DEVELOPMENT OF KNOWLEDGE-BASED MEASUREMENT WITH MONOCULAR VISION

De Xu

Key Laboratory of Complex Systems and Intelligent Science, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, P. R. China
Email: sdxude@yahoo.com

Abstract: In this tutorial, knowledge-based visual measure methods with monocular vision system are investigated. The approaches of visual measure based on knowledge known in advance can be classified to four basic categories according to the types of knowledge, such as point position, line, size or shape knowledge, and motion knowledge. The principle for each category is shortly introduced. Furthermore, visual measure methods based on environment information are also discussed. Finally, conclusion and outlook for the development of knowledge-based visual measure are presented.

Keywords: Visual measure, positioning, pose estimation, knowledge-based method, monocular vision, robot vision

1. INTRODUCTION

Visual measure has been widely applied to many fields such as navigation and localization of mobile robots, object positioning and manipulation of industrial robots, and etc. Generally, visual measure is realized with stereovision method, which relies on the intrinsic and extrinsic parameters of the cameras. The coordinates of a point in Cartesian space is measured via geometric relation between the projective lines formed with the principal points of the cameras and the imaging points.

Obviously, the knowledge of the objects to be measured is not utilized in stereovision method. It is sure that the knowledge of the objects or environment known in advance is very helpful to estimate the objects' positions. For example, when we watch TV or see film, the projective relations and the environment knowledge such as building, car, street, person and etc. help us to have the relative positions of different objects in flat pictures. Another example also gives us evidence for the importance of knowledge in visual positioning. If someone stood on the top of a hill at dark night, it would be very difficult to correctly estimate the distance of a lamp far away in village. In this case, human observes the lamp with two eyes at night as same as at daylight. But the environment knowledge is not available. In other words, the position of an object can not be well estimated for human without the help of references.

In this tutorial, the discussion focuses on the investigation of knowledge-based visual measure with monocular vision system. The rest of this tutorial is organized as follows. In Section 2, four basic kinds of visual measure methods based on the knowledge of point position, line, size or shape, and motion are discussed separately, including their principles, developments and applications. In Section 3, visual measure methods based on environment information, such as general information and environment map, are introduced. Conclusion and outlook are given in Section 4.

2. KNOWLEDGE-BASED VISUAL MEASURE

There are many examples to utilize the knowledge of objects or environments to estimate the positions of objects in the frame of vision system. They fall into four basic categories, such as point position, line, size or shape, and motion knowledge-based.

2.1 Based on Position Knowledge

If the positions of n points in Cartesian space are known in advance, and the intrinsic parameters of a camera are also available, the position and orientation of the camera with respect to the scene object formed by n correspondent points can be determined. It is the famous perspective-n-point (PnP) problem (Fishler and Bolles,1981; Horaud *et a*l, 1989).

Assume the distortion in the camera lens is negligible. The four parameters intrinsic model of a camera is described as given in (1).

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c / z_c \\ y_c / z_c \\ 1 \end{bmatrix} \tag{1}$$

where (u, v) are the coordinates of a point in an image; (u_0, v_0) denote the image coordinates of the camera's principal point; (x_c, y_c, z_c) are the coordinates of a point in the camera frame; k_x and k_y are the magnification coefficients from the imaging plane coordinates to the image coordinates.

Assume that the camera frame is denoted as C, and the world frame as W. The transformation from C to W is known as the extrinsic parameters for the camera.

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} {}^c n_{wx} & {}^c o_{wx} & {}^c a_{wx} & {}^c p_{wx} \\ {}^c n_{wy} & {}^c o_{wy} & {}^c a_{wy} & {}^c p_{wy} \\ {}^c n_{wz} & {}^c o_{wz} & {}^c a_{wz} & {}^c p_{wz} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \qquad (2)$$

where $[x_w, y_w, z_w]$ are the coordinates of a point in the world frame W; ${}^c n_w=[{}^c n_{wx}\ {}^c n_{wy}\ {}^c n_{wz}]^{\mathrm{T}}$, ${}^c o_w=[{}^c o_{wx}\ {}^c o_{wy}\ {}^c o_{wz}]^{\mathrm{T}}$, and ${}^c a_w=[{}^c a_{wx}\ {}^c a_{wy}\ {}^c a_{wz}]^{\mathrm{T}}$ are the direction vectors of the X, Y and Z-axis for the frame W expressed in the frame C; ${}^c p_w=[{}^c p_{wx}\ {}^c p_{wy}\ {}^c p_{wz}]^{\mathrm{T}}$ is the position vector of the origin of the frame W expressed in the frame C.

