Czech Pattern Recognition Workshop 2000, Tomáš Svoboda (Ed.)
Peršlák, Czech Republic, February 2–4, 2000
Czech Pattern Recognition Society

Local Planar Model Verification in a Polynocular Image Set

Vít Zýka and Radim Šára

Czech Technical University
Center for Machine Perception
Karlovo nám. 13, CZ 121 35
zyka@cmp.felk.cvut.cz

Abstract *This work deals with verification of local planar 3D geometric model recovered from noisy data. Verification is understood as making decision about consistency of the model with a set of images. To evaluate the consistency, a statistical measure is chosen and a novel procedure for testing statistical dependence among a set of random variables is proposed.*

1 Introduction

In various computer vision tasks it is possible to encounter the problem of verification of a 3D model. The purpose of verification is to assess the consistency between the model and a set of images that are supposed to be views of the scene that is modeled. For example, Baillard et al. [2] describe the reconstruction of 3D urban site models from aerial images. The local candidate models for each building roof are recovered from edge-based stereo. The roof regions are then verified in the input images based on the mutual consistency of image values.

A similar problem has been reported in [4], where local planar surface models recovered in a bottom-up reconstruction process are verified for their pose. Models that are not consistent with a set of images are rejected.

These practical problems prompted us to study the verification itself in greater detail. The general problem will be restricted to local planar models. Then, the verification decision is whether the model (a reconstruction hypothesis) is consistent or not with a set of images taken from general viewpoints. Here we present it as a statistical decision procedure based on the model pose parameters, the projection parameters, and a set of images. The decision should be insensitive to a change in the unknown parameters of image formation (namely surface non-Lambertianity), on the textural properties of the observed object (namely the statistical texture parameters), and on image sensor properties (noise, discretization).

Good statistical decision allows to eliminate local models that are unlikely to correspond to real objects, which significantly improves the performance of subsequent steps in global 3D model reconstruction.

This paper is organized as follows. In the next section we give a general definition of verification, and narrow it down to a useful concept. In Section 3 we deal with suitable image consistency measures. In Section 4 we focus on verification implemented as a statistical decision procedure. A novel approach to testing statistical dependence is developed. An experiment with the verification of a set of local surface models is demonstrated on a test object in Section 5. Section 6 then concludes the paper.

2 Model Verification

The intuition behind model verification is demonstrated in Figure 1. To be precise, we give the following definition.

Definition *A local planar model is verified if its projections to all possible images of the corresponding object are equivalent up to the class of transformations given by the physics of image formation.*

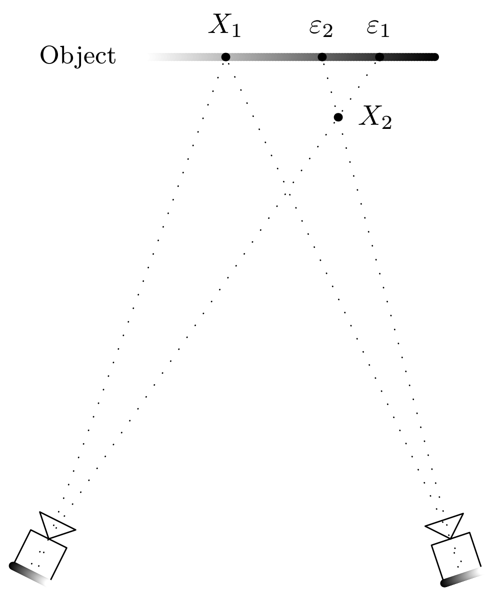

Figure 1: Point model verification. Model X_1 is verified since the image neighborhoods of its projection are mutually compatible. Model X_2 is inconsistent since the left camera perceives it projected on the background ε_1 and the right camera on ε_2, so their images generally differ.

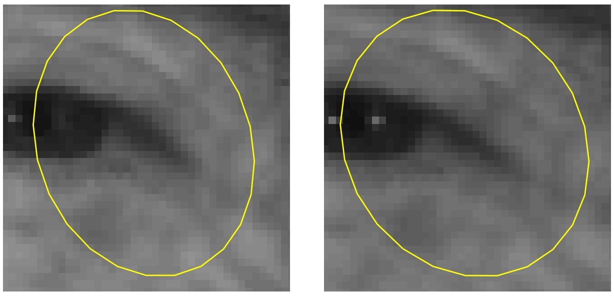

Figure 2: Circular disc model consistent with its images.

