Proceedings of the 1998 IEEE
International Conference on Robotics & Automation
Leuven, Belgium • May 1998

Zoom Tracking

Jeffrey A. Fayman Oded Sudarsky Ehud Rivlin
Department of Computer Science
Technion—Israel Institute of Technology
Haifa 32000, Israel

Abstract

*In this paper we present a new active vision technique called **zoom tracking**. Zoom tracking is the continuous adjustment of a camera's focal length, to keep a constant-sized image of an object moving along the camera's optical axis. Two methods for performing zoom tracking are presented: a closed-loop visual feedback algorithm based on optical flow, and use of depth information obtained from an autofocus camera's range sensor.*

We show that the image stablity provided by zoom tracking improves the performance of algorithms that are scale varient, such as correlation-based trackers. While zoom tracking cannot totally compensate an object's motion, due to the effect of perspective distortion, an analysis of this distortion provides a quantitative estimate of the performance of zoom tracking.

1 Introduction

Over the past several years, the research field of *active vision* has been exploring the benefits and advantages of movable visual sensory systems, as possessed by biological systems, over passive systems. Moving systems have been shown to lead to improved robustness and the elimination of ill-posed conditions in several computer vision problems [1, 2, 3]. The degrees of freedom of interest in active vision systems include both *extrinsic* parameters (eye motion) and *intrinsic* parameters (eye configuration). We explore an interesting use of an intrinsic parameter unique to mechanical active vision systems: adjustable focal length, also known as *zoom*.

Applications of zoom include the ability to image a target with maximum resolution [19], determine depth [9, 14, 15] and minimize view degeneracies [22]. However, to our knowledge, no previous work dealt with *zoom tracking:* the use of zoom to stabilize the image of an object that moves along a camera's optical axis, i.e. to keep the image at a constant size.

An object moving towards a camera produces an expanding image, while an object moving away from the camera produces a contracting image. *Zoom tracking* compensates this expansion or contraction through focal length adjustments, thus stabilizing the object's imaged size.

Zooming can accurately compensate for object translation only for object points lying on one plane, the reference plane. All object points not on this plane will undergo perspective distortion. We provide an analysis and quantification of perspective distortion deriving an upper bound on the residual error of the zoom tracking process.

Our experiments demonstrate the effect that zoom tracking has on a typically scale varient alogorithm: template matching.

The remainder of the paper is organized as follows. Related work is reviewed in section 2. In Section 3 we discuss the motion model, imaging model and optical flow field used in our work. The necessary equations for focal length control are provided in section 4. In Section 5, we derive an upper bound on the error induced by perspective distortion. In section 6, we present experiments and conclusions are given in section 7.

2 Related Work

Since 1985, when active vision first appeared in the literature [1, 2, 3], the topic has received a dramatic increase in interest. Initial work focused on building active vision devices, and on understanding and transferring to these devices capabilities possessed by biological vision systems, such as saccades [6], smooth pursuit [6, 8], fixation [17], attention [20] and prediction [7]. Some work has appeared in the literature which explores uses of the zoom mechanism, a visual parameters not shared by biological vision systems. However, to our knowledge, the ability of the zoom mechanism to stabilize moving objects in the image has not been explored.

Cahn von Seelen et al. [19] present an algorithm that allows an active vision system to track targets at varying scale while decreasing the risk of template drift. Their algorithm is based on an adaptive correlation method that selectively updates the correlation template in response to zoom-induced scale changes. They assume an external agent controls zooming, and update the correlation template to compensate the resulting change of scale. This improves the performance of a template-matching tracker. Our approach is to take control of the focal length, and to adjust it to ensure that the scale of an object's image is minimally affected as the object moves.

0-7803-4300-x-5/98 $10.00 © 1998 IEEE

2.1 Camera models

Research in computer vision (and computer graphics) generally makes use of one of four different camera models: pin-hole, thin-lens, thick-lens, and geometric.

The pin-hole model is the simplest: it assumes that all light rays coming from the object focus through a single point (the focal point) onto the image plane [21, Chapter 1.4]. This geometrically simple model is widely used computer vision. However, it is only accurate if the focal length, i.e. the distance between the image plane and the focal point, is negligible compared to the object distance (e.g. in aerial photography).

