Single View Based Measurement on Space Planes

Guang-Hui Wang, Zhan-Yi Hu, and Fu-Chao Wu

National Laboratory of Pattern Recognition, Institute of Automation, The Chinese Academy of Sciences
Beijing 100080, P.R. China
E-mail: ghwang@nlpr.ia.ac.cn
Received June 20, 2002; revised April 8, 2003.

Abstract The plane metrology using a single uncalibrated image is studied in the paper, and three novel approaches are proposed. The first approach, namely key-line-based method, is an improvement over the widely used key-point-based method, which uses line correspondences directly to compute homography between the world plane and its image so as to increase the computational accuracy. The second and third approaches are both based on a pair of vanishing points from two orthogonal sets of parallel lines in the space plane together with two unparallel referential distances, but the two methods deal with the problem in different ways. One is from the algebraic viewpoint which first maps the image points to an affine space via a transformation constructed from the vanishing points, and then computes the metric distance according to the relationship between the affine space and the Euclidean space, while the other is from the geometrical viewpoint based on the invariance of cross ratios. The second and third methods avoid the selection of control points and are widely applicable. In addition, a brief description on how to retrieve other geometrical entities on the space plane, such as distance from a point to a line, angle formed by two lines, etc., is also presented in the paper. Extensive experiments on simulated data as well as on real images show that the first and the second approaches are of better precision and stronger robustness than the key-point-based one and the third one, since these two approaches are fundamentally based on line information.

Keywords single view metrology, projective geometry, geometrical parameter retrieval, plane homography

1 Introduction

One of the main aims of Computer Vision is to take measurements of the environment and reconstruct its 3D model. Using vision to measure world distance has attracted a lot of attention and found wide applications in recent years[1-10], such as architectural and indoor measurement, reconstruction from paintings, forensic measurement and traffic accident investigation[11-5]. Traditional approach to measurement is to take all the distances manually by using metric tapes or rulers or by some special devices such as ultrasonic devices, laser range finder, etc. These approaches are time consuming, prone to errors and invasive. With computer vision based methods, what one needs to do is only to take several pictures, then all measurements can be done offline with higher accuracy, flexibility and efficiency. There are several potential advantages for this kind of approaches. First, it is user friendly. Once the images are acquired, users can take measurements desktoply and store them in a database. Second, the data acquisition process is rapid, simple and minimally invasive, since it only involves a camera to take pictures of the environment to be measured. Third, the acquired data are stored digitally in a disk ready for reuse at any time negating to go back to the original scene when new measurements are needed. Finally, the hardware involved is cheap and easy to use.

Generally speaking, the methods of computer vision based measurements in the literature may be broadly divided into two categories. The classical method is to reconstruct the metric structure of the scene from 2 or more images by stereo vision techniques[1,9-11]. If we can obtain the Euclidean reconstruction of a scene, then any geometrical information about the scene can be retrieved accordingly. However, this is a hard task due to the problem of seeking for correspondences from different views. Besides, the precision of reconstruction is closely related with that of correspondences and camera calibration.

The other one is to directly use a single uncalibrated image[1-8]. It is well known that only one image cannot provide enough information for a complete 3D reconstruction. However, some metrical quantities can be inferred directly from one image under the knowledge of some geometrical scene constraints such as planarity of points and
parallelism of lines and planes. In [1, 2], a key-
point-based approach to calculate the Euclidean
distance of two points on a world plane was pro-
posed. In this method, at least the coordinates of
four control points on the world plane and their
.corresponding image points should be known be-
forehand. In [1, 3], the authors described another
approach to compute 3D affine measurement from
a single perspective image. It is assumed that the
vanishing line of a reference plane in the scene as
well as a vanishing point in a reference direction
(not parallel to the plane) can be determined from
the image, then three canonical type measurements
(i.e., distances between any planes which are par-
allel to the reference plane, area and length ratio
on these planes and the camera’s position) can be
computed. In [4-7], some methods were also in-
vestigated for object reconstruction from measure-
ments in a single view in both computer vision com-
.munity and photogrammetric community. These
methods were based on the constraints of the ob-
ject to be reconstructed, such as edges, coplanarity,
parallelism, perpendicularity, etc.

Since distance measurement on a plane is of
great importance and has wide applications, in this
paper, we introduce three novel single-view-based
methods for distance measurement as well as a brief
description about the retrieval of other geometrical
parameters on a planar scene. In addition, a com-
parative study of the proposed methods, together
with the widely used key-point-based method, is
carried out. The paper is organized as follows. In
Section 2, some preliminaries on homography, van-
ishing point and cross ratio are introduced briefly.
Then the three novel approaches are elaborated in
Section 3. In Section 4, a brief description about
how to retrieve other geometrical parameters is
presented. In Section 5, all the methods for dis-
tance measurement are compared experimentally
with both simulated data and real images. Some
conclusions are given at the end of this paper.

2 Some Preliminaries

In order to facilitate our discussions in the subse-
quent sections, some preliminaries on homography,
vanishing point and cross ratio are presented here.
The key-point-based method for distance measure-
ment is also outlined. In this paper, the following
denotation is used: a 3D or 2D column vector is de-
noted by $x$, while a homogeneous (augmented by
adding $w$ or 1) vector is denoted by $\tilde{x} = [x^T, w]^T$,
or $\tilde{x} = [x^T, 1]^T$.

2.1 Plane to Plane Homography

Under the pinhole camera model, a 3D point $\tilde{x}$
in space is projected to an image point $\tilde{m}$ via a
3 x 4 projection matrix $P$ as:

$$s\tilde{m} = P\tilde{x} = [p_1, p_2, p_3, p_4]\tilde{x}$$

where, $\tilde{m}$ and $\tilde{x}$ are homogeneous coordinates in
the form of $\tilde{m} = (u, v, w)^T$, $\tilde{x} = (X, Y, Z, W)^T$,
and $s$ is a nonzero scalar. For 3D coplanar points,
without loss of generality, we assume $Z = 0$, then

$$s \begin{bmatrix} u \\ v \\ w \end{bmatrix} = P \begin{bmatrix} X \\ Y \\ 0 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ W \end{bmatrix}$$

Hence, mapping between corresponding points
on the plane and its image is

$$s\tilde{m} = H\tilde{x}$$

where, $H = [p_1, p_2, p_3]$ is called plane to plane ho-
mography. Usually, $H$ is a non-singular 3 x 3 ho-
mogeneous matrix (degeneracy occurs if and only
if camera center is on the reference plane) with 8
degrees of freedom because it can only be defined
meaningfully up to a scale factor. According to
(2), each image to world point correspondence can
give rise to two linear constraints on the 9 elements
of homography, thus, given $N$ ($N = 4$) coplanar
space points in general position (no three points are
collinear) and their correspondences in image, the
homography matrix can be uniquely determined.
If $N > 4$, the matrix is over determined. For non-
perfect data, $H$ can be estimated by a suitable
minimization scheme[1,6]. Once the homography
matrix between the world and image planes is de-
termined, an image point can be back projected to
a point on the world plane via $H^{-1}$, hence the dis-
tance between two points on the world plane can
then be simply computed from the Euclidean dis-
tance between their back-projected images. This
is the basic principle of the key-point-based plane
measurement[1,2]. It is clear that the accuracy of
this method depends greatly on that of the selec-
tion of key points and the detection of their corre-
sponding image points. In real applications, if we
directly use the point correspondences to compute
the homography, it may be subject to a loss of ac-
curacy due to the noise in extracted image points.