

Absolute Orientation from Uncertain Point Data: A Unified Approach

Yaacov Hel-Or * Michael Werman

Institute of Computer Science
The Hebrew University of Jerusalem
91904 Jerusalem, ISRAEL

published in CVPR 92

Abstract

A general and flexible method for fusing and integrating different 2D and 3D measurements for pose estimation is proposed. The 2D measured data is viewed as 3D data with infinite uncertainty in a particular direction. The method is implemented using Kalman filtering, it is robust and easily parallelizable. (Keywords: sensor fusion, Kalman filter, parallel algorithms, pose estimation, model based)

1 Introduction

In model-based pose determination the position of a known object is determined from different types of surface measurements (for example: [3, 7, 9, 1]). In this paper we deal with model and measurements consisting of features points (such as maximum curvature of a curve, segment endpoints and corners) where the measured data is non exact. The aim is to determine the correct transformation of model points to measured points. This subject is known as *absolute orientation* in photogrammetry and is classified into two major categories according to the type of measurements:

1. 3D to 3D correspondence: both model and measurements supply information about the 3D location of features (measurements as from range data, stereo, etc.).
2. 2D to 3D correspondence: model is 3D while the available measurements supply projected information in 2D. The projection can be perspective or orthographic.

Methods to compute the absolute orientation have been presented, most of which exploit least-square techniques in either closed [2, 1, 9] or iterative form

[2, 7, 3]. However, each method can be applied to only one of the categories described above.

In this paper we suggest a uniform framework to compute the absolute orientation, where the measured data can be both 2D and 3D. Unification of the different types of measurements is done by associating an uncertainty matrix with each measured feature. Uncertainty depends both on the measurement noise and on the type of measurement: when a measurement is a projection into 2D, we regard it as a measurement in 3D with infinite uncertainty in the direction of the projection. This representation unifies the two categories of the absolute orientation problem into a single problem that varies only in the uncertainty values associated with the measurements. With this paradigm we obtain a uniform mathematical formulation of the problem and can fuse different kinds of measurements to obtain a better solution. The algorithm we describe has additional advantages of simple parallelization, supplying a certainty measure of the estimate and enabling an efficient matching strategy.

2 General Overview of the Problem

A *model* M of a 3D object is represented by a set of points:

$$M = \{\mathbf{u}_i\}_{i=1\dots n}$$

where \mathbf{u}_i is a 3 dimensional object-centered vector associated with the i^{th} point.

A *measurement* of a 3D object is represented by M' which, similar to the model representation, is a collection:

$$M' = \{(\hat{\mathbf{u}}'_j, \Lambda_j)\}_{j=1\dots m} \quad , \quad m \leq n \quad .$$

$\hat{\mathbf{u}}'_j$ - is a synthesized noise-contaminated measure of the real location-vector \mathbf{u}'_j associated with the j^{th} measured point. $\hat{\mathbf{u}}'_j$ is a 3 dimensional vector, even

*Y.H. was supported by the Leibniz Center for Research in Computer Science. This work was supported by a grant from AFIRST through the Israeli NCRD.

though the measurements are not necessarily 3D, and is represented in a viewer-centered frame of reference. Λ_j - is the covariance matrix depicting the uncertainty in the sensed vector $\hat{\mathbf{u}}'_j$. The dimensionality of the measurements is encoded in the covariance matrix, therefore, the values of the matrix entries depend on the measurement noise and the projection (if this measure is projected data).

A *matching* between the model M and the measurement M' is a collection of pairs of the form

$$\text{matching} = \{\mathbf{u}_k, (\hat{\mathbf{u}}'_k, \Lambda_k)\}_{k=1\dots m}$$

which represents the correspondence between the model points to the measured points.

The problem:

Given a model M , a measurement M' and a matching as above, estimate a transformation T which optimally maps the points \mathbf{u}_i of the model onto the corresponding measured points $(\hat{\mathbf{u}}'_i, \Lambda_i)$. The estimated transformation T describes the position of the measured object M' in the 3D scene.

The method described below fuses the information from all the measured points and estimates the transformation T by incremental refinement using Kalman-filter tools. At each step a new measured point is accepted and a new updated solution is produced.

