On Robust Regression in Photogrammetric Point Clouds

Konrad Schindler and Horst Bischof

Institute of Computer Graphics and Vision
Graz University of Technology, Austria
{schindl,bischof}@icg.tu-graz.ac.at

Abstract. Many applications in computer vision require robust linear regression on photogrammetrically reconstructed point clouds. Due to the modeling process from perspective images the uncertainty of an object point depends heavily on its location in object space w.r.t. the cameras. Standard algorithms for robust regression are based on distance measures from the regression surface to the points, but these distances are biased by varying uncertainties. In this paper a description of the local object point precision is given and the Mahalanobis distance to a plane is derived to allow unbiased regression. Illustrative examples are presented to demonstrate the effect of the statistically motivated distance measure.

1 Introduction

Several photogrammetric applications require robust surface fitting in 3D point clouds obtained by dense image matching, as for example architectural reconstruction [1] or reverse engineering [2].

Robust fitting algorithms are mostly based on sampling strategies: Hypotheses are generated and their support is measured in the point cloud. Examples for this strategy are the least-median-of-squares (LMS) estimator [3] and different variants of the RANSAC principle [4]. This class of algorithms requires to compute the *support* of each point for a given hypothesis, i.e. the probability that the point is explained by the hypothesis. The simplest form is to assign the probability 1 to all points within a certain threshold distance and 0 to all other points (*voting*). Another popular strategy is to assign each point a probability which is inversely proportional to its distance to the plane (*linear weighting*). The statistically correct procedure is to use the value of the probability density function (*pdf*) at the given distance. In any case computing the support requires a suitable distance measure[1].

We argue that this distance measure must take into account the individual uncertainties of the object points, because they are highly inhomogeneous in

[1] It should be mentioned that the same problem exists for robust regression methods which operate in parameter space, namely clustering methods and the Hough transform, where we need a threshold for the clustering radius in the parameter space.

B. Michaelis and G. Krell (Eds.): DAGM 2003, LNCS 2781, pp. 172–178, 2003.

photogrammetric point clouds (see Figure 1). In this paper we treat the basic case of (orthogonal) linear regression. The task to be solved is *optimally fitting an unknown number of planes to a cloud of 3D points with known, individually different variances and covariances.* However the presented ideas are also valid for regression with other parametric surfaces (in fact the distance from a point to a higher-order surface is the distance from the point to the tangent plane through the closest surface point).

NB: Least-squares fitting as a postprocessing step may in many cases alleviate, but not solve the problem, because it requires a correct partitioning of the point cloud into inliers and outliers, and this partitioning is based on the support computed during robust regression. It has been suggested to use a generous threshold for the inliers and compute the fit with an M-estimator [5]. However it seems an awkward strategy to approximately detect inliers with a robust method, re-include some outliers and fit with another robust method, instead of cleanly dividing the task into two steps, one for robust detection of the correct point set and one for optimal fitting.

The paper is organized as follows: in section 2 we briefly review the uncertainty propagation of the photogrammetric reconstruction process. In section 3 we derive the Mahalanobis distance to a plane, and in section 4 we illustrate the difference between the use of the Mahalanobis distance and the geometric distance.

2 Uncertainty of Photogrammetric Points

This section is a brief recapitulation of the uncertainty propagation in the photogrammetric reconstruction process. For lack of space we refer to photogrammetric textbooks, e.g. [6] for details. A measured image point is described by its coordinates and the covariance matrix

$$\mathbf{x} = [x, y]^\mathsf{T} \quad , \quad \mathtt{S}_{\mathbf{xx}} = \begin{bmatrix} s_{xx} & s_{xy} \\ s_{xy} & s_{yy} \end{bmatrix} \tag{1}$$

Different image points are assumed to be statistically independent. A 3D point $\mathbf{u}$ is constructed from N image points by intersecting their viewing rays, i.e. solving the $2 \times N$ collinearity equations for $\mathbf{u}$. This gives an overdetermined (non-linear) equation system which is linearized with the Jacobian $\mathtt{A}$ of the collinearity equations

