

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# An Approach to Perception Enhancement in Robotized Surgery using Computer Vision

Agustín A. Navarro<sup>1</sup>, Albert Hernansanz<sup>1</sup>, Joan Aranda<sup>2</sup> and Alícia Casals<sup>2</sup>

<sup>1</sup>*Centre of Bioengineering of Catalonia, Automatic Control Department, Technical University of Catalonia  
Barcelona, Spain*

<sup>2</sup>*Institute for Bioengineering of Catalonia - Technical University of Catalonia  
Barcelona, Spain*

## 1. Introduction

The knowledge of 3D scene information obtained from a video camera allows performing a diversity of action-perception tasks in which 2D data are the only inputs. Exterior orientation techniques are aimed to calculate the position and orientation of the camera with respect to other objects in the scene and perceptually contribute to control action in this kind of applications. In Minimally Invasive Surgery (MIS), the use of a 2D view on a 3D world is the most common procedure. The surgeon is limited to work without direct physical contact and must rely heavily on indirect perception (Healey, 2008). To effectively link action to perception it is necessary to assure the coherence between the surgeon body movements with the perceived information. Therefore, locating the instruments with respect to the surgeon body movements through computer vision techniques serves to enhance the cognitive mediation between action and perception and can be considered as an important assistance in MIS.

The 2D-3D pose estimation technique serves to map this relation by estimating the transformation between reference frames of surgical instruments and the endoscopic camera. There are several methods proposed to estimate the pose of a rigid object. The first step of those algorithms consists in the identification and location of some kind of features that represent an object in the image plane (Olensis, 2000); (Huang & Netravali, 1994). Most of them rely on singular points and apply closed-form or numerical solutions depending on the number of objects and image feature correspondences (Wrobel, 2001); (Harris, 1992). In this approach lines are the features selected as the most appropriate. As higher-order geometric primitives, lines can describe objects where part of its geometry is previously known. These kinds of features have been incorporated to take advantage of its inherent stability and robustness to solve pose estimation problems (Dornaika & Garcia, 1999); (Christy & Horaud, 1997). A diversity of methods has been proposed using line correspondences, parting from representing them as Plücker lines (Selig, 2000), to their combination with points (Horaud et al., 1995), or sets of lines (Park, 2005).

In the case of image sequences, motion and structure parameters of a scene can be determined. Motion parameters are calculated by establishing correspondences between selected features in successive images. The specific field of computer vision which studies features tracking and their correspondence is called dynamic vision (Dickmanns, 2004). The use of line correspondences, increases robustness. Nevertheless, the benefit from the gained stability introduces some disadvantages: more computationally intensive tracking algorithms, low sampling frequency and mathematic complexity (Rehbinder & Ghosh, 2003). Therefore, some early works have chosen solutions based on sets of nonlinear equations (Yen & Huang, 1983), or iterated Kalman filters through three perspective views (Faugeras et al., 1987). Recently, pose estimation algorithms have combined sets of lines and points for a linear estimation (Park, 2005), or used dynamic vision and inertial sensors (Rehbinder & Ghosh, 2003). The uniqueness of the structure and motion was discussed for combinations of lines and points correspondences, and their result was that three views with a set of homologue features, two lines and one point, or two points and one line give a unique solution (Holt & Netravali, 1996).

This approach focuses on the analysis of changes in the image plane determined by line correspondences. These changes are expressed as angular variations, which are represented differently depending on their orientation with respect to the camera. They are induced applying transformations to an object line. Some properties of these motions are useful to estimate the pose of an object addressing questions as the number of movements or motion patterns required which give a unique solution. Using a monocular view of a perspective camera, some of these questions are answered in this chapter by the development of a specific methodology. It is inspired in those monocular cues used by the human visual system for spatial perception, which is based on the proper content of the image. This algorithm focuses on the distortion of geometric configurations caused by perspective projection. Thus, the necessary information is completely contained within the captured camera view (Navarro, 2009).

The content of this chapter relies on the concept of biologically inspired vision methods as an appropriate tool to overcome limitations of artificial intelligence approaches. The main goal is highlighting the capacity of analysis of spatial cues to enhance visual perception, as a significant aid to improve the mediation between action and perception in MIS. The remainder of this chapter introduces the benefits of using computer vision for assistance in robotized surgery. It is followed by an analysis of the angular variation as a monocular cue for spatial perception, with a mathematical description of the proposed algorithm and experimental results the article finalizes with some concluding remarks.

## **2. Mediating action and perception in MIS**

The introduction of minimally invasive surgery (MIS) as a common procedure in daily surgery practice is due to a number of advantages over some open surgery interventions. In MIS the patient body is accessed by inserting special instruments through small incisions. As a result tissue trauma is reduced and patients are able to recover faster. However, the nature of this technique forces the surgeon to work physically separated from the operation area. This fact implies a significant reduction of manipulation capabilities and a loss of

direct perception. For this reason, robotic and computer-assisted systems have been developed as a solution to these restrictions to help the surgeon.

Some solutions have been proposed to overcome those limitations concerning the constrained workspace and the reduced manipulability restrictions. Approaches dedicated to assist the surgeon are basically aimed to provide an environment similar to conventional procedures. In this sense, robotic surgery developments are especially focused on the enhancement of dexterity, designing special hand-like tools or adding force-feedback through direct telerobotic systems (Grimberger & Jaspers, 2004); (Mayer, 2004). Other systems aid the surgeon through auxiliary robotic assistants, as is the case of a laparoscopic camera handler (Muñoz et al., 2004); (Hurteau et al., 1994). Nevertheless, though the limitation of the visual sense has been tackled by robotic vision systems capable of guiding the laparoscopic camera to a desired view (Doignon et al., 2007); (Casals et al., 1995), 3D perception and hand-eye coordination reduction in terms of cognitive mediation have not been extensively developed.

## 2.1 Computer assistance in MIS

The visual sense in MIS environments is limited because it imposes a 2D view of the operative site. Therefore, approaches focused to assist the surgeon are fundamentally based on image content recognition and presentation. As an example of this computer assistance, there are some approaches focused on surgical tool tracking (Dutkiewicz, 2005), the study of the distribution of markers to accurately track the instruments (Sun et al., 2005), the establishment of models of lens distortion (Payandeh, 2001). These examples constitute emergent techniques to assist the surgeon by the enhancement of the image content. The work in which this approach is addressed, however, is based on the integration of visual and motion information to perceptually locate the instruments with respect to the surgeon.

Healey in (Healey, 2008) describes the mediation between action and perception in MIS environments and states the necessity of effectively linking action to perception in egocentric coordinates. In this approach, it is suggested that the integration of egocentric information, as visual and limb movements, can be provided with the capacity of locating surgical instruments at a desired position in the operation scene and the knowledge of their orientation with respect to the laparoscopic camera. As a result, the surgeon perception is enhanced by a sense of presence. Thus, computer vision issues such as the 2D-3D pose estimation and exterior orientation, deal with this problem and can be applied to aid the surgeon in this kind of procedures.