In fact, the PnP problem is to calculate ${}^c n_w$, ${}^c o_w$, ${}^c a_w$, and ${}^c p_w$ with the intrinsic parameters u_0, v_0, k_x, k_y and the given points' data including the image and Cartesian coordinates. In PnP problem, P3P, P4P and P5P attract most attention of the researchers in the visual measure field since it is believed that less given points provide more flexible in applications. Many approaches have been developed to solve the PnP problem (Gao and Chen, 2001; Hu and Wu, 2002; Nister, 2004). However, there exist multiple solutions for P3P, P4P or P5P, which result in difficulty for their applications. The multiple solutions could not be avoided if the given points were not limited to a defined arrangement (Xu *et al*, 2008). To ensure the unique solution for PnP problem, four coplanar given points at least are necessary and any three of them should not be collinear (Xu *et al*, 2008). Of course, more given points are helpful to improve the accuracy of visual measure with PnP method.

Without loss of generality, assume that the world frame is established on the plane containing the four given points. Hence the coordinate z_{wi} is zero for any given points on the plane. Combining equation (1) and (2), we have

$$\begin{cases} x_{wi}\,{}^c n_{wx} + y_{wi}\,{}^c o_{wx} - x_{1ci} x_{wi}\,{}^c n_{wz} - x_{1ci} y_{wi}\,{}^c o_{wz} \\ +{}^c p_{wx} - x_{1ci}\,{}^c p_{wz} = 0 \\ x_{wi}\,{}^c n_{wy} + y_{wi}\,{}^c o_{wy} - y_{1ci} x_{wi}\,{}^c n_{wz} - y_{1ci} y_{wi}\,{}^c o_{wz} \\ +{}^c p_{wy} - y_{1ci}\,{}^c p_{wz} = 0 \end{cases} \qquad (3)$$

where

$$\begin{cases} x_{1ci} = x_{ci} / z_{ci} = (u_i - u_0)/k_x \\ y_{1ci} = y_{ci} / z_{ci} = (v_i - v_0)/k_y \end{cases} \qquad (4)$$

The vectors ${}^c n_w$ and ${}^c o_w$ of the camera relative to the world frame can be obtained via solving a group of equation (3) forming from given points with linear method such as least square algorithm. The vector ${}^c a_w$ is obtained via the cross production of ${}^c n_w$ and ${}^c o_w$.

$${}^c a_w = {}^c n_w \times {}^c o_w \qquad (5)$$

PnP based visual measure has widely applications in robotics and automation, such as in-flight refuelling and auto-assembling.

2.2 Based on Line Knowledge

Parallels are very common in artificial environments. The parallels with specified property can be utilized to estimate the pose of a camera (Guillou *et al*, 2000).

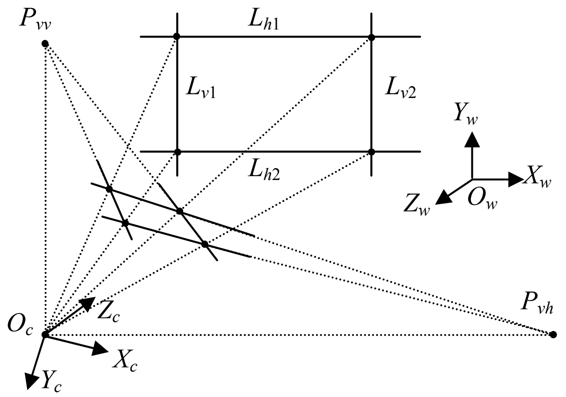

Fig.1 Visual measure based on parallels

The orthogonal parallels are defined as two groups of parallels, any lines in one group are orthogonal to the lines in another group. Symbol L_{hi} denotes a line in one group L_h, and L_{vi} denotes a line in another group L_v. Without loss of generality, two parallels L_{h1} and L_{h2} are selected from the group L_h, another two parallels L_{v1} and L_{v2} are selected from the group L_v. As shown in Fig.1, the world frame is established as follows. Its origin locates at some place. Its X_w axis is selected to parallel to L_{h1} and L_{h2}, and Y_w axis is selected to parallel to L_{v1} and L_{v2}. Two vanishing points can be computed with the image of the four lines, which are denoted as P_{vh} and P_{vv} on the imaging plane. Then their coordinates in the camera frame are calculated as given in (4).

