ESTIMATING EGO MOTION BY FIXATION CONTROL OF MOUNTED ACTIVE CAMERAS

Akihiro Sugimoto

National Institute of Informatics
Tokyo 101-8430, Japan
sugimoto@nii.ac.jp

Wataru Nagatomo, Takashi Matsuyama

Graduate School of Informatics
Kyoto University
Kyoto 606-8501, Japan

ABSTRACT

We propose a method for incrementally estimating ego motion by two mounted active cameras. Our method independently controls the two cameras during the ego motion so that each camera automatically fixates its optical axis to its own fixation point. The correspondence of the fixation point over two frames together with the correspondence of lines nearby the fixation point gives us sufficient constraints to determine the ego motion. Two cameras do not have to share the common field of view in this case whereas in stereo vision they have to do. Namely, our method allows the diverging viewing-lines of two cameras that are prohibited in stereo vision.

1. INTRODUCTION

Computing three-dimensional camera motion from image measurements is one of the fundamental problems in computer vision and robot vision, and it has many applications. In the robot vision, for example, mobile robot navigation and docking require the robot localization, the process of determining and tracking the position (location) of mobile robots relative to their environments [2]. In the wearable computer, on the other hand, understanding where a person was and where the person is/was going is a key issue [1, 6, 7, 9] for providing the person with useful and timely information.

Most successful approaches[1] to estimating the position and motion of a moving robot use landmarks such as ceiling lights, gateways or doors [12, 13, 15] and are usually based on the framework of stereo vision [3, 4, 10, 11]. When we employ the stereo vision algorithm, however, we have to make two cameras share the common field of view and, moreover, establish feature correspondences between the images captured by two cameras. This kind of processing has difficulty in its stability. In addition, we have another problem in using the stereo vision framework. Namely, though the accuracy of the estimation is well known to highly depend on the baseline of two cameras, keeping the baseline wide is hard when we mount cameras on a robot or wear cameras. Therefore, the accuracy of the estimation of motion is limited if we employ the stereo vision algorithm.

In this paper, we propose a method for incrementally estimating ego motion using two mounted active cameras where the *fixation control*, the camera control in which a camera automatically fixates its optical axis to a selected point (called the *fixation point*) in 3D, plays a key role. Our method applies the fixation control independently to each active camera. We call this camera control the *binocular independent fixation control* (Fig. 1). The correspondence of the fixation point over two frames together with the correspondence of lines nearby the fixation point gives us sufficient constraints to determine the ego motion in 3D.

In the binocular independent fixation control, each camera automatically fixates its optical axis to its own fixation point in 3D and two fixation points are not necessarily the same. This indicates that the two cameras need not share the common field of view. The viewing lines of the two cameras are divergent in this case in contrast to stereo vision where convergence is always imposed on two viewing lines. Moreover, in the binocular independent fixation control, the estimation accuracy becomes independent of the baseline of two cameras and is expected to become higher than the case where we use the stereo vision algorithm. This can be understood as follows. If we assume that we set a camera at each fixation point and that the optical axis of each camera

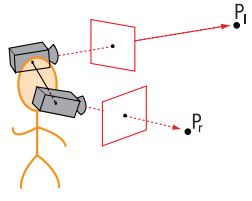


Fig. 1. Binocular independent fixation control.

[1] Several approaches to ego-motion estimation are carefully compared in [14].

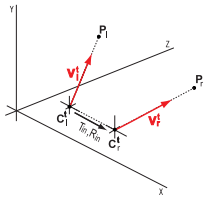


Fig. 2. Relationship between the projection centers and the fixation points at time t.

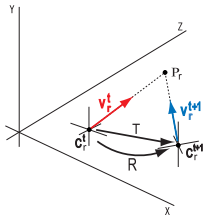


Fig. 3. Geometry based on the fixation correspondence of the right camera.

is toward a robot or a person, then the binocular independent fixation control can be regarded as the situation where we apply the stereo vision framework to estimating the position of the robot or the person from the two fixation points. The baseline in this case is identical with the distance of the two fixation points. This means that the estimation accuracy is independent of the baseline of two mounted cameras and that selecting fixation points as far as possible from each other allows the estimation accuracy to become high.