3 Uncertainty of The Primitives

As previously noted, a major paradigm in the approach is the necessity of evaluating the uncertainty matrix of each and every synthesized point feature. The uncertainty is derived from several factors:

- Uncertainty due to measurement-noise (e.g. digitization, blurring and chromatic aberrations).
- Uncertainty dependent upon the feature detection process. For example, a detected end-point of a line segment will have a low positional uncertainty in the direction perpendicular to the line segment and a high uncertainty in its direction.
- Uncertainty due to the lack of information caused by projections.

In this paper, we will not deal with the modeling of the measurement noise but we present the unique representation of the measured data. We separate our class of measurements into two categories:

3D measured data:

The simplest case is that of a point $q \in M'$ presented by the pair:

$$q = [\hat{\mathbf{u}}', \Lambda]$$

where $\hat{\mathbf{u}}' = (\hat{x}', \hat{y}', \hat{z}')$ is the measured location-vector and Λ is its uncertainty.

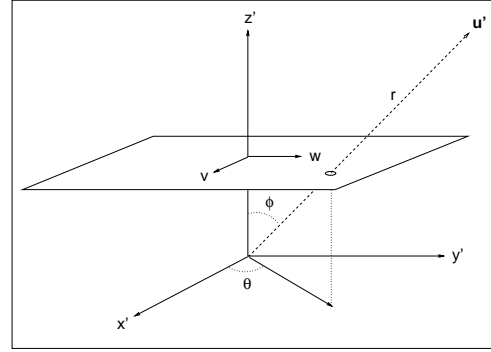


Figure 1: A perspective projection into an image plane.

projected data:

When the measurements are obtained using a projection, we describe the measured data as a measurement in 3D where the uncertainty in the direction of projection is infinite. Assume the measurements are performed on the image plane using the coordinate system (v, w) :

$$\text{proj}(q) = [(\hat{v}, \hat{w}), \Sigma_{vw}] \quad .$$

Σ is a 2×2 covariance matrix describing the uncertainty of the measurement (\hat{v}, \hat{w}) .

In the case where the projection is along the z-axis (orthographic) we represent this data as:

$$q = [(\hat{v}, \hat{w}, \hat{z}'), \begin{pmatrix} \Sigma_{vw} & 0 \\ 0 & \infty \end{pmatrix}]$$

where \hat{z}' is any value or an estimate of the z' coordinate.

In the case of perspective projection, modeling of the uncertainty is a little more complicated. Assume the origin of the viewer-centered frame of reference is at the focal point as shown in Figure 1 and the focal length is equal to one. We aim to transform the measurement given in the image-plane coordinate system into a representation in the Cartesian system $\mathbf{u}' = (x', y', z')$.

Considering the spherical coordinate system (r, ϕ, θ) (Figure 1). The vector (\hat{v}, \hat{w}) determines the angular coordinates (ϕ, θ) but leaves the value of r unknown:

$$\hat{\phi} = \arctan(\sqrt{\hat{v}^2 + \hat{w}^2}) \quad ; \quad \hat{\theta} = \arccos\left(\frac{\hat{v}}{\sqrt{\hat{v}^2 + \hat{w}^2}}\right) \quad .$$

Additionally, the uncertainty of (\hat{v}, \hat{w}) is translated into a covariance matrix in the (ϕ, θ) system as follows:

$$\Lambda_{\phi\theta} = \begin{pmatrix} \frac{\partial(\phi, \theta)}{\partial(v, w)} \end{pmatrix} \Sigma_{vw} \begin{pmatrix} \frac{\partial(\phi, \theta)}{\partial(v, w)} \end{pmatrix}^t$$

where $\frac{\partial(\phi, \theta)}{\partial(v, w)}$ is the Jacobian of the transform from (v, w) to (ϕ, θ) , and the derivative is taken at point (\hat{v}, \hat{w}) . The Jacobian matrix is:

$$\frac{\partial(\phi, \theta)}{\partial(v, w)} = \begin{pmatrix} \hat{v}\kappa^2\psi & \hat{w}\kappa^2\psi \\ -\hat{w}\psi^2 & \hat{v}\psi^2 \end{pmatrix}$$

where $\kappa \equiv \frac{1}{\sqrt{\hat{v}^2 + \hat{w}^2 + 1}}$; $\psi \equiv \frac{1}{\sqrt{\hat{v}^2 + \hat{w}^2}}$.

