COMPUTATIONAL STATISTICS & DATA ANALYSIS

Computational Statistics & Data Analysis 41 (2002) 3–18

www.elsevier.com/locate/csda

Consistent fundamental matrix estimation in a quadratic measurement error model arising in motion analysis

A. Kukush[1], I. Markovsky*, S. Van Huffel

Dept. Elektrotechniek, ESAT-SCD-SISTA, K.U. Leuven, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium

Abstract

Consistent estimators of the rank-deficient fundamental matrix yielding information on the relative orientation of two images in two-view motion analysis are derived. The estimators are derived by minimizing a corrected contrast function in a quadratic measurement error model. In addition, a consistent estimator for the measurement error variance is obtained. Simulation results show the improved accuracy of the newly proposed estimator compared to the ordinary total least-squares estimator. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Consistent estimation; Adjusted least squares; Fundamental matrix; Total least squares; Errors-in-variables

1. Introduction: fundamental matrix estimation

This paper deals with the exploitation of the epipolar constraint information for the construction of the fundamental matrix for uncalibrated images, which once decomposed, solves the structure from motion problem (Cirrincione and Cirrincione, 1999; Mühlich and Mester, 1998; Xu and Zhang, 1996; Cirrincione, 1998).

Given a sequence of images, captured e.g. by one mobile camera (egomotion), the first step is the extraction of the feature image points. These matches are then used for the essential matrix (E) estimation if the camera is calibrated. In the uncalibrated case, by using the same techniques, the fundamental matrix (F) can be recovered. The

* Corresponding author.

E-mail addresses: alexander.kukush@esat.kuleuven.ac.be (A. Kukush), ivan.markovsky@esat.kuleuven.ac.be (I. Markovsky), sabine.vanhuffel@esat.kuleuven.ac.be (S. Van Huffel).

[1] On leave from National Taras Shevchenko University, Vladimirskaya st. 64, 01601 Kiev, Ukraine.

PII: S0167-9473(02)00068-3

essential matrix, after decomposition, yields the motion parameters. Solving for these matrices requires the same approach. In the absence of noise, the fundamental matrix is obtained from the epipolar constraints given below.

Let $u_i=[u_i(1)\ u_i(2)\ 1]^{\mathrm{T}} \in \mathbb{R}^{3\times 1}$ and $v_i=[v_i(1)\ v_i(2)\ 1]^{\mathrm{T}} \in \mathbb{R}^{3\times 1}$, $i=1,\dots,N$, represent the homogeneous pixel coordinates in the first and second image, respectively. The model is

$$v_i^{\mathrm{T}} F u_i = 0 \quad \text{for } i = 1,\dots,N, \tag{1}$$

where $F \in \mathbb{R}^{3\times 3}$ is the *fundamental matrix* which is identical for all pairs of corresponding vectors u_i, v_i, $1 \leqslant i \leqslant N$. We assume that $\operatorname{rank}(F)=2$, and F is a parameter of interest. This set can be solved exactly only in absence of noise, e.g. by using the eight-point algorithm (Hartley, 1997). For noisy images, more matches are needed and a measurement error model (Fuller, 1987) must be considered, because the first two components of the vectors u_i, v_i are observed with errors. We suppose that

$$u_i = u_{0,i} + \tilde{u}_i \quad \text{and} \quad v_i = v_{0,i} + \tilde{v}_i \quad \text{for} \quad i = 1,\dots,N \tag{2}$$

and that there exists $F_0 \in \mathbb{R}^{3\times 3}$, such that

$$v_{0,i}^{\mathrm{T}} F_0 u_{0,i} = 0 \quad \text{for} \quad i = 1,\dots,N. \tag{3}$$

The matrix $F_0 \in \mathbb{R}^{3\times 3}$ is the true fundamental matrix F and $\operatorname{rank}(F_0) = 2$. We assume that F_0 is normalized, i.e., $\|F_0\|_F = 1$. The vectors $u_{0,i}$ and $v_{0,i}$ are the true values of the measurements u_i and v_i, respectively, and $\tilde{u}_i$ and $\tilde{v}_i$ represent the measurement errors.

In Mühlich and Mester (1998) a total least-squares (TLS) (Van Huffel and Vandewalle, 1991) estimator of F_0 is proposed. The idea is to transform (1) in the form

$$(u_i \otimes v_i)^{\mathrm{T}} \operatorname{vec}(F) = 0 \quad \text{for } i = 1,\dots,N \tag{4}$$

and to interpret the observations $a_i \triangleq u_i \otimes v_i$ as

$$a_i = u_{0,i} \otimes v_{0,i} + d_i, \tag{5}$$

where $d_1,\dots,d_N$ are zero mean i.i.d. random vectors. These assumptions justify the application of the TLS method (Van Huffel and Vandewalle, 1991).

The TLS estimator of F_0 is found by solving

$$\min_{f=\operatorname{vec}(F)} \|Af\|_2 = \min \sum_{i=1}^{N} r_i^2 \quad \text{s.t. } f^{\mathrm{T}} f = 1, \tag{6}$$

where $A \triangleq [a_1 \cdots a_N]^{\mathrm{T}}$ and $r_i \triangleq a_i^{\mathrm{T}} f$ is the ith residual. This problem is solved by the eigenvector of $A^{\mathrm{T}}A$ (moment matrix) associated to the smallest eigenvalue or equivalently the right singular vector of A associated to the smallest singular value. The TLS solution is suboptimal, biased, and inconsistent (Van Huffel and Vandewalle, 1991) because the perturbations in the a_i^{T} rows are not Gaussian distributed as their elements involve the product of two spatial coordinates. Even if the combined vector of measurement errors $[\tilde{u}_i^{\mathrm{T}}\ \tilde{v}_i^{\mathrm{T}}]^{\mathrm{T}}$ is zero mean i.i.d., d_i is not i.i.d. It can be shown that

$$\mathbf{E}[d_i d_i^{\mathrm{T}}] = V_{\tilde{u}} \otimes (v_{0,i} v_{0,i}^{\mathrm{T}}) + (u_{0,i} u_{0,i}^{\mathrm{T}}) \otimes V_{\tilde{v}} + V_{\tilde{u}} \otimes V_{\tilde{v}},$$

where $\mathbf{E}[\tilde{u}_i \tilde{u}_i^{\mathrm{T}}] \triangleq V_{\tilde{u}}$ and $\mathbf{E}[\tilde{v}_i \tilde{v}_i^{\mathrm{T}}] \triangleq V_{\tilde{v}}$.

A lot of techniques have been tried in order to improve the accuracy of the eight-point algorithm in the presence of noise (Cirrincione and Cirrincione, 1999; Cirrincione, 1998; Chaudhuri and Chatterjee, 1996; Torr and Murray, 1997; Hartley, 1997; Mühlich and Mester, 1998; Leedan and Meer, 2000). In case of large images, the condition number of $A^{\mathrm{T}}A$ worsens because of the lack of homogeneity in the image coordinates. In order to avoid this problem, several scalings of the point coordinates have been proposed with good results (Hartley, 1997). One way of scaling is to normalize the input vectors. Chaudhuri and Chatterjee (1996) use this preprocessing before ordinary TLS (this approach yields very bad results). Another preprocessing used in the literature is the statistical scaling of Hartley (1997) which requires a centering and a scaling (either isotropic or non-isotropic) of the image feature points. This preprocessing has found a theoretical justification in the paper of Mühlich and Mester (1998) limited to the assumption of noise confined only in the second image. These authors only justify the isotropic scaling in the second image while accepting the two scalings in the first image, and propose the use of the mixed LS-TLS algorithm (Van Huffel and Vandewalle, 1991). However, these assumptions are also not realistic.