2. GEOMETRIC CONSTRAINTS ON EGO MOTION

We here derive geometric constraints on ego motion based on information obtained during the binocular independent fixation control. Between two mounted cameras, i.e., a right camera and a left camera, we set the right camera is the base. Moreover, for simplicity, we assume that the orientation of the camera coordinates does not change even though we change pan and tilt of the camera for the fixation control. This means that only the ego motion causes changes in orientation and translation of the camera coordinates. We also assume that the ego motion is identical with the motion of the base-camera coordinates. We thus develop a method to estimate the motion of the right-camera coordinates.

We assume that the extrinsic parameters between the two cameras as well as the intrinsic parameters of each camera are calibrated in advance. Namely, we let the translation vector and the rotation matrix to make the left-camera coordinates identical with the right-camera coordinates be $\boldsymbol{T}_{\rm in}$ in the left-camera coordinates and $R_{\rm in}$ in the right-camera coordinates, respectively. $\boldsymbol{T}_{\rm in}$ and $R_{\rm in}$ are both assumed to be known.

2.1. Constraints from fixation correspondence

The fixation control gives us the correspondence of the viewing lines of a camera toward the fixation point over time-series frames. We call this correspondence a *fixation correspondence*. The fixation correspondence enables us to derive a constraint on the ego motion.

Let the projection centers of the left camera and the right camera be C_ℓ^t and $C_{\rm r}^t$ in 3D at time t. We assume that the both cameras have their own fixation points P_ℓ and $P_{\rm r}$. We denote by $\boldsymbol{v}_{\rm r}^t$ the unit vector from $C_{\rm r}^t$ to $P_{\rm r}$ in the right-camera coordinates at time t. We see that $\boldsymbol{v}_{\rm r}^t$ represents the viewing line of the right camera toward the fixation point at time t. We also denote by $\boldsymbol{v}_\ell^t$ the unit vector from C_ℓ^t to P_ℓ in the left-camera coordinates at time t (Fig. 2).

We first focus on the right camera. We assume that the projection center of the right camera moves from $C_{\rm r}^t$ to $C_{\rm r}^{t+1}$ in 3D due to the ego motion from time t to $t+1$ (Fig. 3). We also assume that the rotation and the translation of the right-camera coordinates incurred by the ego motion are expressed as rotation matrix R in the right-camera coordinates at time t and translation vector $\boldsymbol{T}$ in the world coordinates. We remark that the orientation of the world coordinates is assumed to be obtained by applying rotation matrix R_0^{-1} to the orientation of the right-camera coordinates at time t. Our aim here is to derive constraints on R and $\boldsymbol{T}$.

It follows from the fixation correspondence of the right camera that

$$\lambda R_0 \boldsymbol{v}_{\rm r}^t \;\; = \;\; \lambda' R_0 R \boldsymbol{v}_{\rm r}^{t+1} + \boldsymbol{T},$$

where λ and λ' are positive unknown constants. This equation is rewritten by

$$\det \left[\; R_0 \boldsymbol{v}_{\rm r}^t \;\middle|\; R_0 R \boldsymbol{v}_{\rm r}^{t+1} \;\middle|\; \boldsymbol{T} \; \right] \;\; = \;\; 0, \qquad (2.1)$$

which gives the constraint on the ego motion, R and $\boldsymbol{T}$, derived from the fixation correspondence of the right camera.

