

Performance of Constraint Based Pose Estimation Algorithms

Bodo Rosenhahn, Yiwen Zhang, Gerald Sommer

Institut für Informatik und Praktische Mathematik
Christian-Albrechts-Universität zu Kiel
Preußerstrasse 1-9, 24105 Kiel, Germany
bro,yz,gs@ks.informatik.uni-kiel.de

Abstract. The paper concerns the performance of 2D-3D pose estimation algorithm in the algebraic language of kinematics. The pose estimation problem is modelled on the base of several geometric constraint equations. The dynamic measurements of these constraints are either points or lines. Instead of using matrix based LMS optimization, the development of special extended Kalman filters is proposed. The experiments aim to compare the use of different constraints and different methods of optimal estimating the pose parameters.

1 Introduction

The paper describes the estimation of pose parameters of known rigid objects in the framework of kinematics. The aim is to experimentally verify the advantages of extended Kalman filter approaches versus linear least squares optimizations. Pose estimation is a basic visual task. In spite of its importance it has been identified for a long time (see e.g. Grimson [3]), and although there is published an overwhelming number of papers with respect to that topic [7], up to now there is no unique and general solution of the problem. Pose estimation means to relate several coordinate frames of measurement data and model data by finding out the transformations between, which can subsume rotation and translation. Since we assume our measurement data as 2D and model data as 3D, we are concerned with a 2D-3D pose estimation problem. The problem can be linearly represented in motor algebra [6] or dual quaternion algebra [5]. We are using implicit formulations of the geometry as geometric constraints. We will demonstrate that geometric constraints are well conditioned (in contrast to invariances) and thus, behave more robust in case of noisy data.

The paper is organized as follows. In section two we will introduce the motor algebra as representation frame for either geometric entities, geometric constraints, and Euclidean transformations. In section three we introduce the geometric constraints and their changes in an observation scenario. In section four we compare the performance of different algorithms for constraint based pose estimation.

2 The motor algebra in the frame of kinematics

A geometric algebra $\mathcal{G}_{p,q,r}$ is a linear space of dimension 2^n , $n = p + q + r$, with a rich subspace structure, called blades, to represent so-called multivectors as higher order algebraic entities in comparison to vectors of a vector space as first order entities. A geometric algebra $\mathcal{G}_{p,q,r}$ results in a constructive way from a vector space \mathbb{R}^n , endowed with the signature (p, q, r) , $n = p + q + r$ by application of a geometric product. To make it concretely, a motor algebra is the 8D even algebra $\mathcal{G}_{3,0,1}^+$, derived from \mathbb{R}^4 , i.e. $n = 4$, $p = 3$, $q = 0$, $r = 1$, with

basis vectors γ_k , $k = 1, \dots, 4$, and the property $\gamma_1^2 = \gamma_2^2 = \gamma_3^2 = +1$ and $\gamma_4^2 = 0$. Because $\gamma_4^2 = 0$, $\mathcal{G}_{3,0,1}^+$ is called a degenerate algebra also the unit pseudoscalar, $\mathbf{I} = \gamma_1\gamma_2\gamma_3\gamma_4$, squares to zero, i.e. $\mathbf{I}^2 = (\gamma_1\gamma_2\gamma_3\gamma_4)^2 = 0$. Remembering that the hypercomplex algebra of quaternions \mathbb{H} represents a 4D linear space with one scalar and three vector components, it can simply be verified that $\mathcal{G}_{3,0,1}^+$ is isomorphic to the algebra of dual quaternions $\widehat{\mathbb{H}}$ [9]. In a general sense, motors are called all the entities existing in motor algebra. They are constituted by bivectors and scalars. Thus, any geometric entity as points, lines, and planes have a motor representation. Changing the sign of the scalar and bivector in the real and the dual parts of the motor leads to the following variants of a motor

$$\begin{aligned} \mathbf{M} &= (a_0 + \mathbf{a}) + \mathbf{I}(b_0 + \mathbf{b}) & \widetilde{\mathbf{M}} &= (a_0 - \mathbf{a}) + \mathbf{I}(b_0 - \mathbf{b}) \\ \overline{\mathbf{M}} &= (a_0 + \mathbf{a}) - \mathbf{I}(b_0 + \mathbf{b}) & \widetilde{\overline{\mathbf{M}}} &= (a_0 - \mathbf{a}) - \mathbf{I}(b_0 - \mathbf{b}) . \end{aligned}$$