The schematic of an application where exterior orientation is used and presented through enhanced visual information to assist the surgeon is shown in Fig. 1. This presentation is commonly performed using augmented reality. There have been early approaches in which this type of resource is used in different kinds of applications (Milgram et al. 1993), others, more specialized in surgery, recognize objects seen by the endoscope in cardiac MIS (Devernay et al., 2001), or design a system for surgical guidance (Pandya & Auner, 2005), being a visual enhancement which serves as a human-machine interface. In this approach, the position and orientation of surgical instruments is the information to be imposed over the image of the surgical scene. It serves to integrate egocentric information, as vision and

limb movements, to provide a sense of presence and relate it with the external environment to help in becoming immersed in the working scenario. Nevertheless, the camera-tool calibration must be calculated. This problem can be tackled by computer vision techniques, as the perspective distortion model presented in this chapter. Thus, this computer assisted system can be expressed as a closed loop process, as shown in Fig. 2.

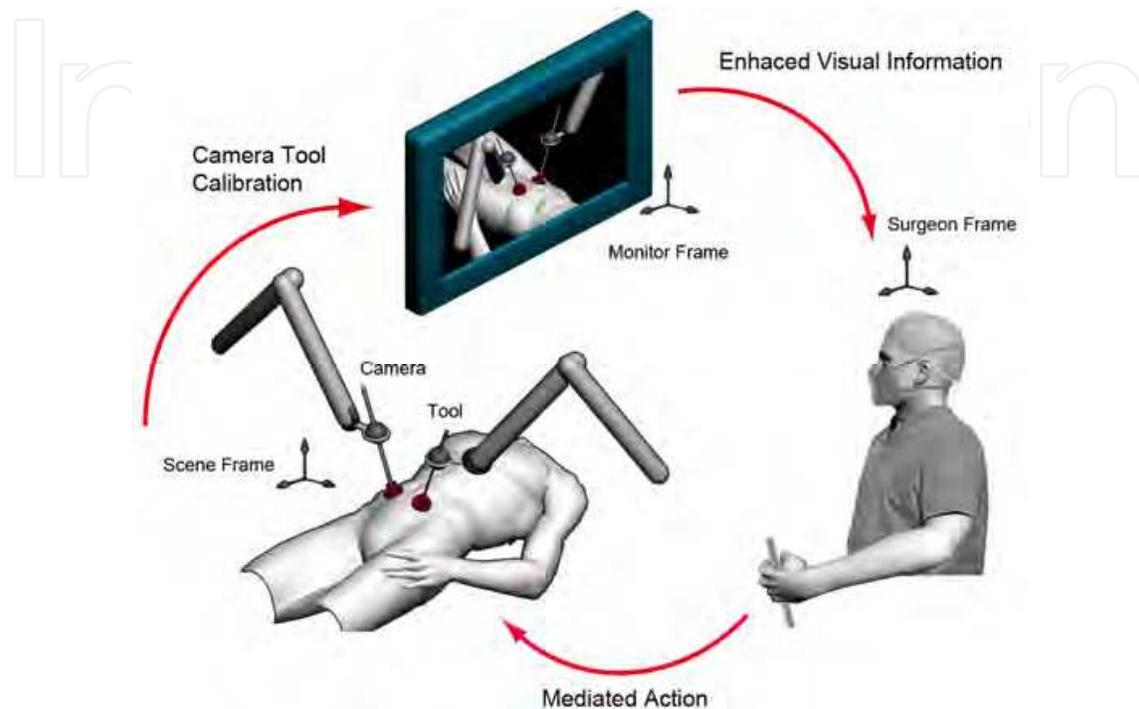


Fig. 1. Schematics of an application assisted surgery in MIS using computer vision.

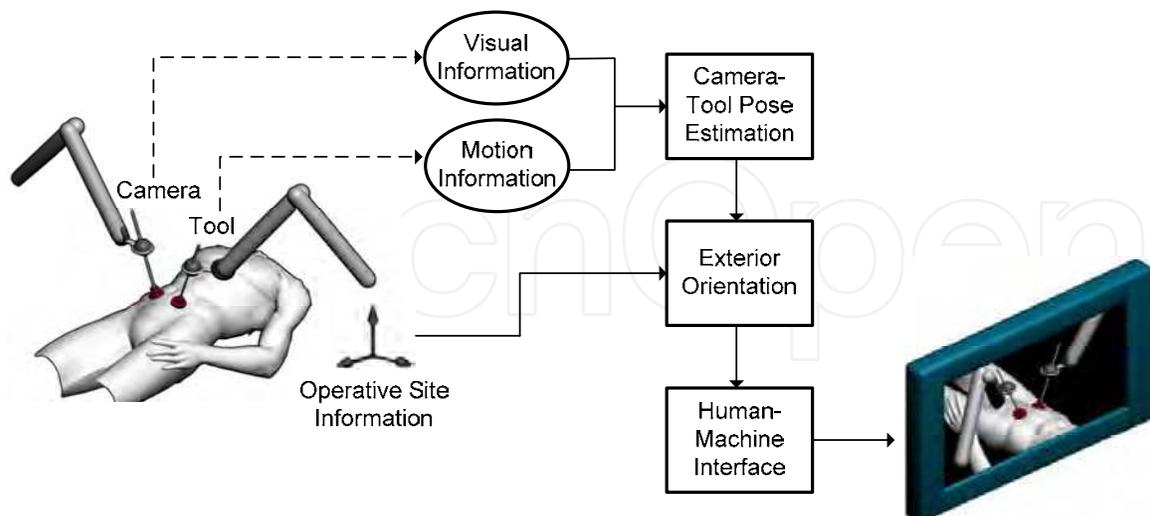


Fig. 2. Visual enhancement process to assist the surgeon. Motion and visual information is related to calibration of a surgical instrument with respect to the camera.

## 2.2 Robotic assistance in MIS

Assistant robots, as specialized handlers of surgical instruments, have been developed to facilitate surgeons' performance in MIS. Since a patient body is accessed by inserting surgical instruments through small incisions, passive wrists have been incorporated for free compliance through the entry port. With such wrist configuration, it is only possible to locate accurately an instrument tip, if its entry port or fulcrum point is known. The fulcrum is a 3D point external to the robotic system and though it has a significant influence on the passive wrist robot kinematics, its absolute position is uncertain.

A number of approaches focus on the control and manipulability of surgical instruments in MIS through robotic assistants. As can be seen in Fig. 3, 3D transformations are applied to produce pivoting motions through the fulcrum point. The more accurately estimate this entry port is, the more accurately the instrument tip is positioned at a desired location. Some approaches evade this difficulty by the incorporation of special mechanisms with actuated wrists (Taylor et al., 1995). Others, based on passive wrists tackle the problem through inverse kinematics approaches. Thus, error minimization methods are applied to determine the outside penetration (Ortmaier & Hirzinger, 2000) (Funda et al., 1995), and compensate the fulcrum location imprecision (Muñoz, 2004).