Cirrincione (Cirrincione, 1998; Chaudhuri and Chatterjee, 1996) further improved the (Mühlich and Mester, 1998) method by means of a robust constrained TLS (CTLS) technique, which solves (6) by taking into account the algebraic dependencies between the errors. Also Leedan and Meer (2000) applied a similar approach using a generalized TLS techniques (Van Huffel and Vandewalle, 1989). Despite these improvements the CTLS estimation remains inconsistent and biased. The same applies to all other estimates mentioned above under the conditions of models (2) and (3).

In this paper we derive a consistent estimator for the fundamental matrix F_0 by taking more realistic assumptions. Instead of (5), we give assumptions on the errors $\tilde{u}_i$ and $\tilde{v}_i$ in (2).

(i) The error vectors $\{\tilde{u}_i, \tilde{v}_i,\ i \geqslant 1\}$ are independent with $\mathbf{E}[\tilde{u}_i] = \mathbf{E}[\tilde{v}_i] = 0$, for $i \geqslant 1$.
(ii) $\mathrm{cov}(\tilde{u}_i) = \mathrm{cov}(\tilde{v}_i) = \sigma_0^2 \cdot \mathrm{diag}(1,1,0),\ i \geqslant 1$, with fixed $\sigma_0 > 0$.

Let $\tilde{u}_i = [\tilde{u}_i(1)\ \tilde{u}_i(2)\ \tilde{u}_i(3)]^{\mathrm{T}}$. Assumption (ii) means that the components of $\tilde{u}_i$ are non-correlated, $\tilde{u}_i(3) = 0$ and $\mathrm{var}(\tilde{u}_i(1)) = \mathrm{var}(\tilde{u}_i(2)) = \sigma_0^2$. The same holds for $\tilde{v}_i$.

Models (2) and (3) are quadratic measurement error models (Fuller, 1987), where the right-hand side is observed without error.

In Section 2, a consistent fundamental matrix estimator is derived assuming that the measurement error variance σ_0^2 is known. Section 3 considers consistent estimator of this measurement error variance if the latter is unknown. The computation of the fundamental matrix is summarized in Section 4 and Section 5 presents simulation results, which confirm the consistency properties of the newly proposed estimator and show its good performance compared to an ordinary TLS estimator.

2. Consistent estimator in the case of known measurement error variance

In this section we suppose that σ_0^2 is known, i.e. the covariance structure of the errors is known. The estimator proposed below is the corrected minimum contrast estimator, considered in Kukush and Zwanzig in a more general context. It is related to the method of corrected score functions a (Carroll et al., 1995, Chapter 6).

We start with the LS objective function

$$q_{\rm LS}(F;u_1,\ldots,u_N;\ v_1,\ldots,v_N)$$
$$\triangleq \sum_{i=1}^{N}(v_i^{\rm T}Fu_i)^2, \quad F\in\mathbb{R}^{3\times 3}, \quad u_i\in\mathbb{R}^{3\times 1}, \quad v_i\in\mathbb{R}^{3\times 1}.$$

Next, we construct an adjusted objective function $q(F;\ u_1,\ldots,u_N;\ v_1,\ldots,v_N)$, such that

$$\mathbf{E}[q(F;u_{0,1}+\tilde{u}_1,\ldots,u_{0,N}+\tilde{u}_N;\ v_{0,1}+\tilde{v}_1,\ldots,v_{0,N}+\tilde{v}_N)]$$
$$= q_{\rm LS}(F;u_{0,1},\ldots,u_{0,N};v_{0,1},\ldots,v_{0,N}) \tag{7}$$

for each $F\in\mathbb{R}^{3\times 3},\ u_{0,i}\in\mathbb{R}^{3\times 1},\ v_{0,i}\in\mathbb{R}^{3\times 1},\ i=1,\ldots,N.$

Note 1. The function $q_{\rm LS}$ is a contrast function in the sense of Kukush and Zwanzig. E.g. it equals 0 (for large enough N) iff F is proportional to the true value matrix. According to the method from Kukush and Zwanzig the $q_{\rm LS}$ function leads through the q function from (7) to a consistent estimating procedure.

At the first stage an estimator $\hat{F}_1$ is defined as the random matrix

$$\hat{F}_1 \in \arg\min q(F;u_1,\ldots,u_N;\ v_1,\ldots,v_N) \quad \text{s.t. } \|F\|_F = 1. \tag{8}$$

(The minimization could have a non-unique solution. See Note 2.) Following Mühlich and Mester (1998), we construct an estimator $\hat{F}$ at the second stage by expanding the current estimator $\hat{F}_1$ to a sum of rank one matrices and suppressing the matrix with the lowest Frobenius norm. Practically, this is done by deleting the smallest singular triplet in the dyadic decomposition of $\hat{F}_1$ (Golub and Van Loan, 1996). For the estimator $\hat{F}$, we have $\text{rank}(\hat{F})=2$ or 1.

Now, we find the solution q of Eq. (7). By assumption (i), it is possible to split the problem and solve the equation

$$\mathbf{E}[c(F,u_0+\tilde{u},v_0+\tilde{v})] = c_{\rm LS}(F,u_0,v_0), \tag{9}$$

$F\in\mathbb{R}^{3\times 3},\ u_0\in\mathbb{R}^{3\times 1},\ v_0\in\mathbb{R}^{3\times 1},\ c_{\rm LS} \triangleq (v_0^{\rm T}Fu_0)^2,$

$$\mathbf{E}[\tilde{u}] = \mathbf{E}[\tilde{v}] = 0, \quad \text{cov}(\tilde{u}) = \text{cov}(\tilde{v}) \triangleq V = \sigma_0^2\,\text{diag}(1,1,0)$$

and $\tilde{u}$ and $\tilde{v}$ are independent.

The function

$$c(F,u,v) \triangleq \text{tr}((vv^{\rm T}-V)F(uu^{\rm T}-V)F^{\rm T}) \tag{10}$$

satisfies Eq. (9) (see Appendix A). Then the solution of (7) is given by

$$q(F; u_1, \dots, u_N;\ v_1, \dots, v_N) = \operatorname{tr}\left(\sum_{i=1}^{N}(v_i v_i^{\mathrm{T}} - V)F(u_i u_i^{\mathrm{T}} - V)F^{\mathrm{T}}\right).$$

We denote $f \triangleq \operatorname{vec}(F)$. Then

$$q(F; u_1, \dots, u_N;\ v_1, \dots, v_N) = f^{\mathrm{T}}\left(\sum_{i=1}^{N}(u_i u_i^{\mathrm{T}} - V) \otimes (v_i v_i^{\mathrm{T}} - V)\right) f.$$

Denote

$$S_N \triangleq \sum_{i=1}^{N}(u_i u_i^{\mathrm{T}} - V) \otimes (v_i v_i^{\mathrm{T}} - V). \tag{11}$$

Let

$$\hat{f}_1 \in \arg\min f^{\mathrm{T}} S_N f \quad \text{s.t. } \|f\| = 1. \tag{12}$$

The matrix S_N is symmetric. From (12) we see that $\hat{f}_1$ is a normalized eigenvector of S_N, associated with the smallest eigenvalue λ_9 of S_N.

