left(\boldsymbol{u}\right)\right]$ minimizes the functional $E\left(\boldsymbol{\gamma}; r_\chi, R_\chi\right)$ subject to the constraint $\Lambda\left(\boldsymbol{\gamma}; R_\chi\right) = A_{R_\chi}$, where $R_\chi = \boldsymbol{\chi}^{-1}\left(R\right)$ and $r_\chi\left(\boldsymbol{u}\right) = \boldsymbol{\chi}^{-1}\left(r\left[\boldsymbol{\chi}\left(\boldsymbol{u}\right)\right]\right)$.*

Proof This can be shown by performing the variable substitution $\boldsymbol{x} = \boldsymbol{\chi}^{-1}\left(\boldsymbol{u}\right)$ in equations (7), (8), and (9). This is omitted due to space restrictions.

Theorem 3.1 states that uniform rectangular sampling is optimal for uniform rectangular epipolar spaces when the area of R is large with respect to the individual epipolar areas, i.e. as the amount of clipping becomes negligible. This holds for

uniform rectangular epipolar spaces of any aspect ratio. It also holds for rectangular sampling of any aspect ratio; that is, the sampling rate in each dimension need not agree. Combining this result with Theorem 3.2 demonstrates that any area preserving sampling function (i.e., it satisfies (9)) that warps nonuniformly sized epipolar spaces into rectangles of uniform size is the minimizing function associated with those nonuniform spaces. In fact, the nonuniform spaces do not have to be warped into rectangles. They need only be mapped into regions of uniform area. The proof of this is omitted for space. In summary, *any area preserving sampling function $\boldsymbol{\gamma}\left(\boldsymbol{u}\right)$ that warps epipolar spaces into regions of uniform area is optimal with respect to their average epipolar area.*

4. OPTIMAL SAMPLING STRATEGY

This section uses the results from previous sections to formulate the optimal sampling schemes for the epipolar spaces defined in (3), (4), and (5). The sampling function that warps these epipolar spaces into regions of nearly equal area, and consequently, minimizes (7) is:

$$\boldsymbol{\gamma}\left(u, v\right) = \left[\beta_u u, \frac{\beta_v \ln v}{\ln c\left(u\right)}\right]^t, \tag{12}$$

where β_u and β_v are constants chosen such that their product satisfies the constraint in (9). The transformations of the epipolar spaces shown in Fig. 2 using (12) are illustrated in Fig. 3. Though the warped boundaries are not identical in shape, they are almost perfectly uniform in area.
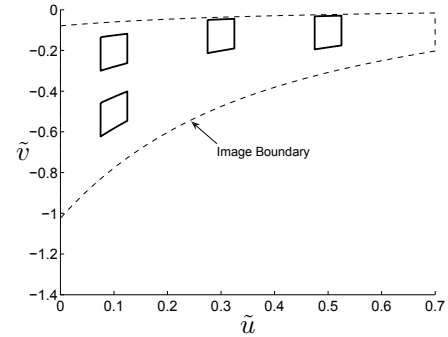

Fig. 3. Epipolar spaces from Fig. 2 after optimal sampling.

As mentioned previously, the constants β_u and β_v in (12) are only restricted in the sense that their product must satisfy (9). This is a consequence of the fact that the average epipolar area is a function of the determinant of the Jacobian of $\boldsymbol{\gamma}\left(\boldsymbol{u}\right)$, and not $\boldsymbol{\gamma}\left(\boldsymbol{u}\right)$ itself. Fig. 4 illustrates the optimal sampling strategies when β_u is set to two different, arbitrarily chosen values: $\beta_u = 1$ and $\beta_u = 4$.

