2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

A New Approach for Active Stereo Camera Calibration

Hyukseong Kwon, Johnny Park and Avinash C. Kak

Abstract— **By active stereo we mean a stereo vision system that allows for independent panning and tilting for each of the two cameras. One advantage of active stereo in relation to regular stereo is the former's wider effective field of view; if an object is too close to the camera baseline, the depth to the object can still be estimated accurately by panning the cameras appropriately. Another advantage of active stereo is that it can yield a larger number of depth measurements simultaneously for each position of the platform on which the camera system is mounted. Panning and tilting over a large angular range, while being the main reason for the advantages of active stereo, also make it more challenging to calibrate such systems. For a calibration procedure to be effective for active stereo, the estimated parameters must be valid over the entire range of the pan and tilt angles. This paper presents a new approach to the calibration of such vision systems. Our method is based on the rationale that an active stereo calibration procedure must explicitly estimate the locations and the orientations of the pan and tilt rotating axes for the cameras through a closed-form solution. When these estimates for the axes are combined with the homogeneous transform relationships that link the various coordinate frames, we end with a calibration that is valid over a large variation in the pan and tilt angles.**

***Index Terms*— active stereo camera, camera calibration, robotic vision.**

I. Introduction

Computer vision has an important role in robotic 3D map building. Over the years, various researchers have explored different approaches to combine information from different cameras for constructing 3D maps of the interior. Some of the earliest approaches fused images from a single camera but using different viewpoints corresponding to the adjacent positions of a mobile platform, and the fusion algorithms used were based on either binocular stereopsis or on the concepts of structure from motion. Later, investigators started mounting stereo camera heads on mobile robots. These cameras were either fixed in position and orientation with respect to the robot platform, or, in more advanced systems, the two cameras were fixed in relation to a pan-and-tilt head. Now potentially 3D information about the environment could be obtained from every location of the robot.

Obviously, mounting the cameras on pan-and-tilt heads allowed for greater freedom in data collection (although, at least with respect to panning, the effect achieved could also be replicated by turning the robot itself). Most of these approaches required the cameras to be carefully calibrated. There were some structure-from-motion studies that tried to generate 3D information without camera calibration by direct matching of the images recorded at the different locations of the robot. More recently, interest has focused on map building tools with stereographic camera heads in which the two cameras can be panned and tilted independently. Stereo vision using such camera heads is generally known as active stereo.

One of the advantages of active stereo includes a wider effective field of view, the larger field of view made possible by the independent panning and the tilting of the two cameras. The same reason also yields a larger number of depth measurements simultaneously for each position of the platform on which the camera system is mounted. Independent panning and tilting of the cameras individually also makes it possible to accumulate multiple depth measurements for the same point in the environment. This then allows the selection of the best possible depth measurement. In this manner, the map building process can exhibit greater immunity to noise and other possible artifacts that may be generated by, say, poor illumination.

However, there is a cost associated with these advantages of active stereo: the camera calibration process becomes more difficult since the output of the calibration procedure must stay valid over a large angular variation in the pan and tilt angles. With traditional camera calibration, we believe that the main reason that the calibration parameters computed for one fixed position of the cameras cease to be valid for other positions is the variability introduced by the unknown locations and the orientations for the pan and tilt axes. As the cameras are panned and tilted, these axes, especially if they are non-intersecting, create varying spatial effects that cannot be captured with a traditional calibration matrix for a fixed pan and tilt angle.

We have therefore devised a new calibration approach that directly estimates the pan and tilt axes. When this explicitly computed information is combined with the estimated homogeneous transforms related the different coordinate frames, we end up with calibration parameters that are valid over a large variation in pan and tilt angles.

The remainder of this paper is organized as follows: In the following section, Section II, we present an overview of the computer vision literature related to active camera calibration. In Section III, our new active stereo camera calibration method is proposed. Then we present experimental results for validating our algorithm in Section IV. Finally, we conclude in Section V.