Now, suppose that $\|\hat{F}_1 - F_0\|_F \leqslant \varepsilon$ with $\hat{f}_1 \triangleq \operatorname{vec}(\hat{F}_1)$. By our conditions, we have $\operatorname{rank}(F_0) = 2$. Therefore for the estimator $\hat{F}$ on the second stage, we have

$$\|\hat{F}_1 - \hat{F}\|_F \leqslant \|\hat{F}_1 - F_0\|_F \leqslant \varepsilon. \tag{13}$$

Then

$$\|\hat{F} - F_0\|_F \leqslant \|\hat{F} - \hat{F}_1\|_F + \|\hat{F}_1 - F_0\|_F \leqslant 2\varepsilon.$$

Thus for consistency of the estimator $\hat{F}$, it is sufficient to show that the estimator $\hat{F}_1$ is consistent. Note that the matrix $(-F_0)$ also satisfies (3), and $\|-F_0\|_F = \|F_0\|_F = 1$. Therefore, we estimate F_0 up to a scalar factor equal to ± 1. Introduce the matrix

$$\mathscr{F}_N \triangleq \frac{1}{N}\sum_{i=1}^{N}(u_{0,i}u_{0,i}^{\mathrm{T}}) \otimes (v_{0,i}v_{0,i}^{\mathrm{T}}). \tag{14}$$

For the vector $f_0 \triangleq \operatorname{vec}(F_0)$, we have, see (3),

$$f_0^{\mathrm{T}} \mathscr{F}_N f_0 = \frac{1}{N}\sum_{i=1}^{N}\operatorname{tr}(v_{0,i}v_{0,i}^{\mathrm{T}} F_0 u_{0,i}u_{0,i}^{\mathrm{T}} F_0^{\mathrm{T}}) = 0,$$

and $\mathscr{F}_N \geqslant 0$. Thus $\lambda_{\min}(\mathscr{F}_N)=0$. We require that there exists N_0 such that $\operatorname{rank}(\mathscr{F}_N)=8$ for $N \geqslant N_0$. Moreover, we need a stronger assumption.

Let $\lambda_1(\mathscr{F}_N) \geqslant \lambda_2(\mathscr{F}_N) \geqslant \cdots \geqslant \lambda_9(\mathscr{F}_N) = 0$ be the eigenvalues of $\mathscr{F}_N$.

(iii) There exist $N_0 \geqslant 1$ and $c_0 > 0$, s.t. for all $N \geqslant N_0$, $\lambda_8(\mathscr{F}_N) \geqslant c_0$.

Note 2. The minimization problem (12) could have a non-unique solution, but due to assumption (iii) for $N > N_0(\omega)$ the smallest eigenvalue of S_N will be unique, and then the estimator $\hat{f}_1$ will be uniquely defined, up to a sign.

The next assumptions are needed for the convergence

$$\frac{1}{N} S_N - \mathscr{F}_N \to 0 \quad \text{as } N \to \infty \text{ a.s.} \tag{15}$$

(iv) $(1/N)\sum_{i=1}^{N} \|u_{0,i}\|^4 \leqslant \text{const}$, and $(1/N)\sum_{i=1}^{N} \|v_{0,i}\|^4 \leqslant \text{const}$.
(v) For fixed $\delta > 0$, $\mathbf{E}[\|\tilde{u}_i\|^{4+\delta}] \leqslant \text{const}$, and $\mathbf{E}[\|\tilde{v}_i\|^{4+\delta}] \leqslant \text{const}$.

For two matrices A and B of the same size define the distance between A and B as the Frobenius norm of their difference,

$$\operatorname{dist}(A,B) \triangleq \|A - B\|_F.$$

Now, we prove the strong consistency of the estimator $\hat{F}_1$, which is defined in (8).

Theorem 1 (Strong consistency). *Assume that assumptions* (i)–(v) *hold. Then*

$$\operatorname{dist}(\hat{F}_1, \{-F_0, +F_0\}) \to 0 \quad \textit{as} \quad N \to \infty \textit{ a.s.} \tag{16}$$

Proof. We divide the proof into several steps.

(a) *Proof of convergence* (15): From (11) and (14) we have

$$\frac{1}{N} S_N - \mathscr{F}_N = \frac{1}{N}\sum_{i=1}^{N}((u_{0,i}u_{i,0}^{\mathrm{T}} + r_i) \otimes (v_{0,i}v_{i,0}^{\mathrm{T}} + q_i) - (u_{0,i}u_{i,0}^{\mathrm{T}}) \otimes (v_{0,i}v_{i,0}^{\mathrm{T}}))$$

with

$$r_i \triangleq (\tilde{u}_i u_{0,i}^{\mathrm{T}} + u_{0,i}\tilde{u}_i^{\mathrm{T}}) + (\tilde{u}_i\tilde{u}_i^{\mathrm{T}} - V), \tag{17}$$

$$q_i \triangleq (\tilde{v}_i v_{0,i}^{\mathrm{T}} + v_{0,i}\tilde{v}_i^{\mathrm{T}}) + (\tilde{v}_i\tilde{v}_i^{\mathrm{T}} - V). \tag{18}$$

Then

$$\frac{1}{N} S_N - \mathscr{F}_N = \frac{1}{N}\sum_{i=1}^{N} r_i \otimes q_i + \frac{1}{N}\sum_{i=1}^{N}((u_{0,i}u_{0,i}^{\mathrm{T}}) \otimes q_i) + \frac{1}{N}\sum_{i=1}^{N}(r_i \otimes (v_{0,i}v_{0,i}^{\mathrm{T}})) \triangleq R_1 + R_2 + R_3. \tag{19}$$

The terms R_1, R_2, and R_3 are average sums of the independent random matrices with zero mean, therefore, we can apply Rosenthal inequality (Rosenthal, 1970).

