A Scanner for Computing Box Dimensions in Real Time

Leandro A. F. Fernandes* Manuel M. Oliveira*
Instituto de Informática - UFRGS

Figure 1: Prototype of the scanner for computing box dimensions.

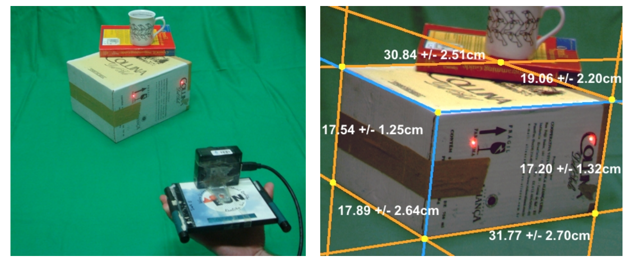

Figure 2: Usage of the scanner (left). Camera's view (from another position) showing the recovered edges, vertices, dimensions and estimated errors computed in real time (right).

Abstract

We describe a new method for computing dimensions of boxes from single perspective projection images in real time. We demonstrate the proposed approach by building a scanner and using it to compute dimensions of real boxes. This can be a useful tool for companies that handle boxes in their day-by-day operations, such as carriers, airline companies, and warehouses. It solves a specific problem: how measure a box that is traveling along a conveyor belt, a situation in which you cannot touch the box physically.

1 Description

We present a completely automatic approach for real-time computation of boxes dimensions from single perspective projection images. Our scanner prototype (Figure 1) computes the box dimensions and estimates the error in the computed measurements simply by pointing it to the target box (Figure 2 and accompanying video). The scanner consists of three modules: a firewire color camera, two parallel laser pointers and some software module. The camera is mounted on a plastic box and the laser pointers are aligned and glued to the sides of the box. The laser beams are (almost) perpendicular to the image plane and the distance between them is known.

Given a box, the dimensions are computed from its silhouette and from the projections of the two parallel laser beams on one of the visible faces of the box. The complete process can be divided into three stages: (i) finding the laser dots and the visible vertices of the box in the image; (ii) eliminating the inherent ambiguity associated with perspective images; and (iii) computing the 3D positions of the box vertices and, from them, the box dimensions.

The projection of the laser beams can be found by looking for the highest luminance pixels in the image. In order to avoid the influence of textures during the identification of the box edges, the scene has a known background color (Figure 2). The system then generates a binary image, which distinguishes the background from the foreground objects. A search starting from the position of one of the laser dots is used to identify the silhouette of the foreground objects that overlap the silhouette of the target box [Fernandes et al. 2006]. Thus, the scene can contain any number of objects, as far as their projected silhouettes do not completely occlude any edge of the target box silhouette. A new and efficient voting scheme for the 2D Hough transform is applied to recover the supporting lines of the silhouettes edges (orange lines in Figure 2). The 2D image coordinates of the silhouette vertices are obtained intersecting pairs of adjacent supporting lines. In the case three faces of the box are visible, they are delimited by computing the internal vertex (intersection of the blue lines in Figure 2). The supporting line of an internal (blue) edge is computed as the line connecting its corresponding vanishing point and the silhouette vertex closest to such a vanishing point. The entire set of automatically recovered vertices and edges is illustrated in Figure 2, which also shows the box dimensions computed from the 3D coordinates of the silhouette vertices.

Before one can compute the 3D coordinates of a box vertex, it is necessary to eliminate the perspective ambiguity. This is solved using the distance in 3D between the projections of the two laser beams onto the box face. Such a distance is obtained using the 2D coordinates of the laser dots in the image, the known distance between the two laser beams and the normal vector of the box face that contains the dots [Fernandes et al. 2006]. The normal vector of a given box face is obtained from the vanishing line of that face [Hartley and Zisserman 2000]. Finally, the plane equation of the face that contains the laser dots is calculated using its normal vector and the 3D position of one of the laser dots (computed assuming a linear camera projection model [Fernandes et al. 2006]). Since all visible faces share some vertices with each other, the plane equations of the other faces can also be obtained, allowing the recovery of the 3D coordinates of the visible vertices.

We have carried out some experiments and computed some statistics using real boxes. An analytical derivation of the uncertainty propagated along the computation chain allows real-time estimation of the error in the computed measurements. The statistics and experimental validation have shown that the proposed approach is accurate and precise. The average relative error in the computed dimensions is 3.75% and is related to the geometry of the boxes, which somewhat differ from parallelepipeds due imperfections (*e.g.*, bent edges and corners, asymmetries, etc.). For a typical scene, our prototype can process video (640*x*480) and compute box dimensions at 29 fps on a 2.8 GHz PC. We are currently investigating ways to eliminate the need of a background with a known color.

References

FERNANDES, L. A. F., et al. 2006. A fast and accurate approach for computing the dimensions of boxes from single perspective images. *J. of the Brazilian Computer Society* 12, 1 (to appear).

HARTLEY, R. I., and ZISSERMAN, A. 2000. *Multiple View Geometry in Computer Vision*. Cambridge University Press.

*{laffernandes,oliveira}@inf.ufrgs.br