Recognition Using Labelled Objects

L. M. Soh, J. Matas, J.Kittler

Centre for Vision, Speech and Signal Processing,
University of Surrey, Guildford, Surrey GU2 5XH, United Kingdom
{l.soh,g.matas,j.kittler}@ee.surrey.ac.uk

Abstract

General object recognition is a difficult problem. In [2], we proposed a novel solution for object recognition in an unconstrained environment. We simplified the recognition problem by attaching a special planar pattern on objects of interest. This approach allows us to determine the pose easily. A robust detector for the pattern was developed for the first stage of the solution (Soh[3]). This paper investigates the next part of the recognition process, i.e. matching of the models using a modified technique based on the Chamfer matching algorithm (Borgefors[1]). The algorithm is enhanced by augmenting the matching process by an additional step which promotes consistency of image gradient directions of the corresponding points.

1. Introduction

The problem of 3D object recognition is considered. We exploit the fact that if the 3D pose (with respect to the viewer) of an object is known, its identity can be confirmed simply by a direct comparison of the image with a model projected onto the image plane. The pose detection is facilitated by placing a readily detectable small pattern of black circles on a white background on the objects of interest. This pattern or chart was based on the design by Tsai[4],[5] for camera calibration and a shareware exists for determining its 3D pose with respect to the camera [6]. A complex but highly robust technique [3] has been developed to handle detection of the chart over a large range of scales, in very difficult illumination conditions and in the presence of partial occlusion. Wherever the chart is (or charts are) detected in the image, the chart or object pose in the camera coordinate system is known.

In this paper we address the issues of the choice of the model, the problem of matching the model and pose optimisation. To identify an object that is tagged by the chart we use the edge map in the vicinity of the chart. In contrast to previous effort, higher level features need not be extracted since the edge map will contain sufficient information about the model in terms of texture, shape and size. Besides that, edges are less sensitive to changes in illumination compared to grey levels. For comparing the two binary edge maps, we modified the Chamfer matching algorithm to get a more reliable correlation score. In particular, we have augmented the matching process by an additional step which promotes the consistency of image gradient directions at the corresponding (match) points. This enhancement upgrades the Chamfer matching algorithm to a more reliable and useful tool for correspondence analysis and matching. The pose estimate of the object is refined iteratively until the compound matching criterion is optimised. The proposed approach reflects more accurately the matching task and enhances the matching result. It has been tested on a large number of images and the test results are promising.

2. The Modified Chamfer Matching Algorithm

In registered views, edges, which are the local maxima of a gradient map, provide stable information about the relative position and gradient direction of the object surface discontinuities and boundary even at different lighting conditions. Hence, the edge map around the chart is used to represent each object.

For matching two binary images, we used a modified technique based on the Chamfer matching algorithm as it is quite insensitive to noise and other disturbances [1]. The Chamfer matching algorithm searches for the best fit of edge points from two different images. The first step of the algorithm requires to generate the distance map from a binary edge map. This distance map associates with each pixel, its distance from the closest edge profile pixel. As the true Euclidean distance is costly to compute, we use a sequential Chamfer Distance Approximation. Two passes over the image are needed for this approximation. The details of the computation of the distance map can be found in Borgefors[1]. By placing another edge map over the distance map, the distances between the two edge maps can be read along the edge profile.

The original Chamfer matching algorithm, as described above, only gives the distance to the nearest edge but the exact position of the nearest point/s is/are not computed. The modification was made to keep track of the nearest point as well. The additional point correspondence information is required to facilitate a more sophisticated matching of image to models which makes use of the edge gradient measurements. The change requires an additional point map to store the nearest point. First, the point map is initialised with each point storing their own position. During the two

passes, the current edgel will not only store the distance of the nearest pixel in the model but also the point map content of the nearest pixel. The final distance and position maps allow the user to read both the position of the nearest point and its distance. This modified Chamfer Matching technique could be applied to other areas where point correspondence based on nearest distance is required to be computed efficiently.