The transformation into spherical coordinates, as a intermediary stage, allows a simple representation of the measurement in $3D$:

$$q = [(\hat{r}, \hat{\phi}, \hat{\theta}), \Lambda_{r\phi\theta}] , \text{ where } \Lambda_{r\phi\theta} = \begin{pmatrix} \infty & 0 & 0 \\ 0 & \Lambda_{\phi\theta} \\ 0 & \Lambda_{\phi\theta} \end{pmatrix}$$

\hat{r} is unknown but an estimation of \hat{r} will be chosen as is elaborated later in this section.

In practice we are interested in representing the measurement in Cartesian coordinates, thus, the measurement is transformed again from the spherical coordinates to Cartesian coordinates \mathbf{u}' as follows:

$$q = [(\hat{x}', \hat{y}', \hat{z}'), \Lambda_{xyz}]$$

where $\hat{x}' = \hat{z}'\hat{v}$; $\hat{y}' = \hat{z}'\hat{w}$; $\hat{z}' = \kappa\hat{r}$

and the covariance matrix is:

$$\Lambda_{xyz} = \left(\frac{\partial(x, y, z)}{\partial(r, \phi, \theta)} \right) \Lambda_{r\phi\theta} \left(\frac{\partial(x, y, z)}{\partial(r, \phi, \theta)} \right)^t$$

The Jacobian is

$$\frac{\partial(x, y, z)}{\partial(r, \phi, \theta)} = \begin{pmatrix} \hat{v}\kappa & \hat{r}\hat{v}\kappa\psi & -\hat{r}\hat{w}\kappa \\ \hat{w}\kappa & \hat{r}\hat{w}\kappa\psi & \hat{r}\hat{v}\kappa \\ \kappa & -\hat{r}\kappa/\psi & 0 \end{pmatrix} .$$

where the derivative is taken at the point $(\hat{r}, \hat{\phi}, \hat{\theta})$. Here too, all values are known except for \hat{r} . Since the solution to the location problem incrementally improves the estimation of T , i.e. at step k there exists an estimate \hat{T}_{k-1} from the previous step. We use this estimate to calculate an estimate of \hat{r} at step k as follows:

$$\hat{r}_k = \|\hat{T}_{k-1}(\mathbf{u})\|$$

where $\mathbf{u} \in M$ is the location of the corresponding point in the model. We emphasize that the uncertainty of this estimate, as expressed in the covariance matrix, is infinite.

It is also possible to associate uncertainty with model-features due to imprecise modeling of the $3D$ object (for example, when modeling faces or other semi-elastic objects). In this paper we assume an exact model with no uncertainties, but generalizing it to include uncertainties is straightforward.

4 The Estimation Process

4.1 The System Definition

The representation of the actual transformation \mathbf{T} is composed of two components:

- The translation component is expressed by a vector t :

$$t = (t_x, t_y, t_z)^t .$$

- The rotation component is described by a quaternion $\tilde{\mathbf{q}}$:

$$\tilde{\mathbf{q}} = (q_0, \mathbf{q}) = (q_0, q_1i + q_2j + q_3k) .$$

The rotation quaternion should satisfy the normality constrains: $\tilde{\mathbf{q}}\tilde{\mathbf{q}}^* = q_0^2 + \|\mathbf{q}\|^2 = 1$, where $\tilde{\mathbf{q}}^*$ is the conjugate of $\tilde{\mathbf{q}}$. In practice we represent the rotation component by the vector $\mathbf{s} \equiv \frac{\mathbf{q}}{q_0}$ from which the quaternion $\tilde{\mathbf{q}}$ can be reconstructed:

$$q_0 = \frac{1}{\sqrt{1 + \mathbf{s}^t\mathbf{s}}} ; \tilde{\mathbf{q}} = (q_0, q_0\mathbf{s})$$

The vector \mathbf{s} is a convenient representation of the rotational component; in addition to being minimal (having 3 parameters) the rotation equation is linear in \mathbf{s} . In order to avoid singularities in the representation when $q_0 = 0$ we always use two object-centered coordinate systems, simultaneously.