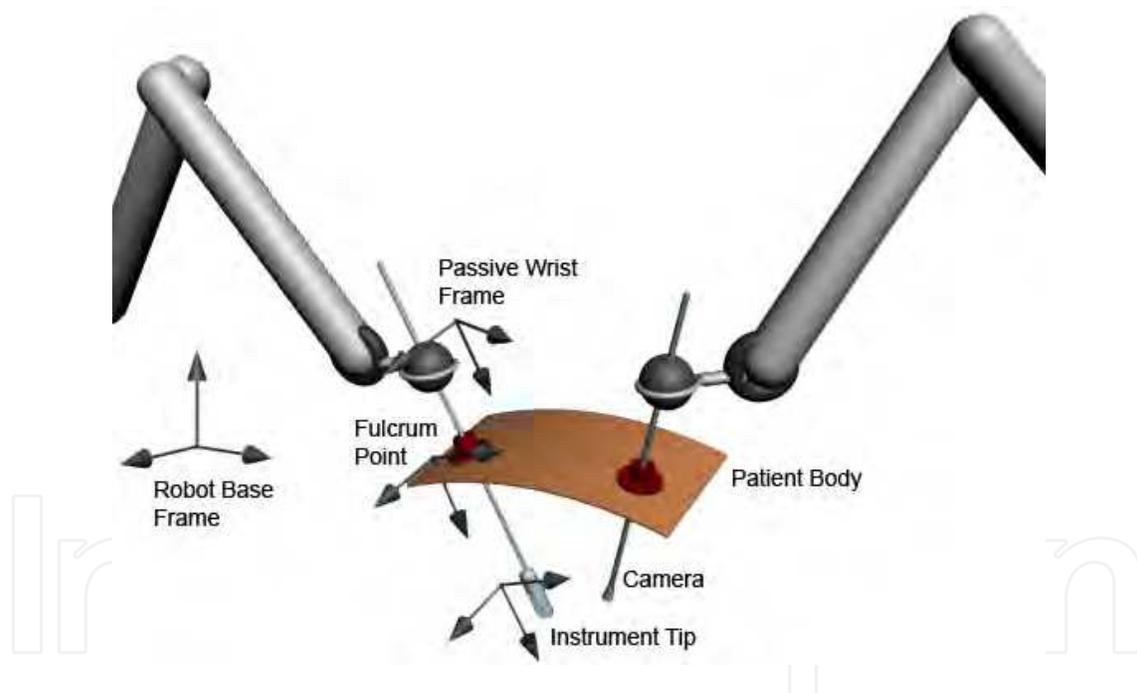


Fig. 3. Minimally invasive robotic surgery systems with passive wrist robotic assistants. The location of the instrument tip depends on the knowledge of the fulcrum point.

A reasonable alternative approach to tackle the passive robot wrist problem is the use of computer vision. Since the laparoscopic camera captures the workspace sight, specialized vision algorithms are capable of estimating 3D geometrical features of the scene. The 2D-3D pose estimation problem serves to map these geometric relations estimating the transformation between reference frames of the instruments and the camera. Several methods have been proposed to estimate the pose of a rigid object. The first step of their algorithms

consists in the identification and location of some kind of features that represent an object in the image plane.

In case of image sequences, motion and structure parameters of a 3D scene can be determined. Correspondences between selected features in successive images must be established, which provide motion information and can be used, as is in this case, to estimate the pose of a moving object. Fig. 4 shows the pivoting motion of a surgical instrument on the fulcrum point. As can be seen, instruments are represented by feature lines and are visualized by the laparoscopic camera. In this example, three views after two different rotations generate three lines in the image plane. Each of them defines a 3D plane called the projection plane of the line. These planes pass through the projection center and their respective lines. Their intersection is a 3D line from the origin of the perspective camera frame to the fulcrum point.

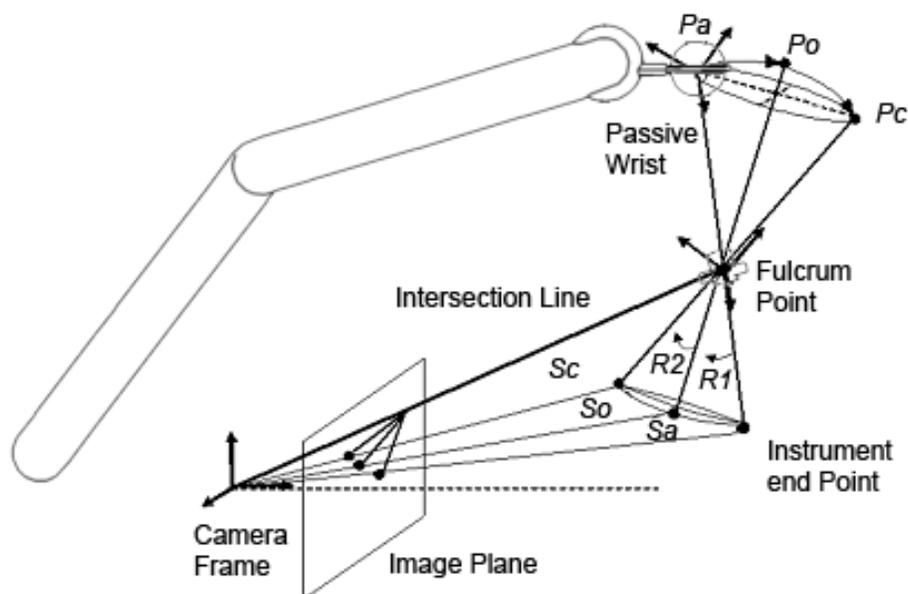


Fig. 4. 3D transformations of the instrument manipulated by a robotic assistant produce a pivoting motion through the fulcrum point. It results in 3D rotations in the workspace captured by the camera through projection planes.

### 3. Exterior orientation estimation based on angular variation

In this approach, the angular variation is considered as a monocular cue, from which spatial information is extracted. The specific monocular cue used to model this spatial perception property is linear perspective. It is related to the appearance of objects under perspective transformations on a flat surface. This bidimensional view generated by the representation of a 3D object is the result of a planar projection, obtained by mapping each point of the object onto a plane through passing lines emanated from a center of projection. Depending on the desired visual effect derived from this representation, this mapping may be different. It could show the general appearance of an object or depict its metric properties. When the center of projection is finite, a perspective projection is obtained, which presents an object as it is seen by the eye, as generally used in computer vision applications.

Perspective projection provides a realistic representation of an object. An impression of depth is created on a 2D surface and its 3D shape can be visualized. However, to provide this impression the geometry of the object is strongly distorted. Different parts are represented at different scales and parallel lines converge at a single point, implying that such a projection is not considered by the principles of Euclidean geometry (Mumford, 2002). Under perspective transformations, distances and angles, which are Euclidean invariants, are not preserved. Nevertheless, different properties of geometric configurations remain unchanged. For instance, a straight line is mapped into a straight line.

### 3.1 Linear perspective and angles

The change of geometric configurations under perspective transformations, particularly the angle between lines, plays an important role in modeling perspective distortion. However, certain configurations invariant to these transformations present essential properties necessary to describe the nature of the projection as well. The invariance of the cross-ratio defines the angular distribution of the projected pencil of uniformly separated coplanar lines. Fig. 5 shows first the 3D rotation of the circle in which the pencil is contained and afterward its resulting projection. There, a tendency of the projected lines to concentrate closer to the axis of rotation with progressive increments of the angular difference between them can be observed. This tendency grows with the applied angle of rotation, having the property of maintaining a constant cross-ratio.