3. The Matching Strategy

The model or the edge map of each object is acquired by taking two or more images with different background and the model is segmented by correlation of the images in the same canonical frame. This method is especially useful when the object is a fixture, e.g. a road sign. First, the edges and their gradient direction are extracted. Next, the Chamfer distance and point map is computed on one of the images using the edge map. Following this, with the pose known, the edge map of the other image is transformed onto the image plane of the distance map. By superimposing the transformed edge map of the other image on the distance map, a set of distances that stand between them can be found along the edge profiles. This set of correlation distances is used to differentiate between the model and the rest of the image.

Further, a gradient direction consistency constraint is used to eliminate random matches and noise. This is achieved by comparing the transformed gradient direction with the associated match point from the point map. It is assumed that the gradient direction would be quite different for a false match. This gradient direction based search helps to maximise the matching function. It is noted that the edge gradient magnitude, as compared to the gradient direction, at the border of the object is more unreliable as it is dependable on the background. Thus, the gradient magnitude of the object was not used as a search constraint.

A problem occurs when the edge profile of one image extends over the other : then, the score does not reflect the best match anymore. To handle this, we match the two images in both ways, i.e. image A against image B then image B against image A. The lowest score is taken.

For matching, a similar method is applied. However, in practice, we use a prestored distance map, gradient direction map and pose from the object database and compare them one by one with the transformed edge map of the test image. Besides this, a geometric transformation is used to warp one image to another in order to minimise the distance between them (see section 4 below), as expressed using a given distance measure. The map attaining the minimum distance between itself and the image, if above a threshold, identifies the tagged object.

4. Optimisation of the Chamfer Algorithm

Initially, the pose estimation based on the special pattern can only transform the image to a region of convergence or close to the plane of the hypothesised model as the marker in relation to the object will be of relatively small size. From experiments, we found that we could not expect the transformation from a small chart to be accurate far away from the chart. However, even slight pose mismatches will make it impossible to compare the scene and model image directly. Thus a pose refinement will have to be attempted first. Using the refined 3D pose parameters, we build a new transformed edge map. This new compensated edge map is superposed on the scene image distance map and a new set of distances is computed. The best match is obtained by finding the set of transformation parameters minimising this set of distances. This process is applied iteratively to achieve the best possible registration.

A small database of images with tagged objects and landmarks has been assembled. The matching and optimisation process are shown in Fig. 1. The pose optimisation brings the initial cue of the transformation nearer to the model. Next, test images of multiple objects in complex scenes are introduced and object recognition takes place. The best match is shown by segmenting the model from the test image. The outcomes of the tests shows a favourable distinction between a mismatch and a hit.

5. Conclusion

We have successfully developed a robust and efficient technique for model matching of binary images using the modified Chamfer matching technique. The results show that the original chamfer matching idea has now been developed into a universally useful edge matching algorithm. The technique facilitates reliable Tagged Object Recognition (TOR), where landmarks and objects are tagged with a special pattern. This provides enabling technology for vision based mobile robot navigation, 3D reconstruction and scene modelling.

References

[1] G. Borgefors. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE PAMI*, 0:849–865, Nov. 1988.

[2] J. Matas, L. M. Soh, and J. Kittler. Object recognition using a tag. *IEEE ICIP'97*, 2:887–891, Oct. 1997.

[3] L. M. Soh, J. Matas, and J. Kittler. Robust recognition of calibration charts. *IEE 6th ICIPA*, 2:487–491, July 1997.

[4] 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, Aug. 1987.

[5] 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.

[6] R. Willson. Tsai camera calibration software. *Web Site: http://www.cs.cmu.edu/People/rgw/TsaiCode.html*, Oct.1995.

Fig. 1   <u>Matching and Optimisation</u>

Image A

Image B

Edge Map

Transform Edge Map
to Same Canonical View

Distance Map

Without Pose
Optimisation

Correlation Using Edge
and Distance Map

After Pose
Optimisation

Fitting

Fitting

Match accepted