(a.1) *Proof of convergence* $R_1 \to 0$ *a.s.*: First, we consider the summand

$$R_{11} \triangleq \frac{1}{N}\sum_{i=1}^{N}(\tilde{u}_i\tilde{u}_i^{\mathrm{T}} - V) \otimes (\tilde{v}_i\tilde{v}_i^{\mathrm{T}} - V).$$

Let δ be a number from assumption (v), $\delta \leqslant 1$. We have

$$\mathbf{E}[\|R_{11}\|^{2+\delta/2}] \leqslant \frac{\text{const}}{N^{2+\delta/2}} \left(\sum_{i=1}^{N} \mathbf{E}[\|(\tilde{u}_i\tilde{u}_i^{\mathrm{T}} - V) \otimes (\tilde{v}_i\tilde{v}_i^{\mathrm{T}} - V)\|_F^{2+\delta/2}] \right.$$

$$+\left(\sum_{i=1}^{N}\mathbf{E}[\|(\tilde{u}_i\tilde{u}_i^{\mathrm{T}}-V)\otimes(\tilde{v}_i\tilde{v}_i^{\mathrm{T}}-V)\|_F^2]\right)^{1+\delta/4}\Bigg)$$

$$\leqslant \frac{\text{const}}{N^{2+\delta/2}}(N+N^{1+\delta/4})$$

$$\leqslant \frac{\text{const}}{N^{1+\delta/4}}$$

and

$$\sum_{N=1}^{\infty}\mathbf{E}[\|R_{11}\|^{2+\delta/2}]<\infty.$$

Therefore by the Chebyshev inequality and Borel–Cantelli lemma (Papoulis, 1991) $R_{11}\to 0$, as $N\to\infty$ a.s.

(a.2) *Proof of convergence* $R_{12}\triangleq(1/N)\sum_{i=1}N(\tilde{u}_i u_{0,i}^{\mathrm{T}})\otimes(\tilde{v}_i\tilde{v}_i^{\mathrm{T}}-V)\to 0$ *a.s.*: We have

$$\mathbf{E}[\|R_{12}\|^{2+\delta/2}]\leqslant\frac{\text{const}}{N^{2+\delta/2}}\left(\sum_{i=1}^{N}\mathbf{E}[\|(\tilde{u}_i u_{0,i}^{\mathrm{T}})\otimes(\tilde{v}_i\tilde{v}_i^{\mathrm{T}}-V)\|_F^{2+\delta/2}]\right.$$

$$\left.+\left(\sum_{i=1}^{N}\mathbf{E}[\|(\tilde{u}_i u_{0,i}^{\mathrm{T}})\otimes(\tilde{v}_i\tilde{v}_i^{\mathrm{T}}-V)\|_F^2]\right)^{1+\delta/4}\right)$$

$$\leqslant\frac{\text{const}}{N^{2+\delta/2}}\left(\sum_{i=1}^{N}\|u_{0,i}\|^{2+\delta/2}+\left(\sum_{i=1}^{N}\|u_{0,i}\|^2\right)^{1+\delta/4}\right)$$

$$=\text{const}\left(\frac{1}{N^{1+\delta/2}}\,\frac{1}{N}\sum_{i=1}^{N}\mathbf{E}[\|u_{0,i}\|^{2+\delta/2}]\right.$$

$$\left.+\frac{1}{N^{1+\delta/4}}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{E}[\|u_{0,i}\|^2]\right)^{1+\delta/4}\right)$$

$$\leqslant\frac{\text{const}}{N^{1+\delta/4}}$$

and

$$\sum_{N=1}^{\infty}\mathbf{E}[\|R_{12}\|^{2+\delta/2}]<\infty,$$

which implies the convergence $R_{21}\to 0$, as $N\to\infty$ a.s.

(a.3) *Proof of convergence* $R_{13} \triangleq (1/N)\sum_{i=1}^{N}(\tilde{u}_i u_{0,i}^{\mathrm{T}}) \otimes (\tilde{v}_i v_{0,i}^{\mathrm{T}}) \to 0$ a.s.: We have

$$
\begin{aligned}
&\mathbf{E}[\|R_{13}\|^{2+\delta/2}] \\
&\leqslant \frac{\text{const}}{N^{2+\delta/2}} \left(\sum_{i=1}^{N} \mathbf{E}[\|(\tilde{u}_i u_{0,i}^{\mathrm{T}}) \otimes (\tilde{v}_i v_{0,i}^{\mathrm{T}})\|_F^{2+\delta/2}] \right. \\
&\qquad \left. + \left(\sum_{i=1}^{N} \mathbf{E}[\|(\tilde{u}_i u_{0,i}^{\mathrm{T}}) \otimes (\tilde{v}_i v_{0,i}^{\mathrm{T}})\|_F^{2}] \right)^{1+\delta/4} \right) \\
&\leqslant \frac{\text{const}}{N^{2+\delta/2}} \left(\sum_{i=1}^{N} \|u_{0,i}\|^{4+\delta} + \sum_{i=1}^{N} \|v_{0,i}\|^{4+\delta} \right. \\
&\qquad \left. + \left(\sum_{i=1}^{N} \|u_{0,i}\|^{2} \right)^{1+\delta/4} + \left(\sum_{i=1}^{N} \|v_{0,i}\|^{2} \right)^{1+\delta/4} \right) \\
&\leqslant \frac{\text{const}}{N^{2+\delta/2}} \left(\left(\sum_{i=1}^{N} \|u_{0,i}\|^{4} \right)^{1+\delta/4} + \left(\sum_{i=1}^{N} \|v_{0,i}\|^{4} \right)^{1+\delta/4} \right. \\
&\qquad \left. + \left(\sum_{i=1}^{N} \|u_{0,i}\|^{2} \right)^{1+\delta/4} + \left(\sum_{i=1}^{N} \|v_{0,i}\|^{2} \right)^{1+\delta/4} \right) \\
&\leqslant \frac{\text{const}}{N^{1+\delta/4}} \left(\left(\frac{1}{N}\sum_{i=1}^{N} \|u_{0,i}\|^{4} \right)^{1+\delta/4} + \left(\frac{1}{N}\sum_{i=1}^{N} \|v_{0,i}\|^{4} \right)^{1+\delta/4} \right. \\
&\qquad \left. + \left(\frac{1}{N}\sum_{i=1}^{N} \|u_{0,i}\|^{2} \right)^{1+\delta/4} + \left(\frac{1}{N}\sum_{i=1}^{N} \|v_{0,i}\|^{2} \right)^{1+\delta/4} \right) \\
&\leqslant \frac{\text{const}}{N^{1+\delta/4}}
\end{aligned}
$$

and this proves that $R_{13} \to 0$, as $N \to \infty$ a.s.

The other summands of R_1 are considered similarly. Thus $R_1 \to 0$, as $N \to \infty$ a.s.

Similarly, it is proved that $R_2 \to 0$ and $R_3 \to 0$, as $N \to \infty$ a.s. Now, convergence (15) follows from expansion (19).

(b) *Proof of convergence* (16): A matrix $\mathscr{F}_N$, which approximates $(1/N)S_N$, has the smallest eigenvalue $\lambda_9(\mathscr{F}_N)=0$, and all remaining eigenvalues are separated from zero, i.e., $\lambda_i(\mathscr{F}_N) \geqslant c_0$, $1 \leqslant i \leqslant 8$, see assumption (iii) (we suppose $N \geqslant N_0$).

We fix $\omega \in \Omega$ (here Ω is the probability space) and $N \geqslant N_0$. Let $\|(1/N)S_N - \mathscr{F}_N\|_F \leqslant \varepsilon$. We want to estimate $\operatorname{dist}(\hat{F}_1(\omega), \{\pm F_0\})$. Recall that $\hat{f}_1(\omega)$ is a normalized eigenvector of $(1/N)S_N(\omega)$ associated with the smallest eigenvalue $\lambda_9((1/N)S_N(\omega))$ and f_0 is a normalized eigenvector of $\mathscr{F}_N$ belonging to $\lambda_9(\mathscr{F}_N)=0$.