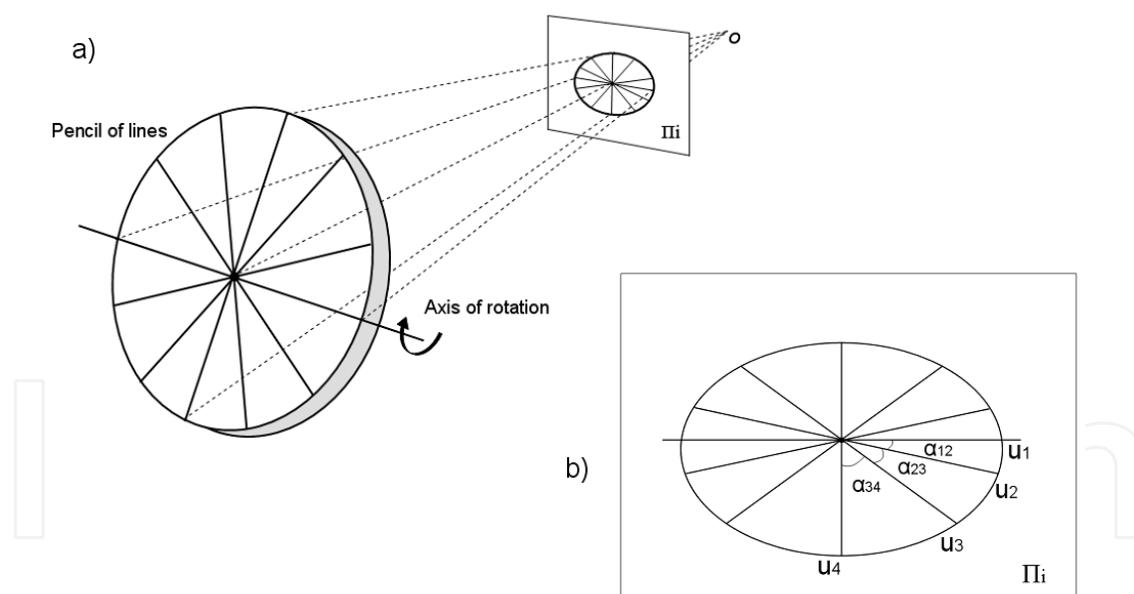


Fig. 5. Perspective projection of a pencil of lines: a) The rotated circle in which the pencil is contained is projected to an ellipse in the image plane; b) An angular variation pattern is produced bringing the lines closer as they approach the major axis of the ellipse.

The result of a projected circle is an ellipse if it is not parallel to the projection plane ( $\Pi_i$ ). It is the locus of points that forms the pencil of lines, confirming the geometric property of the construction of conics from the cross-ratio (Mundy & Zisserman, 1992). The invariance of this property is strongly related to the angular distribution of the projected pencil of lines,

and consequently, with the eccentricity of the ellipse. It provides information about the orientation of the object, knowing that the eccentricity ( $e$ ) of the ellipse can be expressed as a function of the angle of rotation ( $\gamma$ ) in the form  $e = \sin(\gamma)$ .

### 3.2 Perspective distortion model

The angular variation pattern of a projected pencil of lines provides sufficient information to model the perspective distortion. It describes the resulting projection of a determined angle depending on the position of the point of incidence of the pencil and the orientation of the circle it forms. It could be seen as a circle rotated about an axis. In the case the center of projection is aligned with the point of incidence of the pencil, this axis is coplanar to the circle and parallel to the major axis of the resulting ellipse. Therefore, the pose of the pencil with respect to the center of projection is defined by the angle of rotation of the circle ( $\gamma$ ), which is given by the eccentricity of its own projection.

#### 3.2.1 Aligned center model

Aligning the centers is the simplest case to model. This model serves as starting point to analyze the projective properties of a rotated circle. If a sphere, with radius  $Z_0$ , is centered at the center of projection ( $o$ ), the axis of rotation coplanar to a circle  $\Gamma_r$  is tangent to the sphere at the point of incidence  $P_o=(0,0,z_0)$ , as shown in Fig. 6. This implies the possibility of estimating the pose of the pencil of lines by an axis and an angle of rotation defined by the angular variation model. The axis is given by the tangent line, which is where the projected lines concentrate. The angle of rotation ( $\gamma$ ), however, as the eccentricity of the ellipse ( $e$ ), must be calculated from the angular variation model.

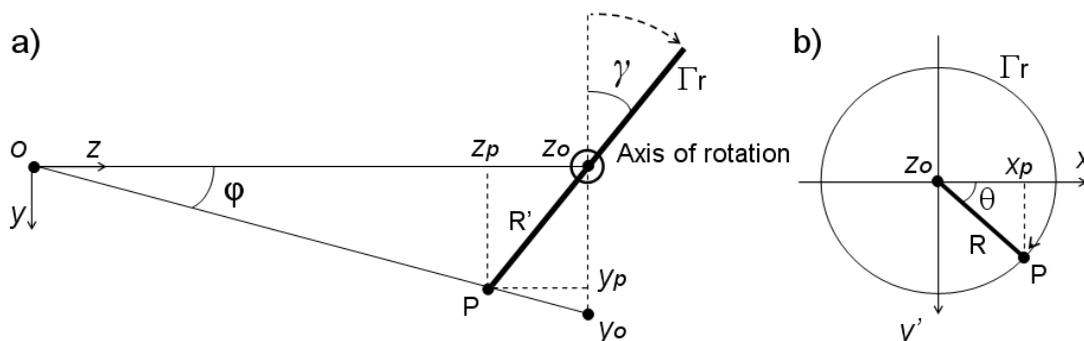


Fig. 6. The alignment of the center of projection with the point of incidence of the pencil of lines describes the distortion model as the rotation of the pencil about an axis: a) Angle of rotation  $\gamma$ , from lateral view; b) Cumulative angle  $\theta$  of the rotated line with length  $R$ , from orthogonal view.

The cross-ratio defines a conic constructed from the series of concurrent projected lines. To be calculated, a minimum of four lines are needed. If they are part of a pencil or form part of a sequence of a rotated line, it is necessary to keep the angle between them constant. The rest of projected angles are calculated knowing the cross-ratio. Thus, parting from the axis of rotation, the angular variation pattern of a projected pencil of lines can be determined. It can be expressed by the relation between the 3D applied angular changes and their projected angular data.

In the case of aligned centers, the axis of rotation is perpendicular to the  $z$  axis, as shown in Fig. 6. Having the rotated circle with radius  $R$ , centered at  $P_o$ ; an angle  $\alpha$  is the projection of the 3D angle  $\theta$  at plane  $Z_o$ , which is the position of the image plane  $\Pi_i$ . This 3D angle  $\theta$  is the rotation angle of a line with length  $R$ , centered in  $P_o$ . Therefore, if the angle  $\theta$  leads to a point  $P=(x_p, y_p, z_p)$  along the circle, the projected angle  $\alpha$  is formed by the angle between the projection of  $P$  at  $Z_o$ , and the axis of rotation of the circle. This can be expressed as:

$$\tan(\alpha) = y_p / x_p = y_0 / x_0 \quad (1)$$

Where, being  $x$  the axis of rotation, the slope of the projected line in  $\Pi_i$  describes the projected angle through  $x_p$  and  $y_p$ , the respective components of  $P$  in the  $xy$  plane. Thus, by the geometry of the configuration, under a rotation  $\gamma$ , in the  $yz$  plane:

$$\tan(\varphi) = R' \cos(\gamma) / (Z_o - R' \sin(\gamma)) \quad (2)$$

Having  $R'$  as the component of  $R$  in the  $y$  axis, as:

$$R' = R \sin(\theta) \quad (3)$$

This leads to the expression of  $y_0$  as:

$$y_0 = \frac{Z_o R \sin(\theta) \cos(\gamma)}{Z_o - R \sin(\theta) \sin(\gamma)} \quad (4)$$

Similarly, in the plane  $xz$  at  $Z_o$ , the component of  $P$  in the  $x$  axis is:

$$x_0 = \frac{Z_o R \cos(\theta)}{Z_o - R \sin(\theta) \sin(\gamma)} \quad (5)$$

It implies the function that describes the angular variation in an aligned center configuration, knowing the relation  $e = \sin(\gamma)$ , is defined as:

$$\tan(\alpha) = \tan(\theta) \sqrt{1 - e^2} \quad (6)$$

This model satisfies the fact that length and scale do not influence in the angular variation pattern of a projected pencil of lines. It is function of the applied 3D angle and the eccentricity of the ellipse. Knowing this, the angular pattern extracted from the projected lines makes possible the calculation of the eccentricity of the ellipse. It is carried out by fitting the angular data to the model. Thus, the orientation of the circle, which corresponds to the plane containing the pencil of lines, is calculated.