Considering the two components, the parameter vector to be estimated is $\mathbf{T} = \begin{pmatrix} \mathbf{s} \\ \mathbf{t} \end{pmatrix}$.

A mathematical relationship between the measured vector and the estimated vector is expressed for each feature i , by a non linear quaternion equation:

$$\tilde{\mathbf{u}}'_i = \tilde{\mathbf{q}}\tilde{\mathbf{u}}_i\tilde{\mathbf{q}}^* + \tilde{\mathbf{t}} \quad (1)$$

where $\tilde{\mathbf{u}}_i, \tilde{\mathbf{u}}'_i, \tilde{\mathbf{t}}$ are quaternions associated with the vectors $\mathbf{u}_i, \mathbf{u}'_i, \mathbf{t}$ respectively. Given that $\tilde{\mathbf{q}}\tilde{\mathbf{q}}^* = 1$, multiplying Equation (1) by $\tilde{\mathbf{q}}$ yields:

$$\tilde{\mathbf{u}}'_i\tilde{\mathbf{q}} = \tilde{\mathbf{q}}\tilde{\mathbf{u}}_i + \tilde{\mathbf{t}}\tilde{\mathbf{q}}$$

Isolating the vector component of this quaternion equation and dividing by q_0 we get the matrix equation:

$$h_i(\mathbf{u}_i, \mathbf{u}'_i, \mathbf{T}) \equiv \langle \mathbf{u}'_i + \mathbf{u}_i \rangle \mathbf{s} + (\mathbf{u}'_i - \mathbf{u}_i) - (I_3 + \langle \mathbf{s} \rangle) \mathbf{t} = \mathbf{0} \quad (2)$$

where $\mathbf{s} \equiv \frac{\mathbf{q}}{q_0}$ as previously defined, I_3 is the 3×3 identity matrix and $\langle \cdot \rangle$ denotes the matrix form of a cross product (i.e. $\langle \mathbf{v} \rangle \mathbf{u} = \mathbf{v} \times \mathbf{u}$).

4.2 The Estimation Control

The estimation process is composed of an incremental process, for which at each step $k - 1$, there exists an estimate $\hat{\mathbf{T}}_{k-1} = \begin{pmatrix} \hat{\mathbf{s}}_{k-1} \\ \hat{\mathbf{t}}_{k-1} \end{pmatrix}$ of the transformation \mathbf{T} and a covariance matrix Σ_{k-1} which represents the "quality" of the estimate $\hat{\mathbf{T}}_{k-1}$. Given a new match

$(\mathbf{u}_k, \hat{\mathbf{u}}'_k)$ the current estimate is updated to be $\hat{\mathbf{T}}_k$ with the associated uncertainty Σ_k . The accuracy of the estimate increases, as additional matches are fused, i.e. $\Sigma_k \leq \Sigma_{k-1}$ ($\Sigma_{k-1} - \Sigma_k$ is nonnegative definite). The process terminates as soon as the uncertainty satisfies our criterion for accuracy or no additional match can be supplied [4].

Fusing the information from a match with the old estimate is performed using the *Extended Kalman filter* (E.K.F.) process which is a generalization of the *Kalman Filter* (K.F.) to non-linear systems [6, 8]. The transition from step $k-1$ to step k is performed by using a linear approximation of the equation h_k defined in (2). Linearization of h_k is obtained by taking the first order Taylor expansion around $(\hat{\mathbf{T}}_{k-1}, \hat{\mathbf{u}}'_k)$:

$$h_k(\mathbf{u}_k, \mathbf{u}'_k, \mathbf{T}) = 0 \approx \hat{h}_k(\mathbf{u}_k, \hat{\mathbf{u}}'_k, \hat{\mathbf{T}}_{k-1}) + \frac{\partial h_k}{\partial \mathbf{u}'_k}(\mathbf{u}'_k - \hat{\mathbf{u}}'_k) + \frac{\partial h_k}{\partial \mathbf{T}}(\mathbf{T} - \hat{\mathbf{T}}_{k-1}) \quad (3)$$

which yields:

$$\langle \hat{\mathbf{s}}_{k-1} \rangle \hat{\mathbf{t}}_{k-1} + \mathbf{u}_k - \hat{\mathbf{u}}'_k \approx [I_3 - \langle \hat{\mathbf{s}}_{k-1} \rangle](\mathbf{u}'_k - \hat{\mathbf{u}}'_k) + [\langle \hat{\mathbf{u}}'_k + \mathbf{u}_k - \hat{\mathbf{t}}_{k-1} \rangle, (\langle \hat{\mathbf{s}}_{k-1} \rangle - I_3)] \mathbf{T} \quad (4)$$

