Object Recognition Using a Tag

J. Matas, L. M. Soh and J. Kittler
CVSSP, Department of Electronic and Electrical Engineering,
University of Surrey, Guildford, Surrey GU2 5XH, United Kingdom
e-mail: {g.matas,l.soh,j.kittler}@ee.surrey.ac.uk

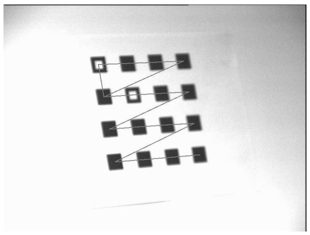
Abstract

We propose a method for object recognition in an unconstrained condition. This includes wide range of illumination, unknown view points and complicated background. We simplified the general problem by placing special design patterns (Tsai[1]'s camera calibration chart) on the object that allows us to solve the pose determination problem easily. After establishing the methodological framework, we experiment with this technique on a wide range of real data to show the reliability of the approach. We argue that with this capability we should be able to a) establish camera position with respect to a landmark b) recognise an object which is tagged with the calibration chart and c) test any camera calibration and 3D pose estimation routines, thus facilitating future research and applications in mobile robots navigations, 3D reconstruction and stereo vision.

1 Introduction

Object localisation and recognition is one of the most challenging problems studied in the field of computer vision. As yet, successful 3D object recognition has only been demonstrated either for a very specific class of objects (eg. polyhedra) or in environments with controlled illumination, viewing geometry, background etc.[6]. On the other hand, if a good approximation of the pose of an object is available (in the camera coordinate frame), verification of its identity may be simply achieved by a direct comparison of the image with a model projected on the image plane.

In order to recognise objects in unconstrained environments we have developed a system based on tagging of objects of interest with a readily detectable small pattern of black squares on a white background. The pattern was designed by Tsai ([1], [2]) for camera calibration and standard procedures exist for determining its 3D pose with respect to the camera. We have developed a complex but extremely robust technique that allows us to detect the Tsai chart over a large range of scales, in very difficult illumination conditions and in the presence of partial occlusion (see fig. 1). A large part of this paper describes the chart detection algorithm.

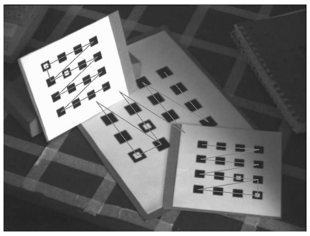


(a) Unfocused input image

(b) Multiple charts with occlusion.

(c) Multiple charts, bad lighting.

(d) Multiple charts, bad lighting.



(e) Multiple charts, occlusion.

(f) Multiple charts, acute angle.

Figure 1: Results on test images with 1. low contrast (c)(d)(f), 2. highly uneven illumination (c)(d)(e), 3. charts appearing at arbitrary orientations and a wide range of scales (b)(c)(d)(e)(f).

Once the chart is (or charts are) detected in the image, the chart pose in the camera coordinate system is known. To identify the object that is tagged by the chart we transform the edge map in the vicinity of the chart to a canonical frame. Next, prestored edge maps from the object database (corresponding to a canonical view) are compared one by one with the transformed edge map and the map attaining a maximum score, if above a threshold, identifies the tagged object. The method can be applied to add navigation of a mobile robot. In this scenario, a Tsai chart attached to a landmark may serve two purposes. Firstly, it is designed specifically to facilitate computations of the relative position of the camera with respect to the landmark. Secondly, by allowing the transformation of the landmark into a canonical frame, landmark recognition can be simplified.

2 Chart Detection

It is often assumed that the detection of charts or markers which are designed to enhance detectability is trivial. However, available implementations of the standard calibration procedure (Willson[3]) do not include detectors, taking image coordinates of points as input. Our experience suggests that generic low level processing approaches (Canny[4]), which do not exploit the specific properties of such standard objects are likely to fail in most but near optimum imaging conditions where the lighting, pose and scale of the object are all favourable. Even if features belonging to the chart are detected, they must be segmented from the background manually.

The design of the calibration chart, as shown in Figure 1a , is simple and the chart can be easily reproduced for use as an object tag or a navigational landmark. We can observe invariant properties that can be exploited to make the chart recognition process robust. Firstly, from the contrast between the black squares and white chart background, we can assume that the image gradient for pixels on the chart is certainly highest locally and not far from the global maximum. Secondly, the compactness, high over a large range of viewing angles, and convexity, invariant under perspective, of the squares can be used to remove most background features satisfying the contrast criterion. Geometrically, we can also take advantage of the relationship between the size and distance of the squares as well as their linear configuration.

The general plan devised to take the full advantage of the above model properties requires several stages of processing. The proposed technique involves a few simple but powerful steps. In order to make the system robust, individual blocks must be able to handle worst case scenarios. Furthermore, each block must provide the best possible output for the next block so that there will not be any need for excessive computation at any stage to cater for inadequacies of the earlier stages of processing.

In the first step, Sobel operator is used for filtering the raw image. A 3x3 mask (ie. no smoothing) is used both to maintain positional accuracy and to facilitate the detection of the smallest feature of interest. This is followed by adaptive thresholding based on local gradient histogram analysis. Next non-maximum suppression and hysteresis thresholding are performed. Subsequently, each square is linked and gap bridging is applied where necessary. In the stage of feature extraction, the centre of gravity of each square is selected as the most stable and accurate feature representing the squares of the chart. Next, constraint checks are performed on the compactness and convexity of the squares.

In the stage of feature extraction, the center of gravity of each square is selected as the most stable and accurate feature representing the squares of the chart. The imprecise nature of corners and edges is caused by the "shrinking" effect as shown in Figure 2.