In fact, line O_cP_{vh} is parallel to L_{h1} and L_{h2} since P_{vh} is their vanishing point. Hence O_cP_{vh} indicates the direction of X_w-axis in the camera frame. The unit vector of X_w-axis in the camera frame is given in (6) (Guillou *et al*, 2000).

$${}^c n_w = \frac{1}{\sqrt{x_{1ch}^2 + y_{1ch}^2 + 1}} \begin{bmatrix} x_{1ch} \\ y_{1ch} \\ 1 \end{bmatrix} \qquad (6)$$

where $(x_{1ch}, y_{1ch}, 1)$ are the coordinates of vanishing point P_{vh} in the camera frame; x_{1ch} and y_{1ch} are computed via formula (4) with the image coordinate of vanishing point P_{vh}.

As same as above, the unit vector of Y_w-axis in the camera frame is as shown in (7). Then the direction of Z_w-axis is obtained with formula (5).

$${}^c o_w = \frac{1}{\sqrt{x_{1cv}^2 + y_{1cv}^2 + 1}} \begin{bmatrix} x_{1cv} \\ y_{1cv} \\ 1 \end{bmatrix} \qquad (7)$$

where (x_{1cv}, y_{1cv}, 1) are the coordinates of vanishing point P_{vv} in the camera frame; x_{1cv} and y_{1cv} are computed from the image coordinate of vanishing point P_{vv} with formula (4).

It is necessary to point out that the vanishing point would be with large error or out of order if the angle formed with the lines on image was very small, especially near to zero. In this case, the lines on image are almost parallel. It means that the corresponding axis of the camera frame is parallel to the group of parallels. Therefore, it is necessary to calculate the angle formed by the image lines of the parallels before their vanishing point is computed. It can be computed with the coordinates of any two image points selected from each line (Xu, Li *et al*, 2006).

In addition, the intrinsic parameters of a camera can be self-calibrated with vanishing points. The orthogonal parallels with two groups of parallels, for example, L_h and L_v as shown in Fig.1, can provide two orthogonal vanishing points. An equation with intrinsic parameters can be deduced from them, as given in (8). The intrinsic parameters can be determined with enough equation as given in (8). For example, four groups of orthogonal parallels are necessary for the calibration of a camera with four intrinsic parameters model (Xu, Li *et al*, 2006).

$$\frac{(u_h-u_0)(u_v-u_0)}{k_x^2}+\frac{(v_h-v_0)(v_v-v_0)}{k_y^2}+1=0 \quad (8)$$

where (u_h, v_h) is the image coordinate of vanishing point P_{vh}, and (u_v, v_v) is the image coordinate of vanishing point P_{vv}.

Line knowledge can be utilized to self-calibrate intrinsic parameters of the camera mounted on an indoor mobile robot. And line knowledge based visual measure can be applied on the navigation and positioning for mobile service robots. It can also be employed in the assembling applications with industrial robots such as manipulators.

2.3 Based on Size Knowledge

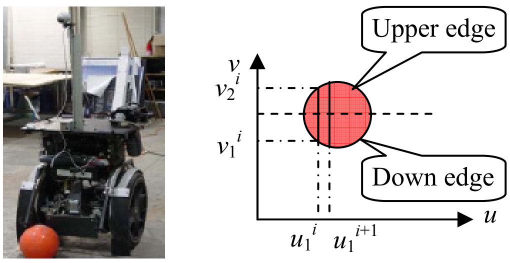

Fig.2 A mobile robot and its manipulating ball, (a) a mobile robot (Browning and Veloso, 2005), (b) the image area of a ball

For a sphere object, it is easy to be recognized with colour-based method (Browning and Veloso, 2005), and its shape is always a circle on the image in the case k_x=k_y=k, even if the pose of the camera is different.