Let us suppose as an example that we have a disk-like 3-D model and a pair of images to which the model projects, as in Figure 2, where the projected model boundary is shown. The image formation introduces geometric distortion (perspective foreshortening) and radiometric distortion due to the surface non-Lambertianity. The former is apparent as the circle-to-ellipse distortion of the model boundary and the latter is visible as a highlight on the cornea, for instance. The class of transformations further includes the discrete image re-sampling and quantization and the optical properties of each of the sensors.

Since the verification is defined in terms of projections of the model to images, we need a measure of mutual consistency among the images (given the model and projection parameters defining the pointwise image correspondences).

For practical considerations, the views used for verification must be reduced to a small number. The smallest possible number of views depends on the directional texture properties, but it is always greater or equal to two, since the verification is based on the relation among the images and not on the model-to-image relation. The likelihood of a successful rejection increases with the number of views involved and reaches certainty for their infinite number. Note that any model is consistent with any number of images if the observed scene is completely texture-free and without any shading. This is not considered a failure. The stronger the texture the easier is the discrimination from a small number of images.

Let the scene be static or all images be captured at the same instant. We assume that in practice the verification is done based on the following:

- on a local planar surface model instance given by its pose and size,
- on a finite number (but at least two) intensity images from different viewpoints but with overlapping visual fields, and
- on known model projection parameters (binding the model and image coordinate systems).

We assume the verification decision is implemented as a statistical test on a suitable consistency measure. Other possibilities are not discussed in this paper. The consistency measure and the statistical decision is dealt with in the following two sections.

3 The Consistency Measure

The measure of local image consistency for a given a model is introduced in this section. We called it *Polynocular Local Image Consistency (PLIC)* [5]. Generally it can be defined as a mapping into real numbers:

$$\mathrm{PLIC} : (\mathrm{model}, \mathrm{images}, \mathrm{projection}) \to \mathbb{R}\,.$$

The higher the PLIC value the better the mutual consistency of the images of the model.

We experimented with three variants of the PLIC measure. They differ in the degree of invariance to the transformation that acts on the images. The evaluation of PLIC is done in the following three steps. First, the set of n spatial points is randomly selected on the model surface. Second, the points are projected to m images. Third, $m \times n$ intensity values L_{ij} are collected at the projected points, where $L_{i\cdot}$ denotes the set of values in image i. Let $\mathcal{P}$ be the set of all pairs among the m cameras. The basic PLIC variants are the following:

Sum of Square Differences assumes that only identity transformation of intensity values acts among the images. All pairwise differences are summed.

$$\mathrm{PLIC}_1 = \frac{-1}{n|\mathcal{P}|} \sum_{(r,s)\in\mathcal{P}} \sum_{j=1}^{n} \left(L_{rj} - L_{sj}\right)^2 . \quad (1)$$

This may be too restrictive in certain applications.

Standard Correlation Coefficient [1] is invariant to linear transformation acting on image values. It is computed as a mean of all pairwise correlation coefficients and has a range of $[-1, 1]$:

$$\mathrm{PLIC}_2 = \frac{1}{|\mathcal{P}|} \sum_{(r,s)\in\mathcal{P}} \frac{\mathrm{cov}(L_{r\cdot}, L_{s\cdot})}{\sqrt{\mathrm{cov}(L_{r\cdot})\,\mathrm{cov}(L_{s\cdot})}} . \quad (2)$$

It allows to suppress the influence of small deviations from surface non-Lambertianity.