These versions will be used to model the motion of points, lines and planes. In line geometry we represent rotation by a rotation line axis and a rotation angle. The corresponding entity is called a unit rotor, \mathbf{R} , and reads as follows

$$\mathbf{R} = r_0 + r_1\gamma_2\gamma_3 + r_2\gamma_3\gamma_1 + r_3\gamma_1\gamma_2 = \cos\left(\frac{\theta}{2}\right) + \sin\left(\frac{\theta}{2}\right) \mathbf{n} = \exp\left(\frac{\theta}{2} \mathbf{n}\right).$$

Here θ is the rotation angle and \mathbf{n} is the unit orientation vector of the rotation axis, spanned by the bivector basis.

If on the other hand, $\mathbf{t} = t_1\gamma_2\gamma_3 + t_2\gamma_3\gamma_1 + t_3\gamma_1\gamma_2$ is a translation vector in bivector representation, it will be represented in motor algebra as the dual part of a motor, called translator \mathbf{T} with

$$\mathbf{T} = 1 + \mathbf{I}\frac{\mathbf{t}}{2} = \exp\left(\frac{\mathbf{t}}{2} \mathbf{I}\right).$$

Thus, a translator is also a special kind of rotor.

Because rotation and translation concatenate multiplicatively in motor algebra, a motor \mathbf{M} reads

$$\mathbf{M} = \mathbf{T}\mathbf{R} = \mathbf{R} + \mathbf{I}\frac{\mathbf{t}}{2}\mathbf{R} = \mathbf{R} + \mathbf{I}\mathbf{R}'.$$

A motor represents a line transformation as a screw transformation. The screw motion equation as motor transformation reads

$$\mathbf{L}' = \mathbf{T}_s \mathbf{R}_s \mathbf{L} \widetilde{\mathbf{R}}_s \widetilde{\mathbf{T}}_s = \mathbf{M} \mathbf{L} \widetilde{\mathbf{M}}.$$

For more detailed introductions see [6]. Now we will introduce the description of the most important geometric entities [6].

A point $\mathbf{x} \in \mathbb{R}^3$, represented in the bivector basis of $\mathcal{G}_{3,0,1}^+$, i.e. $\mathbf{X} \in \mathcal{G}_{3,0,1}^+$, reads $\mathbf{X} = 1 + x_1\gamma_4\gamma_1 + x_2\gamma_4\gamma_2 + x_3\gamma_4\gamma_3 = 1 + \mathbf{I}\mathbf{x}$.

A line $\mathbf{L} \in \mathcal{G}_{3,0,1}^+$ is represented by $\mathbf{L} = \mathbf{n} + \mathbf{I}\mathbf{m}$ with the line direction $\mathbf{n} = n_1\gamma_2\gamma_3 + n_2\gamma_3\gamma_1 + n_3\gamma_1\gamma_2$ and the moment $\mathbf{m} = m_1\gamma_2\gamma_3 + m_2\gamma_3\gamma_1 + m_3\gamma_1\gamma_2$.

A plane $\mathbf{P} \in \mathcal{G}_{3,0,1}^+$ will be defined by its normal \mathbf{p} as bivector and by its Hesse distance to the origin, expressed as the scalar $d = (\mathbf{x} \cdot \mathbf{p})$, in the following way, $\mathbf{P} = \mathbf{p} + \mathbf{I}d$.

In case of screw motions $\mathbf{M} = \mathbf{T}_s \mathbf{R}_s$ not only line transformations can be modelled, but also point and plane transformations. These are expressed as follows.

$$\mathbf{X}' = \mathbf{M} \mathbf{X} \widetilde{\mathbf{M}} \quad \mathbf{L}' = \mathbf{M} \mathbf{L} \widetilde{\mathbf{M}} \quad \mathbf{P}' = \mathbf{M} \mathbf{P} \widetilde{\mathbf{M}}$$

We will use in this study only point and line transformations because points and lines are the entities of our object models.