As an example, a mobile robot and its manipulating ball are given in Fig.2(a), and the ball's image sketch is shown in Fig.2(b).The relation between the image area and its actual projection area can be derived from (1).

$$\begin{aligned} S &= \sum_{i=1}^{M}(y_{c2}^{i}-y_{c1}^{i})(x_{c1}^{i+1}-x_{c1}^{i}) \\ &= \frac{z_c^2}{k^2}\sum_{i=1}^{M}(v_2^i-v_1^i)(u_1^{i+1}-u_1^i)=\frac{z_c^2}{k^2}S_m \end{aligned} \quad (9)$$

where v_2^i and v_1^i are the vertical image coordinates of the points on upper and down edges of the ball; u_1^i and u_1^{i+1} are horizontal image coordinates; y_{c2}^i and y_{c2}^i are the coordinates of the edge points corresponding to the image points (u_1^i, v_2^i) and (u_1^i,v_1^i), on Y_c-axis in the camera frame; x_{c1}^i and x_{c1}^{i+1} are the coordinates of the edge points corresponding to the image point (u_1^i, v_1^i) and its next adjacent point, on X_c-axis in the camera frame; M is the number of the divided samples of the ball as shown in Fig.2(b); S_m is the image area of the ball; S is the actual projection area of the ball; k is an intrinsic parameter of the camera.

Assume that the size of the ball is known in advance. Then the depth of the ball relative to the camera can be given as in (10).

$$z_c = k\sqrt{S/S_m} = kd/d_m \quad (10)$$

where d is the actual diameter of the ball; and d_m is the circle's diameter of the ball on the image.

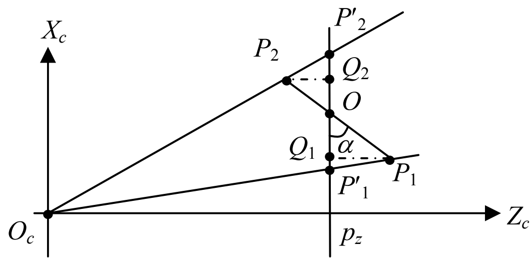

Fig. 3 Space point and its imaging principle

For an object with a plane, the depth can also be estimated with the size knowledge if the pose of the camera relative to the object has been obtained (Xu, Tan *et al*, 2006). Assume that the target frame is established on the plane of the object. The image coordinates of the origin of the target frame is denoted as (u_b, v_b). The relation between the space point and its imaging point is shown in Fig.3. According to the camera's pinhole model in (1), the target point P_1 in the space and the point P'_1 on the plane z_c=p_z share the same image coordinates. Using the imaging coordinates P'_1 and the angle α, the x-coordinate of the point P_1 in the target frame is obtained.

$$P_{1x} = OP_1 = \frac{OP_1'}{\cos\alpha + (P_{1x}'/p_z)\sin\alpha} \quad (11)$$

where α is the angle between projections of the Z axis in the target frame and the Z_c axis in the image frame on the plane $X_cO_cZ_c$; $P'_{1x}/p_z=x_c/z_c$ is obtained from (1) using the image coordinates; OP'_1 is the offset on X_c-axis between the points P'_1 and O on the plane $z_c=p_z$.

Considering $\tan\alpha = {}^c a_{wx}/{}^c a_{wz}$, $P'_{1x}/p_z = (u_1-u_0)/k_x$ and $OP'_1=(u_1-u_b)p_z/k_x$, the formula (11) is rewritten to (12). Similarly, P_{1y}, the coordinate for P_1 on the Y axis in the target frame can be expressed as given in (13).

$$P_{1x} = \frac{(u_1-u_b)\sqrt{({}^c a_{wx})^2+({}^c a_{wz})^2}}{{}^c a_{wz}k_x+{}^c a_{wx}(u_1-u_0)}p_z = m_{1x}p_z \quad (12)$$

$$P_{1y} = \frac{(v_1-v_b)\sqrt{({}^c a_{wy})^2+({}^c a_{wz})^2}}{{}^c a_{wz}k_y+{}^c a_{wy}(v_1-v_0)}p_z = m_{1y}p_z \quad (13)$$

In the target plane, coordinates offset on the axis Y of both top and bottom brims of the rectangle are integrated along the axis X to obtain the area S of the target plane.

$$\begin{aligned} S &= \sum_{i=1}^{M}(P_{2y}^i - P_{1y}^i)(P_{1x}^{i+1}-P_{1x}^i) \\ &= [\sum_{i=1}^{M}(m_{2y}^i - m_{1y}^i)(m_{1x}^{i+1}-m_{1x}^i)]p_z^2 = S_m p_z^2 \end{aligned} \quad (14)$$

where S_m is the computed target area on the normalized focus imaging plane; S is the actual area of the target plane.