Equation (4) can be rewritten as a linear equation:

$$\mathbf{z}_k = H_k \mathbf{T} + \eta_k \quad (5)$$

where

$$\begin{aligned} \mathbf{z}_k &= \langle \hat{\mathbf{s}}_{k-1} \rangle \hat{\mathbf{t}}_{k-1} + \mathbf{u}_k - \hat{\mathbf{u}}'_k \\ H_k &= [\langle \hat{\mathbf{u}}'_k + \mathbf{u}_k - \hat{\mathbf{t}}_{k-1} \rangle, (\langle \hat{\mathbf{s}}_{k-1} \rangle - I_3)] \\ \eta_k &= [I_3 - \langle \hat{\mathbf{s}}_{k-1} \rangle](\mathbf{u}'_k - \hat{\mathbf{u}}'_k) \end{aligned}$$

\mathbf{z}_k represents the new “measurement”, H_k is the matrix denoting a linear connection between the “measurement” and the actual transformation \mathbf{T} . Both \mathbf{z}_k and H_k can be derived from $\hat{\mathbf{u}}'_k$, \mathbf{u}_k , $\hat{\mathbf{T}}_{k-1}$. The term η_k - depicts the noise in the “measurement” \mathbf{z}_k and satisfies:

$$\begin{aligned} E\{\eta_k\} &= 0 \quad ; \quad E\{\eta_i \eta_j^t\} = 0 \quad \forall i \neq j \\ \text{Var}\{\eta_k\} &= [I_3 - \langle \hat{\mathbf{s}}_{k-1} \rangle] \Lambda_k [I_3 - \langle \hat{\mathbf{s}}_{k-1} \rangle]^t = W_k \end{aligned}$$

The K.F. procedure refines the estimate $\hat{\mathbf{T}}_k$ by fusing an additional match $(\mathbf{u}_{k+1}, \hat{\mathbf{u}}'_{k+1})$. The process is of the form:

$$\hat{\mathbf{T}}_{k+1} = f(\hat{\mathbf{T}}_k, \Sigma_k, \mathbf{u}_{k+1}, \hat{\mathbf{u}}'_{k+1}, \Lambda_{k+1}, h_{k+1})$$

Thus, at each stage k , there is no need of retaining any of the previously considered matches. Only the current estimate $\hat{\mathbf{T}}_k$ and its associated uncertainty Σ_k need be retained. The result is optimal in the linear minimal variance criterion. In the case where the measurement noise is a Gaussian process the result is also optimal in the sense of maximum-likelihood criterion [8].

5 Computational Aspects

5.1 Stability

In order to avoid computational instabilities several variations of the K.F. are used during the process according to the quality of information accumulated about the estimated vector. In general, in order to reduce the problem of representing a wide range of values in a limited wordlength and to improve the condition number of Σ it is preferable to use the *square root filter* [8]. With this implementation it is assured that Σ will be a non negative matrix and we double the effective precision of its values (due to the limited range of values of the square root matrix). It is provable that the computation using the square root filter is a stable process [8].

The *inverse square root filter* [8] uses only the Σ^{-1} matrix and not Σ . Thus, in the early stages of the process, when Σ^{-1} is not of full rank, the inverse square root filter should be used. In addition, stability of the process can be increased in the early stages by considering several measurements at once. This is done by grouping equations of a number of measurements into a single vector equation. The recommended number of features to be grouped together should be equal to the number of correspondences that constrain the matching to have a finite set of solutions, which is three points for 3D or perspective projection and four points for orthographic projection. This assures that Σ^{-1} is full ranked at every step. On the other hand, it is important not to use too many measurements at a single step, as this will deteriorate the linear approximation of the non-linear equations used by the E.K.F.

