Analysis of Landmark-based Vision Systems using Analogous Spatial Mechanisms

Wen-Yeuan Chung

Department of Mechanical Engineering, Chinese Culture University

ABSTRACT

A vision system that includes a single camera and several landmarks attached to the objects is studied using an analogous spatial mechanism. The object is equivalent to the moving platform, while the camera and landmarks are modeled into the fixed platform and supporting legs respectively. The proposed analogous model gives landmark-based vision system an alternative interpretation, and makes it easy to solve two primary problems. First is the determination of the minimum number of landmarks so that the pose of an object can be reconstructed correctly. This number is obtained by calculating the degrees of freedom of the mechanism. The second problem is reconstructing the object's pose analytically from a single image, and similar to analyzing the analogous mechanism's configuration. Four landmarks are shown necessary for reconstructing the pose of an object with six degrees of freedom, and the computation becomes straightforward while three of them are collinear. The strategies of attaching landmarks properly and reconstructing pose efficiently for objects being subject to various movement constraints are studied. The results also extend to the application for multi-body systems.

Keywords: analogous spatial mechanism, degrees of freedom, configuration, landmarks, vision system

以對等空間機構解析俱標誌點的視覺系統

鐘文遠

中國文化大學機械工程學系

摘　要

對於含單一攝影機及數個標誌點的視覺系統，提出與其對等類比的空間機構，並進而加以解析。物體與相機如同空間機構的移動平台與基座，標誌點則相似於可伸縮的支撐腿。藉此對等機構，除可對視覺系統提供另一種解讀方式，亦可使兩個主要問題迎刃而解。其一為若欲成功重建物體的姿態，所需標誌點的最低數目；該問題轉化為空間機構的自由度分析。另一為從單一張影像的資訊以重建物體的姿態；此則如同於對機構作構形分析。結果顯示俱六個自由度的物體，最少需四個標誌點始能重建其姿態；且若其中三點共線，計算過程可大幅簡化。對於使用不同數目的標誌點以搭配物體所受的各類拘束，亦做分析及探討。相關模式、概念與成果，亦可擴及至由多個物體相連結的系統。

關鍵詞：對等空間機構，自由度，構形，標誌點，視覺系統

文稿收件日期 100.6.14；文稿修正後接受日期 101.3.26　；*通訊作者
Manuscript received June 14, 2011; revised March 26, 2012; * Corresponding author

Ⅰ. INTRODUCTION

Robot vision is a topic that received intensive interest recently [1, 2]. Stereoscopic vision is found in both human beings and other natural systems. The coordinates of a landmark can be calculated using the intersection of two particular rays [3]. Each ray is defined by the camera center and the projection point on camera screen. This concept has been used by many researchers for visual odometry or tracking [4~7] and other applications [8~10].

Monocular systems have also been devised. Without landmarks, the interpretation of the 2D image is mainly based on the geometry description or geometric constraints of the object, such as coplanarity, collinearity, or points belonging to a parallelogram [11~13]. By estimating frame-to-frame camera motion from successive images, the monocular system can also execute visual odometry [14, 15]. The pose of human body or hand was recovered by direct nonlinear regression of joint angles [16] or various joint constraints [17, 18]. If only two landmarks are used, the vision system can only be used for planar applications [19]. In order to overcome the limitations of monocular systems, laser range finder was also used as assistance [20]. With three landmarks, at least four possible solutions for the object's pose can be obtained [21, 22]. Hence, redundant solutions have to be filtered out if possible. By using four landmarks and one camera, Katsuki et al. adopted numerical methods to reconstruct the pose [23, 24]. However, even the fundamental problem of how many landmarks are needed to reconstruct an object's pose when it is subject to various constraints has not been clarified. The works [21~24] are related to the problem known as Perspective-n-Point (PnP). As described in [25], it has found many applications in computer animation, computer vision, augmented reality, automation, image analysis, automated cartography, robotics, and model-based machine vision systems.

The Stewart-Gough platform is one of the most popular spatial mechanisms. Giving the length of the limbs, the pose of the platform can be determined and has at most 40 configurations [26, 27]. Since direct kinematics is a difficult problem, many special types have been devised [28~31]. The analogous mechanism proposed in this paper is similar to the Stewart-Gough platform. However, finding the object's pose becomes solving the lengths of limbs, and is thus not the same as forward position analysis for the Stewart-Gough platform.

In this work, the specification of a camera and the image displaying all landmarks are treated as given. Thereafter, the vision system and the analogous parallel mechanism refer to the same system. For different number of landmarks, the applicability to reconstructing poses of objects subject to various constraints is analyzed. The results are summarized and then extended to multi-body systems.