By convergence (15), established in part (a) of the proof, we can view $(1/N)S_N$ as a (small) perturbation of $\mathscr{F}_N$. We refer to classical perturbation theory, see e.g. (Golub and Van Loan, 1996, p. 396, Corollary 8.1.6), bounding the eigenvalues of perturbed matrices. For the smallest eigenvalues of $(1/N)S_N$ and $\mathscr{F}_N$ we have

$$\left\|\frac{1}{N}S_N - \mathscr{F}_N\right\|_F \leqslant \varepsilon \Rightarrow \left|\lambda_9\left(\frac{1}{N}S_N(\omega)\right) - \lambda_9(\mathscr{F}_N)\right| \leqslant \varepsilon$$

$$\Rightarrow \left|\lambda_9\left(\frac{1}{N}S_N(\omega)\right)\right| \leqslant \varepsilon. \tag{20}$$

More important, however, is the effect of the perturbation on the corresponding normalized eigenvectors $\hat{f}_1$ and f_0. By making use of the perturbation theorems of eigenvectors, as given in Wedin (1972) and Davis and Kahan (1970), we have

$$\operatorname{dist}(\hat{f}_1(\omega), \pm f_0) \leqslant \frac{\varepsilon}{\lambda_8(\mathscr{F}_N) - \lambda_9((1/N)S_N(\omega))}.$$

By assumption (iii) and inequality (20), we have

$$\operatorname{dist}(\hat{f}_1(\omega), \pm f_0) \leqslant \frac{\varepsilon}{c_0 - \varepsilon}.$$

Then

$$\operatorname{dist}(\hat{F}_1(\omega), \{\pm F_0\}) = \operatorname{dist}(\hat{f}_1(\omega), \{\pm f_0\}) \leqslant L(\varepsilon) \triangleq \frac{\varepsilon}{c_0 - \varepsilon}$$

and $\lim_{\varepsilon\to 0} L(\varepsilon) = 0$. This relation and the convergence $\|(1/N)S_N - \mathscr{F}_N\|_F \to 0$ as $N\to\infty$ a.s. prove convergence (16). Theorem 1 is proved. □

As a consequence we have for the estimator $\hat{F}$, which is obtained at the second stage, that

$$\operatorname{dist}(\hat{F}, \{\pm F_0\}) \to 0 \quad \text{as } N\to\infty \text{ a.s.} \tag{21}$$

Recall that $\operatorname{rank}(F_0)=2$. This and (21) imply that a.s. there exists a random number $N_1 = N_1(\omega)$ such that for all $N > N_1$, $\operatorname{rank}(\hat{F}) = 2$.

3. Consistent estimator in the case of unknown noise covariance

Denote

$$T \triangleq \operatorname{diag}(1,1,0).$$

Then $V = \operatorname{cov}(\tilde{u}_i) = \operatorname{cov}(\tilde{v}_i) = \sigma_0^2 T$. Now, we suppose that σ_0^2 is unknown. We assume the following.

(vi) $\sigma_0^2 \in (0, d^2]$, with known $d > 0$. (d depends on the data. See Note 3.)

We want to construct a consistent estimator $\hat{\sigma}^2$, based on observations $u_i, v_i, 1 \leqslant i \leqslant N$, in models (2) and (3). We strengthen assumption (iii). Introduce a matrix

$$\mathscr{F}_N(\alpha) \triangleq \frac{1}{N}\sum_{i=1}^{N}(u_{0,i}u_{0,i}^{\mathrm{T}} + \alpha T) \otimes (v_{0,i}v_{0,i}^{\mathrm{T}} + \alpha T) \quad \text{for } \alpha \in [-d^2, d^2].$$

(vii) For each $0 < \varepsilon < d$,

$$\liminf_{N\to\infty} \min_{\varepsilon^2 \leqslant \alpha \leqslant d^2} \lambda_{\min}(\mathscr{F}_N(\alpha)) > 0$$

and

$$\liminf_{N\to\infty} \min_{-d^2 \leqslant \alpha \leqslant -\varepsilon^2} |\lambda_{\min}(\mathscr{F}_N(\alpha))| > 0.$$

Assumption (vii) implies that for $0 < \alpha \leqslant d^2$ and large N, $\mathscr{F}_N(\alpha)$ is positive definite, and for $-d^2 \leqslant \alpha < 0$ and large N, $\mathscr{F}_N(\alpha)$ is either positive definite or has a negative eigenvalue. We mention that by assumption (iii), the matrix $\mathscr{F}_N(0) = \mathscr{F}_N$ is positive semidefinite with $\lambda_9(\mathscr{F}_N) = 0$ and $\lambda_8(\mathscr{F}_N) \geqslant c_0,\ N \geqslant N_0$.

We introduce the objective function

$$Q_N(\sigma^2) \triangleq |\lambda_{\min}(S_N(\sigma^2))| \quad \text{for } 0 \leqslant \sigma^2 \leqslant d^2, \tag{22}$$

where

$$S_N(\sigma^2) \triangleq \sum_{i=1}^{N} (u_i u_i^{\mathrm{T}} - \sigma^2 T) \otimes (v_i v_i^{\mathrm{T}} - \sigma^2 T). \tag{23}$$

Note that $S_N(\sigma_0^2) = S_N$ is given in (11). We define an estimator $\hat{\sigma}^2$ as a random variable with

$$\hat{\sigma}^2 = \hat{\sigma}_N^2 \in \arg \min_{0 \leqslant \sigma^2 \leqslant d^2} Q_N(\sigma^2). \tag{24}$$

Note 3. $Q_N(\sigma^2)$ tends to 0, as σ^2 tends to infinity. It is reasonable to define d from assumption (vi), in such a way that for $\sigma \geqslant 2d Q_N(\sigma^2)$ is small, with fixed given threshold.