3 Geometric constraints and pose estimation

First, we make the following assumptions. The model of an object is given by points and lines in the 3D space. Furthermore we extract line subspaces or points

in an image of a calibrated camera and match them with the model of the object. The aim is to find the pose of the object from observations of points and lines in the images at different poses. The method of obtaining the line subspaces is out of scope of this paper. Contemporary we simply got line segments by marking certain image points by hand. To estimate the pose, it is necessary to relate the observed lines in the image to the unknown pose of the object using geometric constraints.

The key idea is that the observed 2D entities together with their corresponding 3D entities are constraint to lie on other, higher order entities which result from the perspective projection. In our considered scenario there are three constraints which are attributed to two classes of constraints:

1. Collinearity: A 3D point has to lie on a line (projection ray) in the space
2. Coplanarity: A 3D point or line has to lie on a plane (projection plane).

With the terms projection ray or projection plane, respectively, we mean the image-forming ray which relates a 3D point with the projection center or the infinite set of image-forming rays which relates all 3D points belonging to a 3D line with the projection center, respectively. Thus, by introducing these two entities, we implicitly represent a perspective projection without necessarily formulating it explicitly. A similar approach of avoiding perspective projection equations by using constraint observations of lines has been proposed in [1].

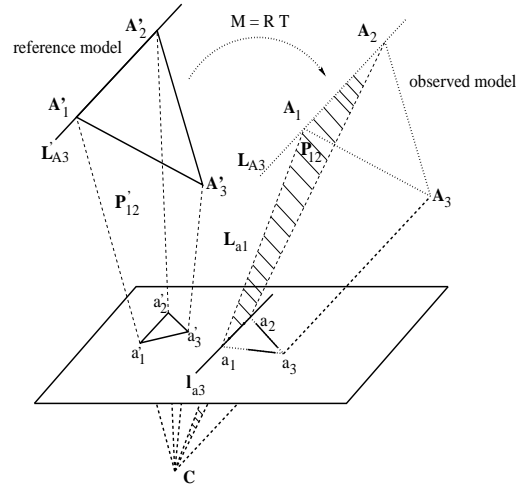


Fig. 1. The scenario. The solid lines at the left hand describe the assumptions: the camera model, the model of the object and the initially extracted lines on the image plane. The dashed lines at the right hand describe the actual pose of the model, which leads to the best fit of the object with the actual extracted lines.

To be more detailed, in the scenario of figure 1 we describe the following situation: We assume 3D points A'_i and lines L'_{A_i} of an object model. Further we extract line subspaces l_{a_i} in an image of a calibrated camera and match them with the model.

Three constraints can be depicted:

1. A transformed point, e.g. A_1 , of the model point A'_1 must lie on the projection ray L_{a_1} , given by C and the corresponding image point a_1 .

constraint	entities	dual quaternion algebra	motor algebra
point-line	point $X = 1 + I\mathbf{x}$ line $L = \mathbf{n} + I\mathbf{m}$	$LX - X\bar{L} = 0$	$XL - \bar{L}X = 0$
point-plane	point $X = 1 + I\mathbf{x}$ plane $P = \mathbf{p} + Id$	$P\bar{X} - X\bar{P} = 0$	$PX - \bar{X}\bar{P} = 0$
line-plane	line $L = \mathbf{n} + I\mathbf{m}$ plane $P = \mathbf{p} + Id$	$LP - P\bar{L} = 0$	$LP + P\bar{L} = 0$

Table 1. The geometric constraints expressed in motor algebra and dual quaternion algebra, respectively.

2. A transformed point, e.g. A_1 , of the model point A'_1 must lie on the projection plane P_{12} , given by C and the corresponding image line l_{a3} .
3. A transformed line, e.g. L_{A3} , of the model line L'_{A3} must lie on the projection plane P_{12} , given by C and the the corresponding image line l_{a3} .

Table 1 gives an overview on the formulations of these constraints in motor algebra, taken from Blaschke [2], who used expressions in dual quaternion algebra.

