

Geometric Vision Tricks for Shape Detection

Chengyuan Peng

VTT Technical Research Centre of Finland
Vuorimiehentie 3, Espoo, Finland
chengyuan.peng@vtt.fi

Abstract - In many our projects, we often met problems of recognition of simple geometric shapes (points, lines) in images where geometric shape boundary representation suffers from various inaccuracies caused by noise in the data. And it was often the first and crucial step for further processing tasks such as camera pose estimation and localization. In this paper, we discuss approaches to solving different problems and their relations to the tasks of the recognition. We present how such simple geometric computer vision based methods facilitate the shape recognition process by presenting several examples. Experiments and implementation show that our methods achieved very good results and can be used to solve common problems in each area.

Keywords: Shape detection, Geometry computer vision, Camera projection model, Epipolar geometry.

1. Introduction

The extraction of geometric shape information from the camera images for particular purposes is needed in various applications such as camera calibration and camera pose estimation. There are several issues related to these tasks, namely, the detection of the object with certain geometric shape, and the localization of the object. However accurate detection and localization suffer from poor quality of the images taken from different environment as listed in the following paragraphs.

Problem 1: Under-water marker recognition and localization. Fig. 3 shows a water quality measurement device Secchi3000 where a mobile phone is on the top of the container (Web-1). The mobile phone camera looks at the water container through a hole. Inside the container there are two marker plates consisting of white, grey and black areas. The tags are located at different depths in order to derive water turbidity (Toivanen et al., 2013). The shape detection task was to obtain accurate positions of white, grey and black rectangles in an image. Image quality taken from mobile phone cameras was noisy, blur, and sometimes rotated with regards to the scene. The approximation becomes poor when marker is deep with respect to its distance from the viewer. The lower marker detection was difficult because the water properties affect the refraction and thus the pixel values, especially in cases of very turbid waters. Detection method based solely on colour would most fail. A purely shape-based approach will not always be robust to this kind of artefacts (Toivanen et al., 2013).

Problem 2: Robust geometric feature points matching between two images. One case is when two images of the same scene were taken from very different positions, by two different types of cameras, and without knowledge of the intrinsic camera parameters (see Fig. 5). Another case is when even if the two images were taken from the same position by the same camera, the extreme change in scene appearance, dominant occlusions, changing of lightning condition, and moving object can still make the matching challenging. Because there are arbitrary rotations of the camera, translation between camera centres, and changes of internal camera parameters, as a result, this generates a lot of incorrect point correspondences and the number of correct matches needed is insufficient.

Problem 3: Tennis court lines detection for sports. The task was to estimate camera pose using four court lines in an image taken from a mobile phone camera. The difficulty was that a mobile phone camera has very narrow field of view and therefore the tennis court lines were not always visible, especially the furthest border lines. That is in the most case the image points only contain three lines (two are parallel lines) (see Fig. 6). The camera pose cannot be estimated uniquely.

2. Methods and Results

In this section we will describe in detail the simple geometric vision approaches used in shape detection and robust geometric feature matching to the above mentioned different problems. In particular we use the properties of camera projections such as perspective projection and weak perspective projection and epipolar geometry.

2. 1. Weak Perspective Camera Model

Fig. 1 gives a cross-sectional view of the Secchi3000 device after filling the container with water. Because of the water refraction, the images (see Fig. 4) formed on image plane were from refracted markers where a scene's relief is small compared to the overall distance from the camera. This process can be modelled as a weak perspective projection in geometry computer vision which is a coarser approximation of the geometry of the image formation process (Forsyth 2012). Fig. 2 illustrates the modelled projection process. First, a near parallel projection followed by isotropic scaling is applied to map marker edge PQ onto edge $P'Q'$ on the frontal plane Π . The frontal plane Π is parallel to the image plane. Then, a perspective projection is employed to map $P'Q'$ to pq on the image plane.

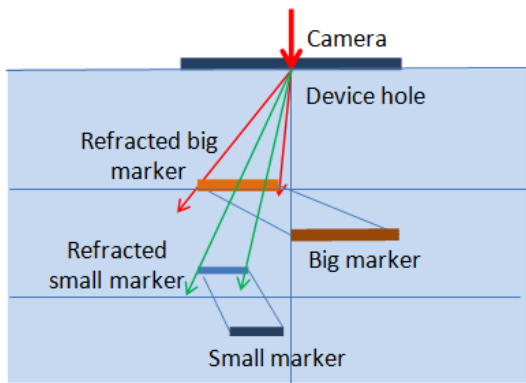


Fig. 1. Cross-sectional view of the device.

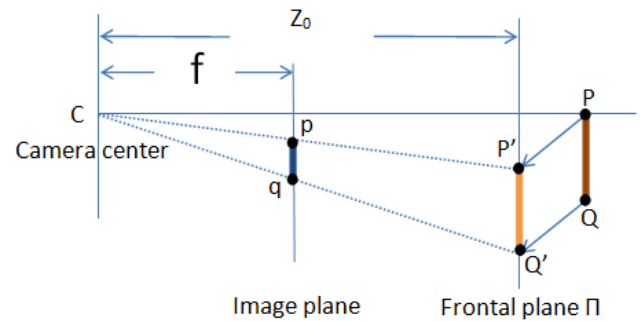


Fig. 2. Weak perspective camera model.

The weak-perspective model is appropriate when the observed markers are far from the camera relative to their sizes. The upper single black/white/grey tag is 13x7mm and the lower single black/white/grey 10x5mm. In this situation the perspective distortions are relatively small. The properties of weak perspective marker representations can be represented with respect to similarity transformation which is an isometry composed with an isotropic scaling. It is a specialization of projective transformation (Hartley 2004).