The thin-lens model [5] assumes an infinitely thin lens, modeled by a plane. Light rays from the object hitting the plane at a particular angle continue after leaving the plane at a modified angle. This model handles aperture effects (such as depth-of-field) better than the pin-hole model. However, it is still inadequate for modeling zoom lenses at a close range.

The thick-lens model represents a lens by two planes called the principle planes. Light rays entering one plane at a specific angle travel parallel to the optical axis through to the second plane, from which they exit at another angle (similar to the thin-lens model). Ignoring aperture effects, this model is equivalent to the pin-hole model with the addition of a virtual axial motion [18], and it is sufficient for modeling zoom lenses, even at close range.

The geometric lens model [10, 14] is based on a complete geometric description of all of the glass elements of the lens. It is the most detailed and accurate of the models, and can be used to design lenses and to model such effects as optical abberations, lens flare etc. However, this model is extremely complicated to implement; the thick-lens model provides a sufficient approximation for our purposes.

2.2 Other uses of zoom

Wilkes et al. [22] use zoom to reduce the probability of view degeneracies. Degenerate views occupy a significant fraction of the viewing sphere surrounding an object. Furthermore, these degeneracies cannot be detected from a single viewpoint. Wilkes et al. choose a focal length that reduces the probability of view degeneracies, improving the performance of systems designed to recognize objects from a single arbitrary viewpoint.

Hosoda et al. [13] develop an interesting controller for robotic systems, which enhances visual servoing using zoom and a manipulator. Their controller is based on the observation that the vision system and the manipulator have complementary characteristics. Zoom can provide a wide range of image changes, but it cannot realize quick motions, and it has only one degree of freedom. On the other hand, the manipulator can move fast ("the hand is quicker than the eye"), but it cannot realize a wide range of image changes. They propose a complementary visual servoing controller of zoom and arm mechanisms.

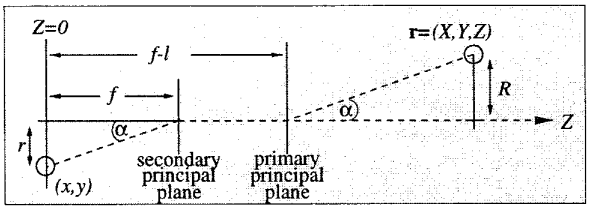

Figure 1: The imaging model

3 Preliminaries

3.1 The Imaging Model

Following the notation of Horn [11, Chapter 17], denote the Cartesian coordinates of a point P on a rigid body by $\mathbf{r} = (X, Y, Z)^T$. The body's motion is composed of a translational component $\mathbf{t} = (U, V, W)^T$ and a rotational component $\omega = (A, B, C)^T$. The velocity of point P is $\mathbf{V} = -\mathbf{t} - \omega \times \mathbf{r}$, or, in component form,

$$\begin{aligned} \dot{X} &= -U - BZ + CY \\ \dot{Y} &= -V - CX + AZ \\ \dot{Z} &= -W - AY + BX \end{aligned} \tag{1}$$

where the dot denotes differentiation with respect to time.

Assume that the camera is static. Define the (X, Y, Z) coordinate frame such that the Z axis coincides with the camera's optical axis and the image plane is at $Z = 0$. As Lavest et al. [14] point out, the pinhole camera model is inadequate for a zoom lens, and the thick-lens model must be used instead; however, the pinhole model can be used if the object is virtually translated along the optical axis by the distance l between the zoom lens's pricipal planes. According to Subbarao [18], this distance is

$$l = f_a + f_b - \frac{f_a f_b}{f} \tag{2}$$

where f is the current focal length of the zoom lens, and f_a, f_b are the focal lengths of the two lens groups corresponding to the two principal planes, determined during calibration. The sign of l is defined to be negative if the image plane is closer to the secondary principal plane than to the primary principal plane. See Figure 1.