Ⅱ. SINGLE LANDMARK

The case with single landmark attached to a moving object is first analyzed. As shown in Fig. 1, a landmark is attached to the object on point A_o. The camera is located so that the camera center lies on point O. The projection of point A_o on the camera screen is defined as A_s. The coordinate of point A_s can be determined based on the screen data and specification of the camera. The point A_o certainly lies on the ray passing through points O and A_s.

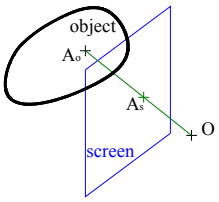


Fig. 1. Camera system with single landmark.

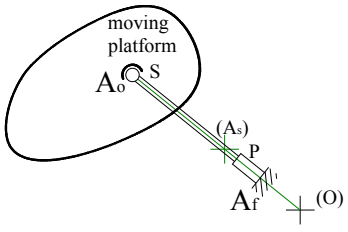


Fig. 2. Analogous mechanism with single landmark.

The present case can be modeled into an analogous mechanism as shown in Fig. 2. The object becomes the moving platform. Because point A_o are restricted on a ray, points O, A_s and A_o are replaced by a spherical-prismatic (SP) leg with points A_o and A_f at both ends.

The point A_o is treated as a spherical joint on the moving platform. In addition, the point A_f is on the fixed platform and can be any point on the ray $\overrightarrow{OA_s}$. Similarly, if there are more landmarks on the object, extra SP legs can be connected between the moving and fixed platforms.

The degrees of freedom, or DOF for brevity, of the analogous mechanism can be calculated easily by applying the Grübler equation which is given by

$$F = 6(n-1) - \sum_j (6 - f_j) \tag{1}$$

The parameter n is the number of links, and f_j is the DOF number of the jth kinematic pair. For present case, there are three links, one prismatic $(f=1)$ and one spherical $(f=3)$ joint. Hence, this mechanism or the moving platform has DOF as 4. Moreover, if an extra SP leg is added, the numbers of links, prismatic joints, and spherical joints all increase by one, and the DOF of analogous mechanism is thus decreased by two.

The DOF of the mechanism is 4 for the case of attaching single landmark. Hence, the pose cannot be reconstructed correctly if the object is not subject to any constraints. Let the object be constrained to move along and rotates about a specified axis. As shown in Fig. 3, the problem is modeled by adding a leg P_fP_o with cylindrical joint between the moving and the fixed platforms. Meanwhile, the moving platform and the leg is solidly welded at point P_o. The mechanism in Fig. 3 thus has zero DOF.

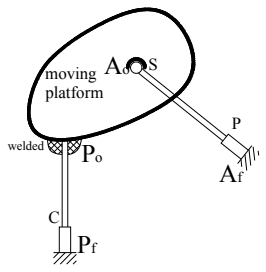


Fig. 3. Cylindrical joint and single landmark.

The problem of analyzing the pose for the object in the vision system is exactly the same as for the moving platform in analogous mechanism. With the cylindrical joint specified, the coordinate of point P_f and unit vector $\overrightarrow{\mathbf{u_P}}$ along line $\overrightarrow{P_fP_o}$ are defined. Let $|P_fP_o| = L_P$, the coordinates of point P_o, represented by boldface $\mathbf{P_o}$, can be written as

$$\mathbf{P_o} = \mathbf{P_f} + L_p\overrightarrow{\mathbf{u_P}} \tag{2}$$

The coordinate of point points O and A_s are obtained from the data of camera screen. The unit vector $\overrightarrow{\mathbf{u_A}}$ is thus

$$\overrightarrow{\mathbf{u_A}} = \frac{\overrightarrow{OA_s}}{|OA_s|} \tag{3}$$

After the fixed point A_f is chosen on the ray $\overrightarrow{OA_s}$ and let $|A_fA_o| = L_A$, $\mathbf{A_o}$ can be written as

$$\mathbf{A_o} = \mathbf{A_f} + L_A\overrightarrow{\mathbf{u_A}} \tag{4}$$

The landmark is attached at $\mathbf{A_o}$ as planned. The point P_o can be any points on the axis of the cylindrical joint, it is thus selected so that $\overrightarrow{A_oP_o}$ is normal to $\overrightarrow{P_fP_o}$ without losing generality. The distance between P_o and A_o is certainly a constant and defined as $|P_oA_o| = R$. Two unknowns L_P and L_A can be solved by combining two equations

$$\left|L_A\overrightarrow{\mathbf{u_A}} - L_p\overrightarrow{\mathbf{u_P}} + \mathbf{A_f} - \mathbf{P_f}\right| = R \tag{5}$$

$$\left(L_A\overrightarrow{\mathbf{u_A}} - L_p\overrightarrow{\mathbf{u_P}} + \mathbf{A_f} - \mathbf{P_f}\right)\cdot\overrightarrow{\mathbf{P_fP_o}} = 0 \tag{6}$$

The orders of both equations are two and one respectively. This implies there are two sets of solutions. On the other hand, present problem can be interpreted geometrically as that the point A_o lies on the intersection points of a cylinder and a straight line defined by A_f and the unit vector $\overrightarrow{\mathbf{u_A}}$. The point P_f and the unit vector $\overrightarrow{\mathbf{u_P}}$ define the center axis of the cylinder, and $|P_oA_o| = R$ is the radius. Consequently, there are two or none intersection points.