### 3.2.2 General case model

Generally the point of incidence of the pencil of lines is not aligned with the center of projection along the  $z$  axis. Its projection can be located at any position in the image plane  $\Pi_i$  and represented, in this case, by a unit director vector  $v_d$  from the center of projection ( $o$ ). Using the concept of the sphere centered at  $o$ , a tangent circle  $\Gamma_p$  is not parallel to the image plane, as would be in the aligned center case and consequently, the axis of rotation of the circle  $\Gamma_r$  is not parallel to the major axis of the projected ellipse. It implies an enhancement on the complexity of the configuration. However, the projective properties are equally satisfied by the angular variation pattern.

The methodology used in this general case approach is based on the aligned center model, on the calculation of the axis and angle of rotation of the circle  $\Gamma_r$  formed by the pencil of lines. Having the angular relation in (6) of two circles representing the rotation of a tangent circle about an angle, the general case configuration can be divided into two simpler aligned center models as shown in Fig. 7. This is due to the fact that in (6) it is assumed that the image plane  $\Pi_i$  is tangent to the sphere, and therefore, perpendicular to the axis of projection defined by  $v_d$ . Thus, the first part is conformed by the rotated circle  $\Gamma_r$  and a tangent circle  $\Gamma_p$ ; and the second part by  $\Gamma_p$  and a circle  $\Gamma_i$  coplanar with  $\Pi_i$ . Both parts are individual aligned center models. It can be seen as the result of the projection of the rotated circle in a tangent plane, and its consecutive projection in the image plane.

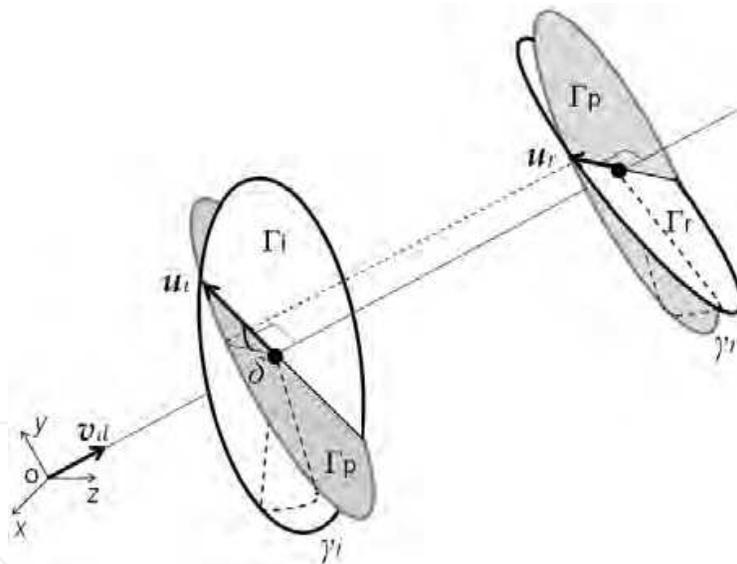


Fig. 7. Division of the general case configuration into two simpler aligned center models. The rotated circle  $\Gamma_r$  is projected over a tangent plane of the sphere, and its consecutive projection in the image plane.

The projection of the tangent circle  $\Gamma_p$  in the image plane could be considered as the first step to define the model, since the initial data is contained in the image. The extraction of the feature lines of the pencil defines the angular pattern of the projection; and the orientation of  $\Gamma_p$  with respect to  $\Pi_i$  through  $v_d$  provides the axis of rotation  $u_i$ . This permits to describe the relation between the angle of rotation  $\theta_p$ , around  $\Gamma_p$ , and the angle  $\theta_i$ , around  $\Gamma_i$ , which is the result of the rotation of  $\Gamma_p$  about a given angle  $\gamma_i$ . Applying the center aligned model, this

relation can be expressed as:

$$\tan(\theta_p) = \tan(\theta_i) \cos(\gamma_i) \quad (7)$$

In the same way, the angular relation between the angle of rotation 3D  $\theta_r$ , around  $\Gamma_r$ , and its projection in the tangent plane  $\theta_p$  can be expressed by the aligned center model having  $\gamma_r$  as the angle of rotation of  $\Gamma_r$  and  $u_r$  as the axis determined from (7).

The two parts of the divided configuration are related to each other by the tangent circle  $\Gamma_p$ . However, the alignment of the two axes of rotation  $u_i$  and  $u_r$  is only a particular case of the model. Therefore, to relate the 3D angle  $\theta_r$  with the angle projected in the image  $\theta_i$ , the difference between these two axes must be taken into account. If this difference is defined by the angle  $\delta$ , the angular variation pattern of  $\Gamma_r$  in the tangent plane is given by:

$$\tan(\theta_p + \delta) = \tan(\theta_r) \cos(\gamma_r) \quad (8)$$

which in conjunction with (7) express the angular variation of the 3D angle in the image plane as:

$$\tan(\theta_i) = \frac{\tan(\theta_r) \cos(\gamma_r) - \tan(\delta)}{\cos(\gamma_i)(1 + \tan(\theta_r) \cos(\gamma_r) \tan(\delta))} \quad (9)$$

This equation serves to model the perspective distortion through angular variations in a general case. This variation depends on the angles of rotation with the respective tangent and image plane, and the difference between their axes of rotation  $\delta$ . If there is no difference between these axes and the image plane is tangent to the sphere, the equation is reduced to the aligned center model. In any other case,  $\gamma_i$  and  $\delta$  are determined from the image data. Thus, through the fitting of the angular data to the model,  $\gamma_r$  can be estimated and consequently, the orientation of the circle with respect to the center of projection can be calculated.

#### 4. Experimental results

The perspective distortion model developed in this approach was validated by using real world data. The experimental setup consisted on a fixed standard analog B/W camera equipped with known focal length optics and a mobile frontal frame able to reach a desired orientation. The tests carried out were aimed to analyze the response of the model to different 3D angle inputs under a range of rotations of the frontal frame. Two sets of tests were employed to analyze the angular variation and its response to the model. The first set was focused on a single rotated 3D angle, while the second on a pencil of lines.

A 3D angle is formed by the intersection of two lines. It can be seen as the difference of the slope of each line lying on a plane, which is how the projected angle in the image plane was calculated. This angle served as input to the model and had to be known, constituting the main drawback of the method. This problem can be solved with the use of a sequence of concurrent lines, as shown further on. In Fig. 8 the angular variation of three different single angles under rotation is shown.