II. Literature Survey

Researchers have used various simplifying assumptions to tackle the problem of active stereo camera calibration. These simplifying assumptions fall in two categories: In

H. Kwon, J. Park, A. C. Kak are with the School of Electrical and Computer Enginerring, Purdue University, West Lafayette, IN. 47907, U.S.A. {hyukseon,jpark,kak}@purdue.edu

the first category belong the methods that are based on the assumption that the center of the camera coordinate frame is either located on the panning axis, or on the tilting axis, or the intersection of both axes. The second category includes the more general methods, that is the methods that do not place this restriction on the location of the origin of the camera coordinate from.

We will start with the work of Basu [1] which is one of the first expositions to show how a single camera could calibrate itself by seeing stable contours in a scene through multiple views at different pan and tilt angles. It has provided mathematical expressions for estimating the focal length and the camera lens center from the images recorded. These formulae use the pan and the tilt angles that are provided by the corresponding encoders attached to the camera head. In a subsequent report, another contribution from the same group [2] showed how the intrinsic parameters of a camera, such as its focal length and the center of the image frame, could be calculated more robustly if we also considered image data at different rolling angles of a camera.

In their work dealing with active stereo camera calibration, Li *et al.* estimates the fundamental matrix [3]. Since the system described allows for the camera to be panned independently (thus allowing for one vergence angle to another), the main problem addressed by their work on calibration is the development of the relationship between a fundamental matrix and the panning angles of the two cameras. From the standpoint of how versatile their calibration method is, central to this work is the assumption that the origin of the coordinate frame for each camera is located on the panning axis. These results in a significant simplification of the calibration procedure because now one can safely assume that the baseline always stays constant. Based on that assumption, they suggest their own analytic form of the fundamental matrix.

A fundamental matrix based framework has also been used by Brooks *et al.* to obtain the calibration parameters from two different recordings of the stereo images, each recording made at a different location of the camera head [4]. The two different recording location yield four separate pairings of the cameras, and thus four separate fundamental matrices. The resulting calibration parameters lead to a final fundamental matrix form for relating the coordinates of the corresponding pixels in the two images of a stereo pair.

In the second category, we have the works of Schluns *et al.* [5] and Fellenz *et al.* [6] that are based on a somewhat more realistic assumption that the panning axis is parallel to the y-axis of the camera coordinate frame, but not located on it. Even though their methods show how we can estimate the real panning angle and the 2-dimensional location of the camera coordinate frame with respect to the panning axis, they do not take into account the possibility of errors caused by the difference between the simplified 2D-location and the real 3D-location of the camera coordinate frame with respect to the panning axis.

As another approach, Davis and Chen have proposed a method in which a pre-calibrated static camera is used to calibrate an active camera system [7]. The calibration procedure proposed by the authors does not require that the panning and the tilting axes intersect at a point; that is, the two axes are allowed to be non-coplanar. By taking advantage of the pre-calibrated static camera network, their method provides the calibration of pan-tilt cameras using the camera projection model onto the image plane. Even though the proposed calibration framework is general, in the sense that the authors allow for non-intersecting pan and tilt axes and do not require that these axes pass through center of the camera coordinate frame, the authors have not fully explained how to go about solving the equations that the above minimization problem can be transformed into.

All of the contributions cited so far have dealt with the case of active stereo consisting of two cameras mounted on a fixed rig. However, a more interesting case — the case that is particularly relevant to the work reported here — consists of mounting a pair of cameras on a stereo head that is free to execute its own panning and tilting motions. Hayman *et al.* has addressed the calibration of such active stereo systems [8]. When stereo cameras are mounted on a stereo head with its own pan and tilt motions, it is best to think of the multiple coordinate frames as constituting a kinematic chain connecting the world frame to the camera frame. The proposed method estimates all the rotation angles associated with each of linkages in the kinematic chain for each camera. From these angles, they can then calculate the verging, panning, and the tilting angles. However their method has a limitation for the exact calibration because it is implicitly based on the assumption that the panning and the tilting axes intersect and that this point of intersection coincides with the center of the coordinate frame. The work of Shih *et al.* [9] also addresses the calibration of active stereo systems in which the camera head is free to execute its own rotations. Their work goes beyond what was reported by [8] in the sense that their cameras are allowed to zoom with motorized focus control. Their camera model includes the dependencies of such intrinsic parameters as the radial distortion on the setting of the focus control. Finally, Collins and Tsin have considered the calibration for an outdoor active camera system [10]. Although this work does not deal directly with active stereo, it is nonetheless relevant to us because the procedure can be extended to the case of stereo. Their system uses optical flows generated by rotating and zooming motions to estimate intrinsic parameters. This work also uses 3D scene landmarks to estimate extrinsic parameters.