The point A_o being on a ray can also be interpreted as on the intersection of two planes described by linear equations. Meanwhile, a cylinder can be expressed as an equation of order two. From the point of view of mathematics, three equations are combined to

obtain two solutions for point A_o. One particular case is that the object's motion is pure 2D translations. The third equation then becomes linear as well. Hence, there is only one solution for point A_o by combining three linear equations.

For constraint of either cylinder joint or pure 2D translation, the DOF of analogous mechanism is zero. Obviously, the ambiguity does not exist only if all equations are linear. The existence of ambiguity can also be explained mathematically. The DOF of the moving platform being zero implies that several equations are used to solve same number of unknowns. The number of solutions can be determined according to the Bézout number [32]. Whenever nonlinear equations exist, at least two solutions can be obtained for the poses of the moving platform. Therefore, the DOF of analogous mechanism should be negative in order to reconstruct the pose without ambiguity. For example, the DOF of analogous mechanism becomes -1 if the object is subject to either prismatic or revolute constraint. The solution for point A_o is the intersection point of two lines or a line and a circle, and can be solved easily and uniquely.

Ⅲ. TWO LANDMARKS

If two landmarks are attached to the object subject to cylindrical constraint, the DOF becomes -2. The correct solution can be filtered out easily with the ray defined by the second landmark.

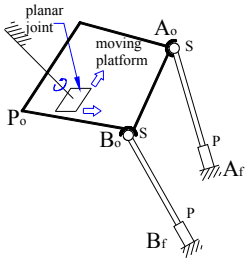


Fig. 4. Planar joint and two landmarks.

Three cases featuring a 3-dof object are considered. The first one is that the object can only make planar movement. The DOF of the analogous mechanism as shown in Fig. 4 becomes -1. $|A_fA_o| = L_A$ and $|B_fB_o| = L_B$ are two unknowns to be solved. Based on the constraint, each point P_o, A_o or B_o can only move on its specific plane. The value of L_A or L_B can be solved uniquely by applying Equation (4) and substituting it into the plane equation. Geometrically, point A_o or B_o is just the intersection of ray $\overrightarrow{A_fA_o}$ or $\overrightarrow{B_fB_o}$ and the plane that it moves on. With the 3D coordinates of points A_o and B_o, that of point P_o and the pose of the object can be calculated. Moreover, present technique can be applied for guiding autonomous mobile vehicle moving on flat terrain.

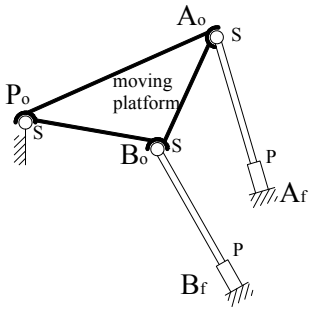


Fig. 5. Spherical joint and two landmarks.

The second case is the object being restricted to rotate around a fixed point P_o. The analogous mechanism is shown in Fig. 5. By observing loop $P_o - A_o - A_f - P_o$, the point A_o lies on the intersection between a sphere and the ray $\overrightarrow{A_fA_o}$. The center of the sphere is point P_o, and $|P_oA_o|$ is the radius. There are two solutions for point A_o. Similarly, two solutions for point B_o can be found from loop $P_o - B_o - B_f - P_o$. Although there are up to four sets of solutions for points A_o and B_o, the correct one must satisfy the length of $|A_oB_o|$.

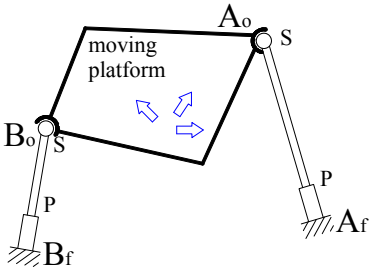


Fig. 6. Pure 3D translations and two landmarks.

The other case features the constraint of pure 3D translations as shown in Fig. 6. Hence, the vector between two landmarks A_o and B_o, $\overrightarrow{A_oB_o}$, is specified. The closure equation for the loop $A_f - A_o - B_o - B_f$ can be written as

$$\mathbf{A_f} + L_A\overrightarrow{\mathbf{u_A}} + \overrightarrow{A_oB_o} = \mathbf{B_f} + L_B\overrightarrow{\mathbf{u_B}} \qquad (7)$$

Equation (7) includes three linear equations and two unknowns L_A and L_B. Both can then be solved to reconstruct the pose of object.

