Identifying spatial point sets

HELMUTH SPÄTH*

Abstract. *Two sets of spatial points are checked whether they can approximately be transformed into each other by applying some unknown translation and some unknown rotation. This problem occurs at least in two dimensions within computational metrology. Numerical methods for two types of objective functions (weighted sum of squared distances and sum of distances) are developed and numerical examples are given.*

Key words: *computational metrology, unknown rotation, unknown translation, identifying spatial point sets*

AMS subject classifications: 65D10, 65C60

Received November 28, 2002 Accepted April 2, 2003

1. Problem description

Let some original point set be given by

$$\boldsymbol{q}_i \in \mathbb{R}^3, \quad i = 1, \ldots, > 6 \tag{1}$$

and a transformed copy by

$$\boldsymbol{p}_i \in \mathbb{R}^3, \quad i = 1, \ldots, m \tag{2}$$

of it (normally with errors).

Let $\boldsymbol{t} \in \mathbb{R}^3$ denote the unknown translation and let

$$D(\boldsymbol{\varphi}) = D_1(\alpha)D_2(\beta)D_3(\gamma), \quad \boldsymbol{\varphi} = (\alpha, \beta, \gamma)^T \tag{3}$$

be the unknown rotation matrix with the property $\det D(\boldsymbol{\varphi}) = +1$ to be determined. In formula (3) the fact is used that each such orthogonal 3×3 matrix $D(\boldsymbol{\varphi})$ can be written as a product of three elementary rotations

$$\begin{gathered} D_1(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D_2(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}, \\ D_3(\gamma) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix}. \end{gathered} \tag{4}$$

*Department of Mathematics, University of Oldenburg, Postfach 2 503, D-26111 Oldenburg, Germany, e-mail: spaeth@mathematik.uni-oldenburg.de

The Euclidean distance of two corresponding points $\boldsymbol{p}_i$ and $\boldsymbol{q}_i$ is

$$d_i = \|\boldsymbol{p}_i - D(\boldsymbol{\varphi})\boldsymbol{q}_i - \boldsymbol{t}\| \geq 0, \quad (i = 1, \ldots, m)\,. \tag{5}$$

Further let be given weights

$$w_i > 0, \quad (i = 1, \ldots, m), \qquad \text{with } w = \sum_{i=1}^{m} w_i\,. \tag{6}$$

Then at least the following three objective functions to be minimized are reasonable for our aim:

$$F(\boldsymbol{t}, \boldsymbol{\varphi}) = \sum_{i=1}^{m} w_i d_i^2 \quad \text{(sum of weighted squared distances)}, \tag{7}$$

$$G(\boldsymbol{t}, \boldsymbol{\varphi}) = \sum_{i=1}^{m} w_i d_i \quad \text{(sum of weighted distances)}, \tag{8}$$

$$H(\boldsymbol{t}, \boldsymbol{\varphi}) = \max_{i=1,\ldots,m} w_i d_i \quad \text{(maximal weighted distance)}. \tag{9}$$

For two dimensions (7) is treated in [3], [5], (8) is treated in [5], and (9) is treated in [1], [2]. We will extend here the numerical algorithms given for two dimensions in [5] to three dimensions; generalizing the method for (9) of [1] from two to three dimensions seems not to be easy.

2. Numerical method for the sum of weighted squared distances

Considering the minimization of (7) we have six unknowns altogether within $\boldsymbol{t}$ and $\boldsymbol{\varphi}$. The necessary conditions for a minimum are

$$\frac{\partial F}{\partial \boldsymbol{t}} = 0\,, \tag{10}$$

$$\frac{\partial F}{\partial \boldsymbol{\varphi}} = 0\,. \tag{11}$$

Condition (10) easily gives

$$\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{\varphi}) = \frac{1}{w} \sum_{i=1}^{m} w_i (\boldsymbol{p}_i - D(\boldsymbol{\varphi})\boldsymbol{q}_i)\,. \tag{12}$$