Consider the uniform grid in Fig. 5(a). This grid represents the uniform sampling of the warped image space re-
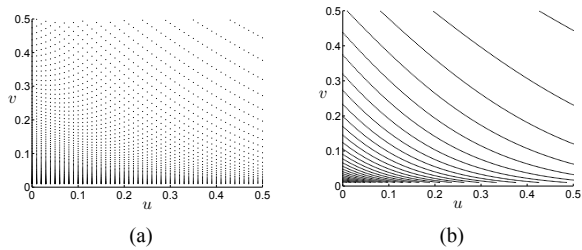
Fig. 4. Examples of optimal sampling scheme.

sulting from the application of the optimal transformation in (12) to a square image plane. Fig. 5(b) illustrates the tessellation resulting from projecting the uniform grid in Fig. 5(a) back into the original square image plane. Each enclosed area in the tessellation is called a super-pixel. Foveating a typical uniformly sampled image requires assigning all the uniform pixels within each super-pixel their average value. In the warped space each uniform pixel is assigned the average of its concomitant super-pixel.
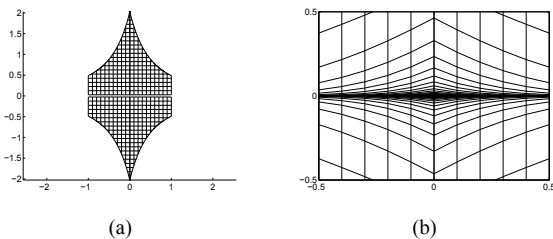

Fig. 5. Tessellations for optimal epipolar sampling.

5. CONCLUSION

In this paper we introduced the concept of an epipolar space. For a point in one image, an 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 eliminate the need for calibration, but at the expense of a greater search space. To reduce this increase we established a foundation of theorems for deriving optimal foveation schemes that minimize the average area per epipolar space. These theorems indicated that any transformation mapping nonuniform epipolar spaces into regions of uniform area is optimal with respect to mean epipolar area. As a demonstration of principle we obtained an optimal sampling scheme and tessellation for the epipolar spaces resulting from a specific range of stereo geometries.

On a final note, we reiterate that for a given set of epipolar spaces there exists an infinite number of optimal sampling strategies. To intelligently select from among these foveation schemes the additional optimization of ancillary metrics, such as 3D discretization error, can also be considered.

6. REFERENCES

[1] J. Aloimonos, I. Weiss, and A. Bandyopadhyay, "Active vision," *Proceedings of the First International Conference on Computer Vision*, pp. 35–54, 1987.

[2] C. F. Weiman, "Binocular stereo via log-polar retinas," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 2488, pp. 309–320, 1995. [Online]. Available: http://dx.doi.org/10.1117/12.211983

[3] F. Li and Y. Zhou, "Disparity estimation based on frequency domain," *Shanghai Jiaotong Daxue Xuebao/Journal of Shanghai Jiaotong University*, vol. 33, no. 5, pp. 516–519, 1999.

[4] S. Shah and J. Aggarwal, "Depth estimation using stereo fish-eye lenses," *Proceedings of the IEEE International Conference on Image Processing*, vol. 2, pp. 740–744, 1994. [Online]. Available: http://dx.doi.org/10.1109/ICIP.1994.413669

[5] R. Manzotti, A. Gasteratos, G. Metta, and G. Sandini, "Disparity estimation on log-polar images and vergence control," *Computer Vision and Image Understanding*, vol. 83, no. 2, pp. 97–117, 2001. [Online]. Available: http://dx.doi.org/10.1006/cviu.2001.0924

[6] K. Schindler and H. Bischof, "The epipolar geometry of the log-polar image plane," *Proceedings of the International Conference on Pattern Recognition*, vol. 4, pp. 40–43, 2004.

[7] A. Basu, "Optimal discretization for stereo reconstruction," *Pattern Recognition Letters*, vol. 13, no. 11, pp. 813–820, Nov. 1992.

[8] A. Elnagar, "Optimal error discretization under depth and range constraints," *Pattern Recognition Letters*, vol. 19, no. 9, pp. 879–888, 1998. [Online]. Available: http://dx.doi.org/10.1016/S0167-8655(98)00062-2

[9] W. N. Klarquist and A. C. Bovik, "Fovea: A foveated vergent active stereo vision system for dynamic three-dimensional scene recovery," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 5, pp. 755–770, 1998. [Online]. Available: http://dx.doi.org/10.1109/70.720351

[10] Y. Zeevi and E. Shlomot, "Nonuniform sampling and antialiasing in image representation," *IEEE Transactions on Signal Processing*, vol. 41, no. 3, pp. 1223–1236, March 1993.