III. Active Stereo Camera Calibration

We first define an "active-stereo coordinate frame." that is attached to the center of the baseline that connects the origins of the two camera coordinate frames for the initial orientations of the cameras, as shown in Fig. 1. The cameras are at their initial orientations when their point along the directions that are used for measuring the pan and the tilt angles. In other words, at initial orientation, the pan and tilt angles are considered to be zero. The positive direction of x-

axis is from the bisecting point O of the baseline to the center O_R of the right rectified camera coordinate frame C_R^{RECT}, that of z-axis is the same viewing direction as both the left and the right rectified camera coordinate frames, and that of y-axis is determined by the right-handed coordinate rule. When the cameras' pan/tilt angles change, the stereo baseline will also change. However the active stereo coordinate frame remains fixed according to the initial orientations for the cameras.

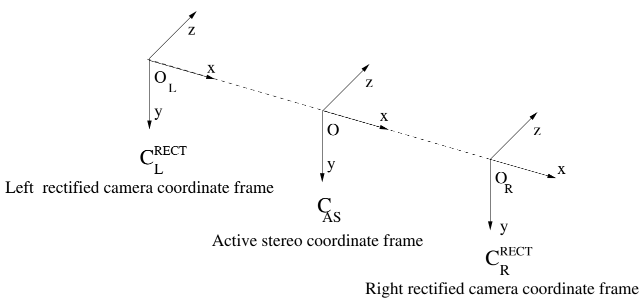

Fig. 1. An active stereo coordinate frame

As mentioned in the Introduction, a key step in our approach to active stereo camera calibration is the estimation of the locations and the orientations of the pan and tilt axes. Towards that end, we analyze the incremental change in the relationship (as represented by a homogeneous transform) between the coordinate frames corresponding to the two cameras for a certain number of predefined pan/tilt angles of the cameras. As we will show in this section, the incremental changes can be translated into a geometrical model for the pan and tilt axes. The model then yields an estimate for the locations and the orientations of the pan and tilt axes. The homogeneous transforms that relate the two camera coordinate frames are estimated using a conventional calibration algorithm. More specifically, we use the "calibration toolbox" functions from the MATLAB library [11].

A. Obtaining homogeneous transforms by estimating the locations and the orientations of rotating axes

We will now describe how we set up the relationship between the coordinate frames for the two cameras taking into account the possible non-intersecting panning and tilting axes for each of the cameras. Our active stereo system has four different types of motions — left panning, left tilting, right panning, and right tilting. Let H_0 denote the currently applicable homogeneous transform matrix that relates the two camera frames. Now let's introduce one or more of the four motions we listed. The question then becomes as to how the homogeneous transform that relates the two camera frames changes as a function of those motions. Each type of motion will be represented by its own homogeneous transform matrix: H_{LP} for left panning, H_{LT} for left tilting, H_{RP} for right panning, and H_{RT} for right tilting. Let H_{INT} stand for the homogeneous transform that relates the two camera frames after some or all of these motions. (See Fig. 2). H_{INT} can be expressed as

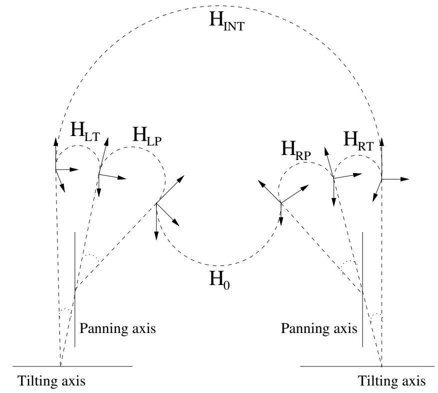

Fig. 2. The integrated homogeneous transform H_{INT}.