The perspective projection is

$$x = \frac{Xf}{Z + l - f}, \quad y = \frac{Yf}{Z + l - f} \tag{3}$$

where $\mathbf{r} = (X, Y, Z)^T$ is an object point and $p = (x, y)$ is the corresponding image point. This implies

$$r = \frac{Rf}{Z + l - f} \tag{4}$$

where $r = \sqrt{x^2 + y^2}$ and $R = \sqrt{X^2 + Y^2}$ are the distances of the image point and the object point, respectively, from the optical axis.

3.2 The Motion Field and the Optical Flow Field

For optical-flow-based zoom tracking, let $(u, v) = (\dot{x}, \dot{y})$ denote the instantaneous velocity of the image point (x, y) under the perspective projection. This velocity can be obtained by taking derivatives of equation 3 with respect to time:

$$\begin{aligned} u &= u_{\text{trans}} + u_{\text{rot}} + u_{\text{zoom}} \\ v &= v_{\text{trans}} + v_{\text{rot}} + v_{\text{zoom}} \end{aligned} \tag{5}$$

where $(u_{\text{trans}}, v_{\text{trans}})$ is the translational component of the optical flow, $(u_{\text{rot}}, v_{\text{rot}})$ is the rotational component and $(u_{\text{zoom}}, v_{\text{zoom}})$ is the zooming component:

$$\begin{aligned} u_{\text{trans}} &= \frac{-Uf + xW}{Z + l - f} \\ u_{\text{rot}} &= \frac{1}{f}\left(Axy - B\left(x^2 + \frac{Zf^2}{Z + l - f}\right) + Cfy\right) \\ u_{\text{zoom}} &= \frac{\dot{f}x}{f}\left(1 + \frac{f^2 - f_a f_b}{f(Z + l - f)}\right) \\ v_{\text{trans}} &= \frac{-Vf + yW}{Z + l - f} \\ v_{\text{rot}} &= \frac{1}{f}\left(A\left(y^2 + \frac{Zf^2}{Z + l - f}\right) - Bxy - Cfx\right) \\ v_{\text{zoom}} &= \frac{\dot{f}y}{f}\left(1 + \frac{f^2 - f_a f_b}{f(Z + l - f)}\right) \end{aligned} \tag{6}$$

The differences between equations 5, 6 and those derived by Horn are due to the different choice of the $Z = 0$ plane, f not necessarily being equal to 1, and the use of the thick-lens model rather than the pinhole model.

Let $I(x, y, t)$ be the image intensity function, where t is time. The time derivative of I can be written as

$$\frac{dI}{dt} = \frac{\partial I}{\partial x}\frac{dx}{dt} + \frac{\partial I}{\partial y}\frac{dy}{dt} + \frac{\partial I}{\partial t} = (I_x, I_y) \cdot \vec{u} + I_t = \nabla I \cdot \vec{u} + I_t \tag{7}$$

where ∇I is the image gradient, the subscripts denote partial derivatives, and $\vec{u} = (u, v)$ is the projected motion field (the optical flow) at the point (x, y). If we assume $dI/dt = 0$, i.e. the image intensity does not vary with time [12], then

$$\nabla I \cdot \vec{u} + I_t = 0. \tag{8}$$

Let $\vec{u} = u_\perp + u_\parallel$ where $u_\perp$ is the normal flow and $u_\parallel$ is perpendicular to $u_\perp$. Because the image gradient ∇I is parallel to $u_\perp$ and perpendicular to $u_\parallel$, only $u_\perp$ can be determined by observing ∇I locally. (This is known as the aperture problem [11, Chapter 12].) Therefore

$$\begin{aligned} \nabla I \cdot \vec{u} + I_t &= \nabla I \cdot (u_\perp + u_\parallel) + I_t \\ &= \nabla I \cdot u_\perp + \nabla I \cdot u_\parallel + I_t \\ &= \nabla I \cdot u_\perp + I_t \\ &= 0. \end{aligned}$$

Consequently

$$u_\perp = -\frac{\partial I}{\partial t}\frac{\nabla I}{\|\nabla I\|^2}. \tag{9}$$

Various techniques [4] have been proposed to solve the aperture problem, that is, to recover $u_\parallel$ and integrate the measurements into a 2D flow field.