Projective invariant features are employed to facilitate the recognition process. That is similarity transform preserves the angles between two lines and the ratio of two line lengths. Similarly a ratio of two areas is an invariant because the scaling (squared) cancels out. In Fig. 2, the magnification (or ratio) m ($-f/Z_0$) can be taken to be constant; the vectors $P'Q'$ and pq are parallel, and $\|pq\| = m\|P'Q'\|$. Angles between lines are not affected by rotation, translation or isotropic scaling. In particular parallel lines are mapped to parallel lines.

Based on the above assumptions and the known sizes of the markers we only need to detect the black tags (because it is easier to detect than other colour tags) and other tags can be derived according to invariant features. The markers captured in very turbid water could be successfully recognized and localized (see Fig. 4). The results were shown that using invariant features for this task the recognition process becomes easier and the error introduced by the use of such an approximation is negligible.



Fig. 3. Secchi3000 device [Web-1].

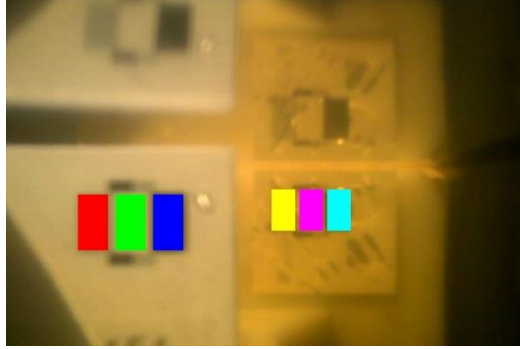


Fig. 4. Under-water marker localization.

2. 2. Epipolar Geometry Method

The epipolar geometry in geometry computer vision is the intrinsic projective geometry between two views. It is independent of scene structure, and only depends on the cameras' internal parameters and relative pose (Hartley 2004). The epipolar geometry method can be used to improve the robustness of image feature point matching. This is based on the fact that given an image taken at an arbitrary location in the area there exist a fundamental matrix between this image and another image of the same scene but captured from a different camera (Lu 2004).



Fig. 5. Robust SIFT feature point matching.

A robust matching method based on RANSAC robust estimation on fundamental matrix constraints without knowing camera intrinsic parameters was employed. Fig. 5 illustrates that a number of robust point correspondences were stably determined for two images across different views of the same scene taken by two different cameras with significant illumination changes. The left side image was captured with Nokia E71 mobile phone. The right side is the panoramic image from Nokia's NAVTEQ data. This geometry was motivated by considering the search for corresponding points in two-view matching. Given a point from the left image in Fig.5, multiplying by the essential matrix, it will tell us which epipolar line to search along the camera view (Hartley 2004). Due to noisy environment it was often happened that more than one point were detected within the specified threshold. According to the properties of epipolar geometry, each point can be found on a certain straight line (i.e. an epipolar line) on image plane.

2. 3. Perspective Projection Invariance

To estimate a mobile camera pose in front of a tennis court, we need to detect four points B1, B2 S1 and S2 in the left side of Fig 6. However, point B2 is in most cases out of the image area. The operation of central projection preserves some geometric properties, that is, lines project to lines. Line intersections can be obtained when intersection points are outside images. Under knowing that line invariance in

perspective projection and using line equation we can obtain point B2. The right side of Fig. 6 illustrates the detection results. Using the estimated pose, the 3-D tennis court model lines were accurately re-projected onto the image plane.

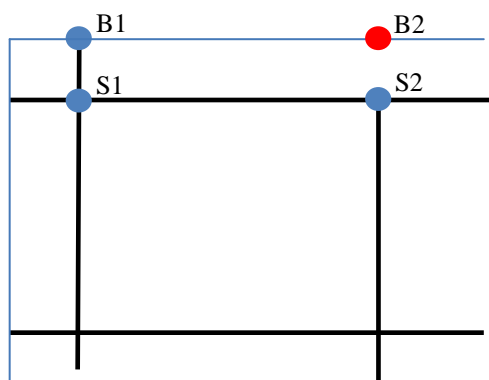


Fig. 6. Tennis court line detection.

3. Conclusions

Shape detection is still one of the most difficult tasks in the field of computer vision. In this paper, perspective approximations were studied for applications to object shape recognition. This paper presented the analysis methods in terms of seeking geometric patterns defined by similarities. These approximations yielded very good results and presented the advantage of simplicity and robustness. And all the methods mentioned in the above section tested and worked in the real projects. Although we studied specific tasks in areas, the methods are applicable to solving more common similar problems.

Acknowledgements

We thank Nokia research centre for providing NAVTEQ data in the project. Special thanks to our colleagues Petri Honkamaa, Mika Suhonen, Alain Boyer and Tuomas Kantonen for their helps during the project work.

References

- Forsyth D.A., Prone J. (2012). "Computer Vision: A Modern Approach Second Edition" PEARSON.
- Hartley R. and Zisserman A. (2004). "Multiple View Geometry in Computer Vision" 2nd edition Cambridge University Press.
- Lu X. and Manduchi R. (2004). Wide Baseline Feature Matching Using the Cross-Epolar Ordering Constraint "Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition," Washington, DC, USA, Jun. 27-Jul. 2, pp. I-16 - I-23 Vol.1.
- Toivanen T., Koponen S., Kotovirta V., Molinier M. and Peng C. (2013). Fluoride removal studies from contaminated ground water by using bauxite. Environmental Systems Research, 2:9.

Web sites:

Web-1: <http://www.syke.fi/download/noname/%7B0D82C656-917B-4E96-A445-7B35BD6467AF%7D/44732> consulted 20 Dec. 2010.