On the other hand, $\boldsymbol{v}_\ell^t$ in the left-camera coordinates at time t is identical with $R_{\rm in}\boldsymbol{v}_\ell^t$ in the right-camera coordinates at time t. The rotation R of the right-camera coordinates from time t to $t+1$ causes the translation $-R_0(R-I)R_{\rm in}\boldsymbol{T}_{\rm in}$ of the left-camera coordinates in the world coordinates where I is the 3×3 unit matrix. This yields

$$\det \left[R_0 R_{\rm in} \boldsymbol{v}_\ell^t | R_0 R R_{\rm in} \boldsymbol{v}_\ell^{t+1} | \boldsymbol{T} - R_0(R-I)R_{\rm in}\boldsymbol{T}_{\rm in} \right] = 0. \qquad (2.2)$$

(2.2) is the constraint on the ego motion derived from the fixation correspondence of the left camera.

(2.1) and (2.2) are the constraints on the ego motion in 3D derived from the fixation correspondences obtained by

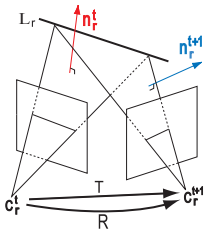


Fig. 4. Geometry based on the line correspondence of the right camera.

the binocular independent fixation control. When we have estimated the ego motion up to time t, we know R_0. Then, the unknown parameters in (2.1) and (2.2) are R and $\boldsymbol{T}$. We see that (2.1) and (2.2) give homogeneous quadratic constraints on R and $\boldsymbol{T}$ respectively.

Naive parameter counting: Suppose first that the 3-D positions of the fixation points are known. Then it is easy to see that the fixation directions of the camera restrict the rig to a 2-parameter locus (consider a simple example here). However, there are two additional degrees of freedom: in practice, we do not know the 3-D positions of the fixation points (we know only their directions), so there is an additional degree of freedom for each fixation point. Hence the rig has four degrees of freedom, which explains why we have only two constraints.

2.2. Constraints from line correspondence

Ego motion has 6 degrees of freedom: 3 for a rotation and 3 for a translation. The number of constraints on the ego motion derived from two fixation correspondences, on the other hand, is two ((2.1) and (2.2)). We therefore need to derive more constraints to estimate the ego motion.

We employ lines nearby the fixation point to obtain other constraints on the ego motion. This is because

(i) we find many lines in the indoor scene, for example, the boundaries between walls and a ceiling, the boundaries of windows and those of doors,

(ii) we can easily and accurately detect lines with less computation by using the Hough transformation,

(iii) we can easily establish line correspondences over time-series frames due to their spatial extents, and

(iv) constraints on the ego motion derived from line correspondences depend only on the rotation as seen in detail below.

We first focus on the right camera. Let the projection center of the right camera be $C_{\rm r}^t$ in 3D. We then assume that we establish the correspondence of images of line $L_{\rm r}$ in 3D over time t and $t+1$, where line $L_{\rm r}$ is selected nearby the fixation point of the right camera. Line $L_{\rm r}$ is called a *focused line* in this paper. We denote by $\boldsymbol{L}_{\rm r}$ the unit direction vector of the focused line $L_{\rm r}$ in the world coordinates[2]. Observing a line in 3D is identical to determining the plane in 3D on which both the projection center at the observation time and the line exist. We thus obtain the unit normal vector of the plane. For the focused line $L_{\rm r}$, this unit vector in the right-camera coordinates at time t is denoted by $\boldsymbol{n}_{\rm r}^t$ (Fig. 4).

From the relationship of the orientations among the world coordinates, the right-camera coordinates at time t and the right-camera coordinates at time $t+1$, we see that $\boldsymbol{n}_{\rm r}^t$ and $\boldsymbol{n}_{\rm r}^{t+1}$ are expressed as $R_0\boldsymbol{n}_{\rm r}^t$ and $R_0R\boldsymbol{n}_{\rm r}^{t+1}$ in the world coordinates. Since $R_0\boldsymbol{n}_{\rm r}^t$ and $\boldsymbol{L}_{\rm r}$ are orthogonal, and $R_0R\boldsymbol{n}_{\rm r}^{t+1}$ and $\boldsymbol{L}_{\rm r}$ are also orthogonal, we obtain the following constraint on the ego motion from the line correspondence over the two frames captured by the right camera:

$$\mu_{\rm r}\boldsymbol{L}_{\rm r} \quad = \quad (R_0\boldsymbol{n}_{\rm r}^t) \times (R_0R\boldsymbol{n}_{\rm r}^{t+1}), \qquad (2.3)$$

where $\mu_{\rm r}$ is an unknown non-zero constant and depends on the focused line.