4 Adjusting the Focal Length

We explore two methods to find the required adjustment of focal length to maintain a constant image size. These methods are based on optical flow and depth from an autofocus camera's range sensor, and are described in sections 4.1 and 4.2, respectively.

Whichever method is chosen, zoom tracking keeps the image stabilized only for object points lying on one plane, called the *reference plane*, which is parallel to the image plane. All other points shift due to the effect of perspective distortion. This distortion is further explored in section 5.

4.1 Optical-flow based zoom tracking

In this section, the optical flow field $\vec{u}$ of section 3 is used in a feedback loop algorithm which adjusts the focal length to keep the size of the object constant in the image. u_{rad} refers to the radial component of the vector $\vec{u}$.

$$\dot{f} = -f\frac{u_{\text{rad}}}{r}\frac{1}{1 + \frac{f}{Z+l-f}\left(1 - \frac{f_a f_b}{f^2}\right)}. \tag{10}$$

The first part of equation 10, i.e. $\dot{f} = -fu_{\text{rad}}/r$, is identical to the equation obtained for zoom tracking based on the pinhole camera model. Using the thick-lens model, the factor $1/\left(1 + (f/(Z + l - f))\left(1 - f_a f_b/f^2\right)\right)$ is introduced. While the pinhole model cannot fully capture the behavior of a zoom lens, it can be used as a reasonable approximation when the object depth Z is unknown. For typical focal lengths, if a 10% error in zoom tracking accuracy is acceptable, then objects can be tracked from a distance of approximately 1 meter.

4.2 Autofocus-sensor based zoom tracking

The required change of focal length can also be calculated based on depth information from an autofocus sensor. Assume the only motion the object is undergoing is translation along the optical axis. Suppose the object is tracked from an initial distance of Z_1 to a final distance of Z_2, as in figure 2. The distances Z_1 and Z_2 are given by the autofocus sensor. Let $\mathbf{r}_1$ be a the initial position of a point P on the object, and let $\mathbf{r}_2$ be the final position of P. We would like to keep the point's image at a constant radial distance r from the image center as P moves from $\mathbf{r}_1$ to $\mathbf{r}_2$. This is accomplished by changing the distances between the image plane and the principal planes from given initial values

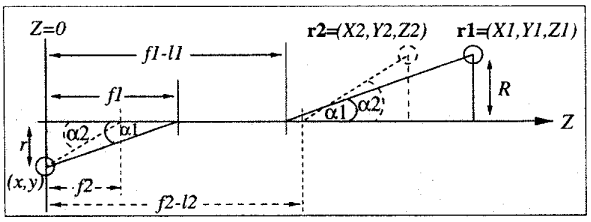

Figure 2: Perspective geometry

of $f_1 - l_1$ and f_1 (for the primary and secondary principal planes, respectively) to final values of $f_2 - l_2$ and f_2.

Let R denote the radial distance of the point from the optical axis. The thick lens model implies the following relationships:

$$r = \frac{f_1 R}{Z_1 + l_1 - f_1} \tag{11}$$

and

$$r = \frac{f_2 R}{Z_2 + l_2 - f_2} \tag{12}$$

where l_1 and l_2 are as in equation 2. Therefore

$$\frac{f_1 R}{Z_1 + l_1 - f_1} = \frac{f_2 R}{Z_2 + l_2 - f_2}. \tag{13}$$

Using equation 2 and solving equation 13 for f_2 and l_2, we obtain

$$\begin{aligned} f_2 &= \frac{f_1(f_a + f_b + Z_2) + \xi}{2(l_1 + Z_1)} \\ l_2 &= \frac{f_1(f_a + f_b)(f_a + f_b + Z_2) + f_b\xi + f_a\gamma}{f_1(f_a + f_b + Z_2) + \xi} \end{aligned} \tag{14}$$