Rank Correlation [3] is the most general since it enables any monotonic transformation among the images. We use Kendall's Rank Concordance [3]. Let R_{ij} be the rank of point i among the values collected in image j, the R_{ij} ranges from 1 to the total number of measurements n. Let $R_i = \sum_{j=1}^{m} R_{ij}$ be the sum of ranks of the same point over all images. The mean value of R_i is $\bar{R} = \frac{1}{2}m(n+1)$. The concordance is *not* calculated pairwise and its range is $[0, 1]$:

$$\mathrm{PLIC}_3 = \frac{S}{S_{\max}} , \quad (3)$$

where S is the sum over all points,

$$S = \sum_{i=1}^{n} \left(R_i - \bar{R}\right)^2 ,$$

and $S_{\max}$ is the maximum possible value for S,

$$S_{\max} = \sum_{i=1}^{n} \left(i\,m - \bar{R}\right)^2 .$$

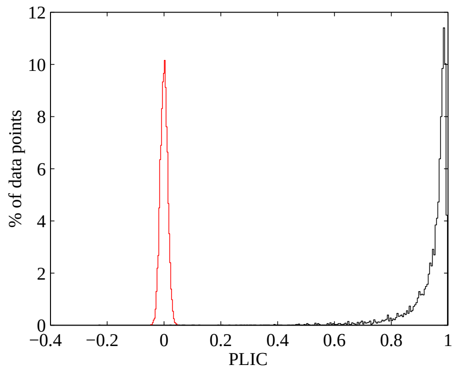

Figure 3: The poor utility of statistical independence test for testing statistical dependence.

The invariance to monotonic image transformations is paid for by the loss of discriminability of PLIC_3. More general transformations of image values are possible, but their form must be explicitly known to construct the congruence measure. The choice is then application-dependent and will not be discussed here. As will be seen in the next section, however, the actual verification is independent on the choice of the measure.

4 The Statistical Decision Procedure

Models are verified by re-projecting n points randomly generated from the model to a set of m camera retinas and computing the mutual congruence of their (cubicly interpolated) images as describe above. The image congruence is computed in k trials and a cumulative histogram of the values is obtained.

The model is accepted based on a statistical test at a given confidence level using the computed image congruence value histogram: If the cumulative histogram value K_α corresponding to the given confidence level α exceeds the *prior* image congruence value K_p, the model is accepted, otherwise it is rejected:

$$K_\alpha > K_p \Rightarrow \text{model is accepted.} \qquad (4)$$

This approach avoids the need for a prior statistical distribution required for the verification decision.

The inadequacy of the statistical independence test is demonstrated in Figure 3. The black curve represents a histogram of PLIC values for a collection of local models reconstructed from images of a scene. Note the mode is very close to the PLIC value of 1. The red curve peaking at $\mathrm{PLIC} = 0$ is a histogram of the PLIC for two statistically independent variables. To base the verification decision on the theoretical distribution under the hypothesis of independence (the red curve in Figure 3), a very high percentile must be chosen as the confidence level (e.g. 0.99999). Clearly, this is not a very intuitive choice, since the confidence level losses its sense. This is the reason we chose the method describe above. Although it requires two more parameters, they have a clear meaning. The K_p is a design parameter whose value must be chosen based on factors like the image signal-to-noise ratio. The k is to be chosen too, but its value is not critical.

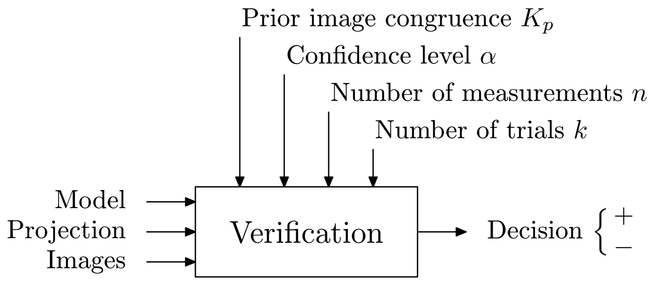

Figure 4: Verification procedure.

The verification decision procedure is summarized by the diagram in Figure 4. The input data comes from the left, the four procedure parameters comes from the top.

5 Experiment

We tested our verification algorithm on the local models obtained from a stereo-reconstruction process. Four synchronized and calibrated cameras were used to capture a ceramic teapot in the distance of approximately 60 cm. The scene was illuminated by random texture pattern from an calibrated texture projector. The parameters of local surface fitting procedure were set in a way that many wrong surface elements were reconstructed. After verification, nearly none of them survived, see Figure 5. In our current setup we use the following values: $K_p \approx 0.6$, $\alpha = 0.9$, $n = 40$, $k = 100$.

