Finding 3D Positions from 2D Images Feasibility Analysis

H. G. Lochana Prematunga
University of Colombo School of Computing,
35, Reid Avenue, Colombo 7, Sri Lanka
lochana.prematunga@ifsworld.com

Anuja T Dharmaratne
University of Colombo School of Computing,
35, Reid Avenue, Colombo 7, Sri Lanka
atd@ucsc.lk

***Abstract*—- In this paper, we prove that it is possible to recover the position (or coordinates) of an object using a single 2D image, given the size and shape of the image. Here we employ a purely mathematical proof to enable guaranteed accuracy. These theories and their derivations can be employed in recovering illumination patterns defined on the actual object using their images.**

Keywords-3D position, 2D image, uniqueness, mathematical modeling.

I. INTRODUCTION

There are situations in computer vision where it is required to determine the position of an object using a single 2D image. In this paper, these positions are expressed as coordinates in a system of coordinates defined by the camera. When an image of a certain object is given, there is a doubt whether there can be multiple object positions that may create this same image. But this paper takes a mathematical approach to prove that, when the shape and the size of an object are given with its image from a certain camera, the coordinates of the object can be determined uniquely. Therefore it is proved that there can be one and only one position for an object with respect to a camera, for a given image.

Also the coordinates of the object with respect to the camera is disclosed in the process as mathematical formulas. Here a rectangular object is taken as the example to prove the uniqueness. But any planer object satisfies the uniqueness condition, since a rectangle drawn on such an object has a unique image inside a given image of that original object.

These discoveries will be of importance in distance detection applications. As an example, an image of a satellite may be used to determine the distance between the satellite and the camera from which the image is taken. Similar approach may be employed to determine the position of a person in an airplane using a photograph of a wing (of the same flight) taken by him in an air plane crash situation.

Distance information may be used as data itself as explained above or it may be an intermediate data used to derive some other important information. For an example, brightness of an image pixel will not yield any brightness information (of the object) if the distance between the camera and the object is unknown.

Also the angle between the object and the reviewer may provide important information in some legal procedures. Another application of the angle matching will be to position the antenna in the most favorable direction for the receiver. This will play an important role in satellite communication and in communicating with space vehicles.

Positioning of 3D objects (surgical equipments) and recovering the position and orientation data of objects (organs) accurately is also a vital step in computer assisted surgeries.

Recovering 3D models from 2D images is also an important step in virtual reality applications.

Finally let us look into the structure of the paper. Next section is devoted to identify the previous work, which is related to the work in this article, done by other researchers. The section after that introduces the camera model used in the rest of the article. Section that follows is devoted to disclose the methodology used to recover the object from the image. And the last section in the paper is allocated to list the conclusions.

II. PREVIOUS WORK

There are attempts to recover 3D images using a series (2 or more) of 2D images of certain objects. One is the research done at the University of Ottawa [2]. This will use 2 images taken from 2 cameras simultaneously and the object is being recovered using the images of some feature points on the 3D model.

In another research there was a successful attempt of recovering a non-rigid 3D model [3] from a sequence of images created by a video stream. Also there are some tools developed to recover the 3D models from images with wide baselines [4]. These tools employ a method that uses a universal camera intrinsic matrix estimation technique to eliminate the need for camera calibration experiments.

In another research [5], there is an attempt to recover smooth objects using image contours that approximate the image with an octree spline structure. Some research work has also been carried out on recovering moving 3D objects [6]. This method consists of integrating the measured 2D motion of the object to recover its 2D-position in the image.

 ISBN: 978-1-61208-184-7

There is an interesting research [7] on recovering 3D object models from a single 2D image. In this method, the matching of corresponding features is employed to recover 3D data. Some research was done to recover the pose of a head [8] including the motion to be mainly used in virtual reality aviators.

An interesting approach to the problem has been presented in [9]. In this paper certain pre-analyzed object classes have been used. Objects belonging to these classes were then extracted from an image. Some work was also done using voting techniques [10] but without feature extraction. This will enable the method to be employed also for smooth objects. Another approach used was to synthesize 3D objects by comparing the image against the data in a pre-stored object library. In one of the studies described in [11] this approach was used in 3D model synthesis to even recover the data in the back (invisible) side of the objects. Some research [12] is also done on the difficult problem of decoupling the relative position recovery and relative orientation recovery.