Now we consider (11), i. e. in turn

$$\frac{\partial F}{\partial \alpha} = 0, \quad \frac{\partial F}{\partial \beta} = 0, \quad \frac{\partial F}{\partial \gamma} = 0\,. \tag{13}$$

The first condition in (13) gives

$$\sum_{i=1}^{m} w_i(\boldsymbol{p}-\boldsymbol{t})^T \frac{dD_1}{d\alpha} D_2 D_3 \boldsymbol{q}_i = \sum_{i=1}^{m} w_i \boldsymbol{q}_i^T D_3^T D_2^T D_1^T \frac{dD_1}{d\alpha} D_2 D_3 \boldsymbol{q}_i . \tag{14}$$

We will show that the right-hand side in (14) is zero, thus dramatically simplifying (14). Consider the matrix

$$D_1^T \frac{dD_1}{d\alpha} = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\sin\alpha & -\cos\alpha & 0 \\ \cos\alpha & -\sin\alpha & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

Multiplying the right-hand side we get

$$D_1^T \frac{dD_1}{d\alpha} = E_1 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

If we define $\boldsymbol{r}_i = D_2 D_3 \boldsymbol{q}_i$, then the right-hand side of (14) results in

$$\sum_{i=1}^{m} w_i \boldsymbol{r}_i^T E_1 \boldsymbol{r}_i = 0$$

because of the special structure of E_1. Thus we end up for (14) with the necessary condition

$$\sum_{i=1}^{m} w_i(\boldsymbol{p}_i-\boldsymbol{t})^T \frac{dD_1}{d\alpha} D_2 D_3 \boldsymbol{q}_i = 0 . \tag{15}$$

Similarly the second and the third conditions in (13) result in

$$\sum_{i=1}^{m} w_i(\boldsymbol{p}_i-\boldsymbol{t})^T D_1 \frac{dD_2}{d\beta} D_3 \boldsymbol{q}_i = 0 , \tag{16}$$

$$\sum_{i=1}^{m} v_i(\boldsymbol{p}_i-\boldsymbol{t})^T D_1 D_2 \frac{dD_2}{d\gamma} \boldsymbol{q}_i = 0 , \tag{17}$$

where we used

$$D_2^T \frac{dD_2}{d\beta} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} , \quad D_3^T \frac{dD_3}{d\gamma} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

and the orthogonality of D_1, D_2, and D_3. Putting temporarily

$$\boldsymbol{u}_i = w_i(\boldsymbol{p}_i-\boldsymbol{t}) , \quad \boldsymbol{v}_i = D_2 D_3 \boldsymbol{q}_i \qquad (i=1,\dots,m)$$

(15) reads

$$\sum_{i=1}^{m} \boldsymbol{u}_i^T \frac{dD_1}{d\alpha} \boldsymbol{v}_i = 0$$

where $\boldsymbol{u}_i, \boldsymbol{v}_i \in \mathbb{R}^3$. We can explicitly solve for α and get

$$\text{tg } \alpha = \text{tg } \alpha(\beta,\gamma) = \frac{\sum_{i=1}^m v_{i1}u_{i2} - v_{i2}u_{i1}}{\sum_{i=1}^m v_{i1}u_{i1} + v_{i2}u_{i2}}\,. \tag{18}$$

Similarly (16) for now

$$\boldsymbol{u}_i^T = w_i(\boldsymbol{p}_i - \boldsymbol{t})^T D_1, \quad \boldsymbol{v}_i = D_3\,\boldsymbol{q}_i \qquad (i = 1,\ldots,m)$$

gives

$$\text{tg } \beta = \text{tg } \beta(\alpha,\gamma) = \frac{\sum_{i=1}^m v_{i1}u_{i3} - v_{i3}u_{i1}}{\sum_{i=1}^m v_{i1}u_{i1} + v_{i3}u_{i3}} \tag{19}$$

and finally for

$$\boldsymbol{u}_i^T = w_i(\boldsymbol{p}_i - \boldsymbol{t})^T D_1 D_2, \quad \boldsymbol{v}_i = \boldsymbol{q}_i \qquad (i = 1,\ldots,m)$$