In the similar way, the line correspondence of the left camera gives us the constraint on the ego motion.

$$\mu_\ell\boldsymbol{L}_\ell \quad = \quad (R_0R_{\rm in}\boldsymbol{n}_\ell^t) \times (R_0RR_{\rm in}\boldsymbol{n}_\ell^{t+1}), \qquad (2.4)$$

where $\boldsymbol{L}_\ell$ denotes the unit direction vector, in the world coordinates, of focused line L_ℓ in the left-camera case and μ_ℓ is an unknown non-zero constant depending on the focused line L_ℓ. $\boldsymbol{n}_\ell^t$ denotes the unit normal vector, in the left-camera coordinates at time t, of the plane determined when the focused line L_ℓ is observed by the left camera.

We see that the translation factors of the ego motion are not involved in the constraints, (2.3) and (2.4), derived from the line correspondence in each camera. We also see that these constraints are linear homogeneous with respect to R and the non-zero constants.

2.3. Estimation of rotation and translation

As investigated in Section 2.2, the constraints derived from line correspondences depend only on the rotation of ego motion. We can thus divide the ego-motion estimation into two steps: the rotation estimation and the translation estimation.

The first step is the rotation estimation of the ego motion. We suppose that we have correspondences of n focused lines over two time-series frames. Then, we have $n+3$ unknowns (n are from scale factors and 3 are from rotation) whereas we have $3n$ constraints in this case. Therefore, we can estimate the rotation of the ego motion if we have correspondences of more than two focused lines. To

[2]We assume here that the unit direction vector of a focused line in the world is known. The vector, however, can be estimated from (2.3) during the motion estimation. Namely, we can compute $\boldsymbol{L}_{\rm r}$ (with $\|\boldsymbol{L}_{\rm r}\| = 1$) from (2.3) because we know R_0 and R if we have estimated the ego motion up to time $t+1$.

be more concrete, we form a simultaneous system of nonlinear equations that consists of the constraints derived from line correspondences and the orthogonality constraints, i.e., $RR^\top = I$, and then apply a nonlinear optimization algorithm such as the Levenberg-Marquart method to solve the system. In general, a nonlinear system has multiple solutions and the local minimum trap problem is serious. In our case, however, employing redundant line correspondences allows us to avoid being trapped in a local minimum. This is because the constraints derived from line correspondences are linear with respect to the unknown parameters (scale factors and R) and because such linear redundancy excludes spurious solutions[3] (spurious solutions do not satisfy additional linear constraints).

When we finish estimating the rotation of the ego motion, we can move to the second step: the translation estimation. The unknowns are now just the translation factors. The constraint derived from the fixation correspondence thus becomes homogeneous linear with respect to unknown parameters. Hence, we can determine the translation up to scale from two fixation correspondences with only linear computation[4].

3. ALGORITHM

Based on the discussion above, we present here the algorithm for estimating ego motion based on the binocular independent fixation control.

Step 0: Detect a fixation point by each camera and select focused lines for each camera. Set $t = 1$.

Step 1: Compute $\boldsymbol{v}_{\mathrm{r}}^t$, $\boldsymbol{v}_{\ell}^t$, $\boldsymbol{n}_{\mathrm{r}}^t$, and $\boldsymbol{n}_{\ell}^t$.

Step 2: For $i = \mathrm{r}, \ell$, do the followings.