III. CAMERA MODELING

The camera model used to establish the mathematical formulas is shown in the figure bellow:

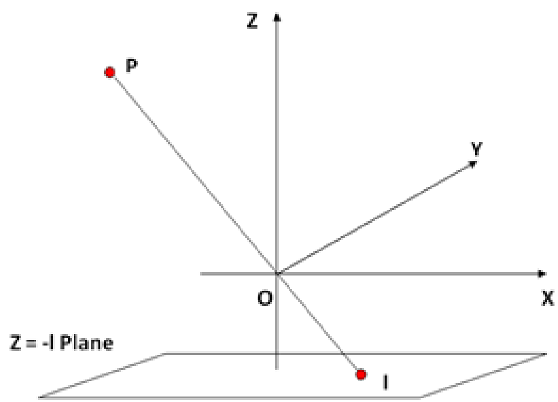

Figure 1: Camera model used in the derivation of mathematical formulas.

Let **O** $(0, 0, 0)^T$ be the center of the camera and the plane **Z = -l** be the imaging surface.

Let **P** $(x, y, z)^T$ be any point that is capable of generating an image on the camera and **I** $(x_I, y_I, -l)^T$ be its image.

Then for a given scalar t observe the relationship:

$$\overrightarrow{OP} = t\,\overrightarrow{IO}$$
$$\boldsymbol{P} = -t\,\boldsymbol{I}$$

That is

$$(x, y, z)^T = -t(x_I, y_I, -l)^T.$$

That is for a given image **I** $(x_I, y_I, -l)^T$ for a point **P**, **P** can be given by $(-tx_I, -ty_I, tl)^T$.

$$P = (-tx_I, -ty_I, tl)^T \text{ ------ (1)}$$

In this case the parameter t should be positive (t>0) in the real world.

Otherwise the point P will be inside the camera or behind it.

IV. RECOVERING THE OBJECT FROM THE IMAGE

Let $P_0P_1P_3P_2$ be a rectangular object with P_0P_2 = m, in the above 3D coordinate system.

And let $I_0I_1I_3I_2$ be its image on the imaging surface z = -l of the camera. Refer to Figure 2.

Let $I_0 = (x_0, y_0, -l)^T$, $I_1 = (x_1, y_1, -l)^T$, $I_2 = (x_2, y_2, -l)^T$ and $I_3 = (x_3, y_3, -l)^T$.

Then, equation (1) implies:

$$P_0 = (-t_0x_0, -t_0y_0, t_0l)^T,$$
$$P_1 = (-t_1x_1, -t_1y_1, t_1l)^T,$$
$$P_2 = (-t_2x_2, -t_2y_2, t_2l)^T,$$
$$P_3 = (-t_3x_3, -t_3y_3, t_3l)^T.$$

Where t0, t1, t2 and t3 are positive scalar (parametric) values.

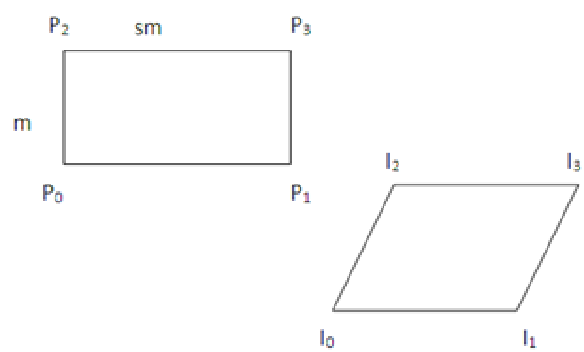

Figure 2: Actual object $P_0P_1P_3P_2$ and its image $I_0I_1I_3I_2$.

