The Photopoint Method

Ben Ward (University of Adelaide)

The process has two stages. First is the scene reconstruction stage, where camera parameters are recovered for each image and a point cloud is generated. Basal area is then estimated for the scene from the images and recovered parameters.

Scene reconstruction:

The input data is a set of images covering 360° around three points, each 2.5m apart. At the centre of the points is cylinder with a calibration pattern.

The scene reconstruction process requires matching feature points (points in the images which can be easily distinguished) between frames.

These matches provide constraints which can be used to determine the relative position and orientation of the cameras which captured the images. With camera parameters determined, 3d points can be recovered from matched 2d points by triangulation. This reconstruction process uses standard computer vision techniques. For details of a reconstruction system using these techniques, see [1].

Selecting frames for point matching is aided by an initial estimate of the camera parameters. QR codes on the calibration object are detected, and the known locations of the point on the codes are used to estimate camera parameters for frames where the object is visible. For each of the three photo sets, matching points in overlapping frames are then detected and used to determine the relative orientation of the cameras in a set. This gives an estimate for parameters of all cameras. These parameters are used to determine cameras with overlapping views of the scene. Features are only matched between cameras with overlapping views, which reduces the chance of false matches.

Reconstruction is then performed for sets of cameras with matching points. This is done for sufficient sets to cover all of the surrounding scene. These sets are then merged to give the final reconstruction.

Known distances between points on the calibration object are used to determine scale for the scene.

This process results in a sparse point cloud, as points are only reconstructed for easily identified features. For a dense reconstruction of the scene, the PMVS software [2] can be applied to sets of images and camera parameters.

Basal Measurement

In this stage cylinders are fitted to trees detected in a set of images covering the surrounding scene. First, rectangular image regions which are likely to correspond to trunks are detected in the images.

For this detection, a Support Vector Machine (SVM) classifier is used. SVMs are a standard machine learning tool. From a set of labelled training vectors, the SVM learns a hyperplane which gives an optimal separation of vectors in different classes. This can then be used to predict the label for subsequent vectors. See [3] for more information. For image classification, the vector contains data extracted for an image region.

In this system, the vector consists of histograms of color and texture data extracted for the region. The SVM is trained from manually-labelled images covering a range of scene types.

In each image, regions are samples covering the likely range of positions, orientations, and scales for trunks. Cylinder fitting is attempted for regions classified by the SVM as trunk. A set of 3D cylinders matching the region boundary is generated along a ray from the camera centre through the centre of the region. Each cylinder is projected into images from the other sets where it is visible. The appearance of the cylinder is compared in a pair of images. If the appearance is consistent this indicates that the cylinder corresponds to an actual object in the scene. A set of patches on the cylinder are compared between images using Normalized Cross Correlation [4]. This gives a measure of consistency which is robust to variation in lighting and exposure. Cylinders are rejected if they have a high proportion of inconsistent patches. From the set of remaining cylinders, a non-overlapping subset is selected. Cylinder radius can then be used for basal area estimation.

References

[1] Marc Pollefeys et al. Visual Modeling with a Handheld Camera, International Journal of Computer Vision, 2004.

[2] http://www.di.ens.fr/pmvs/

[3] http://www.statsoft.com/textbook/support-vector-machines/

[4] http://en.wikipedia.org/wiki/Cross-correlation