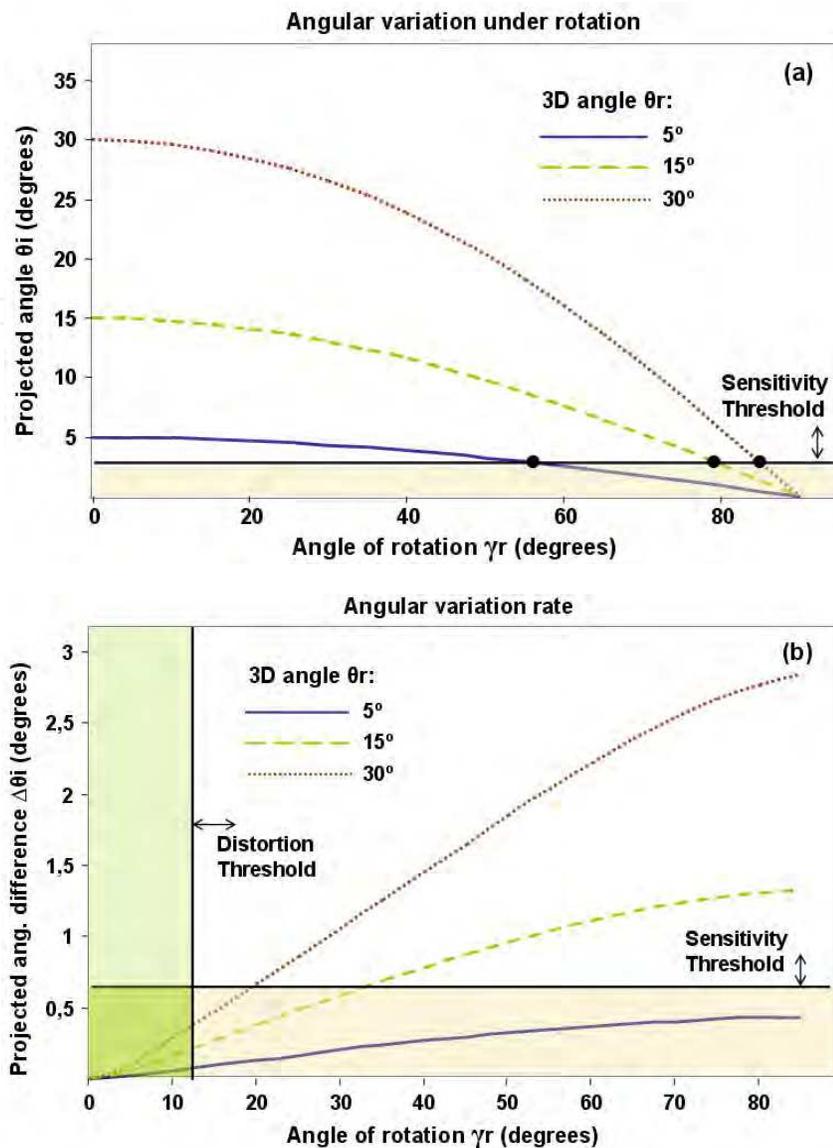


Fig. 8. Angular variation of a single projected angle under rotation. The response of the perspective distortion model is highly affected by line feature errors and the resolution of its acquisition method: a) reduced projected angles caused by long rotations produce unstable performance under a sensitivity threshold; b) the response resolution, dependent on the angular variation rate, misreads data under a distortion threshold.

Parting from the parallelism between the image and the frontal frame, the reduction of the projected angle under rotations can be seen. Likewise, as the resolution of the response depends on the detected angular variation, it does not only increase with higher 3D angles, but it also augments as the angle of rotation increases. As the area of extreme perspective provides enough angular variation with the more accurate response, the area of minor rotations provides negligible variations and consequently a misread response.

The sensitivity of the model is highly affected by the line feature error and the resolution of its acquisition method. Since the model response depends on the projected angle, its

extraction using real world data is prone to inaccurate detections, particularly with small angles. The resulting effect of this line feature error is an unstable and poor performance of the model below a sensitivity threshold. However, a good resolution can be obtained from standard methods, which are enough for an accurate performance requirement. Equally, the negligible variation at minor rotations produces a misread response below a threshold caused by the low perspective distortion. Fig. 9 shows the performance error of the model in the single angle test having real world data as input.

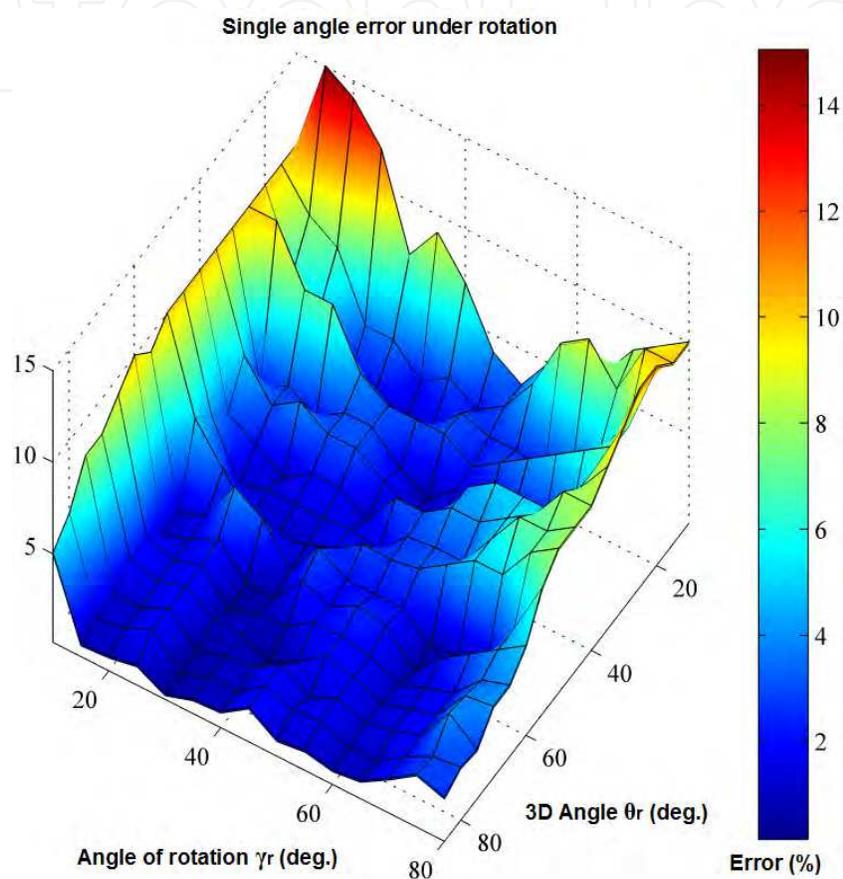


Fig. 9. Performance error of the single angle test. The response of the model is prone to errors with negligible variations at minor rotations and small projected angles. This improves as the 3D angles increases.

In general, the error decreases with increments of the 3D angle applied. Two areas of high error are differentiated along the applied rotations. The first is located below the distortion threshold, where minor rotations were applied. This is an area where the error is uniformly high independently of the 3D angle used. It is caused by the resolution of the line feature extraction method and the low perspective distortion. The second area, on the contrary, is located at long rotations. It decreases as the 3D angle applied increases, caused by the sensitivity of the model with reduced angles at long rotations.

The second set of tests employs pencils of lines of different constant angle separations. The method used to estimate the model parameters is based on fitting the projected measured angles. At least four line samples are required to calculate the cross-ratio or only three when the model is known. This permits to calculate the next angles of the sequence of concurrent lines and, consequently, more samples to obtain an improved fit. According to the invariant property of the pencil of lines, the distribution of the angles is unequal, while the cross-ratio is constant. Fig. 10 presents the performance sensitivity to lines orientation detection. The angular error increases with the amount of angle rotation. As in the previous set of tests, the model response is highly affected by the resolution and the error of the line extraction method, mainly at minor rotations.