The meaning of the constraint equations is immediately clear. They represent the ideal situation, e.g. achieved as the result of the pose estimation procedure with respect to the observation frame. With respect to the previous reference frame, indicated by primes, these constraints read

$$\begin{aligned}
(MX'\widetilde{M})L - \bar{L}(MX'\widetilde{M}) &= 0 \\
P(MX'\widetilde{M}) - \overline{(MX'\widetilde{M})P} &= 0 \\
(ML'\widetilde{M})P + \overline{P(ML'\widetilde{M})} &= 0.
\end{aligned}$$

These compact equations subsume the pose estimation problem at hand: find the best motor M which satisfies the constraint. We will get a convex optimization problem. Any error measure $|\epsilon| > 0$ of the optimization process as actual deviation from the constraint equation can be interpreted as a distance measure of misalignment with respect to the ideal situation of table 1. That means e.g. that the constraint for a point on a line is almost fulfilled for a point near the line. The complete analysis of the constraints can be found in [4].

4 Experiments

In our experimental scenario we took a B21 mobile robot equipped with a stereo camera head and positioned it two meters in front of a calibration cube. We focused one camera on the calibration cube and took an image. Then we moved the robot, focused the camera again on the cube and took another image. The edge size of the calibration cube is 46 cm and the image size is 384×288 pixel. Furthermore, we defined on the calibration cube a 3D object model. Figure 2 shows the scenario. In the first row two perspective views of the 3D object model are shown. In the left image of the second row the calibration is performed and the 3D object model is projected onto the image. Then the camera is moved and corresponding line segments are extracted. To visualize the movement, we also projected the 3D object model on its original position. In these experiments we actually selected certain points by hand and from these the depicted line segments are derived and, by knowing the camera calibration by the cube of the first image, the actual projection ray and projection plane parameters are computed. In table 2 we show the results of different algorithms for pose estimation. In the second column of table 2 EKF denotes the use of an extended Kalman filter. The design of the extended Kalman filters is described in [4]. MAT denotes matrix algebra, SVD denotes the singular value decomposition of

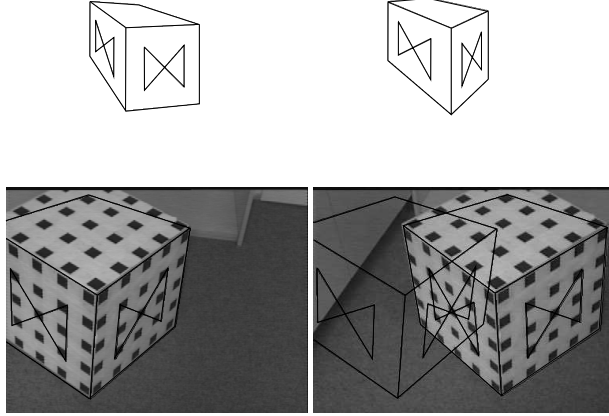


Fig. 2. The scenario of the experiment: In the top row two perspectives of the 3D object model are shown. In the second row (left) the calibration is performed and the 3D object model is projected on the image. Then the camera moved and corresponding line segments are extracted.