Then the depth of the object is computed via (15).

$$p_z = \sqrt{S/S_m} \quad (15)$$

Size knowledge is very helpful for a mobile robot to estimate the distance from the robot to a ball object. Size knowledge based visual measure can be well applied on football match with mobile robots. In addition, formula (10) can be extended to the visual measure of the depth of a cylinder object.

Fig. 4 An indoor scene with TV set and door

The depth estimation with the visual measure method based on area knowledge of the objects with a plane has potential applications on environment exploration, positioning and navigation for mobile robots working in indoor. For example, as shown in Fig.4, many objects such as windows, pictures on wall, TV sets and etc. can serve as natural landmarks for indoor mobile robots. In addition, well-designed artificial landmarks are also popular for the navigation of mobile robots. These objects have common characteristics including multiple kinds of known knowledge such as four points at least, a group of orthogonal parallels and actual area. Hence the visual measure methods from section 2.1 to 2.3 are available to estimate the camera's pose relative to the objects. And the area knowledge is helpful to improve the measuring precision.

2.4 Based on Motion Knowledge

It is well known that many view points can be formed with the movement of a mobile robot. If the motion information is available, the position of an object in Cartesian space can be calculated with the images captured at two view points via stereovision method. Generally, the relative position and orientation of the camera at two view points are not very accurate because of noise influence. They will affect the visual measure precision. Therefore, it is necessary to optimize the relative position and orientation of the camera at two view points with image features as constrains. An index function is designed as (16) to verify the motion information (Xu and Li, 2007).

$$F = \sum_{k=1}^{M_1}\left\|U_k - \hat{U}_k\right\| + \sum_{m=1}^{M_2}\left[\arccos\left|L_m \cdot \hat{L}_m\right|\right]^2 \quad (16)$$

where M_1 and M_2 are the numbers of matched points and line segments in previous and current images; $U_k=(u_k, v_k)$ denotes the coordinates of a feature point P_k; L_m is the orientation of a line segment; $\hat{U}_k$ and $\hat{L}_k$ are the estimated current features derived from the previous features, camera's intrinsic parameters and the motion information.

The motion information between two view points can be optimized with the minimum of the index function in (16) via optimization such as Levenberg-Marquardt method. Then the position can be recalculated with the optimized motion information.

With the motion knowledge of a camera, the visual measure with monocular vision is converted to stereovision problem with optimization. It makes the visual measure of the mobile robot with single camera be much simpler and easier.

3. VISUAL MEASURE BASED ON ENVIRONMENT INFORMATION

General information of the environment is also very important for the positioning and navigation of mobile robots.

3.1 Based on General Environment Information

As known, an omni-directional camera can get much information of the surrounding environment. A visual method to estimate the rotation of a robot with an omni-directional camera on board is presented by Labrosse (Labrosse, 2006). A kind of omni-directional camera is shown in Fig.5. It consists of a normal camera, a hyperbolic reflection mirror and a perspex tube. The camera captures omni-directional scene via the hyperbolic reflection mirror. As an example, an omni-directional image captured with the omni-directional camera is also given in Fig.5.

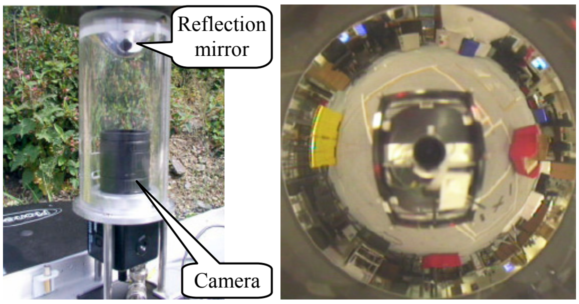

Fig.5 An omni-directional camera and its image (Labrosse, 2006)