$$H_{INT} = \begin{bmatrix} R_{INT} & T_{INT} \\ 0 & 1 \end{bmatrix} = H_{RT} \cdot H_{RP} \cdot H_0 \cdot H_{LP} \cdot H_{LT} \quad (1)$$

The rotation part of the transform that applies after the motions is given by

$$R_{INT} = R_{RT} \cdot R_{RP} \cdot R_0 \cdot R_{LP} \cdot R_{LT} \quad (2)$$

and translation vector by

$$T_{INT} = R_{RT} \cdot R_{RP} \cdot R_0 \cdot R_{LP} \cdot T_{LT} + R_{RT} \cdot R_{RP} \cdot R_0 \cdot T_{LP} + R_{RT} \cdot R_{RP} \cdot T_0 + R_{RT} \cdot T_{RP} + T_{RT} \quad (3)$$

The goal of the calibration procedure is to find H_{INT}. The elements of this matrix will be estimated by considering separately each of the two possible types of rotations for each camera, for a total of four types of rotations. In what follows, we will focus on just one type of rotation. It could, for example, be left-panning rotation.

For the rotation type and the camera selected, the camera is moved to two different positions that will be denoted C_1 and C_2, with C_0 being the initial position of that camera. Obviously, C_0 corresponds to the transform matrix H_0. We need two rotational positions so that cross-product of the difference of the same direction vector in the two positions will yield the desired rotation axis. Each rotation will be through θ degrees (Fig. 3).

Let the origins of the coordinate frames C_0, C_1, and C_2 be O_{C_0}, O_{C_1}, and O_{C_2} (O_{C_0}, O_{C_1}, and O_{C_2} are represented in the world 3-space, not in the homogeneous coordinate). The homogeneous transform that relates C_0 to C_1, on the one hand, and C_1 and C_2, on the other, consists of three consecutive homogeneous transforms: a translation from the location of the current camera coordinate frame to the rotating axis ($H_{T_{-O_{ROT}}}$), a rotation around the rotating axis with a desired angle (H_{R_θ}), and a translation to the new location of the camera coordinate frame ($H_{T_{O_{ROT}}}$). To do this, we need to find the orientation of the rotating axis and the location of the point O_{ROT} which we call the center of the rotating axis in Fig. 3 with respect to the location of the

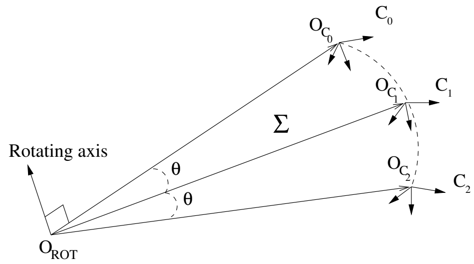

Fig. 3. Movement of the camera coordinate frame by rotation

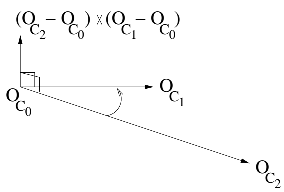

Fig. 4. Estimation of the orientation of the rotating axis using a cross product

origin of the camera coordinate frames. If we represent this homogeneous transform by $H_k(\theta)$ where k is LP, RP, LT, or RT used in Eq. 1, 2, and 3, we can write

$$\begin{aligned} H_k(\theta) &= H_{T_{-O_{ROT}}} \cdot H_{R_\theta} \cdot H_{T_{O_{ROT}}} \\ &= \begin{bmatrix} I & O_{ROT} \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} R_\theta & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} I & -O_{ROT} \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} R_\theta & (I - R_\theta)O_{ROT} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_k & T_k \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{4}$$