The convergence of the estimate to the true solution can be evaluated by studying the qualitative behavior of the covariance matrix Σ . Under the assumption that H and Λ are constant along time, we get from the K.F. equations: $\Sigma_t \sim \frac{(H^t \Lambda^{-1} H)^{-1}}{t}$, i.e. the convergence of the estimate to the true solution decays at a rate of $\frac{1}{t}$.

5.2 Initial Prediction

The K.F. process assumes an initial estimate $\hat{\mathbf{T}}_0$ with uncertainty Σ_0 . This information is not necessarily available, and must be extracted from the data. The importance of choosing a reasonable estimate is due to the linearization of the equations $\{h_i\}$. Since the linear approximation of h_1 is performed around $\hat{\mathbf{T}}_0$, an unreasonable initial estimate may influence the approximation so that the process will not converge to the correct solution.

Following many simulations, it was found that even unreasonable initial estimations converge to the ex-

pected solution using *local iterative K.F.* [6] at the initial stages of the process. The use of local iterations is effective in limiting the negative influence of the linear approximation and induces quick convergence (up to 7 iterations in our algorithm) to a reasonable estimate. At further stages of the process, where the linear approximation is reasonable, the use of local iterations does not significantly improve the transformation estimate and can be discontinued.

In any case, it is possible to give a crude estimate using simple techniques (for example: [1, 9, 7]) utilizing a small number of matched features. The initial uncertainty Σ_0 is chosen to be a diagonal matrix with very large values, so, the influence of the initial estimate $\hat{\mathbf{T}}_0$ on the final solution is marginal.

6 Advantages Using This Framework

6.1 Matching Control

The use of the K.F. process enables us to obtain reasonable matches during the estimation process [2]; at each step i , use the current estimate $\hat{\mathbf{T}}_i$ and its corresponding confidence Σ_i to select “good” matches. The selection is done using goodness of fit tests: Given a model-feature \mathbf{u}_k , let a candidate for a match be the measurement $(\hat{\mathbf{u}}'_k, \Lambda_k)$. According to this hypothesis, $\hat{h}_k(\mathbf{u}_k, \hat{\mathbf{u}}'_k, \mathbf{T}_{k-1})$ (Equation 3) is an independent random variable with a normal distribution which satisfies:

$$E\{\hat{h}_k\} = 0$$

$$\text{Var}\{(h_k)\} = \left(\frac{\partial h_k}{\partial \hat{\mathbf{u}}'_k}\right) \Lambda_k \left(\frac{\partial h_k}{\partial \hat{\mathbf{u}}'_k}\right)^t + \left(\frac{\partial h_k}{\partial \mathbf{T}}\right) \Sigma_k \left(\frac{\partial h_k}{\partial \mathbf{T}}\right)^t \equiv S_k$$

The “goodness” of fit between \mathbf{u}_k and $\hat{\mathbf{u}}'_k$ is then given by the Mahalanobis distance:

$$d(\mathbf{u}_k, \hat{\mathbf{u}}'_k) = (\hat{h}_k) S_k^{-1} (\hat{h}_k)^t = g(\hat{\mathbf{T}}_k, \Sigma_k)$$

where g has χ^2 distribution with $\text{rank}(S_k)$ degrees of freedom. The probability that the match is correct is inverse to g . As the process proceeds, the uncertainty Σ decreases lowering the number of acceptable matches.

6.2 Parallelization of the Process

Assume, the E.K.F. process was performed on 2 separate channels a and b , using n matches in each channel:

$$\begin{aligned} \{\hat{\mathbf{u}}'_i\}_{i=1\dots n} & \xrightarrow{E.K.F.} \hat{\mathbf{T}}_a, \Sigma_a \\ \{\hat{\mathbf{u}}'_j\}_{j=n+1\dots 2n} & \xrightarrow{E.K.F.} \hat{\mathbf{T}}_b, \Sigma_b \end{aligned}$$

Optimal fusion of the $2n$ matches will be performed as follows:

$$\hat{\mathbf{T}}_{ab} = \hat{\mathbf{T}}_a + K(\hat{\mathbf{T}}_b - \hat{\mathbf{T}}_a) \quad ; \quad \Sigma_{ab} = (I - K)\Sigma_a$$

where $K = \Sigma_a(\Sigma_a + \Sigma_b)^{-1}$. This result can be obtained easily from the K.F. equations if we interpret $(\hat{\mathbf{T}}_a, \Sigma_a)$ as an *a-priori* estimation of \mathbf{T} and considered $\hat{\mathbf{T}}_b$ as a new measurement with associated covariance matrix Σ_b which satisfies Equation (5) such that: $\mathbf{z} = \hat{\mathbf{T}}_b$, $H = I$ and $\text{Var}\{\eta\eta^t\} = \Sigma_b$.