(17) gives

$$\text{tg } \gamma = \text{tg } \gamma(\alpha,\beta) = \frac{\sum_{i=1}^m v_{i2}u_{i3} - v_{i3}u_{i2}}{\sum_{i=1}^m v_{i2}u_{i2} + v_{i3}u_{i3}}\,. \tag{20}$$

In (18), (19), (20) we have suppressed the dependence on $\boldsymbol{t}$. If one of the denominators in (18), (19), or (20) becomes zero, then the corresponding angle is $\frac{\pi}{2}$. For each angle $\chi = \alpha, \beta, \gamma$ we have to choose between χ and $\chi + \pi$. In order to get $\frac{\partial^2 F}{\partial \chi^2} > 0$, i. e. to get a minimum of $F = F(\chi)$, χ has to be replaced by $\chi + \pi$ in the algorithm below if $B\cos\chi + A\sin\chi < 0$ where A denotes the denominator and B denotes the nominator in the right-hand side terms of (18), (19) and (20), respectively. Further we will replace χ by $\chi + 2\pi$ if $\chi < 0$ should happen. Using the necessary conditions (12), (19), (19), and (20) we can thus design the following descent algorithm for the objective function (7):

Step 1. Choose starting values $\boldsymbol{\varphi}^{(0)} = (\alpha^{(0)}, \beta^{(0)}, \gamma^{(0)})^T$, e. g. $\boldsymbol{\varphi}^{(0)} = 0$ and set $k = 0$.

Step 2. Use (12) to get

$$\boldsymbol{t}^{(k+1)} = \frac{1}{w}\sum_{i=1}^m w_i(\boldsymbol{p}_i - D(\boldsymbol{\varphi}^{(k)})\boldsymbol{q}_i) \tag{21}$$

Step 3. Solve (18) for

$$\alpha^{(k+1)} = \text{atan } \alpha(\beta^{(k)}, \gamma^{(k)})\,. \tag{22}$$

Step 4. Solve (19) for

$$\beta^{(k+1)} = \text{atan } \beta(\alpha^{(k+1)}, \gamma^{(k)})\,. \tag{23}$$

Step 5. Solve (19) for

$$\gamma^{(k+1)} = \text{atan } \gamma(\alpha^{(k+1)}, \beta^{(k+1)})\,. \tag{24}$$

Step 6. Set $k = k + 1$. If some stopping criterion, e. g. $k < k_{\max}$, is not fulfilled, then go back to Step 2 else STOP.

Note that in (23) and (24) the latest results for α and α, β are used. Thus Steps 2 through 5 will give a descent. Of course, convergence neither to some local minimum nor even to a global minimum is guaranteed. Nevertheless, empirically it will turn out that for arbitrary starting values $\boldsymbol{\varphi}^{(0)}$ the same (most likely) global minimum will be attained, as the following examples (among others) will demonstrate. These test examples were produced in the following way:

Example 1. We started with the data for $\boldsymbol{q}_i$ given by

i	1	2	3	4	5	6	7	8	9	10	11	12	13
q_1	1	3	0	−2	2	0	−1	1	0	0	−2	−2	2
q_2	1	−4	5	0	−1	1	1	0	1	0	−1	3	−3
q_3	1	1	0	3	0	−2	2	0	0	1	1	−1	1

Then we used the orthogonal matrix

$$D = \begin{pmatrix} \frac{1}{4}\sqrt{3} + \frac{1}{2} & \frac{1}{4}\sqrt{3} - \frac{1}{2} & -\frac{1}{4}\sqrt{2} \\ \frac{1}{4}\sqrt{3} - \frac{1}{2} & \frac{1}{4}\sqrt{3} + \frac{1}{2} & -\frac{1}{4}\sqrt{2} \\ \frac{1}{4}\sqrt{3} & \frac{1}{4}\sqrt{3} & \frac{1}{4}\sqrt{2} \end{pmatrix} \tag{25}$$

and $\boldsymbol{t} = (2, 5, -3)^T$ to produce (exact) values $\boldsymbol{p}_i = D\boldsymbol{q}_i + \boldsymbol{t}$ $(i = 1, \ldots, 13)$. The components of all $\boldsymbol{p}_i$ are in the interval $[-4.4, 9.7]$.