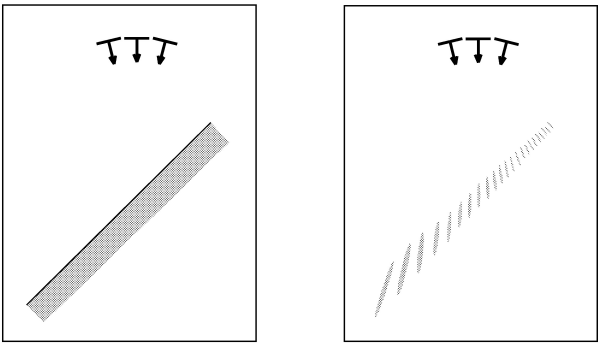
$$\mathtt{A}\mathbf{u} = \mathbf{b} \quad , \quad \mathbf{b} = \begin{bmatrix} \mathbf{x}_1 \ldots \mathbf{x}_n \end{bmatrix}^\mathsf{T} \tag{2}$$

and solved through iterative least-squares adjustment[2]. Error propagation gives a linear approximation $\mathtt{S}_{\mathbf{uu}}$ for the covariance matrix of the estimated 3D point coordinates.

[2] An elegant formulation in homogeneous coordinates has recently been published by Förstner [7].

$$\mathbf{u} = \mathtt{S}_{\mathbf{uu}}\mathtt{A}^{\mathsf{T}}\mathtt{S}_{\mathbf{bb}}^{-1}\mathbf{b} \quad , \quad \mathtt{S}_{\mathbf{uu}} = (\mathtt{A}^{\mathsf{T}}\mathtt{S}_{\mathbf{bb}}^{-1}\mathtt{A})^{-1} \tag{3}$$

From equation 3 we can see that the uncertainty of point $\mathbf{u}$, through the collinearity equations, depends on the camera positions. The uncertainty is low, if the point is close to the cameras and if the intersection angles between different rays are close to 90 degrees. Typically dense point clouds are generated from image sequences with short baselines in order to enable automatic matching with area-based correlation. In such a recording setup the depth has the highest uncertainty. Moreover, if the camera path is (nearly) linear, e.g. in turntable sequences or when recording buildings from ground level, the depth is highly correlated with the direction of camera motion, so that covariances *cannot* be neglected. Figure 1 shows a prototypical setup, which we will use as an example.



(a) (b)

Fig. 1. Typical recording setup for automatic reconstruction. (a) Top view of cameras and recorded object plane. (b) Reconstructed points, given by error ellipsoids. The gray ellipses are the projections of the error ellipsoids onto the x, z-plane, computed for a image measurement accuracy of 1:1000. The error ellipsoids are scaled by a factor 10 to make the figure easier to read.

3 Mahalanobis Distance to a Plane

The point coordinates and covariance matrix describe the point position in space with a 3-dimensional probability distribution. The point is thus given by

$$\mathbf{u} = [u, v, w]^{\mathsf{T}} \quad , \quad \mathtt{S}_{\mathbf{uu}} = \begin{bmatrix} s_{uu} & s_{uv} & s_{uw} \\ s_{uv} & s_{vv} & s_{vw} \\ s_{uw} & s_{vw} & s_{ww} \end{bmatrix} \tag{4}$$

Let the plane hypothesis in homogeneous coordinates be denoted by $\mathbf{p} = (p_1, p_2, p_3, p_4)$, where $h_{\mathbf{p}} = \sqrt{p_1^2 + p_2^2 + p_3^2}$ is the homogeneous scale. Since the goal of the distance measure is to determine the probability that a point belongs to the plane, the distance measure must take into account the uncertainty of each individual point. This is achieved by the Mahalanobis distance [8] – for a more detailed treatment of its role in optimal geometric fitting see [5]. The Mahalanobis distance from an uncertain point $\mathbf{x}$ to a given point $\overline{\mathbf{x}}$ (in Euclidian coordinates) is defined as

$$d_M^2 = (\mathbf{x} - \overline{\mathbf{x}})^{\mathsf{T}} \mathtt{S}_{\mathbf{xx}}^{-1} (\mathbf{x} - \overline{\mathbf{x}}) \tag{5}$$

To get the Mahalanobis distance to a plane, we have to apply a whitening transform and normalization to the covariance matrix $\mathtt{S}_{\mathbf{uu}}$ and the plane (geometrically this means warping the error ellipsoid to a unit sphere). The centering is simply a shift from $\mathbf{u}$ to $(0,0,0)^{\mathsf{T}}$. Since the covariance matrix $\mathtt{S}_{\mathbf{uu}}$ is symmetric, its singular value decomposition (SVD) directly gives the rotation $\mathtt{R}$ and the scale $\mathtt{V}$.

$$\mathtt{S}_{\mathbf{uu}} = \mathtt{R}^{\mathsf{T}} \mathtt{V} \mathtt{R} \quad , \quad \mathtt{V} = diag(a^2, b^2, c^2) \tag{6}$$