Extension of this method can fuse estimates obtained from a greater number of channels (requiring $\log m$ steps for m channels). It is thus possible to decompose the K.F. process into several channels and then fuse the obtained estimates into a single optimal solution. This framework can also be used to fuse information from several kinds of measured primitives, e.g. lines and planes, where each channel is dedicated to one kind of primitive.

Note, that in this process, the initial estimate $\hat{\mathbf{T}}_0$ and the initial uncertainty Σ_0 , are considered in each of the channels. Thus it must be ensured that Σ_0 is very large so that the accumulated error due to repeated consideration of $\hat{\mathbf{T}}_0$, can be neglected. Simulation results of the parallel process show convergence equivalent to the serial process [5].

7 Simulation Results

We tested our paradigm by simulating a model as a collection of points. Points of the model were chosen by random sampling in the space 100^3 . The model points were transformed by a transformation \mathbf{T} composed of a rotation \mathbf{s} and a translation \mathbf{t} limited in length to 200. The measurements of the transformed points were contaminated by white Gaussian noise. The algorithm estimates the transformation \mathbf{T} using three kinds of data types. The algorithm assumes the correspondence between the points of the model and the measured points is known.

A sample of the results obtained by simulations for integrated measurements is presented in Graphs 2-4. Perspective projection, orthographic projection and 3D data measurements are fused in a interleaved series, one measurement type at time. The graphs depict two cases: One case where the s.t.d. for perspective, orthographic and 3D measurements are 0.2, 5.0 and 10.0 respectively, and the other where the s.t.d are 0.3, 5.0 and 5.0.

Graphs 2 and 3 show the convergence of the estimates of the translation $\hat{\mathbf{t}}$ and the rotation $\hat{\mathbf{s}}$ as a function of the number of measurements (matched points). The vertical ordinate represents the normalized error of the estimate:

$$t_i^{error} = \frac{\|\hat{\mathbf{t}}_i - \mathbf{t}\|}{\|\mathbf{t}\|} \quad ; \quad s_i^{error} = \frac{\|\hat{\mathbf{s}}_i - \mathbf{s}\|}{\|\mathbf{s}\|}$$

in Graphs 2 and 3 respectively. It can be seen that the convergence rate of the estimate is quite fast in

the initial stages of the process and corresponds to a decay of rate $\frac{1}{t}$ as expected (see Section 5.1). The convergence rate depends on the amount of noise in the measurements, however, in all cases, following quick initial convergence further improvement of the estimate requires an exponentially increasing number of measurements. When measurements are noise free, convergence to the correct solution is immediate (3-6 matches, see results in [5]).

Graph 4 shows the similarity between the trace of the covariance matrix corresponding to the estimate $\hat{\mathbf{T}}$: $E\{\|\mathbf{T} - \hat{\mathbf{T}}\|^2\}$ and the squared deviation of $\hat{\mathbf{T}}$ from the true transformation: $\|\mathbf{T} - \hat{\mathbf{T}}\|^2$. This similarity indicates that the algorithm indeed exploits the measured information.

8 Conclusion

We have proposed a general framework for model based pose estimation and have shown how it can be used to integrate two and three dimensional information. We plan in the future to extend the results to more complicated features such as line and curved primitives.