where

$$\xi = \sqrt{f_1(-4f_a f_b(l_1 + Z_1) + f_1(f_a + f_b + Z_2)^2)}. \tag{15}$$

and

$$\gamma = -2f_b l_1 - 2f_b Z_1 + \xi \tag{16}$$

Equation 15 gives the focal length required to keep the object at the same visual size as its distance from the camera changes. Also, given a camera with a fixed zoom range $f_{\text{wide}} \rightarrow f_{\text{tele}}$, where f_{wide} is the shortest focal length and f_{tele} is the longest, equation 15 gives the depth range in which an object can be tracked: given the initial distance Z_1 and the initial focal length f_1, the object can be tracked from a minimum distance of

$$d_{\min} = (f_b/f_{\text{wide}} - 1)f_a - f_b + (l_1 + Z_1)f_{\text{wide}}/f_1) \tag{17}$$

to a maximum distance of

$$d_{\max} = (f_b/f_{\text{tele}} - 1)f_a - f_b + (l_1 + Z_1)f_{\text{tele}}/f_1), \tag{18}$$

where l_1 is given by equation 2.

5 Perspective Distortion

Zooming can accurately compensate for object translation only for object points lying on one plane, the reference plane. All object points not on this plane will undergo perspective distortion [16]. An analysis and quantification of perspective distortion provides an upper bound on the residual error of the zoom tracking process.

The thick-lens imaging model is given by equation 4. If point P is on the reference plane that is being tracked, then its image remains constant at $r = Rf_1/(Z_1 + l_1 - f_1) = Rf_2(Z_2 + l_2 - f_2)$. However, if P is not on the reference plane, but is at a distance z from this plane, then its image shifts from $r_1 = Rf_1/(Z_1 + z + l_1 - f_1)$ to $r_2 = Rf_2/(Z_2 + z + l_2 - f_2)$. The perspective distortion of point P is thus given by

$$r_2 - r_1 = \frac{Rf_2}{Z_2 + z + l_2 - f_2} - \frac{Rf_1}{Z_1 + z + l_1 - f_1} \tag{19}$$

If our goal is to stabilize the object's image, e.g. for recognition purposes, then perspective distortion is error or "noise" in the stabilization process (in addition to any error induced by inaccuracies in the calculation of optical flow or autofocus sensing). This error is maximized at points of the object that are farthest away from the reference plane and from the optical axis.

If the object has a bounding box, then the maximum error e is attained at the corner of the box that is farthest away from the reference plane. Let z be the maximum distance of any point in the box from the reference plane, and let R be the maximum distance of any point in the box from the optical axis. By equation 19, the perspective distortion error e of any point on the object is bounded by

$$e \leq R\left|\frac{f_2}{Z_2 + z + l_2 - f_2} - \frac{f_1}{Z_1 + z + l_1 - f_1}\right|. \tag{20}$$

Equation 20 specifies the maximum error that can be expected, given object size and depth difference between two object positions (ignoring zoom range constraints and inaccuracies of flow calculations or autofocus sensing). Alternatively, given a maximum tolerable error and an object size, initial object distance, and initial focal length, equation 20 can be used to calculate the range of object distances that may be tracked without exceeding this error.

6 Experiments

To demonstrate the usefulness of zoom tracking, we show its effect on an algorithm that is scale varient: template matching. We implemented a standard *sum of squared differences* (SSD) tracker. The idea behind template matching is to find the location of a particular object in an image by searching the image for instances of a second, smaller image called a 'template' which contains the object. The template matching algorithm compares the template with the image at different image locations and finds the location in the image which best matches the template.

Correlation provides the basis of template matching. For each image location, a similarity measure is computed indicating how well the template matches the image at that location. The image location which provides the maximal similarity measure is selected as the location of the object in the image.

The SSD tracking method is a classic example of an algorithm that is scale varient. Indeed, researchers have proposed a variety of techniques to handle this problem such as updating the correlation template in response to scale changes in an image sequence [19] and providing templates at different scales and orientations [23].

The tracker, which computes the similarity of patches P in an image $I(x,y)$ to a template $T(x,y)$ is defined as follows:

$$E(i,j) = \sum_{x,y \in P} (T(x,y) - I(x-i, y-j))^2 . \qquad (21)$$

This experiment was conducted in the Intelligent Systems Laboratory at the Technion–Israel Institute of Technology. The system consisted of a Silicon Graphics Indy workstation connected to a Canon VC-C1 communication camera(Figure 3(a)). The Canon camera is a small pan-tilt unit with a serial interface allowing computer control of the extrinsic parameters of pan and tilt as well the intrinsic parameters of focus, zoom and iris.