$\mathtt{R}$ aligns the coordinate system with the ellipsoid axes, while the elements a, b, c of $\mathtt{V}$ compensate for the non-uniform scale along different axes. We can now write the transformation as a product of homogeneous 4×4 matrices

$$\mathtt{R}_h = \begin{bmatrix} \mathtt{R} & \mathbf{0} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix} \quad , \quad \mathtt{T}_h = \begin{bmatrix} \mathtt{I} & \mathbf{0} \\ -\mathbf{u}^{\mathsf{T}} & 1 \end{bmatrix} \quad , \quad \mathtt{A}_h = diag(a, b, c, 1) \tag{7}$$

and transform the plane $\mathbf{p}$ to a new plane $\mathbf{q} = \mathtt{A}_h \mathtt{R}_h \mathtt{T}_h \mathbf{p}$. The Mahalanobis distance is the distance from the transformed plane $\mathbf{q}$ to the center of the unit sphere, which lies in the origin of the new coordinate system (see Figure 2).

$$d_M(\mathbf{u}) = \frac{q_4}{\sqrt{q_1^2 + q_2^2 + q_3^2}} \tag{8}$$

Lacking a more qualified *pdf*, we recur to the common approximation of normally distributed image measurement errors (although it is theoretically questionable). In first-order approximation the distances from the points to the plane, being functions of the image point coordinates, also follow a normal distribution. A points support for the plane is thus determined by the percentile rank of $d_M(\mathbf{u})$ in the normalized Gaussian probability density function:

$$S(\mathbf{u}) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-d_M(\mathbf{u})}^{+d_M(\mathbf{u})} e^{-\frac{t^2}{2}} dt = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{-d_M(\mathbf{u})} e^{-\frac{t^2}{2}} dt \tag{9}$$

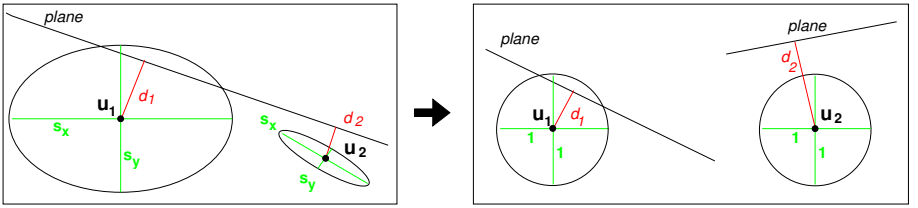


Fig. 2. Comparing distances in the presence of varying point precision. The probability of belonging to the plane is higher for point $\mathbf{u}_1$ than for point $\mathbf{u}_2$ although $d_1 > d_2$. After transforming the error ellipsoids to unit spheres the distance measures correctly reflect the probabilities.

The total support for a plane in a point set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is the sum of the support values $S(\mathbf{u}_i)$.

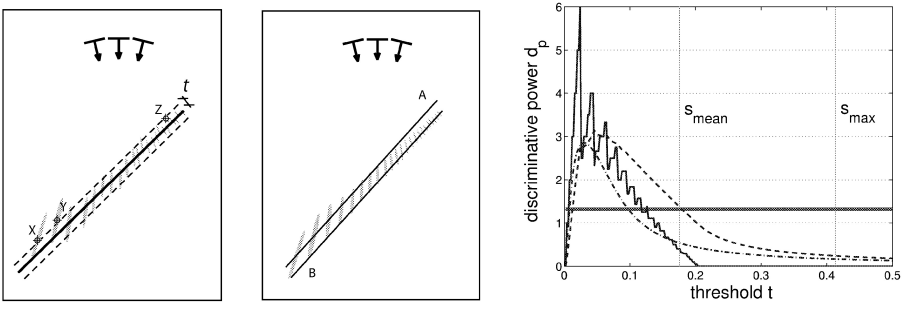
$$S_{\mathbf{p}} = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} \int_{-\infty}^{-d_M(\mathbf{u}_i)} e^{-\frac{t^2}{2}} dt \tag{10}$$

4 Examples

In this section we give two examples of how regression with the Mahalanobis distances $d_M(\mathbf{u})$ differs from regression with the plain geometric distances $d(\mathbf{u})$. We will use the synthetic data set shown in Figure 1. Sampling-based regression algorithms instantiate many planes and test their support in the point set. The instantiation is usually done from a minimal set of 3 points. Note however that this is not relevant for the statistical properties. The planes can be arbitrarily derived and are error-free hypotheses.