Ⅳ. THREE LANDMARKS

We now study the case of adhering three landmarks to an object. The analogous parallel mechanism has three SP legs as shown in Fig. 7, and has zero DOF. The points A_o, B_o and C_o represent three landmarks, and the lengths $|A_oB_o|$, $|B_oC_o|$, and $|C_oA_o|$ are all given. The projection of three landmarks on the camera screen are designated as A_s, B_s and C_s. The 3D coordinates of these points and the camera center can be obtained from the camera system. Three angles θ_a, θ_b and θ_c, as shown in Fig. 7, can then be found from

$$\cos\theta_c = \frac{\overline{OA_s}\cdot\overline{OB_s}}{|OA_s||OB_s|} \qquad (8)$$

$$\cos\theta_b = \frac{\overline{OA_s}\cdot\overline{OC_s}}{|OA_s||OC_s|} \qquad (9)$$

$$\cos\theta_a = \frac{\overline{OC_s}\cdot\overline{OB_s}}{|OC_s||OB_s|} \qquad (10)$$

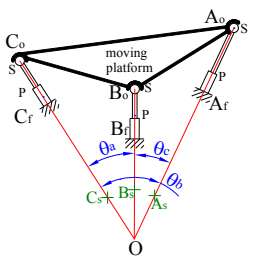


Fig. 7. Analogous mechanism with three landmarks.

Three unknowns $|OA_o| = a$, $|OB_o| = b$, and $|OC_o| = c$ are to be solved. Based on the geometry relationship and after eliminating the variables a and c [22], an equation with b as single unknown can be derived in the form as

$$G_4(b^2)^4 + G_3(b^2)^3 + G_2(b^2)^2 + G_1(b^2) + G_0 = 0 \qquad (11)$$

Coefficients G_4, G_3, G_2, G_1 and G_0 are all constants related to the setup and data of camera screen. Although Equation (11) has order as 8, the order is 4 with respect to the term b^2. This can be understood by treating point O as the mirror point. If (a*, b* and c*) is one set of solution, (-a*, -b* and -c*) is certainly another set of solution. Therefore, at most four nontrivial solutions can be obtained. The application is illustrated by the following example.

Example 1: The distances between three landmarks are $|A_oB_o| = |B_oC_o| = |C_oA_o| = 1$. Three angles are calculated from 2D image and Equations (8)~(10) as $\theta_a = 52.39^o$, $\theta_b = 44.36^o$ and $\theta_c = 45.88^o$. Three unknowns $|OA_o| = a$, $|OB_o| = b$, and $|OC_o| = c$ are to be found.

Sol: By following the processes described in [22] and solving Equation (11), four solutions for b^2 can be solved. Four sets of nontrivial solutions can be obtained and illustrated in Fig. 8. Meanwhile, each figure is plotted by unfolding the tetrahedron $OA_oB_oC_o$.

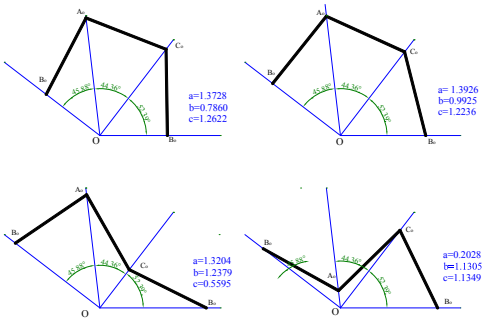


Fig. 8. Solutions for Example 1.

If the pose for a non-constrained object is to be reconstructed by adhering three landmarks, the problem of ambiguity exists since at most four sets of solutions can be found. However, if the object is subject to any movement constraint, the pose can be reconstructed correctly. For example, let the object be held by a serial manipulator that has a universal joint at fixed

point P_F and a spherical joint P_o on the wrist. The number of DOF of this objet is five. Apparently, the point P_o must lie on a sphere with P_F and $|P_F P_o|$ as the center and radius. After obtaining all possible solutions as in Example 1 and checking whether the point P_o is on the sphere, the correct solution and pose of the object can thus be obtained.

A special case is shown in Fig. 9 when triangle $A_oB_oC_o$ degenerates into a line. All points O, A_o, B_o and C_o are then on a plane, and the analogous mechanism becomes a planar linkage. More precisely, it is an Assur kinematic chain with five links. Both ternary links have three prismatic and revolute joints respectively. Therefore, there are only two configurations [33], and both are on each side of point O. Hence, the pose of this line can be reconstructed without ambiguity. However, the pose of object remains uncertain since the object can still rotate about this line. Consequently, the fourth landmark is requisite to correctly reconstruct the object's pose. On the other hand, a line has five DOF or is defined by five parameters. Therefore, we can interpret that the reduction of DOF is five when three landmarks are deployed collinearly.