In the visual measure method proposed by Labrosse, whole valid image area is employed to compute the rotation angle between the compared images. An index function is designed to assess the match, as given in (17). With different drift angle added to previous image, it compares two images pixel by pixel to search best image match. The drift angle resulting in best match is taken as the rotation angle.

$$D(I_i, I_j, \theta) = \sqrt{\sum_{k=1}^{h\times w}\sum_{l=1}^{c}[I_j(\theta,k,l) - I_i(k,l)]^2} \quad (17)$$

where $I_i(k, l)$ is the l-th colour component of the k-th pixel of image I_i; h and w are the height and width of the image after expended to a rectangle; c is the colour components per pixel; θ is the added draft angle to image I_j; D is the index function indicating the distance between two images.

Labrosse pointed out that the value of index D can also represent the translation of the mobile robot between two positions to capture the images when the two images are well matched. The method proposed by Labrosse has the advantage that the mobile robot could be self-located with the images of the environment even if there was no available knowledge of the camera and environment. However, its shortage is also apparent. The computational burden is very heavy since the index in formula (17) is computed pixel by pixel on the whole image many times with different draft angle to find the best match.

In fact, it is can be noticed from Fig.5 that the direction of an object in the environment is very obvious in the mobile robot frame. If an object can be recognized from the environment, its direction in the mobile robot frame can be obtained with the line from the image circle center to the object. On the other hand, the rotation angle of the camera and the mobile robot can be obtained from the direction angles in the mobile robot frame of a specified object on the two compared images of the environment. Although the measuring accuracy of this method is not satisfactory, its real-time performance is very good. For example, Wolf estimated the direction of an object in robot football field via the method above. Furthermore, based on the information obtained from the omni-directional camera and the color knowledge of the field and pillars landmarks as shown in Fig.6, a Monte Carlo based self-localization method was developed (Wolf, 2003).

Fig.6 A scene of robot football (Wolf, 2003)

The visual measure with an omni-directional camera is very useful for the map construction, navigation and positioning of mobile robots. For example, Menegatti *et al* reported that it is possible to create strict link between the spatial semantic hierarchy and the omni-directional images (Menegatti *et al*, 2001). Taiana *et al* applied omni-directional vision on 3D tracking in robot football (Taiana *et al*, 2007). It is a developing technique with promising future for omni-directional vision to be applied on autonomous mobile service robots.

3.2 Based on Map

Map is a very useful source of knowledge. If the map of environment is available, it is sure that the mobile robot can be localized with the current local map and the known global map. The process of match the local and global map is also known as map matching. Generally, the process of map matching as well as the construction of the local map needs heavy computational cost that prevents map-based localization from the applications requiring action in real time (Aguirre *et al*, 2004). To improve the performance in real time, Garulli *et al* selected line feature getting from 2D laser rangefinder to present environment (Garulli *et al*, 2005). Of course, the line features can also be extracted from images captured by a camera. The process of autonomous forming map is also known as Simultaneous Mapping and Localization (SLAM). Using the data of an odometer and a laser range finder mounted on the robot, Araneda *et al* proposed a complete probabilistic representation of the SLAM problem and obtain a Bayesian solution (Araneda *et al*, 2007). Indeed, monocular vision is very popular in SLAM recently (Eade and Drummond, 2006; Mouragnon *et al*, 2006).

On the other hand, the coarse position of the mobile robot relative to the objects recognized from the environment is easy to be determined according to the relations of the objects described in the map. For example, the mobile robot can be considered to be left front of the TV set, and it is farther away from the door than the TV set according to the relations of the door and TV set as shown in Fig.4. This kind of coarse position estimation is very helpful for a mobile service robot working in home.

4. CONCLUSION AND OUTLOOK

The basic kinds of knowledge-based visual measure with monocular vision system are investigated, including the methods based on the knowledge of point position, line, size or shape, and motion. Their principles are introduced and their applications are commented. The omni-directional visual measure based on general information of environment and map-based visual measure are also discussed simply.

In future, the combination of the basic visual measure methods based on knowledge will have promising practice applications on the navigation and positioning of mobile service robots. And environment map can play a very important role in the localization of a mobile robot with monocular system. However, how to acquire the required knowledge is a big problem for knowledge-based visual measure methods. Hence knowledge acquisition has been being a very active research topic. In future, the omni-directional visual measure and the monocular SLAM that remains to be further explored will attract more and more attention in the autonomous robot community.