Now, our issue is to obtain R_θ and O_{ROT}. Note from Fig. 3 that the origins of the three frames C_0, C_1, and C_2 are on the same plane that we will denote by Σ. And the rotating axis is vertical to Σ. Now, we can obtain the unit orientation u_{ROT} of the rotating axis using a cross product because the normal vector for the plane Σ has the same orientation as u_{ROT} as illustrated in Fig. 4. Thus, u_{ROT} can be expressed as:

$$u_{ROT} = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} = \frac{(O_{C_2} - O_{C_0}) \times (O_{C_1} - O_{C_0})}{\|(O_{C_2} - O_{C_0}) \times (O_{C_1} - O_{C_0})\|_2} \tag{5}$$

Now, from the obtained unit vector u_{ROT}, the rotation matrix R_θ in the homogeneous transform $H_k(\theta)$ from Eq. 4 is obtained using the axis-angle rotation as in the following equation ($c\theta$ stands for $\cos\theta$ and $s\theta$ stands for $\sin\theta$):

$$R_\theta = \begin{bmatrix} c\theta + b_{c_x} & u_x \cdot a_{c_y} - u_z \cdot s\theta & u_x \cdot a_{c_z} + u_y \cdot s\theta \\ u_x \cdot a_{c_y} + u_z \cdot s\theta & c\theta + b_{s_y} & u_y \cdot a_{c_z} - u_x \cdot s\theta \\ u_x \cdot a_{c_z} - u_y \cdot s\theta & u_y \cdot a_{c_z} + u_x \cdot s\theta & c\theta + b_{c_z} \end{bmatrix} \tag{6}$$

where $a_{c_i} = u_i \cdot (1 - c\theta)$, $a_{s_i} = u_i \cdot (1 - s\theta)$, $b_{c_i} = u_i^2 \cdot (1 - c\theta)$, and $b_{s_i} = u_i^2 \cdot (1 - s\theta)$ where $i = x, y, z$.

Only information we need more is the location of the point O_{ROT} in order to acquire a translation vector from the camera coordinate to the rotating axis and vice versa. The point O_{ROT} can be calculated from the following derivation.

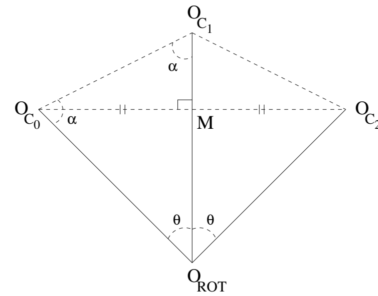

Fig. 5. Estimation of the position of O_{ROT}

From Fig. 5, we have several isosceles triangles which are $\triangle O_{C_0}O_{C_1}O_{C_2}$, $\triangle O_{ROT}O_{C_0}O_{C_1}$, $\triangle O_{ROT}O_{C_1}O_{C_2}$, and $\triangle O_{ROT}O_{C_0}O_{C_2}$. From $\triangle O_{C_0}O_{C_1}O_{C_2}$, we can find M as a middle point of a line segment $\overline{O_{C_0}O_{C_2}}$ and from the right triangle $\triangle O_{C_0}O_{C_1}M$, we can estimate the angles α and $\overline{\theta}$, the estimated angle of θ, using the following equations:

$$\alpha = \cos^{-1}\left(\frac{|\overline{O_{C_0}O_{C_1}}|}{|\overline{O_{C_1}M}|}\right) \tag{7}$$

and

$$\overline{\theta} = \pi - 2 \cdot \alpha \tag{8}$$

The value of $\overline{\theta}$ can be different from the value of θ which is provided by the encoder because of mechanical errors. For the rotation matrix with an arbitrary angle ϕ, we will use the modified angle considering the ratio of $\overline{\theta}$ over θ, which means that a rotating angle ϕ from the encoder will be considered as $(\overline{\theta}/\theta) \cdot \phi$ to use in R_ϕ. Finally, the intersecting point O_{ROT} can be acquired through the following:

$$O_{ROT} = O_{C_1} + (M - O_{C_1}) \cdot \frac{|\overline{O_{ROT}O_{C_1}}|}{|\overline{O_{C_1}M}|} \tag{9}$$

where $\overline{O_{ROT}O_{C_1}} = \frac{\frac{1}{2}\overline{O_{C_0}O_{C_1}}}{\cos\alpha}$ and $\overline{O_{C_1}M} = \overline{O_{C_0}O_{C_1}} \cdot \cos\alpha$. Now, we have the complete $H_k(\theta)$ from Eq. 4 by using the obtained R_θ and O_{ROT}.