6 Conclusion

We have studied the verification of local planar surface model in a small set of intensity images. It is based on mutual image consistency. Several image consistency measures were overviewed. Each of them is invariant to a different class of radiometric image distortions. An algorithm for testing statistical dependence among a set of random variables was proposed. The algorithm is independent on the choice of the consistency measure. The verification method was tested on real data to demonstrate its utility for 3D model reconstruction from stereo.

Acknowledgements

This research was supported by the Czech Ministry of Education, grant VS96049, by the Grant Agency of the Czech Republic, grants 102/97/0480, 102/97/0855, and 201/97/0437, by grant OCAMS No. 4/11/AIP CR, and by the Research Programme J04/98:210000012 Transdisciplinary Research in Biomedical Engineering.

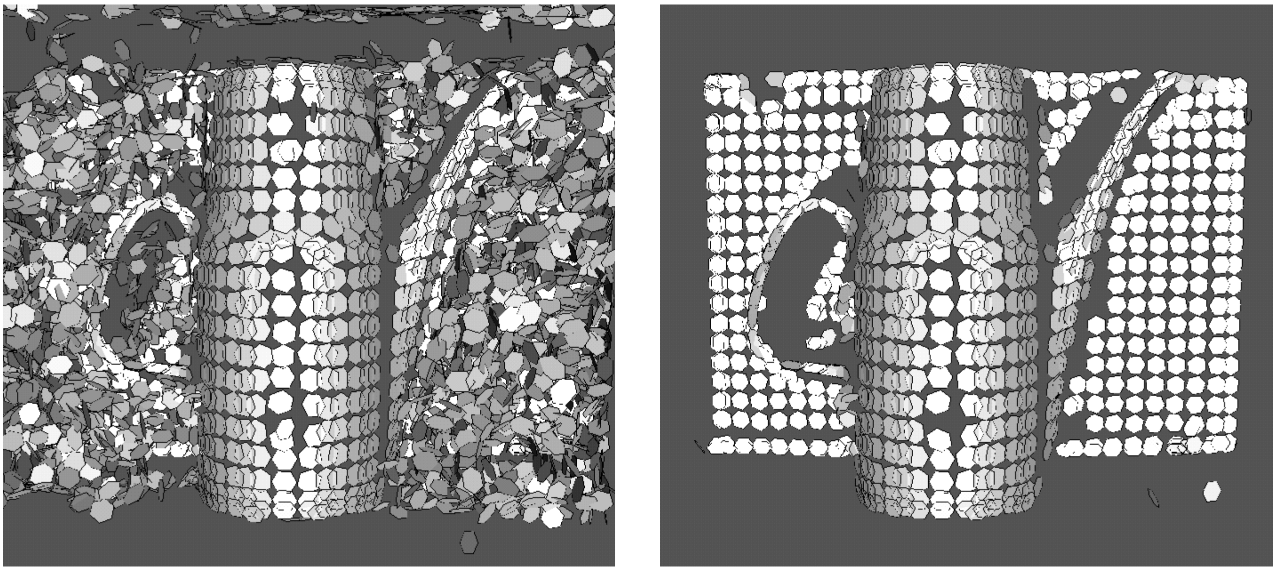

Figure 5: Example of verification. The reconstructed model of a teapot with a large number of misplaced elements is shown left. Right is the same model after verification. (Click on for VRML models.)

References

[1] J. Anděl. *Statistické metody*. Matfyzpress, Praha, 1998.

[2] C. Baillard, C. Schmid, A. Zisserman, and A. Fitzgibbon. Automatic line matching and 3D reconstruction of buildings from multiple views. In Proc. *ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, IAPRS Vol.32, Part 3-2W5*, September 1999.

[3] M. Kendall and J. D. Gibbons. *Rank Correlation Methods*. Edward Arnold, 1990.

[4] R. Šára. Accurate natural surface reconstruction from polynocular stereo. In Proc. *NATO Advanced Research Workshop: Confluence of Computer Vision and Computer Graphics*, NATO ASI Series, The Netherlands, August 1999. IEEE Slovenia Section, Kluwer Academic Publishers.

[5] V. Zýka. Recovering accurate geometric surface model from passive stereo vision. Technical Report K335-98-153, ČVUT FEL, Dept. of Control Engineering, Praha, Czech Republic, Jan. 1998. Thesis proposal, ftp://cmp.felk.cvut.cz/pub/cmp/articles/zyka/tr98-153.ps.gz.