Closed-form solutions to the minimal absolute pose problems with known vertical direction

Zuzana Kukelova[1], Martin Bujnak[2], Tomas Pajdla[1]

[1] Center for Machine Perception, Czech Technical University in Prague
[2] Bzovicka 24, 85107, Bratislava, Slovakia

Abstract. In this paper we provide new simple closed-form solutions to two minimal absolute pose problems for the case of known vertical direction. In the first problem we estimate absolute pose of a calibrated camera from two 2D-3D correspondences and a given vertical direction. In the second problem we assume camera with unknown focal length and radial distortion and estimate its pose together with the focal length and the radial distortion from three 2D-3D correspondences and a given vertical direction. The vertical direction can be obtained either by direct physical measurement by, e.g., gyroscopes and inertial measurement units or from vanishing points constructed in images. Both our problems result in solving one polynomial equation of degree two in one variable and one, respectively two, systems of linear equations and can be efficiently solved in a closed-form. By evaluating our algorithms on synthetic and real data we demonstrate that both our solutions are fast, efficient and numerically stabled [3].

1 Introduction

Cheap consumer cameras become precise enough to be used in many computer vision applications, e.g. structure from motion [2, 22, 23, 15], or recognition [13, 14]. The good precision of these cameras allows to simplify camera models and therefore also algorithms used in these applications. For instance, for the pinhole camera model it is common to set the camera skew to zero, the pixel aspect ratio to one and the principal point to the center of the image. Moreover, since most of the digital cameras put the information about the focal length into the image header (EXIF), it is frequently assumed that this is a good approximation of the whole internal calibration of the camera and therefore the camera is considered calibrated. These assumptions allow to use simpler algorithms, e.g. the 5-point relative pose algorithm for calibrated cameras [19, 24], and as it was shown in [22, 23, 16, 12, 15] they work well even when these assumptions are not completely satisfied.

More and more cameras become also equipped with inertial measurement units (IMUs) like gyroscopes and accelerometers, compasses, or GPS devices

[3] This work has been supported by EC project FP7-SPACE-218814 PRoVisG and by Czech Government under the research program MSM6840770038

and add information from these devices into the image header too. This could be even more observable in modern cellular phones and other smart devices.

Unfortunately, since GPS position precision is in order of meters, it is not sufficient to provide external camera calibration, i.e. the absolute position of the camera. Neither compass accuracy (which could provide cameras yaw) is good enough to provide camera orientation since its sensor is highly influenced by magnetic field disturbances.

On the other hand currently available IMUs (even the low cost ones) provide very accurate roll and pitch angle, i.e. the vertical direction. The angular accuracy of roll and pith angle in low cost IMUs like [26] is about $0.5°$, and in high accuracy IMUs is less than $0.02°$. While most today cameras use information from gyroscopes and IMUs only to distinguish whether the image orientation is landscape or portrait, we show that knowing the camera "up vector" can radically simplify the camera absolute pose problem.

Determining the absolute pose of a camera given a set of n correspondences between 3D points and their 2D projections is known as the Perspective-n-Point (PnP) problem. This problem is one of the oldest problems in computer vision with a broad range of applications in structure from motion [2, 15] or recognition [13, 14].

PnP problems for fully calibrated cameras for three and more than three points have been extensively studied in the literature. One of the oldest papers on this topic dates already to 1841 [7]. Recently a large number of solutions to the calibrated PnP problems have been published [5, 18, 20, 21] including several solutions to the PnP problems for partially calibrated cameras [1, 3, 9, 25]. These solutions assume unknown focal length [1, 3, 25], unknown focal length and radial distortion [9] or unknown focal length and principal point [25].

In this paper we provide new simple closed-form solutions to two minimal absolute pose problems for cameras with known rotation around two angles, i.e. known vertical direction. In the first problem we estimate the absolute pose of a calibrated camera from two 2D-3D correspondences and a given vertical direction. In the second problem we assume camera with unknown focal length and radial distortion and estimate its pose together with the focal length and the radial distortion from three 2D-3D correspondence and a given vertical direction. In both cases this is the minimal number of correspondences needed to solve these problems.

Both these problems result in solving one polynomial equation of degree two in one variable and one respectively two systems of linear equations and can be efficiently solved in a closed-form way. By evaluating our algorithms on synthetic and real data we show that both our solutions are fast, efficient and numerically stable. Moreover, both presented algorithms are very useful in real applications like structure from motion, surveillance camera calibration or applications of registering photographs taken by a camera mounted on a car moving on a plane.

Our work builds on recent results from [10] where the efficient algorithm for relative pose estimation of a camera with known vertical direction from three point correspondences was presented. It was demonstrated that in the presence of good vertical direction information this algorithm is more accurate than the classical 5-point relative pose algorithm [19, 24] for calibrated cameras.

Next we provide the formulation of our two problems and show how they can be solved

2 Problem formulation

In this section we formulate two problems of determining absolute pose of a camera given 2D-3D correspondences and the vertical direction, respectively rotation angles around two axes.

Problem 1. Given the rotation of the calibrated camera around two axes and the images of two known 3D reference points, estimate the absolute position of the camera and the rotation of the camera around the third axis (y-axis).

Problem 2. Given the rotation of the camera with unknown focal length around two axes and the distorted images of three known 3D reference points, estimate the absolute position of the camera, the rotation of the camera around the third axis (y-axis), the unknown focal length and the parameter of radial distortion.

In both problems we use the standard pinhole camera model [8]. In this model the image projection $\mathbf{u}$ of a 3D reference point $\mathbf{X}$ can be written as

$$\lambda \, \mathbf{u} = \mathtt{P} \, \mathbf{X}, \tag{1}$$

where $\mathtt{P}$ is a 3×4 projection matrix, λ is an unknown scalar value and points $\mathbf{u} = [u, v, 1]^\top$ and $\mathbf{X} = [X, Y, Z, 1]^\top$ are represented by their homogeneous coordinates.

The projection matrix $\mathtt{P}$ can be written as

$$\mathtt{P} = \mathtt{K} \, [\mathtt{R} \mid \mathbf{t}], \tag{2}$$

where $\mathtt{R}$ is a 3×3 rotation matrix, $\mathbf{t} = [t_x, t_y, t_z]^\top$ contains the information about camera position and $\mathtt{K}$ is the 3×3 calibration matrix of the camera.

In our problems we assume that we know the vertical direction, i.e. the coordinates of the world vector $[0, 1, 0]^\top$ in the camera coordinate system. This vector can be often obtained from the vanishing point or directly from the IMU by a very simple calibration of the IMU w.r.t. the world coordinate system. It is important that all what has to be known to calibrate the IMU w.r.t. the world coordinate system, is the direction of the "up-vector", i.e. the direction which is perpendicular to the "ground plane", i.e. the plane $z = 0$ in the world coordinate system. This direction is easy to calibrate by reseting the IMU when lied down on the "ground plane". Note that this "ground plane" need not be horizontal. Moreover, very often IMU's are precalibrated to use the gravity vector as the up-vector. Then, the IMU is calibrated when the plane $z = 0$ of the world coordinate system corresponds to the ground plane, i.e. if the coordinates $z = 0$ are assigned to the points on the ground plane.

From the "up-vector" returned by the IMU rotation we can compute the rotation of the camera around two axes, in this case the x-axis (ϕ_x) and the z-axis (ϕ_z). Note that IMU sometimes returns directly these two angles. We will denote this rotation as

$$\mathtt{R}_v = \begin{bmatrix} \cos\phi_z & -\sin\phi_z & 0 \\ \sin\phi_z & \cos\phi_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_x & -\sin\phi_x \\ 0 & \sin\phi_x & \cos\phi_x \end{bmatrix}. \tag{3}$$

Therefore, the only unknown parameter in the camera rotation matrix $\mathtt{R}$ is the rotation angle ϕ_y around the y-axis, i.e. the vertical axis and we can write

$$\mathtt{R} = \mathtt{R}(\phi_y) = \mathtt{R_v}\,\mathtt{R_y}(\phi_y), \tag{4}$$

where $\mathtt{R_v}$ is the known rotation matrix (3) around the vertical axis and $\mathtt{R_y}(\phi_y)$ is the unknown rotation matrix around the y-axis of the form

$$\mathtt{R}_y = \begin{bmatrix} \cos\phi_y & 0 & -\sin\phi_y \\ 0 & 1 & 0 \\ \sin\phi_y & 0 & \cos\phi_y \end{bmatrix}. \tag{5}$$

With this parametrization the projection equation (1) has the form

$$\lambda\mathbf{u} = \mathtt{K}\,[\mathtt{R}(\phi_y)\;|\,\mathbf{t}]\mathbf{X} = \mathtt{K}\,[\mathtt{R_v}\,\mathtt{R_y}(\phi_y)\;|\,\mathbf{t}]\mathbf{X}, \tag{6}$$

For our problems we "simplify" the projection equation (6) by eliminating the scalar value λ and trigonometric functions sin and cos.

To eliminate sin and cos we use the substitution $q = \tan\frac{\phi_y}{2}$ for which it holds that $\cos\phi_y = \frac{1-q^2}{1+q^2}$ and $\sin\phi_y = \frac{2q}{1+q^2}$. Therefore we can write

$$\left(1+q^2\right)\mathtt{R}_y(q) = \begin{bmatrix} 1-q^2 & 0 & -2q \\ 0 & 1+q^2 & 0 \\ 2q & 0 & 1-q^2 \end{bmatrix}. \tag{7}$$

The scalar value λ from the projection equation (6) can be eliminated by multiplying equation (6) with the skew symmetric matrix $[\mathbf{u}]_\times$. Since $[\mathbf{u}]_\times\,\mathbf{u} = 0$ we obtain the matrix equation

$$[\mathbf{u}]_\times\,\mathtt{K}\,[\mathtt{R_v}\,\mathtt{R_y}(q)\;|\,\mathbf{t}]\mathbf{X} = 0. \tag{8}$$

This matrix equation results in three polynomial equations from which only two are linearly independent. This is caused by the fact that the skew symmetric matrix $[\mathbf{u}]_\times$ has rank two.

Polynomial equations (8) are the basic equations which we use for solving our two problems. Next we describe how these equations look for our two problems and how they can be solved.

3 Absolute pose of a calibrated camera with known up direction

In the case of Problem 1 for calibrated camera the calibration matrix $\mathtt{K}$ is known and therefore the projection equation (8) $[\mathbf{u}]_\times \mathtt{K}\,[\mathtt{R_v}\,\mathtt{R_y}\,(q)\mid \mathbf{t}]\mathbf{X} = 0$ has the form

$$\begin{bmatrix} 0 & -1 & v \\ 1 & 0 & -u \\ -v & u & 0 \end{bmatrix} \begin{bmatrix} r_{11}\,(q) & r_{12}\,(q) & r_{13}\,(q) & t_x \\ r_{21}\,(q) & r_{22}\,(q) & r_{23}\,(q) & t_y \\ r_{31}\,(q) & r_{32}\,(q) & r_{33}\,(q) & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = 0, \tag{9}$$

where $\mathbf{u}$ is a calibrated image point and $r_{ij}\,(q)$ are the elements of the rotation matrix $\mathtt{R} = \mathtt{R_v}\,\mathtt{R_y}\,(q)$. Note that these elements are quadratic polynomials in q.

In this case we have four unknowns $t_x, t_y, t_z, q = \tan\frac{\phi_y}{2}$ and since each 2D-3D point correspondence gives us two constraints, two independent polynomial equations of the form (9), the minimal number of point correspondences needed to solve this problem is two.

Matrix equation (9) gives us two linearly independent equations for each point and each from this equations contains monomials $q^2, q, t_x, t_y, t_z, 1$. Since we have two points, we have four of these equations and therefore we can use three from them to eliminate t_x, t_y and t_z from the fourth one. In this way we obtain one polynomial in one variable q of degree two. Such equation can be easily solved in a closed-form and generally results in two solutions for $q = \tan\frac{\phi_y}{2}$. Backsubstituting these two solutions to the remaining three equations give us three linear equations in three variables t_x, t_y and t_z. By solving these linear equations we obtain solutions to our problem.

4 Absolute pose of a camera with unknown focal length and radial distortion and known camera up direction

In Problem 2 we assume that the skew in the calibration matrix $\mathtt{K}$ is zero, the aspect ratio is one and the principal point is in the center of the image. Modern cameras are close to these assumptions and as it will be shown in experiments these assumptions give good results even when they are not exactly satisfied.

The only parameter from $\mathtt{K}$ which cannot be safely set to a known value is the focal length f. Here we assume that the focal length f is unknown and therefore the calibration matrix $\mathtt{K}$ has the form $diag\,[f, f, 1]$. Since the projection matrix is given only up to a scale we can equivalently write $\mathtt{K} = [1, 1, w]$ for $w = 1/f$.

For this problem we further assume that the image points are affected by some amount of radial distortion. Here we model this radial distortion by the one-parameter division model proposed in [6]. This model is given by formula

$$\mathbf{p}_u \sim \mathbf{p}_d / \left(1 + k r_d^2\right), \tag{10}$$

where k is the distortion parameter, $\mathbf{p}_u = [u_u, v_u, 1]^\top$, resp. $\mathbf{p}_d = [u_d, v_d, 1]^\top$, are the corresponding undistorted, resp. distorted, image points, and r_d is the radius of $\mathbf{p}_d$ w.r.t. the distortion center. We assume that the distortion center is in the center of the image and therefore $r_d^2 = u_d^2 + v_d^2$.

Since the projection equation (8) contains undistorted image points $\mathbf{u}$ but we measure distorted ones we need to put the relation (10) into equation (8).

Let $\hat{\mathbf{u}}_i = [\hat{u}_i, \hat{v}_i, 1]$ be the i^{th} measured distorted image point, projection of the 3D point $\mathbf{X}_i = [X_i, Y_i, Z_i, 1]$, and let $\hat{r}_i^2 = \hat{u}_i^2 + \hat{v}_i^2$. Then $\mathbf{u}_i = \left[\hat{u}_i, \hat{v}_i, 1 + k\hat{r}_i^2\right]^\top$ and the projection equation (8) for the i^{th} point has the form

$$\begin{bmatrix} 0 & -1-k\hat{r}_i^2 & \hat{v}_i \\ 1+k\hat{r}_i^2 & 0 & -\hat{u}_i \\ -\hat{v}_i & \hat{u}_i & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & w \end{bmatrix} \begin{bmatrix} r_{11}(q) & r_{12}(q) & r_{13}(q) & t_x \\ r_{21}(q) & r_{22}(q) & r_{23}(q) & t_y \\ r_{31}(q) & r_{32}(q) & r_{33}(q) & t_z \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} = 0, \quad (11)$$

where again $r_{ij}(q)$ are the elements of the rotation matrix $\mathtt{R} = \mathtt{R_v}\,\mathtt{R_y}(q)$.

In this case we have six unknowns $t_x, t_y, t_z, q = \tan\frac{\phi_y}{2}, k, w = 1/f$ and since each 2D-3D point correspondence gives us two constraints the minimal number of point correspondences needed to solve this problem is three.

For three point correspondences the matrix equation (11) gives us nine polynomial equations from which only six are linearly independent. Let denote these nine polynomial equations $g_{1i} = 0, g_{2i} = 0$ and $g_{3i} = 0$ for $i = 1, 2, 3$, where $g_{1i} = 0$ is the equation corresponding to the first row in the matrix equation (11), $g_{2i} = 0$ to the second and $g_{3i} = 0$ to the third row.

Now consider the third polynomial equation from the matrix equation (11), i.e. the equation $g_{3i} = 0$. This equation has the form

$$g_{3i} = c_1\, q^2 + c_2\, q + c_3\, t_x + c_4\, t_y + c_5 = 0, \quad (12)$$

where c_j, $j = 1, \ldots, 5$ are known coefficients. The polynomial equation (12) doesn't contain variables t_z, k and w. Since we have three 2D-3D correspondences we have three equations $g_{3i} = 0$, $i = 1, 2, 3$. Therefore, we can use two from these equations to eliminate t_x and t_y from the third one, e.g. from the $g_{31} = 0$.

In this way we obtain one polynomial equation in one variable q of degree two. Such equation can be easily solved in a closed-form and generally results in two solutions for $q = \tan\frac{\phi_y}{2}$. Backsubstituting these two solutions to the remaining two equations corresponding to the third row of (11), i.e. $g_{32} = 0$ and $g_{33} = 0$, gives us two linear equations in two variables t_x and t_y which can be again easily solved.

After obtaining solutions to q, t_x and t_y we can substitute them to the equations corresponding to the first row of (11). These equations have after the substitution the form

$$g_{1i} = c_6\, t_z\, w + c_7\, w + c_8\, k + c_9 = 0 \quad (13)$$

for $i = i, \ldots, 3$ and known coefficients c_j, $j = 6, \ldots, 9$. Equations 13 can be again solved in a linear way. This is because we can consider the monomial $t_z\, w$ as a new variable and solve a linear system of three equations in three variables. Solving this linear system together with solutions to q, t_x and t_y gives us in general two solutions to our absolute pose problem for camera with unknown focal length and radial distortion given camera rotation around two axis.

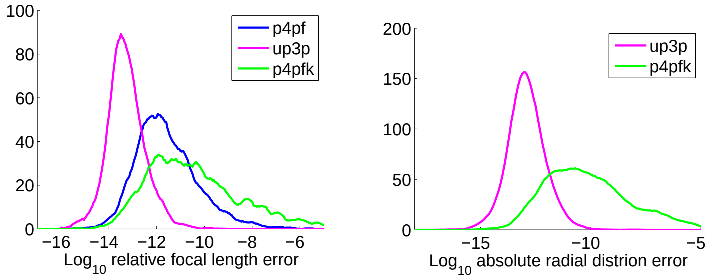

Fig. 1. Log_{10} relative error of the focal length f (Left) and Log_{10} absolute error of the radial distortion parameter k (Right) obtained by selecting the real root closest to the ground truth value $f_{gt} = 1.5$ and $k_{gt} = -0.2$

5 Experiments

We have tested both our algorithms on synthetic data (with various levels of noise, outliers, focal lengths and radial distortions and different angular deviation of the vertical direction) and on real datasets and compared them with the P3P algorithm for calibrated camera presented in [5], the P4P algorithm for camera with unknown focal length from [3] and with the P4P algorithm for camera with unknown focal length and radial distortion [9].

In all experiments we denote our new algorithms as up2p and up3p+f+k, and compared algorithms as p3p [5], p4p+f [3] and p4p+f+k [9].

5.1 Synthetic data set

We initially studied our algorithms on synthetically generated ground-truth 3D scenes. These scenes were generated randomly with the Gaussian distributions in a 3D cube. Each 3D point was projected by a camera, where the camera orientation and position were selected randomly but looking on the scene. Then the radial distortion using the division model [6] was added to all image points to generate noiseless distorted points. Finally, Gaussian noise with standard deviation σ was added to the distorted image points assuming a 1000×1000 pixel image.

Noise free data set In the first synthetic experiment, we have studied the numerical stability of both proposed solvers on exact measurements and compared them with the algorithms [5, 3, 9].

In this experiment 1000 random scenes and camera poses were generated. The radial distortion parameter was set to $k_{gt} = -0.2$ and the focal length to $f_{gt} = 1.5$. In the case of our calibrated up2p algorithm, the p3p algorithm from [5] and the p4p+f algorithm [3] these values were used to precalibrate cameras.

The rotation error was measured as the rotation angle in the angle-axis representation of the relative rotation RR_{gt}^{-1} and the translation error as the angle

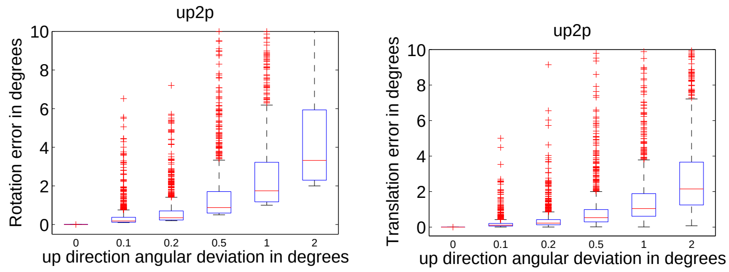

Fig. 2. The influence of the accuracy of the vertical direction on the estimation of rotation (Left) and translation (Right) for our calibrated up2p algorithm.

between ground-truth and estimated translation vector. In this case the rotation and translation errors for both our algorithms were under the machine precision and therefore we were not able to properly display them here.

Figure 1 (Left) shows the log_{10} relative error of the focal length f obtained by selecting the real root closest to the ground truth value $f_{gt} = 1.5$. The log_{10} absolute error of the radial distortion parameter is in Figure 1 (Right).

It can be seen that the new proposed up3p+f+k solver for camera with unknown focal length and radial distortion and given vertical direction (Magenta) provides more accurate estimates of both k and f then the p4p+f+k solver presented in [9] (Green) and the p4p+f solver from [3] (Blue).

Vertical direction angular deviation In the real applications the vertical direction obtained either by direct physical measurement by, e.g., gyroscopes and IMUs or from vanishing points constructed in images is not accurate. The angular accuracy of roll and pith angle in low cost IMUs like [26] is about 0.5°, and in high accuracy IMUs is less than 0.02°.

Therefore, in the next experiment we have tested the influence of the accuracy of the vertical direction on the estimation of rotation (Figure 2 (Left) and 3 (Top left)), translation (Figure 2 (Right) and 3 (Top right)), focal length (Figure 3 (Bottom left)) and radial distortion (Figure 3 (Bottom right)). In this case the ground truth focal length was $f_{gt} = 1.5$ and the radial distortion $k_{gt} = -0.2$.

For each vertical direction angular deviation 1000 estimates of both our algorithms, the calibrated up2p (Figure 2) and the up3p+f+k with unknown focal length and radial distortion (Figure 3), were made. The vertical direction angular deviation varied from 0° to 2°. All results in Figures 2 and 3 are represented by the MATLAB function boxplot which shows values 25% to 75% quantile as a blue box with red horizontal line at median. The red crosses show data beyond 1.5 times the interquartile range.

The rotation error is again measured as the rotation angle in the angle-axis representation of the relative rotation RR_{gt}^{-1} and the translation error as the angle between ground-truth and estimated translation vector. Focal length and radial distortion figures show directly estimated values.

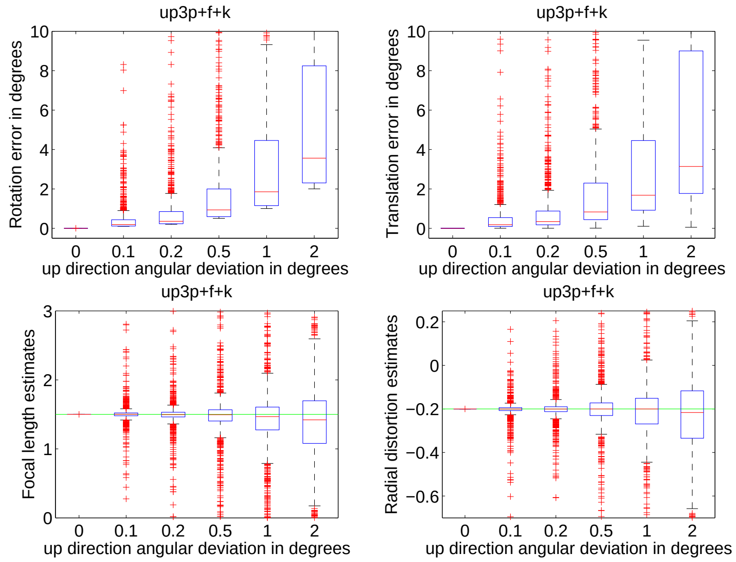

Fig. 3. The influence of the accuracy of the vertical direction on the estimation of rotation (Top left), translation (Top right), focal length (Bottom right), radial distortion (Bottom right) for our up3p+f+k algorithm.

It can be seen that even for relatively high error in the vertical direction the median values of estimated focal lengths and radial distortion parameters are very close to the ground truth values (Green horizontal line). Note that the angular rotation and translation errors cannot be smaller than the angular error of the vertical direction.

Data affected by noise Boxplots in Figure 4 show behavior of both our solvers, the calibrated up2p solver (Cyan) and up3p+f+k solver (Magenta), together with the p3p algorithm [5] (Red), the p4p+f algorithm from [3] (Blue) and the p4p algorithm [9] (Green) in the presence of noise added to image points.

In this experiment for each noise level, from 0.0 to 2 pixels, 1000 estimates for random scenes and camera positions and $f_{gt} = 1.5$, $k_{gt} = -0.2$, were made.

Here it can be seen that since our algorithms use less points than [5, 3, 9] they are little bit more sensitive to a noise added to the image points than the algorithms [5, 3, 9]. The difference is only in the rotation error (Figure 4 (Top left)) where both our new solvers outperform existing algorithms [5, 3, 9]. This might be due to the fact, that the rotation axis is fixed in both our algorithms. However, our new algorithms still provide good estimates also for translation (Figure 4 (Top right)), focal length (Bottom left) and radial distortion (Bottom right) in the presence of noise.

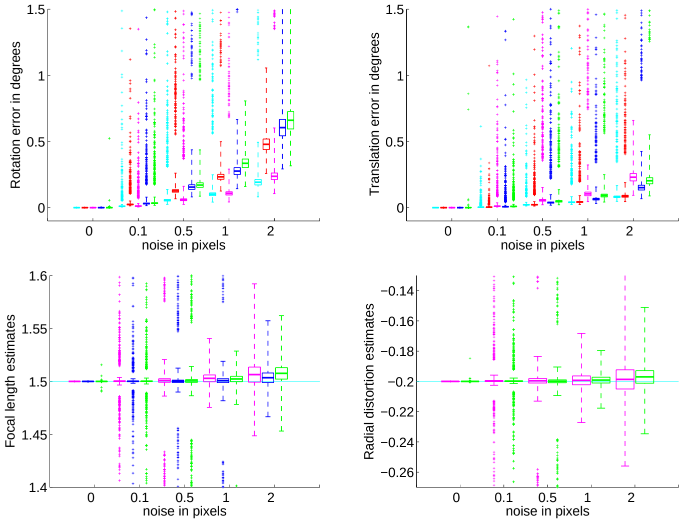

Fig. 4. Error of rotation (Top left), translation (Top right), focal length estimates (Bottom left) and radial distortion estimates (Bottom right) in the presence of noise for our calibrated up2p algorithm (Cyan), the calibrated p3p algorithm from [5] (Red), our p3p+f+k algorithm (Magenta), the p4p+f algorithm presented in [3] (Blue) and the p4p+f+k algorithm [9] (Green)

Ransac experiment The final synthetic experiment shows behaviour of our algorithms within the RANSAC paradigm [5]. Figure 5 shows the mean value of the number of inliers over 1000 runs of RANSAC as a function of the number of samples of the RANSAC. We again compare our calibrated up2p algorithm (Cyan) with the p3p algorithm for calibrated camera presented in [5] (Red), and our up3p+f+k algorithm for camera with unknown focal length and radial distortion (Magenta) with the p4p+f algorithm from [3] (Blue) and the p4p+f+k algorithm [9] (Green). Figure 5 (Left) shows results for 50% outliers, pixel noise $0.5px$, $f_{gt} = 1.5$, $k_{gt} = -0.2$ and the vertical direction angular deviation $0.02°$ and Figure 5 (Right) results for the same configuration but the vertical direction angular deviation $1°$. These vertical direction angle deviations reflect accessible precisions using low cost and high cost IMUs. Vertical direction precision about $1°$ can be received by taking pictures from hand using standard smartphone. As it can be seen both our new algorithms little bit outperform the remaining algorithms [5, 3, 9].

Computation times Since both our problems result in solving one polynomial equation of degree two in one variable and one, respectively two, systems of linear equations and can be solved in a closed-form, they are extremely fast. The Matlab implementation of both our solvers runs about $0.1ms$. For comparison, our Matlab implementation of the p3p algorithm [5] runs about $0.6ms$, the

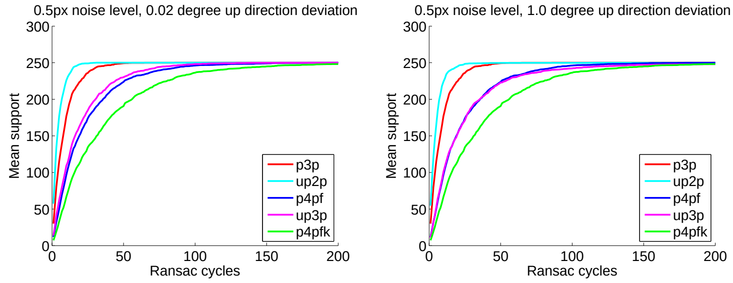

Fig. 5. The mean value of the number of inliers over 1000 runs of RANSAC as a function of the number of samples of the RANSAC.

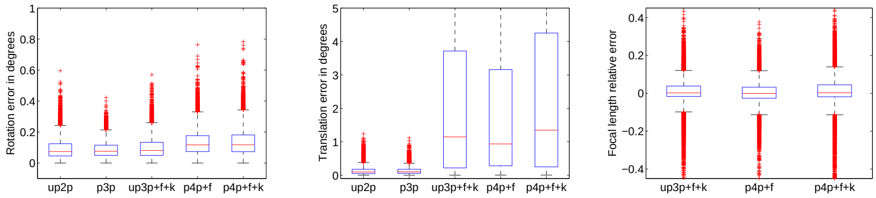

Fig. 6. Results of real experiment for vertical direction randomly rotated with standard deviation 0.5°.

Matlab implementation of the p4p+f algorithm [3] downloaded from [17] about $2ms$ and the original implementation of the p4p+f+k algorithm [9] even about $70ms$.

5.2 Real data

Synthetically generated vertical direction In the real experiment we created 3D reconstruction from a set of images captured by an off-the-shelf camera. We used reconstruction pipeline similar to one described in [23]. We take this reconstruction as a ground truth reference for further comparison even if it might not be perfect. To simulate a measurement from gyroscope we have extracted "vertical direction" from reconstructed cameras and randomly rotated them by a certain angle. This way we have created 1000 random vertical direction with normal distribution for maximal deviation of 0.02, 0.05 and 0.5 and 1 degrees to simulate industry quality, standard and low cost gyroscope measurement. Note that from the 3D reconstruction we have a set of 2D-3D tentative correspondences as well as correct correspondences. To make registration more complicated we added more outliers to the set of tentative correspondences for those cameras where fraction of correct features was greater than 50%. We randomly connected 2D measurements and 3D points to make 50% of tentative correspondences wrong (outliers). Next, for each random vertical direction we executed locally optimized RANSAC [4] with both our new algorithms. Further we also

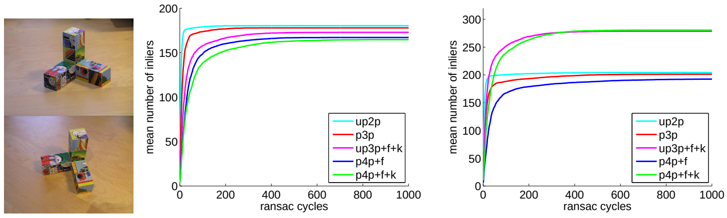

Fig. 7. Real experiment: Example of input images (Left). Results from RANSAC on the image with small radial distortion (Center) and on the image with significant radial distortion (Right)

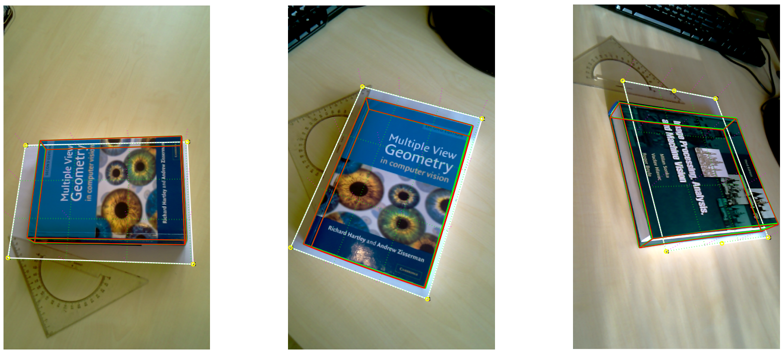

Fig. 8. Results of real experiment for real IMU sensor from HTC HD2.

evaluated 1000 runs of locally optimized ransacs with the calibrated p3p algorithm [5], the p4p+f algorithm presented in [3] and the p4p+f+k solver from [9] and compared all results with ground truth rotation, translation and the focal length. Note that our calibrated up2p algorithm and the p3p algorithm for calibrated camera from [5] require internal camera calibration. Hence we used calibration obtained from the 3D reconstruction.

Results for vertical direction randomly rotated with standard deviation $0.5°$ are in Figure 6. It can be seen that our new algorithms (up2p and up3p+f+k) provide comparable estimates of rotation (Left), translation (Center) and focal length (Right) as algorithms p3p [5], p4p+f [3] and p4p+f+k [9] even for this higher vertical direction deviation. Moreover, our algorithms use less points, give only up to two solutions which have to be verified inside ransac loop and are considerably faster than the algorithms [5, 3, 9]. All these aspects are very important in RANSAC and real applications. The mean value of the number of inliers over 1000 runs of RANSAC as a function of the number of samples of the RANSAC for two cameras and vertical direction deviation $0.5°$ can be seen in Figure 7. It can be seen that both our algorithms outperform existing algorithms [5, 3, 9].

Vertical direction obtained from real IMU In this experiment we used HTC HD2 phone device to obtain real images with IMU data. We took by hand several images of a book placed on a A4 paper. The image size was 800×600 pixels. We assigned 3D coordinates to the paper corners and manually clicked corresponding 2D points in all input images. First, RANSAC with the new solvers was used. Then, a non-linear refinement [8] optimizing focal length, radial distortion, rotation and translation was performed. For our calibrated up2p algorithm we used radial distortion and focal length calculated using non-linearly improved result obtained from up3p+f+k.

Figure 8 shows 3D model of the paper and book back projected to images using calculated camera poses. Green wire-frame model of the book shows estimated camera from minimal sample using our algorithms (up2p (Left), up3p+f+k (Center+Right)), red color represent model after non-linear refinement. Magenta lines represent gravity vector obtained from phone device. As it can be seen we have obtained quite precise results (reprojection error about 1.5px) even for standard images taken by hand and without calibration of relative position of IMU sensor and camera inside the phone.

6 Conclusion

In this paper we have presented new closed-form solutions to two minimal absolute pose problems for the case of known vertical direction, i.e. known rotation of the camera around two axes. In the first problem we estimate absolute pose of a calibrated camera and in the second of a camera with unknown focal length and radial distortion.

We show that information about the vertical direction obtained by direct physical measurement by, e.g. gyroscopes and inertial measurement units (IMUs) or from vanishing points constructed in images, can radically simplify the camera absolute pose problem. In the case of two presented problems the information about the vertical direction leads to very fast, efficient and numerically stable closed-form solutions. For comparison the only existing algorithm for estimating absolute pose of a camera with unknown focal length and radial distortion [9] without the vertical direction leads to LU decomposition of 1134×720 matrix and further eigenvalue computations and is very impractical in real applications.

Experiments show that both our algorithms are very useful in real application like structure from motion. Moreover, since still more and more cameras and smart devices are equipped with IMUs and can save information about the vertical direction, i.e. roll and pitch angles, to the image header, we see the great future potential of presented solvers.

References

1. M. a. Abidi and T. Chandra. A new efficient and direct solution for pose estimation using quadrangular targets: Algorithm and evaluation. *IEEE PAMI*, 17(5):534–538, 1995.

2. 2D3. Boujou. www.2d3.com.
3. M. Bujnak, Z. Kukelova, T. Pajdla, A general solution to the P4P problem for camera with unknown focal length. *CVPR 2008*.
4. O. Chum, J. Matas, and J. Kittler. Locally optimized RANSAC. In *Proc DAGM*, pages 236–243, 2003.
5. M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–95, 1981.
6. A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. *CVPR 2001*, pp. 125–132.
7. J. A. Grunert. Das pothenot'sche problem, in erweiterter gestalt; nebst bemerkungen über seine anwendung in der geodäsie. *Archiv der Mathematik und Physik*, 1:238248, 1841.
8. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
9. K. Josephson, M. Byröd, and K. Åström. Pose Estimation with Radial Distortion and Unknown Focal Length], *Proc. Conference on Computer Vision and Pattern Recognition (CVPR'09)*, Florida, USA, 2009.
10. M. Kalantari, A. Hashemi, F. Jung and Jean-Pierre Guédon. A New Solution to the Relative Orientation Problem using only 3 Points and the Vertical Direction, *Computing Research Repository (CoRR)*, abs/0905.3964, 2009.
11. J. Kelly and G. S. Sukhatme Fast Relative Pose Calibration for Visual and Inertial Sensors *Springer Tracts in Advanced Robotics*, Volume 54, 2009.
12. X. Li, C. Wu, C. Zach, S. Lazebnik and J.-M. Frahm. Modeling and Recognition of Landmark Image Collections Using Iconic Scene Graphs. *ECCV 2008*.
13. B. Leibe, N. Cornelis, K. Cornelis, and L. J. V. Gool. Dynamic 3d scene analysis from a moving vehicle. In *CVPR*, 2007.
14. B. Leibe, K. Schindler, and L. J. V. Gool. Coupled detection and trajectory estimation for multi-object. In *ICCV*, 2007.
15. D. Martinec and T. Pajdla. Robust rotation and translation estimation in multiview reconstruction. In *CVPR*, 2007.
16. Microsoft Photosynth. http://photosynth.net/
17. Minimal Problems in Computer Vision http://cmp.felk.cvut.cz/minimal/
18. F.Moreno-Noguer, V.Lepetit, and P.Fua. Accurate non-iterative o(n) solution to the pnp problems. In *ICCV*, 2007.
19. D. Nister. An efficient solution to the five-point relative pose. *IEEE PAMI*, 26(6):756–770, 2004.
20. L. Quan and Z.-D. Lan. Linear n-point camera pose determination. *IEEE PAMI*, 21(8):774–780, August 1999.
21. G. Reid, J. Tang, and L. Zhi. A complete symbolic-numeric linear method for camera pose determination. In *ISSAC 2003*, pages 215–223, NY, USA, 2003. ACM.
22. N. Snavely, S. M. Seitz, R. Szeliski, "Photo tourism: Exploring photo collections in 3D," *ACM Transactions on Graphics (SIGGRAPH Proceedings)*, 25(3), 2006.
23. N. Snavely, S. M. Seitz, R. Szeliski, "Modeling the world from Internet photo collections," *International Journal of Computer Vision*, Vol 80, Issue 2, pp 189-210, 2008.
24. H. Stewenius, C. Engels, and D. Nister. Recent developments on direct relative orientation. *ISPRS J. of Photogrammetry and Remote Sensing*, 60:284–294, 2006.
25. B. Triggs. Camera pose and calibration from 4 or 5 known 3d points. *In Proc. 7th Int. Conf. on Computer Vision*, Kerkyra, Greece, pages 278284. IEEE Computer Society Press, 1999.
26. Xsens. www.xsens.com.