Since $P_0P_1P_3P_2$ is a rectangle: $\overrightarrow{P_0P_1} = \overrightarrow{P_2P_3}$.

$$\boldsymbol{P_1 - P_0 = P_3 - P_2}$$

$$(-\boldsymbol{t_1x_1}, -\boldsymbol{t_1y_1}, \boldsymbol{t_1l})^{\boldsymbol{T}} - (-\boldsymbol{t_0x_0}, -\boldsymbol{t_0y_0}, \boldsymbol{t_0l})^{\boldsymbol{T}} = (-\boldsymbol{t_3x_3}, -\boldsymbol{t_3y_3}, \boldsymbol{t_3l})^{\boldsymbol{T}} - (-\boldsymbol{t_2x_2}, -\boldsymbol{t_2y_2}, \boldsymbol{t_2l})^{\boldsymbol{T}}$$

Comparing z components: $t_1l - t_0l = t_3l - t_2l$

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-184-7

$$t_1 - t_0 = t_3 - t_2.$$

Thus,

$$t_1 + t_2 - t_3 = t_0 \quad \text{-- (2)}$$

Comparing x components:

$$-t_1x_1 - (-t_0x_0) = -t_3x_3 - (-t_2x_2).$$

Therefore,

$$t_1x_1 + t_2x_2 - t_3x_3 = t_0x_0 \quad \text{-- (3)}$$

Similarly, by comparing y components:

$$t_1y_1 + t_2y_2 - t_3y_3 = t_0y_0 \quad \text{--(4)}$$

By solving these equations:

$$t_3 = t_0 \frac{(x_{10}y_{21} - x_{21}y_{10})}{(x_{31}y_{21} - x_{21}y_{31})} \quad \text{--- (5)}$$

where $x_{ij} = x_i - x_j$.

And

$$t_2 = t_0 \frac{(x_{10}y_{31} - x_{31}y_{10})}{(x_{31}y_{21} - x_{21}y_{31})} \quad \text{--- (6)}$$

That is:

$$t_2 = at_0 \quad \text{---- (7)}$$

And

$$t_3 = bt_0 \quad \text{--- (8)}$$

Where

$$a = \frac{(x_{10}y_{31} - x_{31}y_{10})}{(x_{31}y_{21} - x_{21}y_{31})}$$

and

$$b = \frac{(x_{10}y_{21} - x_{21}y_{10})}{(x_{31}y_{21} - x_{21}y_{31})}.$$

Considering the fact that $P_0P_2 = m$.

Therefore $|P_2 - P_0| = m$

$$|(-t_2x_2, -t_2y_2, t_2l)^T - (-t_0x_0, -t_0y_0, t_0l)^T| = m$$

$$(t_2x_2 - t_0x_0)^2 + (t_2y_2 - t_0y_0)^2 + l^2(t_2 - t_0)^2 = m^2.$$

Using the equation (7):

$$(at_0x_2 - t_0x_0)^2 + (at_0y_2 - t_0y_0)^2 + l^2(at_0 - t_0)^2 = m^2$$

$$t_0^2\{(ax_2 - x_0)^2 + (ay_2 - y_0)^2 + l^2(a-1)^2\} = m^2 \quad \text{---(9)}$$

Therefore we have

$$t_0 = \frac{\pm m}{\sqrt{\{(ax_2 - x_0)^2 + (ay_2 - y_0)^2 + l^2(a-1)^2\}}}$$

But by definition of t_0, we have $t_0 > 0$.

Therefore:

$$t_0 = \frac{m}{\sqrt{\{(ax_2 - x_0)^2 + (ay_2 - y_0)^2 + l^2(a-1)^2\}}} \quad \text{---- (10)}$$

That means t_0 is a unique value for a given m.

Therefore by considering equations (5) and (6) we can uniquely determine t_2 and t_3.

- For a given image $I_0I_1I_3I_2$, we can uniquely determine the points P_0, P_3 and P_2.
- By geometry we can uniquely determine the point P_1.
- For a given image $I_0I_1I_3I_2$ we can uniquely determine the rectangle $P_0P_1P_3P_2$ that created the image.

Now consider the image shown in Figure 3.

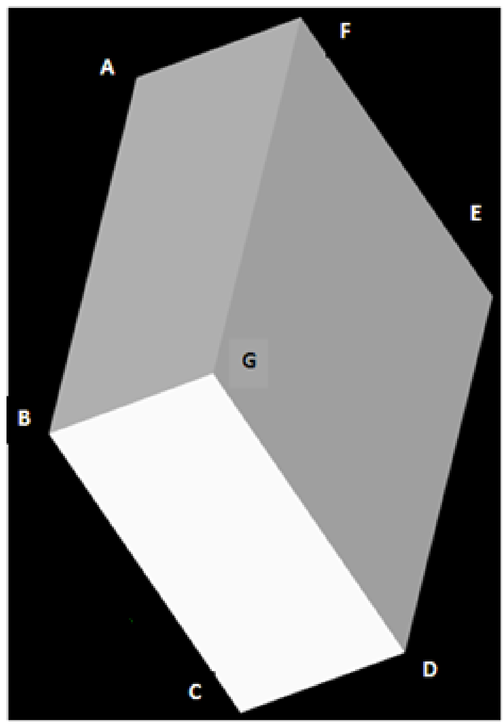

Figure 3: Image of an actual object to be measured.