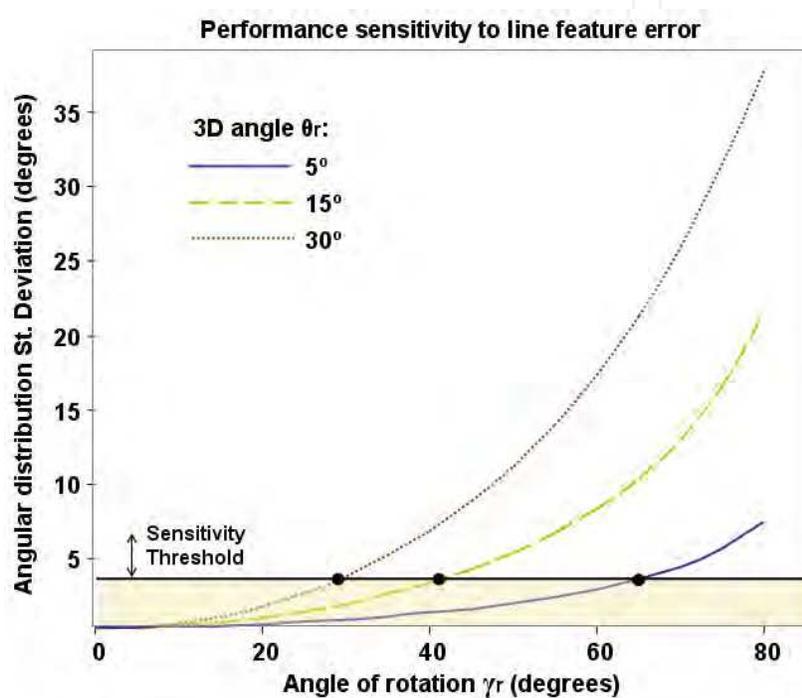


Fig. 10. Angular distribution variation in a rotated pencil of lines. The model response is highly affected under a sensitivity threshold caused by line feature errors at minor rotations.

The standard deviation of the angular distribution of the pencil indicates a superior performance of the model with pencils of higher 3D angle separation. It implies that the model response is prone to errors below a sensitivity threshold. This is depicted in Fig. 11, where real world data was used. Only one area presents a notable high error. It is located where minor rotations were applied. In contrast to the previous set of tests, this error is caused by the low angular distribution variation.

The change in this area is not only negligible, which depends on the resolution of the line feature extraction method, it is also highly sensitive to line feature errors since the pencil angles are similarly distributed.

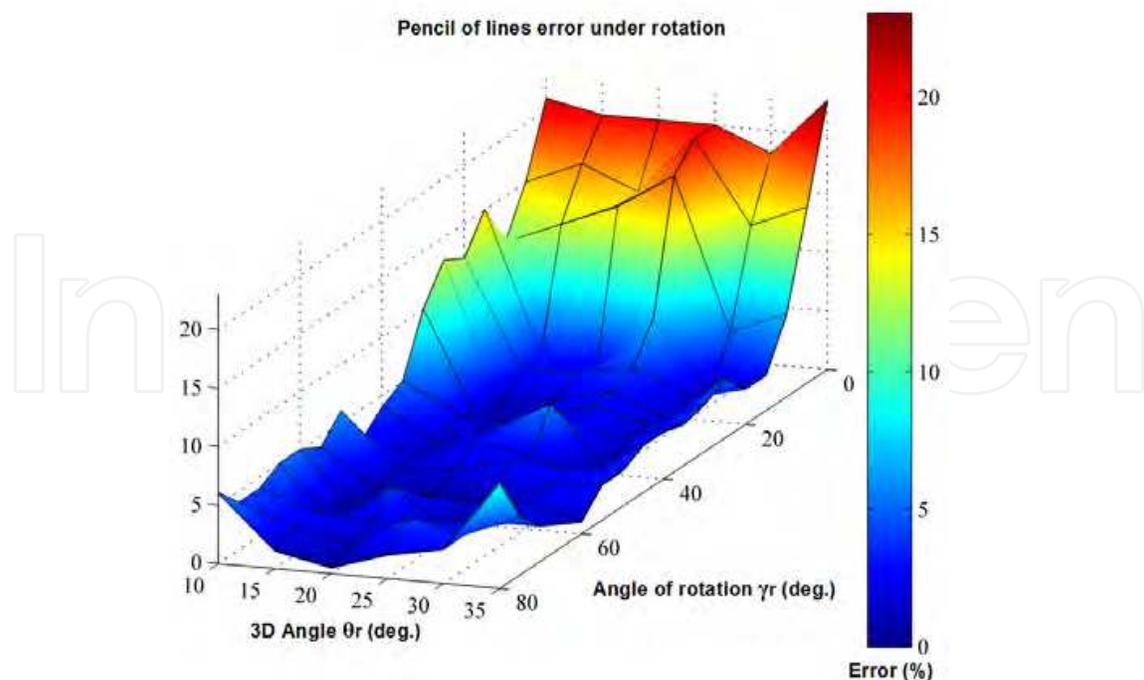


Fig. 11. Performance error of the pencil of lines test. The poor performance of the model at long rotations is improved, while the 3D separation angle of the pencil is limited.

In general the use of pencil of lines improves the performance of the model. It does not require the angular shift from the axis of rotation and provides a robust response at high 3D angle separations. Nevertheless, as the estimation is based on the fitting of sampled data, this 3D angle separation is limited due to the low number of measured samples. It is also suitable for real time applications, where a moving line feature forms the pencil of lines and its angular variation is modeled by its perspective distortion.

## 5. Conclusion

The capacity of the human visual system to perceive 3D space from a monocular view is characterized by a set of stimulus and processes. As human's perception of depth on a flat surface is enhanced by the use of linear perspective, a set of geometric properties from the projective transformation provide enough 3D information to effectively calculate the pose of an object from 2D images. In this work, a model of the variance of a geometric configuration under perspective transformation has been developed. Particularly, this geometric configuration is the angular variation caused by perspective distortion. Since the resulting projected variation depends on the position of the observer, the exterior orientation of the camera can be determined by the model.

In the presented approach the image plane contains the whole required information to solve the exterior orientation problem. Thus, there is no necessity to measure angular samples from determined 3D structures.

Experimental results show the efficiency and robustness of the method, in this case using a set of concurrent lines. Its accuracy can be affected by the reliability of the line feature

extraction technique applied. The potential of this novel approach is prominent in various applications, parting from simple calibration tasks to presence enhancement in immersive teleoperation.