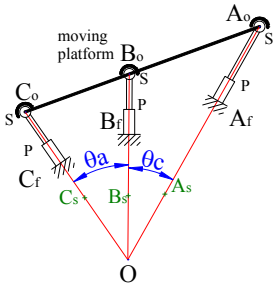


Fig. 9. Analogous mechanism with three collinear landmarks.

V. FOUR LANDMARKS

The problem of ambiguity still exists by attaching three landmarks to an object. If four landmarks are used, the DOF of the analogous mechanism become -2 and the pose can be reconstructed without ambiguity. Let four landmarks be attached to the object on points A_o, B_o, C_o and D_o. Any two tetrahedrals can be chosen to reconstruct the object's pose. For example, four solutions for b^2 can be obtained from tetrahedral $OA_oB_oC_o$ and the other four from $OA_oB_oD_o$. Two of all eight solutions should be the same and the correct one. The other unknowns, a, c, and d, can then be solved.

Deploying three among four landmarks collinearly is an alternate suggestion. As described in previous section and Fig. 9, the pose of the line through points A_o, B_o and C_o can be reconstructed uniquely. The fourth landmark D_o is then restricted on a circle and the ray $\overrightarrow{OD_o}$, and its coordinates can be found easily. The following example shows how the pose is computed from a 2D picture with four landmarks.

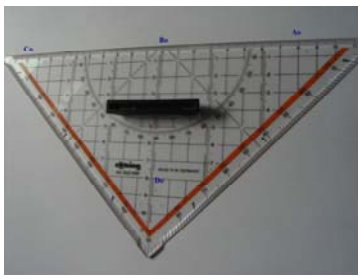


Fig. 10. Picture of triangular ruler.

Example 2: A picture of triangular ruler shown in Fig. 10 is used as an example for illustration. Three points A_o, B_o and C_o are collinear. They along with point D_o work as four landmarks. The pose of this ruler or the 3D coordinates of four landmarks are to be reconstructed.

Sol: Four steps, calibration of the camera, data conversion, 3D coordinate reconstruction and verification, are executed and explained as follows.

Step 1: A line with length L is used for calibration. It is placed parallel to the camera screen and its projection is located in the middle of the screen. Fig. 11 is the top view for illustration. Point O is the camera center. The line is firstly fixed at A_1B_1, and the length on the picture is measured as s. The line is then translated by distance d to A_2B_2, and the length on the second picture is t. Based on the triangle OA_1B_1 and the first picture, the following relationship can be derived

$$\frac{L}{d+f+g}=\frac{s}{f} \qquad (12)$$

Similarly, the triangle OA_2B_2 and the second picture lead to

$$\frac{L}{f+g}=\frac{t}{f} \tag{13}$$

Both Equations (12) and (13) can be combined to get

$$f=\frac{d}{L}\frac{s\cdot t}{t-s} \tag{14}$$

The resolution of the camera used in this experiment is 2816 by 2112 pixels. As shown in Fig. 11, both lengths s and t are proportional to f. In other words, the conversion between pixels and length depends on the placement of the pictures or distance f. In this example, the pictures are assumed to be placed or zoomed so that 100 pixels represent 1 mm. After substituting data into Equation (14), the distance between the camera center and pictures is obtained as f=32.08 mm.

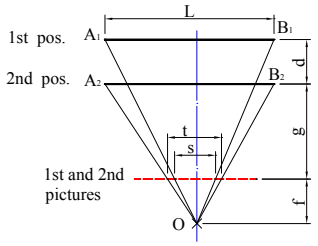


Fig. 11. Illustration of camera calibration.

Executing the calibration described above is just one way to find the distance between the camera center and pictures. If the specifications, such as field of view and angle of view, of the camera are available, they should be used to derive the distance and the precision can thus be improved.

Step 2: The projection of four landmarks on the camera screen are designated as A_s, B_s, C_s and D_s. A frame is set so that point O is the origin and z axis is perpendicular to camera screen. The data from the picture, shown in Fig. 10, can then be transferred to three dimensional coordinate as reported in Table 1.

The landmarks can be illuminants to be identified easily. In addition, lots of symbols, e.g. ×, +, ‹, ›, ⊢, ⊥, and ┼, can be adopted. The intersection point representing the landmark can thus be recognized with image processing tools.