Point	Pixel X	Pixel Y	Actual X	Actual Y
A	103	58	2.682292	1.510417
B	33	340	0.859375	8.854167
C	186	562	4.84375	14.63542
D	316	515	8.229167	13.41146
E	387	231	10.07813	6.015625
F	233	10	6.067708	0.260417
G	163	293	4.244792	7.630208

Here the ratio 38.4 was taken to convert pixel values to centimeters. Considering the geometry, it is safe to take l = 1, without losing the generality. Considering the rectangle BCDG, and with BG = 1cm, we got $t_0 = 0.0865161$.

That means the actual coordinates of B is given by (-0.07435, -0.76603, 0.0865161) where all the coordinates are given in cm. Similarly other coordinates can also be calculated.

V. CONCLUSION

Given the shape and size of an object with an image of it, there can be one and only one position for the object with respect to the camera from which the image is taken. This implies that for the given class of objects, (planar objects with an identifiable rectangular shape on them) when an image is given, and the 3D positions are calculated, there is no need to find out whether there are any more possible object instances that we need to consider. This will greatly reduce the complexity of successive steps in a system where object extraction is an intermediate step. Otherwise it would be required to apply the same algorithm to multiple possible objects and validate each other to select the best appropriate.

This simple observation not only reduces the time and complexity of the resulting systems, but also reduces the possibility of errors.

References

[1] Eric Marchand. Control Camera and Light Source Positions using Image Gradient Information. IEEE Int. Conf. on Robotics and Automation, ICRA'07, Roma, Italia, April 2007.

[2] Lavoie P., Lonescu D., Petriu E.M.. 3D object model recovery from 2D images using structured light. Instrumentation and Measurement, IEEE Transactions.

[3] Christoph Bregler, Aaron Hertzmann, Henning Biermann. Recovering Non-Rigid 3D Shape from Image Streams.

[4] Yuzhu Lu, Shana Smith. A Comprehensive Tool for Recovering 3D Models From 2D Photos With Wide Baselines. J. Comput. Inf. Sci. Eng. December 2006.

[5] Lavallee S, Szeliski R. Recovering the position and orientation of free-form objects from image contours using 3D distance maps. IEEE Pattern Analysis and Machine Intelligence, IEEE Transactions on Apr 1995.

[6] Cretual A, Chaumette F, Bouthemy P. Complex object tracking by visual servoing based on 2D image motion. IEEE Pattern Recognition, 1998, Proceedings, Fourteenth International Conference on 16-20 Aug 1998.

[7] D Danial Sheu, Alan H Bond. A Generalized Method for 3D Object Location from Single 2D Images. Pergamon Press Ltd, 1992 Pattern Recognition Society.

[8] M D Cordea, E M Petriu, N D Georganas, D C Petriu, T E Whalen. 3D Head Pose Recovery for Interactive Virtual Reality Avatars. IEEE Instrumentation and Measurement Technology Conference, Budapest, Hungary, May 21-23, 2001.

[9] Han-Pang Chiu, Huan Liu, Leslie Pack Kaelbling, Tom´as Lozano-P´erez. Class-Specific Grasping of 3D Objects from a Single 2D Image. The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems October 18-22, 2010, Taipei, Taiwan.

[10] Kenpo Tsuchiya, Shuji Hashimoto, Toshiaki Matsushima. A Pixel Voting Method to Recover 3D Object Shape from 2D Images. MVA'94 IAPR Workshop on Machine Vision Applications Dec. 13-1 5, 1994, Kawasaki.

[11] Tal Hassner, Ronen Basri. Example Based 3D Reconstruction from Single 2D Images. Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06) IEEE.

[12] Rodrigo L Carceroni, Christopher M Brown. Decoupling Orientation Recovery from Position Recovery with 3D-2D Point Correspondences. University of Rochester, Computer Science Department, Rochester, NY-14627-USA.