Lemma 2. *Assume that assumptions* (i)–(vii) *hold. Then* $\hat{\sigma}^2 \to \sigma_0^2$ *as* $N \to \infty$ *a.s.*

Proof. First we observe that

$$\frac{1}{N} S_N(\sigma^2) = \frac{1}{N} \sum_{i=1}^{N} (u_i u_i^{\mathrm{T}} - V + (\sigma_0^2 - \sigma^2) T) \otimes (v_i v_i^{\mathrm{T}} - V + (\sigma_0^2 - \sigma^2) T)$$

is a quadratic function of $(\sigma_0^2 - \sigma^2)$, $\sigma_0^2 - \sigma^2 \in [-d^2, d^2]$. Similar to the proof of (15), it is easy to show that

$$\varepsilon_N(\omega) \triangleq \sup_{0 \leqslant \sigma^2 \leqslant d^2} \left\| \frac{1}{N} S_N(\sigma^2) - \mathscr{F}_N(\sigma_0^2 - \sigma^2) \right\|_F \to 0 \quad \text{as } N \to \infty \text{ a.s.} \tag{25}$$

We have

$$\left| \lambda_{\min} \left(\frac{1}{N} S_N(\hat{\sigma}^2) \right) \right| \leqslant \left| \lambda_{\min} \left(\frac{1}{N} S_N(\sigma_0^2) \right) \right| \leqslant \varepsilon_N(\omega) \tag{26}$$

and

$$\left| \lambda_{\min} \left(\frac{1}{N} S_N(\hat{\sigma}^2) \right) \right| \geqslant |\lambda_{\min}(\mathscr{F}_N(\sigma_0^2 - \hat{\sigma}^2))| - \varepsilon_N(\omega). \tag{27}$$

We fix such $\omega \in \Omega$, for which $\varepsilon_N(\omega) \to 0$, as $N \to \infty$. The sequence $\{\hat{\sigma}_N^2(\omega),\ N \geqslant 1\}$ belongs to the interval $[0, d^2]$. Consider any convergent subsequence $\{\sigma_{N(m)}^2(\omega),\ m \geqslant 1\}$, $\sigma_{N(m)}^2(\omega) \to \sigma_\infty^2$ as $m \to \infty$. Suppose that $\sigma_\infty^2 \neq \sigma_0^2$. Then for certain $N_1 = N_1(\omega)$ and $\delta = \delta(\omega) > 0$ we have for all $N_{(m)} > N_1$

$$|\lambda_{\min}(\mathscr{F}_{N_{(m)}}(\sigma_0^2 - \hat{\sigma}^2))| \geqslant \min_{\delta^2 \leqslant |\alpha| \leqslant d^2} |\lambda_{\min}(\mathscr{F}_{N_{(m)}}(\alpha))|. \tag{28}$$

From (26)–(28), we have for $N \geqslant N_1$

$$\min_{\delta^2 \leqslant |\alpha| \leqslant d^2} |\lambda_{\min}(\mathscr{F}_N(\alpha))| \leqslant 2\varepsilon_N(\omega) \to 0 \quad \text{as } N \to \infty.$$

But this contradicts assumption (vii). Therefore $\sigma_\infty^2 = \sigma_0^2$. Thus each convergent subsequence of $\{\hat{\sigma}_N^2(\omega),\ N \geqslant 1\}$ converges to σ_0^2, therefore $\hat{\sigma}_N^2(\omega) \to \sigma_0^2$, as $N \to \infty$. We fixed ω from a set Ω_0 of probability one, therefore $\hat{\sigma}_N^2 \to \sigma_0^2$ a.s. Lemma 2 is proved. □

Now, the estimator $\hat{f}_1$ is defined as a normalized eigenvector belonging to the minimal eigenvalue of $S_N(\hat{\sigma}^2)$, and $\hat{F}_1$ is a matrix with $\operatorname{vec}(\hat{F}_1) = \hat{f}_1$.

Theorem 3. *Under assumptions* (i)–(vii), $\operatorname{dist}(\hat{F}_1, \{\pm F_0\}) \to 0$, as $N \to \infty$ *a.s.*

Proof. Due to the quadratic structure of $S_N(\sigma^2)$, we have

$$\sup_{N \geqslant 1} \sup_{\substack{0 \leqslant \sigma_1^2 \leqslant d^2,\ 0 \leqslant \sigma_2^2 \leqslant d^2 \\ |\sigma_1^2 - \sigma_2^2| \leqslant \delta}} \left\| \frac{1}{N} S_N(\sigma_1^2) - \frac{1}{N} S_N(\sigma_2^2) \right\|_F \to 0 \quad \text{as } \delta \to 0 \quad \text{a.s.}$$

This means that the function $\{S_N(\sigma^2),\ \sigma^2 \in [0, d^2];\ N \geqslant 1\}$ is equicontinuous, a.s. Therefore, see Lemma 2,

$$\begin{aligned}
\left\| \frac{1}{N} S_N(\hat{\sigma}_N^2) - \mathscr{F}_N(0) \right\|_F &\leqslant \left\| \frac{1}{N} S_N(\hat{\sigma}_N^2) - \frac{1}{N} S_N(\sigma_0^2) \right\|_F \\
&\quad + \left\| \frac{1}{N} S_N(\sigma_0^2) - \mathscr{F}_N(0) \right\|_F \\
&\leqslant \sup_{N \geqslant 1} \sup_{\substack{0 \leqslant \sigma^2 \leqslant d^2 \\ |\sigma^2 - \sigma_0^2| \leqslant |\hat{\sigma}_N^2 - \sigma_0^2|}} \left\| \frac{1}{N} S_N(\sigma^2) - \frac{1}{N} S_N(\sigma_0^2) \right\|_F \\
&\quad + \left\| \frac{1}{N} S_N(\sigma_0^2) - \mathscr{F}_N(0) \right\|_F \\
&\to 0 \quad \text{as } N \to \infty \quad \text{a.s.}
\end{aligned}$$

Recall that $\hat{f}_1$ is an eigenvector of $(1/N) S_N(\hat{\sigma}_N^2)$ and f_0 is an eigenvector of $\mathscr{F}_N(0)$, and both correspond to the minimal eigenvalue. Then like in part (b) of the proof of Theorem 1, we obtain that $\operatorname{dist}(\hat{F}_1, \{\pm F_0\}) \to 0$, as $N \to \infty$.

Now, the estimator $\hat{F}$ at the second stage is obtained from $\hat{F}_1$ by expanding the current estimate $\hat{F}_1$ to a sum of rank-one matrices and suppressing the matrix with the lowest Frobenius norm. As a consequence of Theorem 3, we have convergence (21) for the estimator $\hat{F}$.

4. Algorithm

For clarity of exposition, we outline here the computational procedure for computing the ALS estimator of the quadratic measurement error model defined by (2) and (3), as described in the previous sections.

Given: N pairs of observations $u_i \in \mathbb{R}^{3\times 1}$, $v_i \in \mathbb{R}^{3\times 1}$, $1 \leqslant i \leqslant N$ and upper bound d^2 satisfying assumption (v).

Stage 1: Computation of $\hat{F}_1$, $\|\hat{F}_1\|_F = 1$.

Compute $\hat{\sigma}^2 = \arg\min_{0 \leqslant \sigma^2 \leqslant d^2} |\lambda_{\min}(S_N(\sigma^2))|$ with

$$S_N(\sigma^2) \triangleq \sum_{i=1}^{N} (u_i u_i^{\mathrm{T}} - \sigma^2 T) \otimes (v_i v_i^{\mathrm{T}} - \sigma^2 T), \quad T = \mathrm{diag}(1, 1, 0).$$

Compute the eigenvector $\hat{f}_1$ corresponding to $\lambda_{\min}(S_N(\hat{\sigma}^2))$.

Set

$$\hat{F}_1 = \begin{bmatrix} \hat{f}_1(1) & \hat{f}_1(4) & \hat{f}_1(7) \\ \hat{f}_1(2) & \hat{f}_1(5) & \hat{f}_1(8) \\ \hat{f}_1(3) & \hat{f}_1(6) & \hat{f}_1(9) \end{bmatrix}.$$

Stage 2: Computation of $\hat{F}$, $\mathrm{rank}(\hat{F}) = 2$.