6.1 Performance with and without zoom tracking

We measured the performance of our sum of squared difference tracker as it tracked a moving picture of a light bulb. In order to effect the motion, we attached the picture to a robotic manipulator. The manipulator effected translation of the picture along the optical axis of the camera such that the distance from the camera to the picture ranged from 400mm to 700mm (a span of 300mm). The template was captured at the halfway location of 550 with a focal length of 21mm. The image of the light bulb picture at the near, middle, and far distances illustrated in figure 3(a,b,c respectively).

Figure 4(a) presents the results of the tracker when no focal length adjustments are made. Figure 4(b) shows the results of the tracker when zoom tracking is used. In the case of no zoom tracking, we expect that the closer the target location is to the location that the template was captured, the lower the error will be. This is indeed the case as can be seen in the graph. The smallest error occurs when the distance between the current target location and the location of template capture is close to zero. When we introduce zoom tracking to the scenario, we find that the performance of the sum of squared difference tracker improves as shown in figure 4(b). (The spikes in both graphs are due to the auto focus mechanism on the Canon camera).

7 Conclusions

Zoom tracking exploits the adjustable focal length parameter in an active vision system to compensate the scale changes of a target moving along a camera's optical axis. We presented two methods by which zoom tracking can be performed: a closed-loop feedback algorithm that measures the optical flow's radial component and then adjusts the focal length to negate it; and utilization of depth information from an autofocus sensor.

Zoom tracking yields exact results for object points lying on a reference plane. Points not lying on this plane shift due to perspective distortion. A quantitative measure of this distortion was derived, and used as an upper bound on the residual error of zoom tracking.

Experiments demonstrated that zoom tracking improves the performance of the scale-variant algorithm, template matching for objects whose scale changes due to motion along the cameras optical axis.

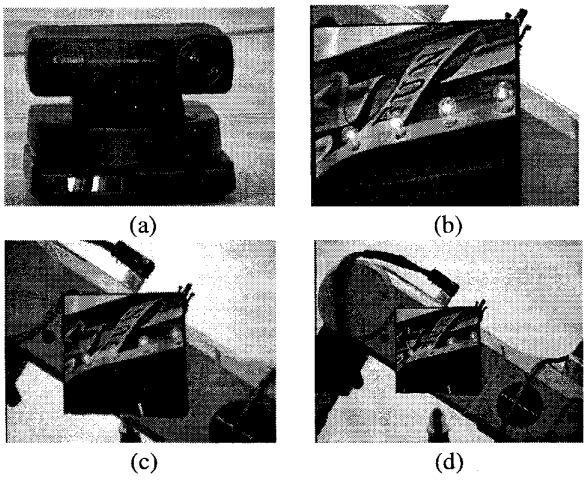

Figure 3: (a) The Canon VC-C1 communication camera; (b) View of light bulb picture at 400mm; (c) View of light bulb picture at 550mm; (d) View of light bulb picture at 700mm.

References

[1] Y. Aloimonos, I. Weiss, and A. Bandyopadhyay. Active vision. In *International Journal on Computer Vision*, pages 333–356, 1987.

[2] R. Bajcsy. Active perception vs passive perception. In *Proceedings of the Third IEEE Workshop on Computer Vision*, pages 55–59, Bellaire, Michigan, 1985.

[3] D.H. Ballard. Animate vision. *Elsevier Artificial Intelligence*, 48:57–86, 1991.

[4] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. In *Proc. DARPA Image Understanding Workshop*, pages 121–130, 1981.