The calibration procedure based on the above calculations is summarized in Fig. 6. When we calibrate at the initial positions of the cameras, the initial homogeneous transform H_0 is estimated using the conventional stereo calibration method as mentioned earlier. The coordinates of the frame origins O_{C_0}, O_{C_1} and O_{C_2} that are needed for u_{ROT} are also estimated through conventional stereo calibration.

B. Active Stereo Coordinate Frame

As mentioned earlier, though the pan/tilt angles of cameras are changed, the location and the orientation of the active stereo coordinate frame is fixed at the center of the baseline of the initial positions of the both cameras. From the relation described in the above subsection, we can transform a coordinate vector for a scene point in any camera coordinate

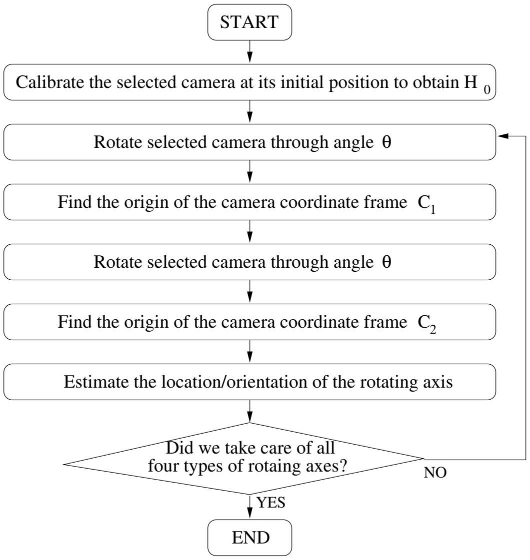

Fig. 6. Overview of our active stereo calibration procedure

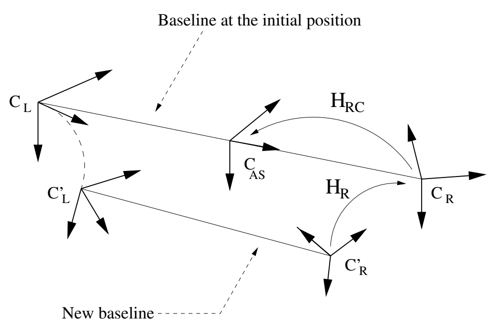

Fig. 7. Homogeneous transforms to the active stereo coordinate frame

frame to another coordinate vector for the same scene point in the active stereo coordinate frame. As shown in Fig. 7, we have a coordinate vector $v_{C'_R}$ in the moved right camera coordinate frame C'_R, then we can obtain the coordinate vector $v_{C_{AS}}$ in the active stereo coordinate frame F_{AS} using the following equation:

$$\begin{bmatrix} v_{C_{AS}} \\ 1 \end{bmatrix} = H_{RC} \cdot H_R \cdot \begin{bmatrix} v_{C'_R} \\ 1 \end{bmatrix} \qquad (10)$$

where H_{RC} is a homogeneous transform from C_R to C_{AS} and H_R is a homogeneous transform from C'_R to C_R.

IV. Experimental Results

All experiments presented in this paper were carried out with two Sony EVI-D100 cameras, each attached to its own pan-tilt motor control.

We will now show the estimated location and orientation of each rotating axis in relation to the corresponding camera coordinate frame. Next, we will show the accuracy of our calibration method by comparing the 3D locations of the intersection points on the grid of a calibration pattern with the results obtained by other stereo calibration methods. By "other methods" we mean methods that make assumptions about the locations and the orientations of the rotating axes.