Compute the SVD of $\hat{F}_1$: $\hat{F}_1 = USV^{\mathrm{T}}$ with $UU^{\mathrm{T}} = I = V^{\mathrm{T}}V$, $U \in \mathbb{R}^{3\times 3}$, $V \in \mathbb{R}^{3\times 3}$, $S = \mathrm{diag}(s_1, s_2, s_3)$ and $s_1 \geqslant s_2 \geqslant s_3$.

Set $\hat{F} = U\hat{S}V^{\mathrm{T}}$ with $S = \mathrm{diag}(s_1, s_2, 0)$.

End

If the noise variance σ_0^2 is known then the computation in Stage 1 reduces to the computation of the smallest eigenpair $(\lambda_9, \hat{f}_1)$ of $S_N(\sigma_0^2)$.

5. Experimental results

In this section, we present numerical results for the derived estimators $\hat{F}$ and $\hat{\sigma}^2$.

The data are simulated. The fundamental matrix F_0 is a randomly chosen rank-two matrix with unit Frobenius norm. The true coordinates $u_{0,i}$ and $v_{0,i}$ have third components equal to one, and the first two components are randomly chosen vectors in $\mathbb{R}^{2\times 1}$ with unit norm and random direction. The perturbations $\tilde{u}_i$ and $\tilde{v}_i$ are selected according to the assumptions stated in the paper, i.e., the third components $\tilde{u}_i(3)$ and $\tilde{v}_i(3)$ are zeros for all $i=1,\ldots,N$ and the set $\{\tilde{u}_i(j),\ \tilde{v}_i(j),\ i=1,\ldots,N,\ j=1,2\}$ form a set of i.i.d random variables, zero mean normally distributed with variance σ_0^2. In each

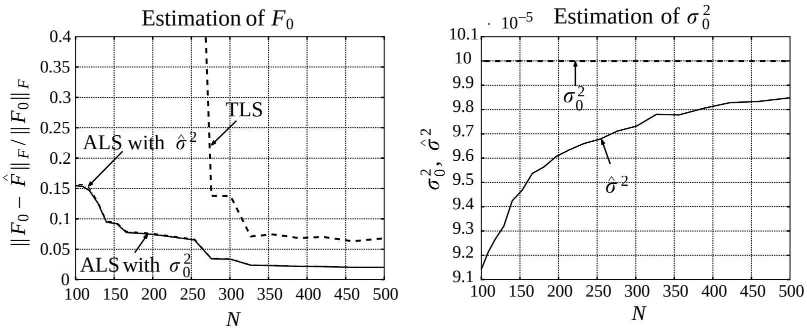

Fig. 1. Left: relative error of estimation $\|F_0 - \hat{F}\|_F / \|F_0\|_F$ as a function of the sample size N, Right: convergence of the noise variance estimate $\hat{\sigma}^2$ to the true value σ_0^2.

experiment, the estimation is repeated a number of times with the same true data and different noise realizations. The presented results (except for Fig. 3) are the average for 1000 repetitions.

The true value of the parameter F_0 is known, which allows evaluation of the results. We compare three estimators: (a) the TLS estimator $\hat{F}_{\text{TLS}}$, (b) the ALS estimator $\hat{F}$ using the true noise variance σ_0^2 (see Section 2), and (c) the ALS estimator $\hat{F}$ using the estimated noise variance $\hat{\sigma}^2$ (see Section 3). The TLS estimator is obtained as the normalized, best rank-two approximation of any solution of the following optimization problem

$$\min_F q_{\text{LS}}(F; u_1, \ldots, u_N; v_1, \ldots, v_N) \quad \text{s.t. } \|F\|_F = 1.$$

This is equivalent to solving the set $Af \approx 0$, see (4), in TLS sense (Van Huffel and Vandewalle, 1991), i.e. $\hat{f}_1$ is given by the right singular vector corresponding to the smallest singular value of A. The TLS solution then results from the truncated rank two SVD (Golub and Van Loan, 1996) of $\hat{F}_1$ constructed from $\hat{f}_1$ (by rearranging the elements of $\hat{f}_1$ column by column in a 3×3 matrix).

Fig. 1 shows the relative error of estimation $\|F_0 - \hat{F}\|_F / \|F_0\|_F$ as a function of the sample size N, on the left plot, and the convergence of the estimate $\hat{\sigma}^2$ on the right plot. Fig. 2, left plot, shows the convergence of the first stage estimator $\hat{F}_1$ to the set of rank-deficient matrices. This empirically confirms inequality (13). The right plot in Fig. 2 confirms the convergence of $(1/N)S_N \to \mathscr{F}_N$, as $N \to \infty$, see (15).

Fig. 3 shows the function $S_N(\sigma^2)$ used in the estimation of σ_0^2 for $N = 500$ on the left plot and for $N = 30$ on the right plot. These results are not averaged, i.e. they are for fixed noise realization. In general, $S_N(\sigma^2)$ is a non-convex, non-differentiable function with many local minima. However, we observed empirically that the number of local minima roughly decreases as N increases. For larger sample sizes and smaller noise variance the function $S_N(\sigma^2)$ becomes unimodal.

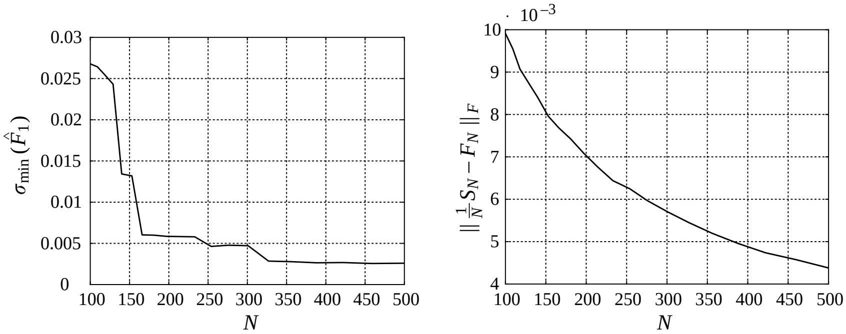

Fig. 2. Left: distance of $\hat{F}_1$ to the set of rank deficient matrices, Right: convergence of $(1/N)S_N$ to $\mathscr{F}_N$.

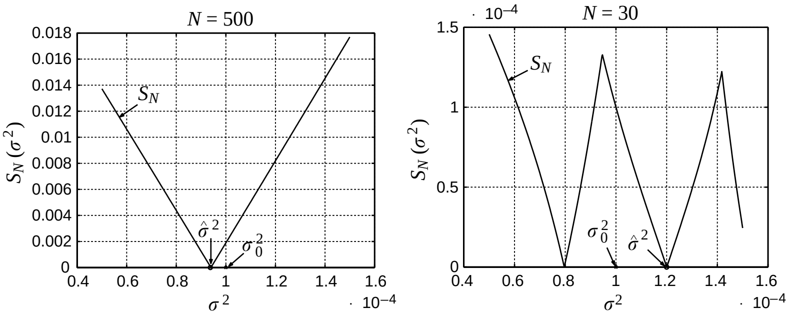

Fig. 3. The function $S_N(\sigma^2)$ used for the estimation of σ_0^2; Left: large sample size, Right: small sample size.