The first example in Figure 3(a) illustrates the task of separating inliers from outliers. The continuous line marks the plane hypothesis we want to evaluate. If we use the geometric distance, all points within the threshold t are considered inliers, the rest are outliers. The two points marked X and Y are classified as outliers, while the point marked Z is classified as inlier. Statistically this is not correct – the probability of being incident to the plane is higher for X and Y than for Z. Thresholding the Mahalanobis distance is equivalent to a χ^2-test and correctly divides the point set into inliers and outliers according to the probabiliy of being incident to the plane.

The second example illustrates the problem of choosing the correct threshold. The two planes depicted in Figure 3(b) by continuous lines are two possible plane hypothesis. Plane A has been randomly instantiated from the leftmost three points, while plane B has been instantiated from the rightmost 3 points. One can clearly see that plane B is a more probable estimate and should be preferred, as it explains more points within their uncertainty. The discriminative power of the sampling procedure, i.e. the ability to discriminate the better solution from



(a) (b) (c)

Fig. 3. Geometric distance vs. Mahalanobis distance. See text for explanation.

the worse one when comparing them, depends on the thresholds. It is a critical task to choose a threshold, which selects a good estimate without discarding too many inliers. Figure 3(c) shows the discriminative power as a function of the threshold for different weighting functions. The evaluated functions are

- voting (continuous line): $S(\mathbf{u}) = \{\ 1 \ldots \text{if } d(\mathbf{u}) \leq t\ ,\ 0 \ldots \text{else}\ \}$
- linear weighting (dashed): $S(\mathbf{u}) = \max(\ (1 - d(\mathbf{u})/t)\ ,\ 0\ )$
- Gaussian weighting (dotted): $S(\mathbf{u}) = 2\ cdf_{\text{Gauss},\mu=0,\sigma=t}(-d(\mathbf{u}))$

The discriminative power is defined as $p_d = (S_B/S_A - 1)$, values close to 0 mean that no reliable discrimination is possible. The horizontal line denotes the constant discriminative power of the Mahalanobis distance, which does not need a threshold. The two vertical lines indicate the mean standard deviation s_{mean} estimated from $S_{\mathbf{uu}}$ (the minimum threshold to make sure that more than half of the inliers are detected) and the maximum standard deviation s_{max} (the threshold we would have to choose to include all inliers). One can clearly see that a threshold with a reasonable discriminative power $d_p > 1$ leads to an incomplete inlier set and thus to a biased fit.

5 Concluding Remarks

We have investigated the influence of uncertainties on sampling-based regression methods in photogrammetric point clouds and derived the Mahalanobis distance from a point to a plane in order to correctly take the uncertainties of individual points into account during regression. Two prototypical examples have been shown to demonstrate the implications of the statistical nature of the point cloud.

We have assumed Gaussian noise of the image measurements. This is a common assumption, however there is no theoretical foundation for it, even less, if the points are derived automatically through a matching procedure. Furthermore it

would be desirable to investigate the influence of systematic errors, which are introduced by the smoothness and ordering constraints [9], [10] of dense matching algorithms.

Acknowledgments. This work has been supported by the European Commission under contract No. IST-1999-20273.

References

1. Werner, T., Zisserman, A.: New techniques for automated architecture reconstruction from photographs. In: Proc. 7th ECCV, Copenhagen. (2002)
2. Varady, T., Martin, R., Cox, J.: Reverse engineering of geometric models - an introduction. Computer Aided Design **29** (1997) 255–268
3. Rousseeuw, P., Leroy, A.: Robust Regression and Outlier Detection. John Wiley and Sons (1987)
4. Fischler, M., Bolles, R.: RANSAC random sampling concensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of ACM **26** (1981) 381–395
5. Triggs, B.: A new approach to geometric fitting. Available from http://www.inrialpes.fr/movi/people/Triggs (1998)
6. Kraus, K.: Photogrammetrie, Band 1. Dümmler Verlag (1994)
7. Förstner, W.: Algebraic projective geometry and direct optimal estimation of geometric entities. In: Proc. 25th AAPR Workshop, Berchtesgaden. (2001) 67–86
8. Duda, R.O., Hart, P., Stork, D.: Pattern Classification. John Wiley and Sons (2001)
9. Ohta, Y., Kanade, T.: Stereo by intra- and inter-scanline search. Pattern Analsis and Machine Intelligence **7** (1985) 139–154
10. Yuille, A., Poggio, T.: A generalized ordering constraint for stereo correspondence. MIT, A.I. Memo 777 (1984)