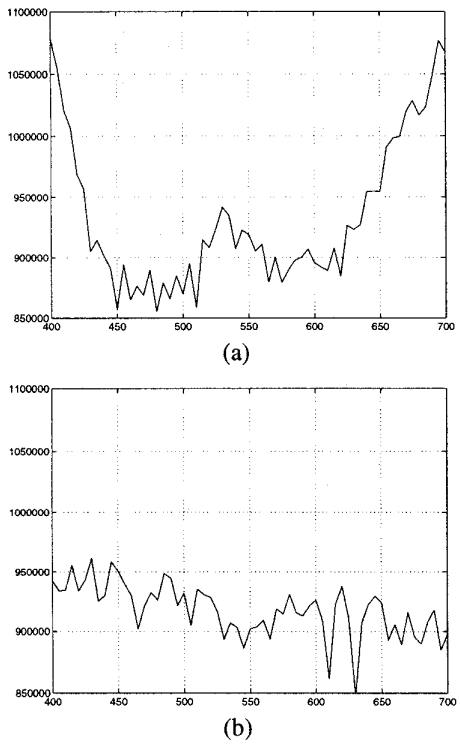

Figure 4: (a) Template matching without zoom tracking; (b) Template matching with zoom tracking. The horizontal axis represents object depth and the vertical axis represents matching error.

[5] M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, Oxford, 1993.

[6] K.J. Bradshaw, P.F. McLauchlan, I. D. Reid, and D.W. Murray. Saccade and pursuit on an active head/eye platform. *Image and Vision Computing*, 12(3):155–163, April 1994.

[7] C. Brown. Prediction and cooperation in gaze control. *Biological Cybernetics*, 63(1):61–70, May 1990.

[8] D. Coombs and C. Brown. Real-time smooth pursuit tracking for a moving binocular head. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 23–28, Champaign, Illinois, June 1992.

[9] C. Delherm, J.M. Lavest, M. Dhome, and J.T. Lapresté. Dense reconstruction by zooming. In *4th European Conference on Computer Vision*, pages 427–438, Cambridge, UK, April 1996.

[10] W. Heidrich, P. Slusallek, and H.P. Seidel. An image-based model for realistic lens systems in interactive computer graphics. In *Proceedings Graphics Interface '97*, pages 68–75, Kelowna, B.C., 1997.

[11] B.K.P. Horn. *Robot Vision*. MIT Press, 1986.

[12] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:189–203, 1981.

[13] K. Hosoda, H. Moriyama, and M. Asada. Visual servoing utilizing zoom mechanism. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 178–183, Nagoya, Aichi, Japan, May 1995.

[14] J.M. Lavest, G. Rives, and M. Dhome. Three-dimensional reconstruction by zooming. *IEEE Transactions on Robotics and Automation*, 9(2):196–207, April 1993.

[15] J. Ma and S.I. Olsen. Depth from zooming. *Journal of the American Optical Society*, 7(10):1883–1890, October 1990.

[16] B.G. Mobasseri. Focal length and compression of space. In *Proceedings of the 11th International Conference on Pattern Recognition*, pages 686–687, 1992.

[17] K. Pahlavan, T. Uhlin, and J.-O. Eklundh. Dynamic fixation. In *Fourth International Conference on Computer Vision*, pages 412–419, Berlin, Germany, May 1993.

[18] M. Subbarao. Parallel depth recovery by changing camera parameters. In *Proceedings of the International Conference on Computer Vision*, pages 149–155, Tampa, Florida, 1988.

[19] U.M. Cahn von Seelen and R. Bajcsy. Adaptive correlation tracking of targets with changing scale. Technical report, GRASP Laboratory Technical Report, June 1996.

[20] W.Y.K. Wai and J.K. Tsotsos. Directing attention to onset and offset of image events for eye-head movement control. In *Proceedings of the IEEE Workshop on Visual Behaviors*, pages 79–84, Seattle, Washington, June 1994.

[21] A. Watt. *Fundamentals of Three-Dimensional Computer Graphics*. Addison Wesley, 1989.

[22] D. Wilkes, S. Dickinson, and J. Tsotsos. A quantitative analysis of view degeneracy and its use for active focal length control. In *Proceedings of the International Conference on Computer Vision*, Cambridge, Massachusetts, 1995.

[23] S. Yoshimura and T. Kanade. Fast template matching based on the normalized correlation by using multiresolution eigenimages. In *International Conference on Intelligent Robots and Systems*, pages 2086–2093, Munich, Germany, September 1994.