ACKNOWLEDGMENTS

The author would like to thank the National Natural Science Foundation of China, and Beijing Natural Science Foundation for their support to this work under grant 60672039 and 4082032, respectively.

REFERENCES

M. A. Fishler and R. C. Bolles (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartomated cartography. *Comm. ACM*, **vol. 24**, no. 6, pp. 381-395.

R. Horaud, B. Conio, and O. Leboulleux (1989). An analytic solutions for the perspective 4-point problem, *Computer Vision, Graphics, Image Processing*, **vol. 47**, no. 1, pp. 33-44.

X. S. Gao and H. F. Chen (2001). New algorithms for the perspective-three-point problem, *Journal of Computer Science and Technology*, **vol. 16**, no. 3, pp. 194-207.

Z. Y. Hu and F. C. Wu (2002) A Note on the number of solutions of the noncoplanar P4P problem, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **vol. 24**, no. 5, pp.550-555.

D. Nister (2004). A minimal solution to the generalised 3-point pose problem, *Proc. of the 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, **vol. 1**, pp. 560-567.

D. Xu, Y. F. Li, M. Tan (2008). A general recursive linear method and unique solution pattern design for the perspective-n-point problem, *Image and Vision Computing*, in press.

E. Guillou, D. Meneveaux, E. Maisel, K. Bouatouch (2000). Using vanishing points for camera calibration and coarse 3D reconstruction from a single image, *Visual Computer*, **vol.16**, no.7, pp. 396-410.

D. Xu, Y. F. Li, Y. Shen, M. Tan (2006). New pose detection method for self-calibrated cameras based on parallel lines and its application in visual control system, *IEEE Transactions on System, Man & Cybernetics—Part B: Cybernetics*, **vol. 36**, no. 5, pp. 1104-1117.

B. Browning and M. Veloso (2005). Real-time, adaptive color-based robot vision, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3871-3876, Edmonton, Canada.

D. Xu, M. Tan, Z. Jiang, H. Hu (2006). Use of colour and shape constraints in vision-based valve operation by a humanoid robot, *International Journal of Advanced Robotic Systems*, **vol. 3**, no. 3, pp. 267-274.

D. Xu, Y. F. Li (2007). A new pose estimation method based on inertial and visual sensors for autonomous robots, *IEEE International Conference on Robotics and Biomimetrics*, pp. 405-410, Sanya, China.

F. Labrosse (2006). The visual compass: performance and limitations of an appearance-based method, *Journal of Field Robotics*, **vol. 23**, no. 10, pp. 913-941.

J. Wolf (2003). Omnidirectional vision system for mobile robot localization in the RoboCup environment, *Institute of Electrical Measurement and Measurement Signal Processing, Technical University Graz*, Master Dissertation.

E. Menegatti, M. Wright, E. Pagello (2001). A new omnidirectional vision sensor for the spatial semantic hierarchy, *Proceedings of IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp.93-98, Como, Italy

M. Taiana, J. Gaspar, J. Nascimento, A. Bernardino, P. Lima (2007). 3D tracking by catadioptric vision based on particle filters, *Proceedings of RoboCup2007 Symposium*, Atlanta, GA, USA

E. Aguirre, A. Gonzalez, R. Munoz (2004). Mobile robot map-based localization using approximate locations and the extended Kalman filter, *10th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pp. 191-198, Perugia, Italia.

A. Garulli, A. Giannitrapani, A. Rossi, A. Vicino (2005). Mobile robot SLAM for line-based environment representation, *44th IEEE Conference on Decision and Control and 2005 European Control Conference*, pp. 2041- 2046.

A. Araneda, S. E. Fienberg, A. Soto (2007). A statistical approach to simultaneous mapping and localization for mobile robots, *The Annals of Applied Statistics*, **vol. 1**, no. 1, pp. 66-84

E. Eade, T. Drummond (2006). Edge landmarks in monocular SLAM, *In Proc. British Machine Vision Conference*, Edinburgh, UK.

E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, P. Sayd (2006). Monocular Vision Based SLAM for Mobile Robots, *The 18th International Conference on Pattern Recognition*, Hong Kong.