References

- [1] K.S. Arun, T.S. Huang, and S.D. Blostein. Least squares fitting of two 3D point sets. *IEEE Trans. on PAMI*, 9(5):698-700, Sept. 1987.
- [2] O.D. Faugeras and M. Hebert. The representation, recognition, and positioning of 3D shapes from range data. In A. Rosenfeld, editor, *Techniques for 3D Machine Perception*, pages 13-51. Elsevier Science, 1986.
- [3] Robert M. Haralick, Hyonam Joo, Chung-Nan Lee, Xinhua Zhuang, Vinay G. Vaidya, and Man Bae Kim. Pose estimation from corresponding point data. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1426-1445, Nov/Dec 1989.
- [4] Y. Hel-Or, A. Shmuel, and M. Werman. Active feature localization. In *Active Perception and Robot Vision*. Springer Verlag, 1991.
- [5] Y. Hel-Or and M. Werman. Absolute orientation from uncertain data: A unified approach. Technical Report TR-92-08, Hebrew University, 1992.
- [6] A.H. Jazwinski. *Stochastic process and filtering theory*. Academic Press, 1970.
- [7] Y. Liu, T.S. Huang, and O.D. Faugeras. Determination of camera location from 2D to 3D line and point correspondences. In *CVPR*, pages 82-88, 1988.
- [8] P.S. Maybeck. *Stochastic Models, Estimation, and Control*, volume 1. Academic Press, 1979.
- [9] S. Shekhar, O. Khatib, and M. Shimojo. Object localization with multiple sensors. *International Journal of Robotics Research*, 7(6):34-44, Dec. 1988.

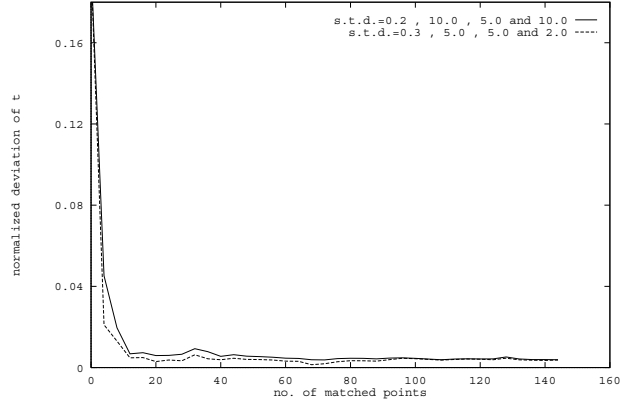


Figure 2: Integrated case: Normalized deviation of the translation estimate $\hat{\mathbf{t}}$.

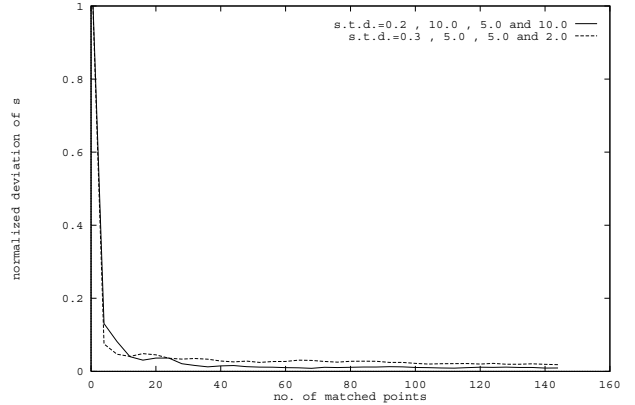


Figure 3: Integrated case: Normalized deviation of the rotation estimate $\hat{\mathbf{s}}$.

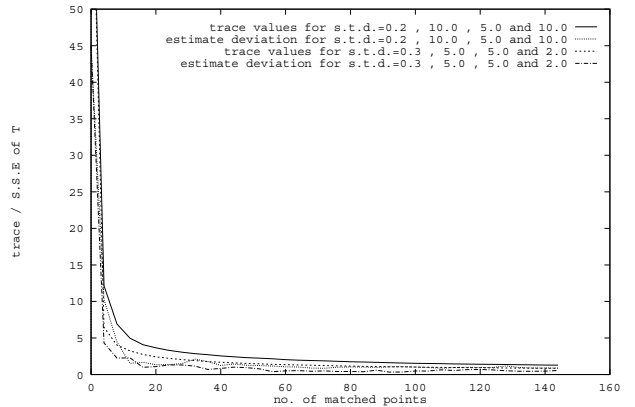


Figure 4: Integrated case: Comparison between the trace of Σ and the squared deviation of the estimate from the real transformation.