(a) Control camera i, and compute $\boldsymbol{v}_i^{t+1}$ and $\boldsymbol{n}_i^{t+1}$. If camera i cannot capture its own fixation point, go to Step (b). Otherwise, goto Step 3.

(b) Detect a next fixation point and select new focused lines, and then return to Step 1.

Step 3: Estimate the camera rotation from (2.3), (2.4) and $RR^\top = I$.

Step 4: Estimate the camera translation from (2.1) and (2.2).

[3] In addition to this, we have another reason to employ the redundancy in line correspondence. In the case where ego motion is just a translation and where the projection center moves on the the plane defined by a focused line and the projection center, the constraints derived from the line correspondence become the identical equation. Namely, the constraints do not make sense and no independent constraint on the ego motion is obtained. Employing the redundancy in line correspondence prevents us from falling into such cases.

[4] Whenever we estimate the translation of the ego motion over two frames, we have one unknown scale factor. The trilinear constraints [5] on corresponding points over three frames enable us to adjust the unknown scales with only linear computation.

Fig. 5. Active vision sensor.

(a) with a wide view (b) with a top view

Fig. 6. Ego-motion trajectory.

Step 5: Set $t = t + 1$, and return to Step 2.

4. EXPERIMENTS

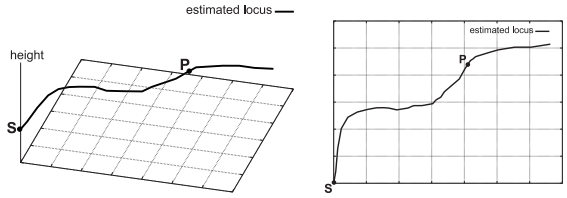
We employed two off-the-shelf cameras (EVI-G20 from Sony) as active cameras. To verify the potential applicability of the proposed method, we applied to them the binocular independent fixation control to estimate ego motion.

We set up an active vision sensor where two cameras with the baseline of about 27cm were mounted on the stage of a tripod (Fig. 5). We then calibrated the intrinsic and extrinsic parameters of the two cameras with the method proposed by Zhang [16]. The size of images captured by each camera was 640×480 pixels.

We moved the active vision sensor in the scene. The trajectory of the right-camera motion is shown in Fig. 6. The length of the trajectory was about 6m. We marked 35 points on the trajectory and regarded them as samples during the ego motion. (In other words, 35 points were sampled during the ego motion of about 6m.) We then applied the binocular independent fixation control only to the samples to estimate the ego motion.

(a) left-camera image (b) right-camera image

Fig. 7. Example of images captured by the two cameras during the ego motion.



(a) 3D representation (b) top-view representation

Fig. 8. Estimated trajectory of the ego motion.

In each image captured by each camera at the starting point of the motion, we manually selected a point to set as the fixation point. During the estimation, we updated fixation points 8 times. This updating was also conducted by hand. We used two focused lines for each camera (we thus used four focused lines in total). In detecting lines, we applied the Hough transformation to the edges detected from each image. Fig. 7 shows an example of image pair captured by the right and left cameras at a marked point. We see that little field of view of the two cameras is common[5]. We remark that the fixation point (the black circle) and two focused lines (the black thick lines) are overlaid onto the images in Fig. 7. In this example, the focused lines accidentally go through the fixation point in each image.

Under the above conditions, we estimated the right-camera motion at each marked point. Fig. 8 shows the trajectory of the right-camera motion that was obtained by concatenating the estimated motions at the marked points. We note that S means the starting point of the motion.

The height from the floor was almost accurately estimated over the trajectory. In fact, the estimated height was almost constant. As for the component parallel to the floor, however, the shape of the former part (from S to P in Fig. 8) of the estimated trajectory fairly coincides with that of the actual trajectory whereas the latter part (after P) of the estimated trajectory has great aberration from the actual trajectory. We have two reasons that may cause this aberration. One is the incorrect estimation of the motion at P and the other is the effect of the estimation error at P upon the subsequent estimations. In other words, since the motion is incrementally estimated, the accumulation of estimation errors and an incorrect estimation at just one marked point cause aberration. The estimation error can be caused by errors in the fixation correspondence or errors in the line detection. Calibration errors of the two cameras also may cause estimation errors.