Table 1. Data of projection of four landmarks

	2D coordinate on camera screen (in pixels)		3D coordinate (in mm)		
A_s	931	-799	9.31	-7.99	32.08
B_s	-148	-753	-1.48	-7.53	32.08
C_s	-1215	-682	-12.15	-6.82	32.08
D_s	-226	292	-2.26	2.92	32.08

Step 3: Three collinear landmarks A_o, B_o and C_o, as shown in Fig. 9, are considered. The distances are $|A_oB_o|=80$ mm and $|B_oC_o|=100$ mm. Based on the 3D coordinates of point O, A_s, B_s and C_s, two angles are found as $\theta_a=17.80^o$ and $\theta_c=18.32^o$. By applying the law of sines to triangles ΔOA_oB_o and ΔOC_oB_o, two equations can be written as

$$\left|\frac{OC_o}{\sin\angle OB_oC_o}\right|=\left|\frac{C_oB_o}{\sin\theta_a}\right|=\frac{100}{\sin 17.80}=327.18\text{ mm}$$

$$\left|\frac{OA_o}{\sin\angle OB_oA_o}\right|=\left|\frac{A_oB_o}{\sin\theta_c}\right|=\frac{80}{\sin 18.32}=254.55\text{ mm}$$

Since three landmarks are collinear and thus $\sin\angle OB_oC_o=\sin\angle OB_oA_o$, the ratio between $|OA_o|=a$ and $|OC_o|=c$ can be calculated as $a/c=0.778$ by combining both equations. From triangle ΔOA_oC_o, the law of cosines gives

$$\cos(\theta_a+\theta_c)=\frac{a^2+c^2-|A_oC_o|^2}{2\cdot a\cdot c}$$

After substituting $a/c=0.778$, the variable c can be solved as $c=305.01$ mm and thus $a=237.30$ mm. The other variable b can found easily as $b=254.23$ mm. Hence, the coordinates of points A_o, B_o and C_o are obtained, and the pose of line $\overline{A_oC_o}$ is reconstructed.

The fourth landmark D_o is then restricted on a circle. The intersection of the circle and ray through points O and D_s leads to point D_o. However, the intersection point may be nonexistent even the inaccuracy is very tiny from the geometrical point of view. Mathematically, too many equations are used to

solve the intersection point. Hence, it is suggested to focus on the plane on which the circle lies on. For present example, this plane includes point B_o and its normal is along vector $\overrightarrow{A_oC_o}$. The only intersection point between this plane and the ray leads to point D_o. The 3D coordinates of all four points are reported in Table 2.

Table 2: Data of four landmarks

	3D coordinate (in mm)		
A_o	64.330	-55.209	221.645
B_o	-11.408	-58.042	247.252
C_o	-105.966	-59.481	279.759
D_o	-16.733	21.620	237.501

Table 3: Verification of reconstruction

Distance between Points	Experimental values (in mm)	Theoretical values (in mm)	error
D_o and B_o	80.43	80.00	0.54%
D_o and A_o	112.81	113.14	-0.29%
D_o and C_o	127.77	128.06	-0.23%

Step 4: For the example at hand, the results are verified by distances between points. The results as listed in Table 3 show that the reconstruction of 3D coordinates of four landmarks has high accuracy. In other words, an object's pose is successfully reconstructed from a single 2D image, as shown in Fig. 10, without using numerical methods [23, 24].

VI. TWO OR MORE OBJECTS

In previous Sections, only one object is considered and some rules can be concluded as

(a). The point where a landmark is attached can move along a ray.

(b). The number of DOF of the analogous mechanism is reduced by 2 by adding a landmark.

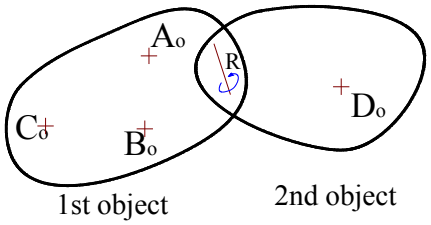
(c). The number of DOF of the analogous mechanism can be zero to reconstruct the pose correctly if only linear equations involved. Otherwise, the number of DOF must be negative.

(d). The pose of an object can be reconstructed by using four non-collinear landmarks.

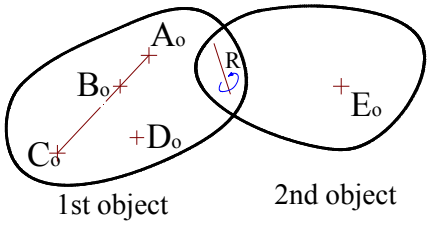
(e). The pose of an object has two or four solutions by using three non-collinear landmarks.

(f). By using three collinear landmarks, the pose of the line can be reconstructed uniquely and the number of DOF of the analogous mechanism is reduced by 5.