	Location(O_{ROT}, unit:mm)	Orientation (u_{ROT})
Left Panning	(6.2060, 0.9195, -3.9420)	(0.1529, -0.9882, 0.0102)
Right Panning	(3.0158, 0.3785, 1.1890)	(0.0756, -0.9895, 0.1231)
Left Tilting	(1.3527, 0.5648, -5.6028)	(0.9671, -0.1264, 0.2207)
Right Tilting	(-0.3265, -0.5959, -1.5231)	(0.9622, 0.1098, -0.2492)

TABLE I
ESTIMATION OF THE LOCATION AND THE ORIENTATION OF EACH ROTATING AXIS WITH RESPECT TO THE CORRESPONDING CAMERA COORDINATE FRAME.

For obtaining the calibration information for the left panning axis, we used two different values for θ, $5°$ and $10°$. In accordance with our earlier explanation, when θ is $5°$, the camera is panned through $5°$ and $10°$. By the same token, when θ is $10°$, the camera is panned through $10°$ and $20°$. As for the right panning axis, we use the same angles, but in the opposite directions. This was done purely for convenience, as it allowed us to keep the calibration pattern in the same place for the two cameras.

For obtaining the calibration information related to the tilting axis, we again carried out the experiments for two different values of θ, $-5°$ and $5°$. For the first choice, each camera was rotated through $-5°$ and $-10°$, and for the second choice through $5°$ and $10°$. The choice of the two different values used for θ, $-5°$ and $5°$, was dictated by the fact that the vertical-plane view angle is narrower than the horizontal-plane view angle for the cameras. The θ angles chosen allowed us to keep the calibration pattern at the same place for both values.

The estimated locations and the orientations for all four rotating axes are presented in TABLE I. So how do we know that these results are accurate, let alone more accurate than what would be produced by other calibration procedures for active stereo? To answer this question, we compared the 3D coordinates of the world points as produced by the calibration matrix using the parameters shown in TABLE I vis-a-vis the 3D coordinates for the same world points as obtained by a direct calibration of the cameras for many different pan and tilt angles. It is important to bear in mind that storing the calibration information for a large number of pan and tilt angles is not a viable option for real-world applications of an active stereo system. We are using direct calibration here merely to test our method for active stereo calibration and to analyze its accuracy.

In order to assess the accuracy of our calibration procedure and to also compare it with the previously proposed calibration procedures for active stereo systems, we will use the following notation for the methods used for the calculation of the 3D world coordinates:

- M_0 : a method based on a direct calibration of the cameras for various pan and tilt angles;
- M_1 : a method based on the most strict assumption that the center of the camera coordinate frame, the panning axis, and the tilting axis all intersect at the

		M_1	M_2	M_3
Left Panning	x	0.5968	1.2053	0.4862
	y	1.2411	0.8944	0.9053
	z	1.3481	0.6312	0.7419
Right Panning	x	6.2828	6.7913	1.6137
	y	6.6030	6.4352	2.7308
	z	1.8943	1.3247	0.7288
Left Tilting	x	1.3814	1.3725	0.7042
	y	8.1757	7.2187	0.9987
	z	0.6273	0.6301	0.2168
Right Tilting	x	6.6589	6.6002	3.5025
	y	3.4626	1.8439	1.8182
	z	1.4575	1.4553	0.6671

TABLE II
COMPARISON OF THE AVERAGES OF ABSOLUTE PERCENTAGE ERRORS AMONG M_1, M_2, AND M_3 (UNIT: %).

same point;

M_2 : a method based on the assumption that the panning axis is parallel to the x-axis of the camera coordinate frame and the tilting axis is parallel to the y-axis of the camera coordinate frame, but we will not assume that the panning and the tilting axes intersect;

M_3 : a method based on the new calibration procedure proposed by us; the calibration matrix explicitly takes into account the calibration parameters shown in TABLE I.

In order to compare errors among M_1, M_2, and M_3, we estimated the locations of the intersections of the grid of the calibration pattern in the world coordinate, then calculated the absolute percentage error e_{abs} for each such point using the following formula:

$$e_{abs} = \left| \frac{v_{M_i} - v_{M_0}}{v_{M_0}} \right| \times 100 \tag{11}$$

where v_{M_i} is a chosen coordinate for a specific scene point as measured by the method M_i for M_i $(i = 1, 2, 3)$ and v_{M_0} is the value for the same coordinate as measured by the direct calibration based method M_0. TABLE II shows the averages of the absolute percentage errors in each coordinate for each type of rotation. In this experiment, we tested with 24 stereo pairs of calibration pattern images.