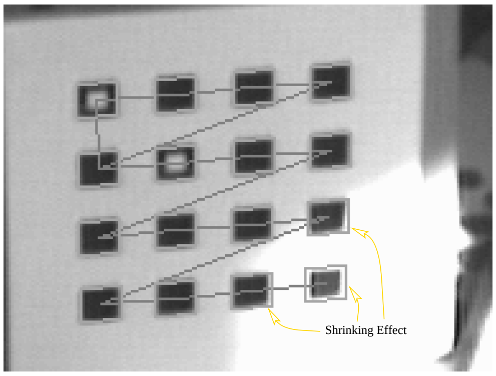


Figure 2: The "shrinking" effect results in inaccurate corners and edges as compared to the center of mass.

To segment the calibration chart, Attributed Relational Graph matching (Bunke[5]) between the features extracted and the model is performed. Various model constraints are utilised to minimise the search space. This transforms the segmentation into a Maximum Weight Matching problem in the graph space. Finally, missing features (occluded or undetected) are predicted based on prior knowledge of the chart.

The proposed chart detection procedure has been extensively tested. A sample set of test images was selected to cover a wide range of conditions with which the system should be able to cope. Thus the set includes charts of different sizes, with complicated background, occlusion, bad lighting conditions, extreme perspective effects and multiple charts. Figure 1 shows results from the sample sets under various test conditions. A 100% detection rate was achieved on this sample set of more than 100 test images. Furthermore, the system was successfully run continuously, processing images grabbed on-line.

The processing time required is between 0.22 to 0.35 seconds for a 256x256 image and 1 to 1.5 seconds for a 768x576 image depending mainly on the level of noise and the degree of difficulties. The distribution of the computation time is quite even with the Sobel filtering and segmentation taking approximately 20% each. All the above tests were conducted on a Silicon Graphics (SGI) Power Challenge.

3 Recognition and Landmark Detection

This part of the project is still in progress. A small database of images with tagged objects and landmarks has been assembled. Example images are shown in figs. 3, 4a and 5a. In first experiments edge maps taken from the vicinity of the detected chart were transform to a coordinate frame corresponding to a frontal (canonical) view. As shown in images 4b and 4d (5b and 5d respectively) the transformed maps can be compared by direct correlation. At present a large database is being assembled and recognition performance tested.

4 Summary and Conclusion

We have developed a robust, model based scheme for the detection of calibration charts. The technique will facilitate a) reliable computation of camera position with respect to a landmark b) recognition of objects which are tagged with the calibration chart and c) testing of any camera calibration and 3D pose estimation routines, thus providing enabling technology for vision based mobile robot navigation, 3D reconstruction and stereo and motion analysis.

References

[1] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 3:323–344, 1987.

[2] R. Y. Tsai. Synopsis of recent progress on camera calibration for 3D machine vision. In O. Khatib et al., editor, *The Robotics Review*. MIT Press, 1989.

[3] R. Willson. Tsai Camera Calibration Software. Web Site *http://www.cs.cmu.edu/People/rgw/TsaiCode.html*, Oct. 1995.

[4] J. Canny. A Computational Approach to Edge Detection. *IEEE PAMI-8*, pp.679, Jan. 1986.

[5] H. Bunke and B.T. Messmer. Efficient Attributed Graph Matching and its Applications to Image Analysis. *ICIAP*, pp.45-55, Sep. 1995

[6] W. Eric L. Grimson. *Object Recognition*. MIT Press, 1990.

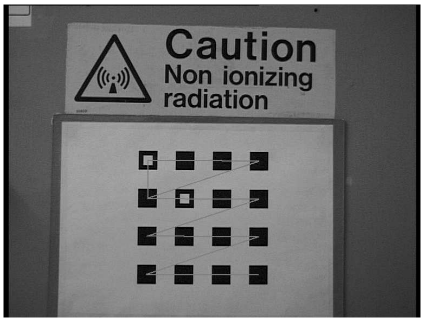
(a)

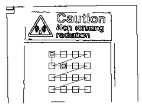
(b)

(c)

(d)

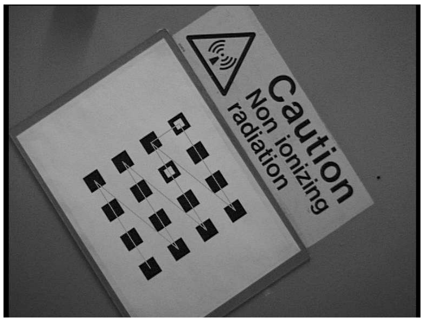
Figure 3: Sample images of tagged objects (a),(b) and landmarks (c),(d).

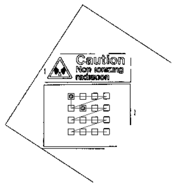




(a) reference view

(b) canonical view

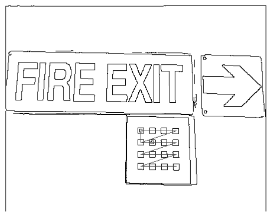




(c) test view

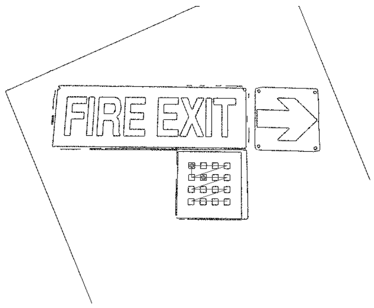
(d) canonical view

Figure 4: Mapping edge maps into canonical view.



(a) reference view

(b) canonical view



(c) test view

(d) canonical view

Figure 5: Mapping edge maps into canonical view.