The rules listed above are applicable to a single object and can be extended to multi-objects jointed together. A simple case is when two objects are connected by a revolute joint. The DOF is $6 \cdot 2 - 5 = 7$. Based on rule (c), at least four landmarks are necessary. One option is attaching three landmarks to an object as shown in Fig. 12(a) to get several possible poses according to rule (e). For each pose, the landmark D_o attached to the second object must lie on a circle due to the revolute joint. In addition, D_o must lie on a ray according to rule (a). If the circle can intersect with the ray, the possible pose is the correct one and the poses of both objects are then reconstructed. On the other hand, using five landmarks may expedite the reconstruction process. Four landmarks are attached to the first object and three of them are collinear as shown in Fig. 12(b). The pose of the first body can be reconstructed with little effort as in Example 2. The fifth landmark E_o is again the intersection of a circle and the corresponding ray. The coordinates of this landmark and pose of the second object can then be obtained.



(a) Using four landmarks



(b) Using five landmarks

Fig. 12. Two objects connected by revolute joint.

Table 4: Minimum number of landmarks for two connected objects

Case	Type of joint	DOF of two connected objects	minimum number of landmarks	DOF of analogous mechanism
1	revolute	7	4	-1
2	prismatic	7	4	-1
3	cylindrical	8	5	-2
4	universal	8	5	-2
5	spherical	9	5	-1

In addition to revolute joints, two objects can be connected by other joints, e.g. prismatic, cylindrical, universal and spherical. The related analysis and the minimum number of landmarks are all listed in Table 4. If minimum number of landmarks is preferred, attaching three landmarks to the first object is suggested. Based on two or four possible solutions, rule (e), the poses of both objects can then be figured out by considering rays, constraints of joints, and geometry constraints between landmarks.

The results in Table 4 give the minimum number of landmarks for two connected objects. The number of landmarks can be increased and the deployment of landmarks can be well designed to expedite the process of pose reconstruction. For example, three landmarks being collinear can simplify the mathematical operation as shown in Example 2. On the other hand, similar techniques can be applied when further objects are jointed.

VII. CONCLUSIONS

Reconstructing the pose of an object attached with several landmarks is presented and proved feasible by using a single camera or even from a single image. The analogous mechanism shows very useful in clarifying the minimum number of landmarks needed to reconstruct the pose of object correctly. The main concern is the DOF of the analogous mechanism. The analysis reveals that adding each landmark reduces the DOF of the system by two. Although the resultant DOF can be zero to reconstruct pose correctly if all equations are linear,

it must be less than zero whenever nonlinear equations involved. Consequently, four landmarks are needed for object with six DOF. In addition, systems as multi-objects connected by various joints can be analyzed similarly.

Reconstructing the pose for an object is the same as analyzing configuration for the analogous mechanism. The most difficult mathematical problem encountered is to solve a polynomial of order four when 3D coordinates of three landmarks are to be found simultaneously. If three landmarks are collinear, the mechanism becomes a special planar Assur kinematic chain with five links and the single solution or configuration is easy to attain.

REFERENCES

[1] Hartley, R. and Zisserman, A., Multiple View Geometry in Computer Vision, Cambridge University Press, 2003.
[2] Cyganek, B. and Siebert, J. P., An Introduction to 3D Computer Vision Techniques and Algorithms, Wiley, 2009.
[3] Florczyk, S., Robot vision, Wiley, 2005.
[4] Howard, A., "Real-time stereo visual odometry for autonomous ground vehicles," IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France, 2008.
[5] Cumani, A. and Guiducci A., "Fast stereo-based visual odometry for rover navigation," WSEAS Transactions on Circuits and Systems, Volume 7 , Issue 7, pp. 648-657, 2008.
[6] Furgale P., Barfoot T., and Ghafoor N., "Rover-Based Surface and Subsurface Modeling for Planetary Exploration," Proceedings of Field and Service Robotics (FSR), Cambridge, MA, 2009.
[7] Milella, A., Nardelli, B., Paola, D. D., and Cicirelli, G., "Robust Feature Detection and Matching for Vehicle Localization in Uncharted Environments," Iros 2009 3rd Workshop: Planning, Perception and Navigation for Intelligent Vehicles, 2009.
[8] Bien Z., Kwon H. Y., Youn J., Suh I. H., “A closed form 3D self-positioning algorithm for a mobile robot using vision and guide-marks,” Robotica, Vol. 9, pp 265-274, 1991.
[9] Chen, C. F., Chou W. S., and Chen Y. C., “The mark-based approaches for motion estimation from 3-D point sets,” Journal of the Chinese Institute of Engineers, Vol. 17,

No.1, pp.43-53, 1994.

[10]Mark, W., Heuvel, J., and Groen F., "Stereo based Obstacle Detection with Uncertainty in Rough Terrain," IEEE Intelligent Vehicles Symposium, pp. 1005-1012, 2007.

[11]Maganto, A. L., Mene'ndez, J. M., Salgado, L., Rendon E., and Garcia N., "A Monocular Vision System For Autonomous Vehicle Guidance," IEEE International Conference on Image Processing, Vol. 1, pp. 236-239, 2000.