no.	$\mathcal{R} - t$	Constraint	Experiment 1		Error
1	RtEKF — RtEKF	XL-XL	$\mathcal{R} = \begin{pmatrix} 0.987 & 0.089 & -0.138 \\ -0.117 & 0.969 & -0.218 \\ 0.115 & 0.231 & 0.966 \end{pmatrix}$	$t = \begin{pmatrix} -58.21 \\ -217.26 \\ 160.60 \end{pmatrix}$	5.2
2	SVD — MAT	XL-XL	$\mathcal{R} = \begin{pmatrix} 0.976 & 0.107 & -0.191 \\ -0.156 & 0.952 & -0.264 \\ 0.154 & 0.287 & 0.945 \end{pmatrix}$	$t = \begin{pmatrix} -60.12 \\ -212.16 \\ 106.60 \end{pmatrix}$	6.7
3	RtEKF — RtEKF	XP-XP	$\mathcal{R} = \begin{pmatrix} 0.987 & 0.092 & -0.133 \\ -0.118 & 0.973 & -0.200 \\ 0.111 & 0.213 & 0.970 \end{pmatrix}$	$t = \begin{pmatrix} -52.67 \\ -217.00 \\ 139.00 \end{pmatrix}$	5.5
4	RtEKF — MAT	XP-XP	$\mathcal{R} = \begin{pmatrix} 0.986 & 0.115 & -0.118 \\ -0.141 & 0.958 & -0.247 \\ 0.085 & 0.260 & 0.962 \end{pmatrix}$	$t = \begin{pmatrix} -71.44 \\ -219.34 \\ 124.71 \end{pmatrix}$	3.7
5	SVD — MAT	XP-XP	$\mathcal{R} = \begin{pmatrix} 0.979 & 0.101 & -0.177 \\ -0.144 & 0.957 & -0.251 \\ 0.143 & 0.271 & 0.952 \end{pmatrix}$	$t = \begin{pmatrix} -65.55 \\ -221.18 \\ 105.87 \end{pmatrix}$	5.3
6	SVD — MAT	LP-XP	$\mathcal{R} = \begin{pmatrix} 0.976 & 0.109 & -0.187 \\ -0.158 & 0.950 & -0.266 \\ 0.149 & 0.289 & 0.945 \end{pmatrix}$	$t = \begin{pmatrix} -66.57 \\ -216.18 \\ 100.53 \end{pmatrix}$	7.1
7	MEKF — MEKF	LP-LP	$\mathcal{R} = \begin{pmatrix} 0.985 & 0.106 & -0.134 \\ -0.133 & 0.969 & -0.208 \\ 0.107 & 0.229 & 0.969 \end{pmatrix}$	$t = \begin{pmatrix} -50.10 \\ -212.60 \\ 142.20 \end{pmatrix}$	2.9
8	MEKF — MAT	LP-LP	$\mathcal{R} = \begin{pmatrix} 0.985 & 0.106 & -0.134 \\ -0.133 & 0.968 & -0.213 \\ 0.108 & 0.228 & 0.968 \end{pmatrix}$	$t = \begin{pmatrix} -67.78 \\ -227.73 \\ 123.90 \end{pmatrix}$	2.7
9	SVD — MAT	LP-LP	$\mathcal{R} = \begin{pmatrix} 0.976 & 0.109 & -0.187 \\ -0.158 & 0.950 & -0.266 \\ 0.149 & 0.289 & 0.945 \end{pmatrix}$	$t = \begin{pmatrix} -80.58 \\ -225.59 \\ 93.93 \end{pmatrix}$	6.9

Table 2. The experiment 1 results in different qualities of derived motion parameters, depending on the used constraints and algorithms to evaluate their validity.

a matrix. In the third column the used constraints, point-line (XL), point-plane (XP) and line-plane (LP) are indicated. The fourth column shows the results of the estimated rotation matrix \mathcal{R} and the translation vector t , respectively. Since the translation vectors are in mm, the results differ at around 2-3 cm. The fifth column shows the error of the equation system. Since the error of the equation system describes the Hesse distance of the entities, the value of the error is an approximation of the squared average distance of the entities. It is easy to see, that the results obtained with the different approaches are close to each other, though the implementation leads to different algorithms. Furthermore the EKF's perform more stable than the matrix solution approaches.

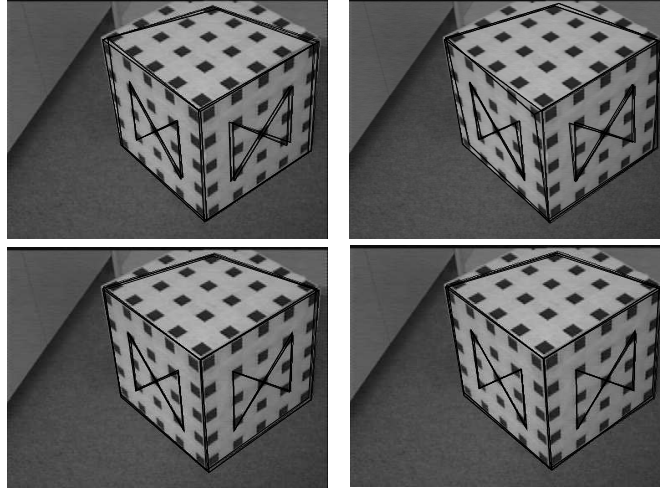


Fig. 3. Visualization of some errors. We calculate the motion of the object and project the transformed object in the image planes. The extracted line segments are also shown. In the first and second row, the results of nos. 5, 3 and nos. 7, 8 of table 2 are visualised respectively.