As we see above, we may conclude that though the function to reduce the accumulation errors in estimation should be incorporated into the proposed method for the accurate estimation, our initial experiment demonstrates the potential applicability of the proposed method.

[5]It is therefore hard to apply the stereo vision algorithm to such input image pairs. In addition, since there were not enough textures in the scene, only few features could be detected. This also causes difficulty in application of the stereo vision algorithm.

5. DISCUSSION ON FIXATION CONTROL

To realize the fixation control, the computer should autonomously select a point in 3D as the fixation point of a camera and then control the camera so that the camera automatically fixates its optical axis to the point. Moreover, updating fixation points during the estimation is required. We discuss here these problems.

5.1. Fixation-point detection

In the static scene, properties listed below are required for a point that is selected as the fixation point of a camera. The point satisfying the properties is suitable for a fixation point and the computer has to automatically detect such a point. How to formalize the criterion in selecting such a point is the central issue. The method to find landmarks for mobile robot navigation [13] may be helpful for this problem.

- Actual existence in 3D. (Two twisted lines in 3D, for example, form a point in the image as the intersection of their image lines. Such a point, however, should not be selected as the fixation point since it does not exist in 3D.)
- Easiness in identification in the image. A fixation point should have a distinguished property to identify in the image for the accurate fixation control.
- Having margins to fixate in the physical control of the camera. The point should not easily disappear from the field of view of the camera during the fixation control.
- Having as a large distance as possible from the other fixation point. As addressed above, the estimation accuracy depends on the distance between two fixation points in the binocular independent fixation control.

5.2. Camera control by template matching

When we select a point as the fixation point in the current frame, to realize the fixation control we should first identify the position where the point is in the next frame, and then head the optical axis of the camera toward the new position of the point. The template matching enables the computer to effectively conduct these procedures.

When template-matching based tracking is applied to the fixation point, appearance changes around the fixation point due to ego motion may cause failure in accurate fixation control. With allowing affine warping of the template, matching may become robust against such changes.

In capturing images during ego motion, motion blur may occur. How to stabilize images is also an important issue. Moreover, treating time-lag in camera action to realize a

smooth fixation control is indispensable. The method proposed by [8] is promising for this problem.

5.3. Updating fixation point

To estimate ego motion in the scene, the binocular independent fixation control should continue without any interruption. When a robot or a person widely moves in the scene the case occurs during the motion where the camera cannot capture the current fixation point due to its physical constraint, i.e., the angle limitation of pan and tilt of the camera. In such a case, a new fixation point should be selected: updating the fixation point is necessary.

The point that is newly selected as the fixation point should also satisfy the properties listed in Section 5.1. We remark that in the binocular independent fixation control, each camera selects its new fixation point independently at different time. This is because two cameras are independently controlled.

In the implementation, before a camera cannot capture the current fixation point due to its physical constraint, we keep a point that can be a new fixation point in the image of the camera. We replace the current fixation point by the point to obtain a new fixation point as soon as the camera loses the current fixation point. We then apply the fixation control with respect to the new fixation point. In this way, the estimation of ego motion continues without any interruption even though a robot or a person widely moves in the scene.

6. CONCLUDING REMARKS

We proposed a method, the binocular independent fixation control, for incrementally estimating ego motion by two mounted active cameras. Our method independently controls the two active cameras so that each camera automatically fixates its optical axis to its own fixation point. The correspondence of the fixation point over two frames together with the correspondence of lines nearby the fixation point gives us sufficient constraints to determine the ego motion in 3D.