[12]Wilczkowiak, M., Boyer, E., and Sturm, P., "Camera Calibration and 3D Reconstruction from Single Images Using Parallelepipeds," The 8th International Conference on Computer Vision, Vancouver, Canada, Vol. 1, pp. 142-148, 2001.

[13]Wilczkowiak, M., Sturm, P., and, Boyer, E. "Using Geometric Constraints Through Parallelepipeds for Calibration and 3D Modelling," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, pp. 194-207, 2005.

[14]Scaramuzza, D. and Siegwart, R., "Appearance-Guided Monocular Omnidirectional Visual Odometry for Outdoor Ground Vehicles," IEEE Transactions on Robotics, Vol. 1–12, No. 5, pp. 1015-1026, 2008.

[15]Tardif, J.-P., Pavilidis, Y., and Danniilidis, K., "Monocular visual odometry in urban environments using an omnidirectional camera," IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008.

[16]Agarwal, A. and Triggs, B., "Recovering 3D Human Pose from Monocular Images", IEEE Transactions on Pattern Analysis and Machine Intelligence archive, Vol. 28, pp. 44–58, 2006.

[17]Eriksson, M. and Carlsson S., "Monocular Reconstruction of Human Motion by Qualitative Selection," Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 863- 868, 2004.

[18]Lee, S. and Cohen I., "3D Hand Reconstruction from a Monocular View," Proceedings of the 17th International Conference on Pattern Recognition, Vol.3, pp. 310- 313, 2004.

[19]Nian, C. Y. and Tarng, Y. S., "An auto-alignment vision system with three-axis motion control mechanism," Int. J Advanced Manufacture Technology, Vol. 26, pp. 1121-1131, 2005.

[20]Nara, S.; Miyamoto, D., and Takahashi, S. "Vision-based Tracking Control using Mark Recognition", Advanced Motion Control, pp. 619-623, 2006.

[21]Haralick, B. M.; Lee, C. N.; Ottenberg K.; and Nölle, M. "Review and analysis of solutions of the three point perspective pose estimation problem," International Journal of Computer Vision, Vol. 13, pp. 331-356, 1994.

[22]Chung, W. Y. and Yang, S. H. "Reconstructing object pose by using single camera and planned marks," Hua Kang Journal of Engineering Chinese Culture University, Vol. 22, pp. 153-161, 2008.

[23]Katsuki, R.; Ota, J.; Mizuta, T.; Kito, T.; Arai, T.; Ueyama, T.; and Nishiyama, T. "Design of an artificial mark to determine 3D pose by monocular vision," Proceedings of the 2003. IEEE International Conference on Robotics and Automation, Vol. 1, pp.995-1000, 2003.

[24]Katsuki, R., Ota, J., Arai, T., and Ueyama, T. "Proposal of Artificial Mark to Measure 3D Pose by Monocular Vision," J of Advanced Mechanical Design, Systems, and Manufacturing, Vol. 1, pp. 155-169, 2007.

[25]Kneip, L., Scaramuzza, D., and Siegwart, R., "A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation," IEEE Conference on Computer Vision and Pattern Recognition, pp. 2969 – 2976, 2011.

[26]Raghavan, M., "The Stewart Platform of General Geometry Has 40 Configurations," ASME Journal of Mechanical Design, Vol. 115, pp.277-282, 1993.

[27]Tsai, L. W., Robot Analysis, John Wiley &Sons, 1999.

[28]Husain, M. and Waldron, K. J., "Direct Position Kinematics of the 3-1-1-1 Stewart Platform," ASME J. Mech. Design, Vol. 116, pp. 1102-1107, 1994.

[29]Lin, W., Crane, C. D., and Duffy, J., "Closed Form Forward Displacement Analysis of the 4-5 In-Parallel Platforms," ASME J. Mech. Design, Vol. 116, pp. 47-53, 1994.

[30]Tavolieri, C., Ottaviano, E., and Ceccarelli M., “Pose Determination for a Rigid Body by Means of CaTraSys II (Cassino Tracking System),” the first European Conference on Mechanism Science, 2006.
[31]Pierrot, F., Nabat, V., Company, O., Krut, S., and Poignet P., "Optimal Design of a 4-dof Parallel Manipulator: From Academia to Industry," IEEE Transactions on Robotics , Vol. 25, pp. 213-224, 2009.
[32]Orzech, G. and Orzech, M., Plane Algebraic Curves, M. Dekker, New York, pp. 174-178 ,1981.
[33]Chung, W. Y., "The Position Analysis of Assur Kinematic Chain with Five Links," Mechanism and Machine Theory, Vol. 40, No. 9, pp. 1015-1029, 2005.