6. Conclusion

Consistent estimation and computation of the rank-deficient fundamental matrix, yielding all informations on motion or relative orientation of two images in two-view motion analysis, is considered here. It is shown that a consistent estimator can be derived by minimizing a corrected contrast function in a quadratic measurement error model. In addition, a consistent estimator of the measurement error variance is derived. The proposed adjusted least-squares estimator is computed in three steps: (1) estimate the measurement error variance, (2) construct a preliminary matrix estimate and (3) project that estimate into the space of singular matrices.

Numerical simulation results confirm that the newly proposed estimator outperforms the ordinary TLS based estimator.

Acknowledgements

A. Kukush is supported by a postdoctoral research fellowship of the Belgian office for Scientific, Technical and Cultural Affairs, promoting Scientific and Technical Collaboration with Central and Eastern Europe.

S. Van Huffel is an associate professor with the Katholieke Universiteit Leuven.

I. Markovsky is a research assistant with the Katholieke Universiteit Leuven.

This paper presents research results of the Belgian Programme on Interuniversity Poles of Attraction (IUAP phase V-10-29), initiated by the Belgian State, Prime Minister's Office—Federal Office for Scientific, Technical and Cultural Affairs, of the Brite-Euram Programme, Thematic Network BRRT-CT97-5040 'Niconet', of the Concerted Research Action (GOA) projects of the Flemish Government MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology), of the IDO/99/03 project (K.U. Leuven) "Predictive computer models for medical classification problems using patient data and expert knowledge", of the FWO projects G.0078.01, G.0200.00 and G.0270.02.

The scientific responsibility is assumed by its authors.

Appendix A

We show that

$$c(F,u,v) \triangleq \mathrm{tr}((vv^{\mathrm{T}} - V)F(uu^{\mathrm{T}} - V)F^{\mathrm{T}})$$

satisfies

$$\mathbf{E}[c(F,u_0+\tilde{u},v_0+\tilde{v})] = c_{\mathrm{LS}}(F,u_0,v_0), \quad c_{\mathrm{LS}}(F,u_0,v_0) \triangleq (v_0^{\mathrm{T}}Fu_0)^2,$$

under the assumptions that $\mathbf{E}[\tilde{u}] = \mathbf{E}[\tilde{v}] = 0$, $\mathrm{cov}(\tilde{u}) = \mathrm{cov}(\tilde{v}) \triangleq V$ and $\tilde{u}$ and $\tilde{v}$ are independent.

$$\begin{aligned}\mathbf{E}&[c(F,u_0+\tilde{u},v_0+\tilde{v})]\\ &=\mathbf{E}[\mathrm{tr}(((v_0+\tilde{v})(v_0+\tilde{v})^{\mathrm{T}} - V)F((u_0+\tilde{u})^{\mathrm{T}}(u_0+\tilde{u})^{\mathrm{T}} - V)F^{\mathrm{T}})]\\ &=\mathbf{E}[\mathrm{tr}((v_0v_0^{\mathrm{T}} + 2v_0\tilde{v}^{\mathrm{T}} + (\tilde{v}\tilde{v}^{\mathrm{T}} - V))F(u_0u_0^{\mathrm{T}} + 2u_0\tilde{u}^{\mathrm{T}} + (\tilde{u}\tilde{u}^{\mathrm{T}} - V))F^{\mathrm{T}})].\end{aligned}$$

After expanding the right-hand side and applying the expectation operator to the summands, the assumptions imply that all summands except for the first one are equal to zero. Thus

$$\mathbf{E}[c(F,u_0+\tilde{u},v_0+\tilde{v})] = \mathrm{tr}((v_0v_0^{\mathrm{T}})F(u_0u_0^{\mathrm{T}})F^{\mathrm{T}}).$$

But

$$\mathrm{tr}((v_0v_0^{\mathrm{T}})F(u_0u_0^{\mathrm{T}}))F^{\mathrm{T}}) = (u_0^{\mathrm{T}}F^{\mathrm{T}}v_0)(v_0^{\mathrm{T}}Fu_0) = (v_0^{\mathrm{T}}Fu_0)^2 = c_{\mathrm{LS}}(F,u_0,v_0).$$

References

Carroll, R.J., Ruppert, D., Stefanski, A., 1995. Measurement error in nonlinear models, no. 63 in Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, London, Boca Raton.

Chaudhuri, S., Chatterjee, S., 1996. Recursive estimation of motion parameters. Comput. Vision Image Understanding 64 (3), 434–442.

Cirrincione, G., 1998. A neural approach to the structure from motion problem. Ph.D. Thesis, LIS INPG Grenoble.

Cirrincione, G., Cirrincione, M., 1999. Robust neural approach for the estimation of the essential parameters in computer vision. Int. J. Artif. Intell. Tools 8 (3), 255–274.

Davis, C., Kahan, W.M., 1970. The rotation of eigenvectors by a perturbation III. SIAM J. Numer. Anal. (7), 1–46.

Fuller, W.A., 1987. Measurement Error Models. Wiley, New York.

Golub, G.H., Van Loan, C.F., 1996. Matrix Computations, 3rd Edition. Johns Hopkins University Press, Battimore, MD.

Hartley, R.I., 1997. In defence of the eight-point algorithm. IEEE Trans. Pattern Anal. Mach. Intell. 19 (6), 580–593.

Kukush, A., Zwanzig, S., On adaptive minimum contrast estimators in the implicit nonlinear functional regression models. Ukrain. Math. J. 53 (9).

Leedan, Y., Meer, P., 2000. Heteroscedastic regression in computer vision: problems with bilinear constraint Int. J. Comput. Vision 37 (2), 127–150.

Mühlich, M., Mester, R., 1998. The role of total least squares in motion analysis. in: Burkhardt, H. (Ed.), Proceedings of the European Conference on Computer Vision (ECCV'98). Springer Lecture Notes on Computer Science, Springer, Berlin, pp. 305–321.

Papoulis, A., 1991. Probability, Random Variables, and Stochastic Processes. McGraw-Hill, New York.

Rosenthal, H.P., 1970. On the subspaces of $L^p(p > 2)$ spanned by sequences of independent random variables. Israel J. Math. (8) 273–303.

Torr, P.H.S., Murray, D.W., 1997. The development and comparison of robust methods for estimating the fundamental matrix. Int. J. Comput. Vision 24 (3), 271–300.

Van Huffel, S., Vandewalle, J., 1989. Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in A are subject to error. SIAM J. Matrix Anal. 10 (3), 294–315.

Van Huffel, S., Vandewalle, J., 1991. The Total Least Squares Problem: Computational Aspects and Analysis. SIAM, Philadelphia, PA.

Wedin, P.A., 1972. Perturbation bounds in connection with the singular value decomposition. BIT (12) 99–111.

Xu, G., Zhang, Z., 1996. Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach. Kluwer Academic Publishers, Dordrecht.