In the binocular independent fixation control, two cameras need not share the common field of view because each camera fixates its optical axis to its own fixation point in 3D and because two fixation points are not necessarily the same. In using binocular cameras, only the framework of stereo vision has been studied for decades where the viewing lines of two cameras are convergent. In contrast, the binocular independent fixation control stands in the other framework where the viewing lines of two cameras are divergent. We believe that the binocular independent fixation control will open a new door to the diverging viewing-lines paradigm in using multiple cameras.

Developing a fully automatic system that realizes the binocular independent fixation control is the urgent future work. Eliminating accumulation errors in estimating ego motion and improving the accuracy of the estimation are also included in the future work.

Acknowledgements This work is supported by by Grant-in-Aid for Scientific Research of the Ministry of Education, Culture, Sports, Science and Technology of Japan under the contraction of 13224051 and 14380161.

7. REFERENCES

[1] H. Aoki, B. Schiele and A. Pentland: *Realtime Personal Positioning System for Wearable Computers*, Vision and Modeling Technical Report, TR-520, Media Lab. MIT, 2000.

[2] J. Borenstein, B. Everentt and L. Feng: *Navigating Mobile Robots: Systems and Techniques*, A. K. Peters, Ltd., Wellesley, MA, U.S.A., 1996.

[3] A. J. Davison and D. W. Murray: Mobile Robot Localisation Using Active Vision, *Proc. of ECCV*, Vol. 2, pp. 809–825, 1998.

[4] A. J. Davison and D. W. Murray: Simultaneous Localization and Map-Building Using Active Vision, *IEEE Trans. on PAMI*, Vol. 24, No. 7, pp. 865–880 (2002).

[5] R. Hartley and A. Zisserman: *Multiple View Geometry in Computer Vision*, Cambridge Univ. Press, 2000.

[6] M. Kourogi, T. Kurata and K. Sakaue: A Panorama-Based Method of Personal Positioning And Orientation And Its Real-Time Applications for Wearable Computers, *Proc. of Int. Symposium on Wearable Computers*, pp.107-114, 2001.

[7] S. W. Lee and K. Mase: Incremental Motion-Based Location Recognition, *Proc. of Int. Symposium on Wearable Computers*, pp. 123–130, 2001.

[8] T. Matsuyama, S. Hiura, T. Wada, K. Murase and A. Yoshikawa: Dynamic Memory: Architecture for Real Time Integration of Visual Perception, Camera Action and Network Communication, *Proc. CVPR*, pp. 728–735, 2000.

[9] W. W. Mayol, B. Trdoff and D. W. Murray: Wearable Visual Robots, *Proc. of Int. Symposium on Wearable Computers*, pp. 95–102, 2000.

[10] N. Molton and M. Brady: Practical Structure and Motion from Stereo When Motion is Unconstrained, *Int. J. of Computer Vision*, Vol. 39, No. 1, pp. 5–23 (2000).

[11] D. W. Murray, I. D. Reid and A. J. Davison: Steering and Navigation Behaviours Using Fixation, *Proc. of British Machine Vision Conference*, 1996.

[12] R. Sim and G. Dudek: Mobile Robot Localization from Learned Landmarks, *Proc. of IEEE/RSJ Conf. on Intelligent Robots and Systems*, 1998.

[13] S. Thrun: Finding Landmarks for Mobile Robot Navigation, *Proc. of IEEE Int. Conf. on Robotics and Automation*, pp. 958–963, 1998.

[14] T. Y. Tian, C. Tomasi and D. J. Heeger: Comparison of Approaches to Egomotion Computation, *Proc. of CVPR*, pp. 315–320, 1996.

[15] M. Werman, S. Banerjee, S. D. Roy and M. Qiu: Robot Localization Using Uncalibrated Camera Invariants, *Proc. of CVPR*, Vol. 2, pp. 353–359, 1999.

[16] Z. Zhang: A Flexible New Technique for Camera Calibration, *IEEE Transactions on PAMI*, Vol. 22, No. 11, pp. 1330–1334 (2000).