The visualization of some errors is done in figure 3. We calculated the motion of the object and projected the transformed object in the image plane. The extracted line segments are overlaid in addition. Figure 3 shows in the first and second row, the results of nos. 5, 3 and nos. 7, 8 of table 2 respectively.

In a second experiment we also compared the noise sensitivity of the algorithms. The experiment is organized as follows. We took the point correspondences of the first experiment and estimated both \mathcal{R} and t . Then we added a Gaussian noise error on the extracted image points. The error varied from 0 to 16 Pixels in 0.25 steps and we estimated \mathcal{R}' and t' for each step. Then we calculated the error between \mathcal{R}' and \mathcal{R} and between t' and t . The results are shown in figure 4. Since \mathcal{R} and \mathcal{R}' are rotation matrices, the absolute value of the error differs in the range $0 \leq \epsilon_{\mathcal{R}} \leq 1$. The error of the translation vector is evaluated in mm. So the error of the translation vector differs by using the matrix solution approach at around $0 \leq \epsilon_t \leq 10$ cm, while using the Kalman filter the corresponding range is $0 \leq \epsilon_t \leq 5$ cm. The matrix based solutions look all very similar. Compared with the EKF results they are very sensitive to noise and the variances between the noise steps are very high. Though the order of sensitivity is very similar, it is important to notice that the XL based solution requires as an extra calculation the estimation of the intersection of the line segments, whereas the LP based solution is fully separated: The LP constraint can be partitioned in one constraint on the real part of the motor and one constraint on the dual part of the motor. The real part can be used to generate equations with the parameters of the rotation as the only unknowns. The constraint on the dual part can then be used to determine the unknown translation. So it is possible to sequentially separate equations on the unknown rotation from equations on the unknown translation without the limitations, known from the embedding of the problem in Euclidean space [5]. This is also very useful, since the two smaller equation systems are easier and faster to solve than one larger equation system. The EKF based solutions perform all very stable and the behavior of

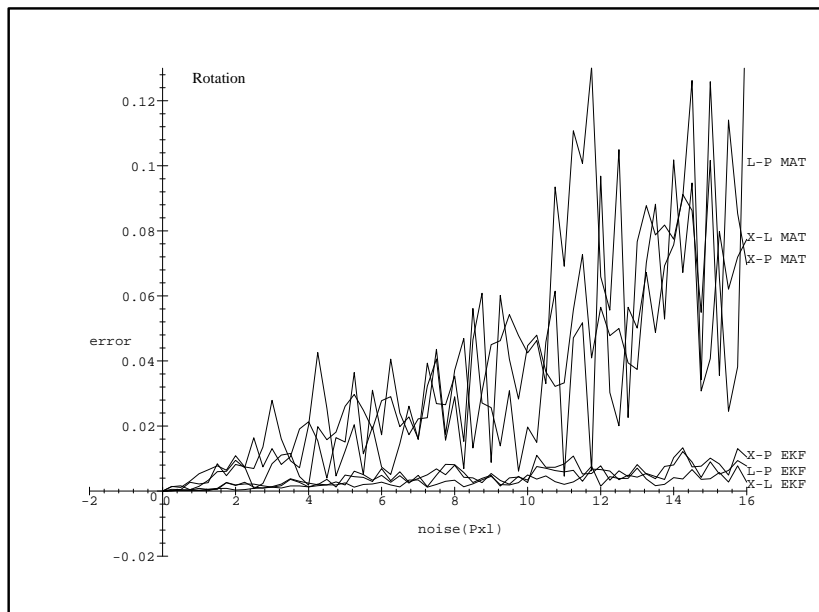
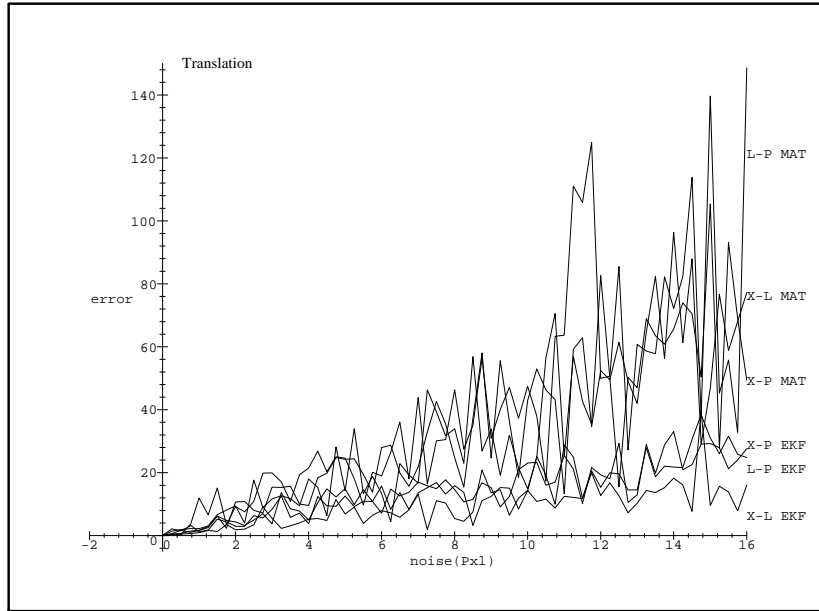