As we can see from TABLE II, our method provides the best performance among the tested methods for most cases because our approach estimates the homogeneous transforms without geometrical assumptions pertaining to the locations and orientations of the rotation axes of the cameras.

V. CONCLUSIONS AND DISCUSSIONS

In this paper, we have presented a new active stereo camera calibration method that makes no assumptions regarding the locations and/or orientations of the panning and tilting axes of the cameras. The calibration procedure estimates these axes independently for each of the cameras. By expressing the overall homogeneous transform (that describes the coordinate frame in which the two cameras are embedded) in terms of the component transforms attached to the individual panning and tilting motions, we showed how to obtain a calibration procedure that explicitly takes into account the locations and the orientations of the panning and tilting axes. Although the decomposition itself is not novel, the new contribution we have made lies in showing how by rotating each camera through two different angles around each of the axes, we can separately obtain the locations and the orientations of the pan and tilt axes for the cameras. That strategy, as we have shown, results in a more accurate calibration of the whole system. The active camera stereo calibration procedures in the past have made assumptions regarding the locations and the orientations of the panning and tilting axes. We believe that those assumptions reduce the accuracy of those procedures. The superior experimental accuracy we have demonstrated with our new procedure bear this out. Also, the closed form solution proposed in this paper allows our calibration method to be computationally much more efficient compared to other methods based on iterative optimization procedures.

VI. ACKNOWLEDGMENT

We gratefully acknowledge the contribution of the Ford Motor Company.

REFERENCES

[1] A. Basu, Active Calibration of Cameras: Theory and Implementation, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 25, no. 2, 1995, pp 256-265.

[2] A. Basu and K. Ravi, Active Camera Calibration Using Pan, Tilt and Roll, *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, vol. 27, no. 3, 1997, pp 559-566.

[3] F. Li and M. Brady and C. Wiles, Fast Computation of the Fundamental Matrix for an Active Stereo Vision System, *4th European Conference on Computer Vision*, Cambridge, UK, Apr. 1996, pp 157-166.

[4] M. J. Brooks and L. de Agapito and H. Q. Huyynh and L. Baumela, Towards Robust Metric Reconstruction via a Dynamic Uncalibrated Stereo Head, *Image and Vision Computing*, vol. 16, no. 14, 1998, pp 989-1002.

[5] K. Schiluns and W. Fellenz and A. Koschan and M. Teschner, A Modular 10-DOF Vision System for High-Resolution Active Stereo, The University of Auckland, Tamaki Campus, Computing and Information Technology Research Computer Vision Unit, no. CITR-TR-2/CS-TR-141, 1997.

[6] W. Fellenz and K. Schluns and A. Koschan and M. Teschner, An Active Vision System for Obtaining High Resolution Depth Information, *7th International Conference on Computer Analysis of Images and Patterns*, Kiel, Germany, Sep. 1997, pp 726-733.

[7] J. Davis and X. Chen, Calibrating Pan-Tilt Cameras in Wide-Area Surveillance Networks, *IEEE International Conference on Computer Vision*, Beijing, China, Oct. 2003, pp 144-149.

[8] E. Hayman and J. Knight and D. W. Murray, Self-Alignment of an Active Head from Observations of Rotation Matrices, *IEEE International Conference on Pattern Recognition*, Barcelona, Spain, Sep. 2000, pp 1080-1084.

[9] S.-W. Shih and Y.-P. Hung and W.-S. Lin, Calibration of an Active Binocular Head, *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans*, vol. 28, no. 4, 1998, pp 426-442.

[10] R. Collins and Y. Tsin, Calibration of an Outdoor Active Camera System, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Ft. Collins, CO, Jun. 1999 pp 528-534.

[11] J.-Y. Bouguet, Camera Calibration Toolbox for Matlab, Aug. 2005, Available from http://www.vision.caltech.edu/bouguetj/calib_doc/.