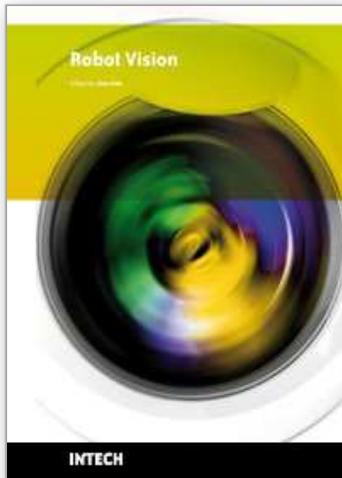
## 6. References

- Casals, A., Amat, J., Prats, D. & Laporte, E. (1995). Vision guided robotic system for laparoscopic surgery, *IFAC Int. Cong. on Advanced Robotics*.
- Christy, S. & Horaud, R. (1997). Fast and reliable object pose estimation from line correspondences, *Proc. Int. Conf. Computer Analysis Images Patterns*, pp. 432-439.
- Devernay, F., Mourgues, F. & Coste-Maniere, E. (2001). Towards endoscopy augmented reality for robotically assisted minimally invasive cardiac surgery, *IEEE Proc. Int. Workshop on Medical Imaging and Augmented Reality*, pp. 16-20.
- Dickmanns, E.D. (2004). Dynamic vision-based intelligence, *AI Magazine*, Vol. 25, No. 2, pp. 10-30.
- Doignon, C., Nageotte, F., Maurin, B. & Krupa, A. (2007). Pose estimation and feature tracking for robot assisted surgery with medical imaging, in: *Unifying Perspectives in Computational and Robot Vision*, D. Kragic and V. Kyrki (Eds.), Springer-Verlag, Berlin, pp. 1-23.
- Dornaika, F. & Garcia, C. (1999). Pose estimation using point and line correspondences, *Real-Time Imag.*, Vol. 5, pp. 215-230.
- Dutkiewicz, P., Kielczewski, M., Kowalski, M., & Wroblewski, W. (2005). Experimental verification of visual tracking of surgical tools, " *Fifth Int. Workshop on Robot Motion and Control*.
- Faugeras, O., Lustran, F. & Toscani, G. (1987). Motion and structure from point and line matches, *Proc. Int. Conf. Computer Vision*.
- Funda, J., Gruben, K., Eldridge, B., Gomory, S. & Taylor, R. (1995). Control and evaluation of a 7-axis surgical robot for laparoscopy, *IEEE Proc. Int. Conf. on Robotics and Automation*, pp. 1477-1484.
- Grimberger, C.A. & Jaspers, J.E. (2004). Robotics in minimally invasive surgery, *Proc. IEEE Int. Conf. on Systems, Man and Cybernetics*.
- Harris, C. (1992). Geometry from visual motion, In: *Active Vision*, A. Blake and A. Yuille (Eds.), Chapter 16, MIT Press, Cambridge, Mass.
- Healey, A. (2008). Speculation on the neuropsychology of teleoperation: implications for presence research and minimally invasive surgery, *Presence: Teleoperators and Virtual Environments*, MIT Press, Vol. 17, No. 2, pp. 199-211.
- Holt, J.R. & Netravali, A.N. (1996). Uniqueness of solution to structure and motion from combinations of point and line correspondences, *Journal of Visual Communication and Image Representation*, Vol. 7, No. 2, pp. 126-136.
- Horaud, R., Phong, T.Q. and Tao, P.D. (1995). Object pose from 2-d to 3-d point and line correspondences, *Int. J. Computer Vision*, Vol. 15, pp. 225-243.
- Huang, T.S. & Netravali, A.B. (1994). Motion and structure from feature correspondences: A review, *Proc. IEEE*, vol. 82, pp. 252-268.
- Hurteau, R., DeSantis, S., Begin, E. & Gagner, M. (1994). Laparoscopic surgery assisted by a robotic cameraman: concept and experimental results, *Proc. IEEE Int. Conf. on Robotics and Automation*.

- Mayer, H., Nagy, I., Knoll, A., Schirmbeck, E.U. & Bauernschmitt, R. (2004). The Endo[PA]R system for minimally invasive robotic surgery, *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*.
- Milgram, P., Zhai, S., Drascic, D. & Grodski, J. (1993). Applications of augmented reality for human-robot communication, *IEEE/RSJ Proc. on Intelligent Robots and Systems*, pp. 1467-1472.
- Mumford, D., Fogarty, J. & Kirwan, F. (2002) *Geometric Invariant Theory*, 3<sup>rd</sup> ed. Springer.
- Mundy, J.L. & Zisserman, A. (1992). *Geometric Invariance in Computer Vision*, The MIT Press, Cambridge, Mass.
- Muñoz, V.F., Garcia-Morales, I., Gomez-DeGabriel, J.M., Fernandez Lozano, J. & Garcia-Cerezo, A. (2004). Adaptive Cartesian motion control approach for a surgical robotic cameraman, *Proc. IEEE Int. Conf. on Robotics and Automation*.
- Navarro, A.A. (2009), *Angular Variation as a Monocular Cue for Spatial Perception*, PhD Dissertation, UPC, Barcelona, Spain.
- Olensis, J. (2000). A critique of structure-from-motion algorithms, *Computer Vision and Image Understanding*, Vol. 80, pp. 172-214.
- Ortmaier, T., & Hirzinger, G. (2000). Cartesian control issues for minimally invasive robotic surgery, *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 465-571.
- Pandya, A. & Auner, G. (2005). Simultaneous augmented and virtual reality for surgical navigation, *IEEE Annual Meeting of the North American Fuzzy Information Processing Society*.
- Park, J.S. (2005). Interactive 3D reconstruction from multiple images: a primitive-based approach, *Pattern Recognition Letters*, Vol. 26, No. 16, pp. 2558-2571.
- Payandeh, S., Xiaoli, Z. & Li, A. (2001). Application of imaging to the laparoscopic surgery, *IEEE Int. Symp. Computer Intelligence in Robotics Automation*, pp. 432-437.
- Rehbinder, H. & Ghosh, B.K. (2003). Pose estimation using line-based dynamic vision and inertial sensors, *IEEE Trans. Automatic Control*, Vol. 48, No. 2, pp. 186-199.
- Selig, J.M. (2000). Some remarks on the statistics of pose estimation, *Technical report SBU-CISM-00-25*, South Bank University London.
- Sun, J., Smith, M., Smith, L. & Nolte, L.P. (2005). Simulation of an optical-sensing technique for tracking surgical tools employed in computer assisted interventions, *IEEE Sensors Journal*, Vol. 5, No. 5.
- Taylor, R., Funda, J., Eldridge, B., Gomory, S., Gruben, K., LaRose, D., Talamini, M., Kavoussi, J. & Anderson, J. (1995). A telerobotic assistant for laparoscopic surgery, *IEEE Engineering in Medicine and Biology Magazine*, Vol. 14, No. 3, pp. 279-288.
- Wrobel, B. (2001). Minimum solutions for orientation, In: *Calibration and Orientation of Cameras in Computer Vision*, A. Gruen and T. Huang (Eds.), Chapter 2, Springer-Verlag.
- Yen, B.L. & Huang, T.S. (1983). Determining 3-D motion and structure of a rigid body using straight line correspondences, *Image Sequence Processing and Dynamic Scene Analysis*, Springer-Verlag.

IntechOpen

IntechOpen



## **Robot Vision**

Edited by Ales Ude

ISBN 978-953-307-077-3

Hard cover, 614 pages

**Publisher** InTech

**Published online** 01, March, 2010

**Published in print edition** March, 2010

The purpose of robot vision is to enable robots to perceive the external world in order to perform a large range of tasks such as navigation, visual servoing for object tracking and manipulation, object recognition and categorization, surveillance, and higher-level decision-making. Among different perceptual modalities, vision is arguably the most important one. It is therefore an essential building block of a cognitive robot. This book presents a snapshot of the wide variety of work in robot vision that is currently going on in different parts of the world.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Agustin A. Navarro, Albert Hernansanz, Joan Aranda and Alicia Casals (2010). An Approach to Perception Enhancement in Robotized Surgery Using Computer Vision, *Robot Vision*, Ales Ude (Ed.), ISBN: 978-953-307-077-3, InTech, Available from: <http://www.intechopen.com/books/robot-vision/an-approach-to-perception-enhancement-in-robotized-surgery-using-computer-vision>

# **INTECH**

open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen