Robust Estimation of Camera Rotation, Translation and Focal Length at High Outlier Rates

Onay Urfalıoḡlu
University of Hannover
Information Technology Laboratory
Hannover, Germany
urfaliog@lfi.uni-hannover.de

Abstract

A new approach for robust estimation of camera rotation, translation and focal length at high outlier rates from feature correspondences between two perspective views is introduced. Many computer vision algorithms use RANSAC based methods to achieve robustness. One disadvantage of these methods is the required computational effort at high outlier rates. The proposed approach is based on a new evolutionary global optimization method called Repulsive Particle Swarm Optimization (RPSO). It is shown that the proposed approach requires less computational effort at high outlier rates compared to RANSAC based robust estimation.

1. Introduction

The estimation of camera parameters in structure-from-motion algorithms is based on correspondences between feature points in multiple views. Finding correspondences is error prone due to camera noise and similarities between detected features. Furthermore only correspondences of static features can be used for camera parameter estimation. The generation of correspondences can yield a high rate of inliers (correct correspondences) if the captured scene contains many unique static features. Scenes containing human built structures like houses, streets or other artificial buildings are likely to yield high inlier rates. But there are also scenes which yield low inlier rates (high outlier rates), e.g. scenes with quasi periodic structures or scenes with many moving objects.

For robust estimation outliers have to be detected and eliminated. Many conventional approaches for outlier detection and elimination are based on the RANSAC [2] algorithm. There have been several enhancements and modifications of RANSAC [3, 4, 5, 6, 7, 8] which are all based on the common random sampling scheme. One disadvantage of random sampling based algorithms is the computational expense at high outlier rates because the number of iterations required increases very rapidly with the rate of outliers. This increase makes the use of RANSAC unpractical at high outlier rates.

This paper presents an algorithm for robust estimation of camera rotation, translation and focal length (at given intrinsic camera parameters) from a given correspondence set which may have a high outlier rate. A new evolutionary global optimization method, called Repulsive Particle Swarm Optimization (RPSO) is introduced. By control of a cost function, RPSO finds the global optimum which is consistent with the estimation of the unknown parameters and so enables the distinction of the correspondences in inliers and outliers. Compared to conventional approaches the new algorithm requires less computational expense at high outlier rates.

In the following section the conventional estimation strategy is briefly presented. In section 3 the computational expense of the RANSAC based outlier detection is discussed. Section 4 presents the new approach for robust estimation of camera rotation, translation and focal length. In section 5 results of the experiments are shown and in the last section the paper is concluded.

2. Conventional Estimation Strategy

Most approaches for robust estimation of camera rotation and translation parameters at given intrinsic camera parameters are structured as follows:

1. Detect feature points in frame 1:

$$\mathbf{p}_{(i)} := \begin{pmatrix} x_{(i)} & y_{(i)} \end{pmatrix}^\top$$

   and in frame 2:

$$\mathbf{p}'_{(j)} := \begin{pmatrix} x'_{(j)} & y'_{(j)} \end{pmatrix}^\top$$

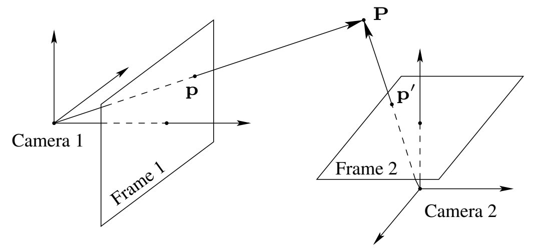

Figure 1. *Feature points* $\mathbf{p}$, $\mathbf{p}'$ *by projection of* $\mathbf{P}$ *into the camera plane*

as shown in Fig. 1

2. Determine the correspondence set C by finding the corresponding $\mathbf{p}'$ for each $\mathbf{p}$
3. Apply RANSAC to detect outliers and determine an initial *fundamental matrix* $\mathbf{F}$ [11] using the 7-point algorithm [12]. The F matrix can be written as:

$$\mathbf{F} := \mathbf{K}^{'-1}\mathbf{T}\mathbf{R}\mathbf{K}^{-1} \tag{1}$$

with $\mathbf{K}$ and $\mathbf{K}^{'}$ describing camera intrinsics [10]:

$$\mathbf{K} := \begin{pmatrix} \frac{f}{p_x} & s & c_x \\ 0 & \frac{f}{p_y} & c_y \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{K}' := \begin{pmatrix} \frac{f'}{p_x} & s & c_x \\ 0 & \frac{f'}{p_y} & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{2}$$

where f is the focal length of the camera generating frame 1, f' is the focal length of the camera generating frame 2, s is the skew of a pixel, c_x, c_y are coordinates of the principal point and p_x, p_y describe the pixel size. Except f', all intrinsic parameters are assumed to be known.
$\mathbf{R}$ is the rotation matrix:

$$\mathbf{R} := \begin{pmatrix} s_\varphi s_\vartheta s_\rho + c_\varphi c_\rho & s_\varphi s_\vartheta c_\rho - c_\varphi s_\rho & s_\varphi c_\vartheta \\ c_\vartheta s_\rho & c_\vartheta c_\rho & -s_\vartheta \\ c_\varphi s_\vartheta s_\rho - s_\varphi c_\rho & c_\varphi s_\vartheta c_\rho + s_\varphi s_\rho & c_\varphi c_\vartheta \end{pmatrix} \tag{3}$$

with φ, ϑ and ρ rotation angles and

$$\begin{aligned} s_\varphi &:= \sin(\varphi),\ s_\vartheta := \sin(\vartheta)\ s_\rho := \sin(\rho) \\ c_\varphi &:= \cos(\varphi),\ c_\vartheta := \cos(\vartheta)\ c_\rho := \cos(\rho). \end{aligned} \tag{4}$$

The components of camera translation are $T_1,\ T_2,\ T_3$ and $\mathbf{T}$ is defined as:

$$\mathbf{T} := \begin{pmatrix} 0 & T_3 & -T_2 \\ -T_3 & 0 & T_1 \\ T_2 & -T_1 & 0 \end{pmatrix}. \tag{5}$$

The magnitude of the camera translation can not be determined by the estimation so $\mathbf{T}$ can be parameterized by 2 values. With 3 parameters for the rotation, 2 parameters for the translation and one parameter for the focal length, $\mathbf{F}$ is entirely determined by 6 parameters. $\mathbf{F}$ fulfills the epipolar condition [10]:

$$\bar{\mathbf{p}}_{\text{hom}}^{'\top}\mathbf{F}\,\bar{\mathbf{p}}_{\text{hom}} = 0 \tag{6}$$

where

$$\bar{\mathbf{p}}'_{\text{hom}} := \begin{pmatrix} \bar{x}' \\ \bar{y}' \\ 1 \end{pmatrix} \text{ and } \bar{\mathbf{p}}_{\text{hom}} := \begin{pmatrix} \bar{x} \\ \bar{y} \\ 1 \end{pmatrix} \tag{7}$$

are noise free homogeneous feature points

4. Extract initial translation, rotation and focal length from the F matrix [10]

3. Computational Expense of RANSAC

The algorithm for outlier detection based on RANSAC comprises 4 steps :

1. Randomly choose 7 correspondences out of C
2. Calculate potential F matrix
3. Determine the inlier rate induced by the current F matrix
4. Go to 1 and repeat until an appropriate inlier rate is achieved

For a given correspondence of detected feature points $\mathbf{p}$,$\mathbf{p}'$ and a given F matrix the decision whether the correspondence is an inlier or an outlier is made by specifying a threshold τ for the distance between the *epipolar line* [11]

$$\mathbf{l}^\top := \mathbf{p}_{\text{hom}}^{'\top}\mathbf{F} \tag{8}$$

and the appropriate point $\mathbf{p}$:

$$\text{distance}(\mathbf{l}^\top, \mathbf{p}) \overset{?}{<} \tau \tag{9}$$

If the distance is below the specified threshold the correspondence is decided to be an inlier, otherwise it is marked as an outlier. A reasonable value for the threshold depends on the noise of the feature point coordinates. Assuming that the noise is Gaussian with zero mean and standard deviation σ a reasonable value for τ is $\tau \sim 10\sigma$.

RANSAC is an iterative algorithm and the number of iterations required to find a correct sample of 7 correspondences out of C depends on the number of correspondences N and on the number of inliers k. The outlier rate is defined as:

$$\beta := \frac{N-k}{N} \tag{10}$$

In the following the probability for a successful search after r iterations is determined.
For a given correspondence set the probability of picking up a sample with 7 inliers is:

$$P_7 := \frac{k}{N} \cdot \frac{k-1}{N-1} \cdots \frac{k-6}{N-6} \tag{11}$$

The probability for at least *one* of the 7 randomly chosen correspondences being an outlier is (resulting in a wrong F matrix) :

$$P_{\text{fail}} := 1 - P_7 \tag{12}$$

If the selection process is repeated r - times the probability that *each* trial fails is:

$$P_{\text{fail}}(r) := (1 - P_7)^r \tag{13}$$

So the probability $P_{\text{succ}}(r)$ for choosing a sample of 7 inliers in r iterations is:

$$P_{\text{succ}}(r) := 1 - P_{\text{fail}}(r) = 1 - (1 - P_7)^r \tag{14}$$

In this paper $P_{\text{succ}}(r)$ is used as a measure for the robustness of the estimation. Fig. 2 shows probability $P_{\text{succ}}(r)$ for $N = 400$ and various outlier rates.

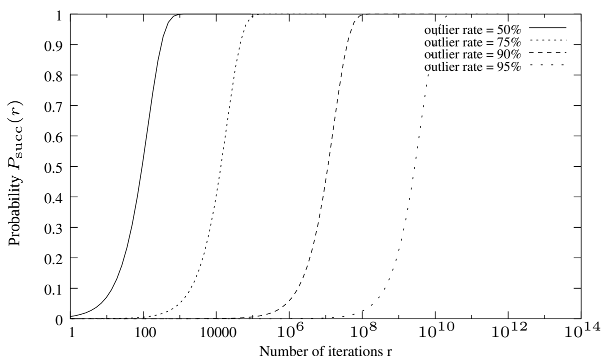

Figure 2. *Probability for a successful search with RANSAC*

The computational expense of RANSAC is defined by the number of iterations r required for a given robustness P_g. To derive the number of required iterations, P_7 is approximated by

$$\hat{P}_7 := (\frac{k}{N})^7 \tag{15}$$

and $P_{\text{fail}}(r)$ is approximated by

$$\hat{P}_{\text{fail}}(r) := (1 - \hat{P}_7)^r \tag{16}$$

With $\hat{P}_7 > P_7$ follows:

$$\begin{aligned} & \hat{P}_{\text{fail}}(r) < P_{\text{fail}}(r) \\ \Rightarrow \quad & \hat{P}_{\text{succ}}(r) := 1 - \hat{P}_{\text{fail}}(r) > P_{\text{succ}}(r) \end{aligned} \tag{17}$$

From (17) and

$$\hat{P}_{\text{succ}}(r) \stackrel{!}{=} P_g \tag{18}$$

follows:

$$\begin{aligned} r & \approx \frac{\ln(1 - P_g)}{\ln(1 - (1 - \frac{N-k}{N})^7)} \\ & \approx \frac{\ln(1 - P_g)}{\ln(1 - (1 - \beta)^7)} \end{aligned} \tag{19}$$

Fig. 3 shows number of iterations r versus outlier rate β. The number of iterations required exceeds an exponential growth with increasing outlier rate.
In the next section it is shown that the new proposed algorithm for estimation of camera rotation, translation and focal length requires less iterations and in consequence less computational expense at high outlier rates.

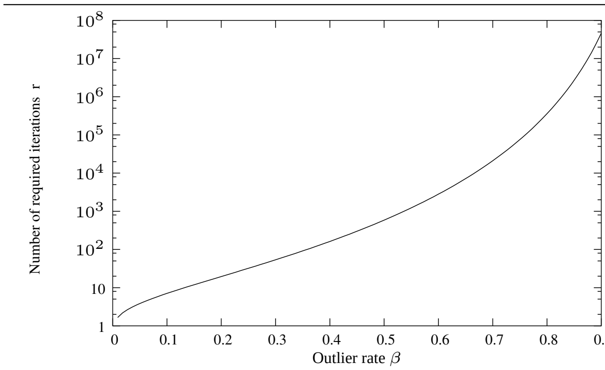

Figure 3. *Iterations r required by RANSAC for $P_g = 0.99$*

4. Estimation of Rotation, Translation and Focal Length by Global Optimization

The proposed new approach searches for the optimal F matrix given a correspondence set C with outliers by control of a cost function. The F matrix is constructed from 6 estimated parameters. Rotation is parameterized by 3 parameters:

$$\varphi, \vartheta, \rho \in [-0.2\,\text{rad}, 0.2\,\text{rad}] \tag{20}$$

The translation is parameterized by 2 parameters:

$$\begin{aligned} \zeta, \eta &\in [0, \pi] \\ T_1 &= \sin(\zeta)\,\cos(\eta) \\ T_2 &= \sin(\zeta)\,\sin(\eta) \\ T_3 &= \cos(\zeta) \end{aligned} \tag{21}$$

where ζ and η parameterize the orientation of a vector in 3D-space (spherical coordinates). For the generation of frame 1 it is assumed, without loss of generality, that the focal length f is set to 1 so $\mathbf{K}$ is completely known. The focal length f' of the camera which generates frame 2 is unknown and must be estimated. With an estimated $\hat{f}'$, $\mathbf{K}'$ is completely determined so the F matrix is constructed by equation (1). The search space is 6 dimensional and the task of the optimization is to find the estimate $\hat{\mathbf{F}}(\hat{\varphi}, \hat{\vartheta}, \hat{\rho}, \hat{\zeta}, \hat{\eta}, \hat{f}')$ (of the true $\bar{\mathbf{F}}(\bar{\varphi}, \bar{\vartheta}, \bar{\rho}, \bar{\zeta}, \bar{\eta}, \bar{f}'))$ yielding a globally optimal cost.

This problem can be solved by global optimization (GO). The GO technique proposed is a new approach called *Repulsive Particle Swarm Optimization* (RPSO) which is a modified *Particle Swarm Optimization* (PSO) [13, 14] and belongs to the class of stochastic global optimizers.

The search method of the new algorithm is fundamentally different compared to the RANSAC algorithm. The RANSAC algorithm constructs solution candidates by randomly chosen samples whereas the GO-based algorithm constructs solution candidates by interactions within a population of different solution candidates and thus makes use of more information (e.g. by sharing ’experiences’ within the population). This is one of the reasons for this new approach outperforming RANSAC at high outlier rates.

4.1. Cost Function

The detection of feature points is error prone. It is assumed that the error distribution of each feature point coordinate is Gaussian with zero mean and uniform standard deviation σ for inliers [3]. The probability density function of the noise perturbated inlier is:

$$p^{(\mathrm{ML})}_{\mathrm{inlier},(i)} := (\frac{1}{\sqrt{2\pi}\sigma})^4 \ \exp(-\frac{d^2_{(i)} + d'^2_{(i)}}{2\sigma^2}) \qquad (22)$$

with (as shown in Fig. 4)

$$\begin{aligned} \bar{\mathbf{F}} &: \text{true F matrix} \\ (\mathbf{p}_{(i)}, \mathbf{p}'_{(i)}) &: \text{detected correspondence} \\ (\bar{\mathbf{p}}_{(i)}, \bar{\mathbf{p}}'_{(i)}) &: \text{true correspondence} \\ d_{(i)} &:= \text{distance}(\bar{\mathbf{p}}_{(i)}, \mathbf{p}_{(i)}) \\ d'_{(i)} &:= \text{distance}(\bar{\mathbf{p}}'_{(i)}, \mathbf{p}'_{(i)}) \end{aligned} \qquad (23)$$

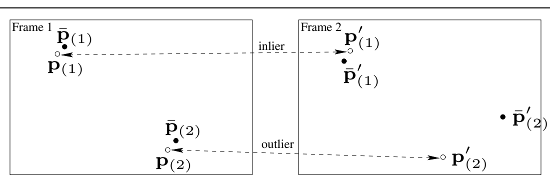

Figure 4. *Example for feature points and detected correspondences (inlier and outlier)*

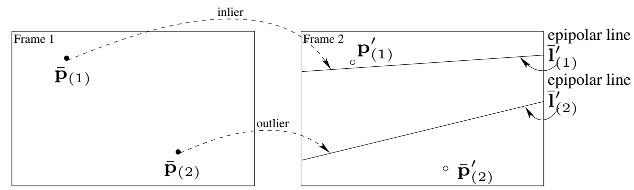

Figure 5. *Epipolar lines and corresponding feature points (inlier and outlier)*

Assuming a uniform probability distribution for the error of point coordinates of outliers the probability distribution of the error of an outlier is (assuming a mismatch in frame 2):

$$p^{(\mathrm{ML})}_{\mathrm{outlier},(i)} := (\frac{1}{\sqrt{2\pi}\sigma})^2 \ \exp(-\frac{d^2_{(i)}}{2\sigma^2})\frac{1}{v^2} \qquad (24)$$

where v is the width (and height) of the window within the outlier may occur. For a correspondence set C with an outlier rate of β the probability density function for a correspondence is:

$$p^{(\mathrm{ML})}_{\mathrm{mix},(i)} := (1-\beta)\, p^{(\mathrm{ML})}_{\mathrm{inlier},(i)} + \beta\, p^{(\mathrm{ML})}_{\mathrm{outlier},(i)} \qquad (25)$$

$p^{(\mathrm{ML})}_{\mathrm{mix},(i)}$ is the probability for the error of a correspondence picked up from a correspondence set with an outlier rate of β. The error model for the correspondence set C has the following probability density distribution:

$$p^{(\mathrm{ML})}_{\mathrm{mix},C}(C|\bar{\mathbf{F}}) := \prod_{i=1}^{N} p^{(\mathrm{ML})}_{\mathrm{mix},(i)} \qquad (26)$$

Since $\bar{\mathbf{F}}$ (6 parameters) and $(\bar{\mathbf{p}}_{(i)},\ \bar{\mathbf{p}}'_{(i)})$ ($4N$ parameters) are unknown this results in an optimization problem with $4N+6$ parameters to be estimated. It is needed to simplify the problem (at the expense of violating the ML-property) to enable a computationally efficient global search.

Therefore, it is assumed that the detected feature points in frame 1 are noise free. In result, the epipolar lines $\bar{\mathbf{l}}'_{(i)}$ in frame 2 obtained by the true F matrix and $\bar{\mathbf{p}}_{(i)}$

$$\bar{\mathbf{l}}_{(i)} := \bar{\mathbf{F}}\bar{\mathbf{p}}_{\mathrm{hom}(i)} \qquad (27)$$

are error free (Fig. 5). So the considered distances are:

$$\begin{aligned} & d_{(i)} = 0 \ \text{ and} \\ & d'_{(i)} \rightarrow d'_{\mathrm{epi},(i)}(\hat{\mathbf{F}}) := \text{distance}(\hat{\mathbf{F}}\mathbf{p}_{\mathrm{hom}(i)}, \mathbf{p}'_{(i)}) \end{aligned} \qquad (28)$$

with $\hat{\mathbf{F}}$ any estimated F matrix. There are 6 free parameters remaining : 3 rotation parameters, 2 translation parameters and 1 parameter for the focal length. The simplified error model has the following probability density distribution for a correspondence:

$$p_{\mathrm{mix},(i)} := \frac{(1-\beta)}{2\pi\sigma^2} \ \exp(-\frac{d'^2_{\mathrm{epi},(i)}(\hat{\mathbf{F}})}{2\sigma^2}) + \frac{\beta}{v^2} \qquad (29)$$

The simplified error model for the correspondence set has the following probability density distribution:

$$p_{\mathrm{mix},C}(C|\hat{\mathbf{F}}) := \prod_{i=1}^{N} p_{\mathrm{mix},(i)} \tag{30}$$

Instead of maximizing $p_{\mathrm{mix},C}(C|\hat{\mathbf{F}})$ the negative logarithmic likelihood $-L$ can be minimized:

$$-L(C|\hat{\mathbf{F}}) := -\ln(p_{\mathrm{mix},C}(C|\hat{\mathbf{F}})) \tag{31}$$

In this approach following cost function has to be minimized by the global optimizer:

$$\xi(C|\hat{\mathbf{F}}) := -L(C|\hat{\mathbf{F}}) \\ = -\sum_{i=1}^{N} \ln\Bigg[\frac{(1-\beta)}{2\pi\sigma_{cf}^2}\ \exp(-\frac{d'^2_{\mathrm{epi},(i)}(\hat{\mathbf{F}})}{2\sigma_{cf}^2}) + \frac{\beta}{v^2}\Bigg] \tag{32}$$

with σ_{cf} the specified variance.

4.2. Repulsive Particle Swarm Optimization

The cost function described is $\xi : \mathbf{U} \subset \mathbb{R}^6 \rightarrow \mathbb{R}$. Unfortunately, this mapping has many local optima making it hard to find the global optimum and so the optimal F matrix. One possibility is to use population based evolutionary global optimization methods to solve this problem. One method, the particle swarm optimization (PSO) was chosen due to its high convergence speed but it proved to be not sufficiently robust for this problem. PSO rather got stuck in local optima than to find the desired global optimum. Therefore, a modified PSO which is called repulsive particle swarm optimization (RPSO) has been developed. In RPSO, there is a population Ψ of solution candidates, called *particles*, having positions $\mathbf{x}$ and velocities $\mathbf{v}$ in the search space. In this case the position $\mathbf{x}^{(r)}$ of a particle at iteration step r is determined by 6 parameters:

$$\mathbf{x}^{(r)} := (\varphi^{(r)}, \vartheta^{(r)}, \rho^{(r)}, \zeta^{(r)}, \eta^{(r)}, f'^{(r)})^\top$$

The velocity $\mathbf{v}^{(r)}$ of a particle determines the next position $\mathbf{x}^{(r+1)}$ after an iteration step:

$$\begin{aligned} \mathbf{v}^{(r)} &:= (\Delta\varphi^{(r)}, \Delta\vartheta^{(r)}, \Delta\rho^{(r)}, \Delta\zeta^{(r)}, \Delta\eta^{(r)}, \Delta f'^{(r)})^\top \\ \mathbf{x}^{(r+1)} &:= \mathbf{x}^{(r)} + \mathbf{v}^{(r)} \end{aligned} \tag{33}$$

Each particle knows its best position $\hat{\mathbf{x}}$ it has achieved so far measured by the cost function. The main difference between PSO and RPSO results from the different definition for the velocity of a particle. In PSO, the equation used to calculate the velocity of a particle for the next iteration is:

$$\begin{aligned} \mathbf{v}^{(r+1)} = \quad & \omega\ \mathbf{v}^{(r)} \\ + \quad & a_{\mathrm{PSO}}\ \chi_1^{(r)}\ (-\mathbf{x}^{(r)} + \hat{\mathbf{x}}^{(r)}) \\ + \quad & b_{\mathrm{PSO}}\ \chi_2^{(r)}\ (-\mathbf{x}^{(r)} + \hat{\mathbf{x}}_\Psi^{(r)}) \end{aligned} \tag{34}$$

In contrast, the equation for the velocity $\mathbf{v}^{(r+1)}$ of a particle in RPSO for the next iteration is:

$$\begin{aligned} \mathbf{v}^{(r+1)} = \quad & \omega\ \mathbf{v}^{(r)} \\ + \quad & a_{\mathrm{RPSO}}\ \chi_1^{(r)}\ (-\mathbf{x}^{(r)} + \hat{\mathbf{x}}^{(r)}) \\ + \quad & b_{\mathrm{RPSO}}\ \chi_2^{(r)}\ \omega(-\mathbf{x}^{(r)} + \hat{\mathbf{y}}^{(r)}) \\ + \quad & c_{\mathrm{RPSO}}\ \chi_3^{(r)}\ \omega\ \mathbf{z}^{(r)} \end{aligned} \tag{35}$$

with

- $\chi_1^{(r)}$, $\chi_2^{(r)}$, $\chi_3^{(r)}$: random numbers $\in [0,\ 1]$
- ω : inertia weight $\in [0.01,\ 0.7]$
- $\hat{\mathbf{x}}^{(r)}$: best position of a particle
- $\hat{\mathbf{y}}^{(r)}$: best position of a randomly chosen other particle from Ψ
- $\mathbf{z}^{(r)}$: a random velocity vector
- $\hat{\mathbf{x}}_\Psi^{(r)}$ the globally best particle within the population
- $a_{\mathrm{PSO}} = b_{\mathrm{PSO}} = 2$
- $a_{\mathrm{RPSO}} = 1.5$, $b_{\mathrm{RPSO}} = -1.5$, $c_{\mathrm{RPSO}} = 0.5$

The term with the a_{RPSO}-scalar on the right side of (35) leads to a motion of the particle towards its best position. The term with the b_{RPSO}-scalar leads to a repulsion between the particle and the best position of a randomly chosen other particle in order to explore new areas in the search space and to prevent the population to get stuck in a local optimum. The term with the c_{RPSO}-scalar generates noise in the velocity of a particle to enhance the exploration to new areas in the search space. The components of the velocity $\mathbf{v}$ and the components of the position $\mathbf{x}$ are constrained to lie within the allowed bounds of corresponding variables. In case of a bound infringement the new value for a component of particle position is calculated by a 'reflection' : e.g.

$$x_1 = 1.2 \notin [0,1] \Rightarrow x_1 \rightarrow 1 - (1.2 - 1) = 0.8 \tag{36}$$

and in case of a bound infringement the new value for a component of the velocity is calcualted by a 'cut off' : e.g.

$$v_1 = 1.2 \notin [0,1] \Rightarrow v_1 \rightarrow 1. \tag{37}$$

RPSO enables the use of more information compared to PSO due to repulsion between particles. In average this leads to a more balanced occupation of the search space through the population where each particle tends to have its own region. The search space is then clustered more homogeneously.

The inertia weight ω decreases in steps of 0.05 from 0.7 to 0.01 when no progress at all is encountered for Δr_ω iteration steps. The search is aborted when a given inlier rate is achieved or a maximum number of iterations is reached.

5. Experimental Results

This approach is tested on synthetic data and on real image pairs.

5.1. Synthetic Data Tests

A virtual camera with rotation angles restricted to ± 0.2 rad and with a maximum focal length variation of 10 % between two following frames is used.
At first a 3D point cloud is randomly constructed. The virtual camera determines the first projection of the 3D points into the camera plane to generate frame 1. After translating, rotating and changing the focal length of the virtual camera frame 2 is generated. The corresponding feature points are determined. A Gaussian noise is added to the feature point coordinates to simulate real world conditions. To simulate outliers the coordinates of some of the correspondences are randomly changed. With $N = 400$ and varying outlier rates the required number of iterations for a 99% successful search is experimentally determined.
The coordinates of the feature points lie within $[-0.5u, 0.5u]$ where u is any length unit. The true variance $\bar{\sigma}^2$ of the Gaussian noise is $\bar{\sigma}^2 = 5 \cdot 10^{-7} u^2$. Assuming that a real camera has 720x576 pixel resolution this is equivalent to a Gaussian error with a standard deviation of $\sigma \sim 0.4$ pel.
It proves that the variance σ^2_{cf} specified in the cost function may be greater than the true variance $\bar{\sigma}^2$ without loosing significant accuracy of the estimated parameters. The reason for specifying a greater variance in the cost function is the resulting search space which enables a faster global convergence. Tab. 1 shows the parameters of the RPSO algorithm depending on the outlier rate.

β	$\sigma^2_{cf}[u^2]$	Δr_ω	size of Ψ	$\tau[u]$
0.5	$11 \cdot 10^{-6}$	25	10	$7 \cdot 10^{-3}$
0.6	$9 \cdot 10^{-6}$	35	10	$7 \cdot 10^{-3}$
0.7	$7 \cdot 10^{-6}$	40	10	$7 \cdot 10^{-3}$
0.8	$5 \cdot 10^{-6}$	55	10	$7 \cdot 10^{-3}$
0.9	$3 \cdot 10^{-6}$	60	25	$5 \cdot 10^{-3}$

Table 1. *Synthetic data tests : Setup of RPSO parameters*

Fig. 6 shows the results of the synthetic data tests where the required number of iterations is plotted.

Compared to RANSAC based estimation the new approach requires less iterations at outlier rates above 70%. At 80% outlier rate the required number of iterations is reduced by a factor of 17. At 90% it is reduced by a factor of 800.

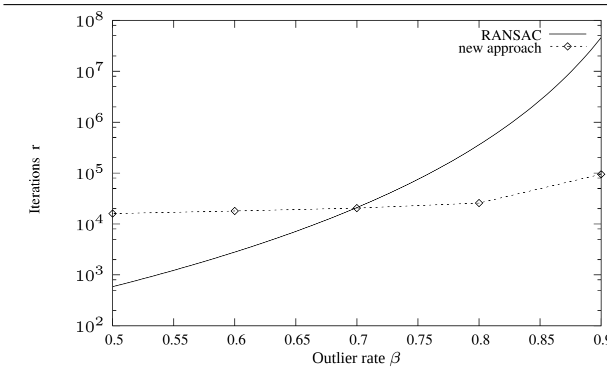

Figure 6. *Mean iterations required for 99% successful search : RANSAC (theoretically derived) vs. RPSO (synthetic data tests)*

5.2. Real Data Tests

The proposed appraoch was tested on two outdoor image pairs. Fig. 7 shows the first image pair named 'leaves' (A). 594 correspondences were detected where RANSAC produced an outlier rate of $\beta = 0.78$. In these images the leaves are moving and make it hard for the correspondence generation step to produce good results.
Fig. 8 shows the second image pair named 'berries' (B) having a static building in the background and a shrub in the foreground. 710 correspondences were detected and RANSAC produced an outlier rate of $\beta = 0.84$. This image pair was even harder for the correspondence generation step due to many moving objects in the foreground.
Both image pairs were processed 800 times. Though RPSO was set up to stop at given inlier rate $1 - \beta$, the algorithm often stopped by reaching a higher inlier rate. Tab. 2 shows the setup parameters of RPSO. In tab. 3 the results of the new approach applied on the two image pairs are given where

- β_{RPSO} is the mean detected outlier rate
- r_{RPSO} is the mean required number of iterations
- P_g is the confidence rate and
- r_{RANSAC} is the theoretically derived mean required number of iterations by RANSAC.

In both cases, the RPSO algorithm is superior to the RANSAC algorithm and the required number of iterations is consistent with the synthetic data experiments.
In order to compare the runtime efficiency the time t needed for 10000 iterations was determined. In tab. 4 the measured

time in seconds for both image pairs and at various numbers of detected correspondences is given. It shows that the runtime efficiency of RPSO is superior to RANSAC by a factor of 8.

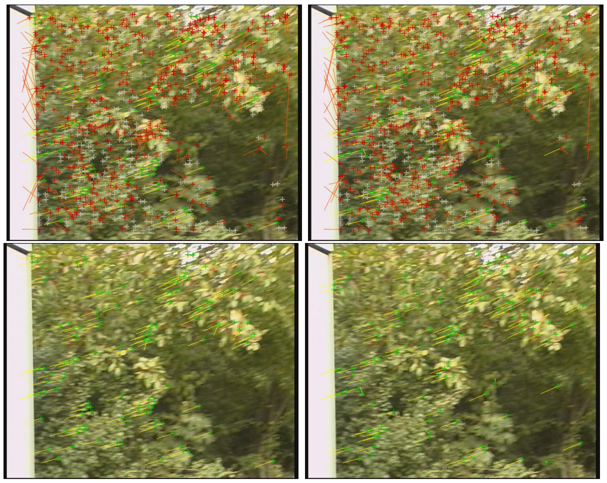

Figure 7. *Detected inliers (yellow) and outliers (red) from image pair 'leaves'. Top left : RANSAC inliers+outliers, top right : RPSO inliers+outliers, bottom left : RANSAC inliers, bottom right : RPSO inliers*

Figure 8. *Detected inliers (yellow) and outliers (red) from image pair 'berries'. Top left : RANSAC inliers+outliers, top right : RPSO inliers+outliers, bottom left : RANSAC inliers, bottom right : RPSO inliers*

	β	$\sigma_{cf}^2[\text{pel}^2]$	Δr_ω	size of Ψ	$\tau[\text{pel}]$
A	0.78	0.9	65	10	0.8
B	0.84	0.7	40	25	0.8

Table 2. *Real data tests: Setup of RPSO Parameters*

	β_{RPSO}	r_{RPSO}	P_g	r_{RANSAC}
A	0.77	$2.89 \cdot 10^4$	0.9963	$2.72 \cdot 10^5$
B	0.83	$6.87 \cdot 10^4$	0.9963	$2.80 \cdot 10^6$

Table 3. *Real data test results: Mean required number of iterations and corresponding confidence rate*

	A (leaves)		B (berries)	
N	694	1193	710	1379
t_{RANSAC} [s]	82	160	94	183
t_{RPSO} [s]	10	20	12	23

Table 4. *Real data tests: Runtime efficiency measurements*

6. Conclusions

A technique for robust estimation of camera rotation, translation and focal length at high outlier rates is presented. The global optimizer in combination with the specified cost function requires less computational expense compared to the RANSAC algorithm. When the runtime efficiency is taken into account, the proposed estimator is superior at outlier rates of 59% and above. Though the proposed technique requires more iterations than the RANSAC algorithm at outlier rates below 70%, the upper bound for the required computational effort of the estimation of camera rotation, translation and focal length is significantly decreased. The real time estimation at high outlier rates can be facilitated. Another advantage of the proposed estimator is that it can be configured to estimate any subset of parameters, e.g. only rotation and one direction translation, so it is possible to benefit from any constraints in the camera parameters.

References

[1] J. Weng, T. S. Huang and N. Ahuja. Motion and Structure from Two Perspective Views : Algorithms, Error analysis and Error Estimation. *IEEE Trans. PAMI*, 11(5), pp. 451-476, 1989.

[2] M.A Fischler and R. C. Bolles. Random sample consensus : A paradigm for model fitting with applications to image analysis

and automated cartography. *Comm. Assoc. Mach.*, 24(6):281-395, 1981.

[3] P.H.S. Torr and A. Zisserman. MLESAC:A New Robust Estimator with Application to Estimating Image Geometry. *Computer Vision and Image Understanding*, (78):138-156, 2000.

[4] O. Chum and J. Matas. Randomized RANSAC with $T_{d,d}$ test. *Proc. of the British Machine Vision Conference*, volume 2, pp. 448-457, 2002.

[5] O. Chum, J. Matas and J. Kittler. Locally Optimized RANSAC. *Proc. DAGM-Symposium*, pp. 236-243, 2003.

[6] B. Tordoff and D. Murray. Guided Sampling and Consensus for Motion Estimation. *Proc. European Conference on Computer Vision, Springer Lecture Notes on Computer Science*, 2350:82-96, 2002.

[7] D. Nister. Preemptive RANSAC for Live Structure from Motion Estimation. *Proc. of the Ninth IEEE International Conference on Computer Vision*, pp. 199-206, 2003.

[8] D. Myatt, P. Torr, S. Nasuto, J. Bishop and R. Craddock. Napsac: High noise, high Dimensional robust estimation - it's in the bag. *BMVC*, volume 2, pp. 458-467, 2002.

[9] H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133-135, September 1981.

[10] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN 0-521-62304-9, 2000.

[11] Q. Luong and O. Faugeras. The fundamental matrix : Theory, algorithms and stability analysis. *International Journal of Computer Vision*, 17(1):43-76, 1996.

[12] R. I. Hartley. Projective reconstruction and invariants from multiple images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(10):1036-1041, October 1994.

[13] J. Kennedy and R. C. Eberhart. Particle swarm optimization, *Proc. of the Fourth IEEE International Conference on Neural Networks*, pp. 1942-1948, 1995.

[14] Y. Shi and R. Eberhart. Parameter selection in particle swarm optimization. *Evolutionary Programming VII:Proc. EP98*, New York: Springer Verlag, pp. 591-600, 1998.