Example 2. We took the same data for $\boldsymbol{q}_i$ but dropped $\boldsymbol{p}_i$ after two decimal digits in the fixed point representation, i. e. each component is now of the form $d_0.d_1d_2$ and has errors.

Example 3. Numbers $\boldsymbol{q}_i$ were again not changed, but now all digits of $\boldsymbol{p}_i$ after the fixed point were cancelled. The resulting integer values of $\boldsymbol{p}_i$ are

i	1	2	3	4	5	6	7	8	9	10	11	12	13
p_1	2	4	1	0	3	2	0	2	1	1	0	0	3
p_1	5	0	9	4	3	6	5	4	5	4	3	8	1
p_3	−1	−2	−1	−1	−2	−4	−1	−2	−2	−2	−3	−3	−2

Example 4. Further keeping fixed $\boldsymbol{q}_i$ we randomly added ± 1 to $\boldsymbol{p}_i$ of *Example 3*.

Example 5. Keeping fixed $\boldsymbol{p}_i$ of *Example 4* we randomly also added ± 1 to $\boldsymbol{q}_i$ from *Example 1*.

For each of those examples we used two starting values $\boldsymbol{\varphi}^{(0)} = (0, 0, 0)^T$ and $\boldsymbol{\varphi}^{(0)} = (-2, 2, -2)^T$. We got identical results presented in the following table. $k_{\max}$ is the number of iterations for a 5-digit accuracy and $r_{\max}$ means the maximal

residuum in the absolute value of all components of all vectors $\boldsymbol{r}_i = \boldsymbol{p}_i - D\boldsymbol{q}_i - \boldsymbol{t}$.

Ex.	$k_{\max}$	t_1	t_2	t_3	α	β	γ	F	$r_{\max}$
1	11	2.0000	5.0000	−3.0000	−.0717	.3613	.3876	0.0000	0.0000
2	11	1.9976	4.9958	−2.9938	−.0710	.3613	.3871	0.0002	0.0085
3	13	1.5303	4.3571	−2.6012	−.0946	.3746	.3693	4.4843	1.4441
4	11	1.3978	4.0360	−2.5907	.0015	.1541	.2026	26.2925	1.4372
5	13	1.1987	3.6818	−2.9990	.1733	.0447	.1245	43.3154	3.0190

In *Example 1* $\boldsymbol{t}$ is reproduced, the orthogonal matrix D of (25) is now given as a product $D = D_1(\alpha)D_2(\beta)D_3(\gamma)$, and F and $r_{\max}$ are zero. With increasing perturbations introduced for *Examples 2* through *5*, the final values for $\boldsymbol{t}$ and $\boldsymbol{\varphi}$ deviate more and more from those of *Example 1* as to be expected. Note that with α, β, γ also e.g. $\alpha - \pi$, $\pi - \beta$, $\gamma - \pi$ give the same solution.

3. Numerical method for the sum of deviations

For simplicity we here assume w_i in the objective function (9) equal to one, i. e. we consider

$$G(\boldsymbol{t}, \boldsymbol{\varphi}) = \sum_{i=1}^{m} d_i \tag{26}$$

to be minimized. This is without loss of generality, because we may suppose that d_i in (26) have been multiplied by the former w_i. Introducing

$$\widetilde{w}_i = \widetilde{w}_i(\boldsymbol{t}, \boldsymbol{\varphi}) = \frac{1}{d_i(\boldsymbol{t}, \boldsymbol{\varphi})} \quad (i = 1, \ldots, m) \tag{27}$$

(assuming $d_i \neq 0$) we can write (26) as

$$G(\boldsymbol{t}, \boldsymbol{\varphi}) = \sum_{i=1}^{m} \widetilde{w}_i \, d_i^2 \, .$$

The difference to the former objective F from (8) is that the weights now depend on the unknowms $\boldsymbol{t}$ and $\boldsymbol{\varphi}$, too. The following algorithm is the well-known iterative least squares method (ILS), see e. g. [4]:

Step 1. Give starting values for $\widetilde{w}_i^{(0)}$, e. g. $\widetilde{w}_i^{(0)} = 1$ $(i = 1, \ldots, m)$. Set $j = 0$. Initialize $\boldsymbol{\varphi}^{(0)}$.

Step 2. Inner iterations: Solve the problem

$$\sum_{i=1}^{m} \widetilde{w}_i^{(j)} d_i^2 \longrightarrow \min$$

approximately by the method of the last section. This gives values $\boldsymbol{t}^{(j)}$ and $\boldsymbol{\varphi}^{(j)}$.

Step 3. Outer iterations: Set

$$w_i^{(j+1)} = \frac{1}{d_i(\boldsymbol{t}^{(j)}, \boldsymbol{\varphi}^{(j)})}$$

(stop for $d_i(\boldsymbol{t}^{(j)}, \boldsymbol{\varphi}^{(j)}) = 0$ or set $d_i(\boldsymbol{t}^{(j)}, \boldsymbol{\varphi}^{(j)}) = \varepsilon$ with e. g. $\varepsilon = 1.E-6$), increase j by one, and, if some stopping criterion like $j < J$ is not fulfilled, go back to Step 2 using $\boldsymbol{\varphi}^{(j)}$ as starting value there.

Convergence results for this algorithm are also poor, but empirically we always found reasonable results. This can be seen when comparing the following results with those in *Section 2*. Again we used $\boldsymbol{\varphi}^{(0)} = (0,0,0)^T$ and alternatively $\boldsymbol{\varphi}^{(0)} = (-2,2,-2)^T$ as starting values and fixed the number of inner iterations to 10. The row titled $j_{\max}$ contains the number of outer iterations to receive five digits accuracy:

Ex.	$j_{\max}$	t_1	t_2	t_3	α	β	γ	G	$r_{\max}$
1	1	2.0000	5.0000	−3.0000	−.0717	.3614	.3876	0.0000	0.0000
2	14	1.9968	4.9965	−2.9934	−.0714	.3611	.3871	0.0417	0.0092
3	19	1.3887	4.3388	−2.6013	−.0885	.3695	.3706	5.6060	1.5709
4	16	1.3413	4.0990	−2.6658	.0877	.2808	.1970	17.8042	2.3342
5	21	1.3413	3.4905	−2.9492	.1611	.1575	.0646	21.6767	3.3297

The results for $\boldsymbol{t}$ and $\boldsymbol{\varphi}$ are quite similar to the former ones. It is reasonable that for the final values F and G it is valid that $F < G$ if $r_{\max}$ is small and $F > G$ for larger values of $r_{\max}$.

References

[1] W. BIHLMAYER, W. FORST, *Approximation by circles*, ZOR **33**(1989), 315–339.

[2] W. BIHLMAYER, W. FORST, A. WECKENMANN, *Numerische Verfahren zur Prüfung von Werkstücken unter Berücksichtigung der Maximum-Material-Bedingung im ebenen Fall*, Technisches Messen**56** (1989), 23–31.

[3] W. KRABS, *Einführung in die lineare und nichtlineare Optimierung für Ingenieure*, B.G. Teubner 1983, 137–141.

[4] H. SPÄTH, *Mathematical Algorithms for Linear Regression*, Academic Press 1992.

[5] H. SPÄTH, *Zur Identifizierung von Lochmustern*, Technisches Messen, accepted.