Fig. 4. Performance comparison of different methods in case of noisy data. With increasing noise the EKF performs with more accurate and more stable estimates than the matrix based methods.

the different constraints are also very similar. These results are in agreement with the well known behavior of error propagation in case of matrix based rotation estimation. This is a consequence of the estimator themselves and of the fact that in our approach rotation is represented as rotors. The concatenation of rotors is more robust than that of rotation matrices. It is obvious, that the results of these experiments are affected by the method to obtain the entities in the image. In this experiment we selected certain points directly by hand and derived from these the line subspaces. So the quality of the line subspaces is directly connected to the quality of the point extraction. For comparison purposes between the algorithms this is necessary and reasonable. But for real applications, since the extraction of lines is more stable than that of points, the XP or LP algorithms should be preferred.

5 Conclusions

In this paper we have described a framework for 2D-3D pose estimation. The aim of the paper is to compare several pose modelling approaches and estimation methods with respect to their performance. The main contribution of the paper is to formulate 2D-3D pose determination in the language of kinematics as a problem of estimating rotation and translation from geometric constraint equations. There are three such constraints which relate the model frame to an observation frame. The model data are either points or lines. The observation frame is constituted by lines or planes. Any deviations from the constraint correspond the Hesse distance of the involved geometric entities. From this starting point as a useful algebraic frame for handling line motion, the motor algebra has been introduced. The use of the motor algebra allows to subsume the pose estimation problem by compact equations. The experiments show some advantages of that representation and of the EKF approach in comparison to normal matrix based LMS algorithms, all applied within the context of constraint based pose estimation.

References

1. Shevlin F. Analysis of orientation problems using Plücker lines. *International Conference on Pattern Recognition, Brisbane*, 1: 685–689, 1998.
2. Blaschke W. Mathematische Monographien 4, Kinematik und Quaternionen. *Deutscher Verlag der Wissenschaften*, 1960.
3. Grimson W. E. L. Object Recognition by Computer. *The MIT Press, Cambridge, MA*, 1990.
4. G. Sommer, B. Rosenhahn and Y. Zhang Pose Estimation Using Geometric Constraints *Technical Report 2003, Institut für Informatik und Praktische Mathematik, Christian-Albrechts-Universität zu Kiel*
5. Daniilidis K. Hand-eye calibration using dual quaternions. *Int. Journ. Robotics Res*, 18: 286–298, 1999.
6. Bayro-Corrochano E. The geometry and algebra of kinematics. In *Sommer G., editor, Geometric Computing with Clifford Algebra. Springer Verlag*, to be published, 2000.
7. Carceroni R. L. and C. M. Brown. Numerical Methods for Model-Based Pose Recovery. *Techn. Rept. 659, Comp. Sci. Dept., The Univ. of Rochester, Rochester, N. Y.*, August 1998.
8. Walker M. W., L. Shao, and R. A. Volz. Estimating 3-D location parameters using dual number quaternions. *CVGIP: Image Understanding*, 54: 358–367, 1991.
9. Hestenes D., Li H. and A. Rockwood. New algebraic tools for classical geometry. In *Sommer G., editor, Geometric Computing with Clifford Algebra. Springer Verlag*, to be published, 2000.

This article was processed using the L^AT_EX macro package with LLNCS style