http://dx.doi.org/10.1109/ICRA.2011.5980380

PTZ Camera Pose Estimation by Tracking a 3D Target

Stefan Hrabar, Peter Corke and Volker Hilsenstein

Abstract— **We present a technique for estimating the 6 DOF pose of a PTZ camera by tracking a single moving target in the image with known 3D position. This is useful in situations where it is not practical to measure the camera pose directly. Our application domain is estimating the pose of a PTZ camera so that it can be used for automated tracking and filming of UAV flight trials. We present results which show the technique is able to localize a PTZ after a short vision-tracked flight, and that the estimated pose is sufficiently accurate for the PTZ to then actively track a UAV based on position data.**

Index Terms— **Visual tracking, PTZ camera, camera pose estimation, numerical optimization**

I. INTRODUCTION

It is often necessary to know the 6 DOF pose of a Pan-Tilt-Zoom (PTZ) camera, however in certain applications it is not practical to measure the pose directly. We present a technique for estimating the camera pose which relies on knowing the 3D position of a moving target, the corresponding image coordinates as well as the pan and tilt angles of the camera. We also assume the camera intrinsic parameters have been determined through calibration. Our application domain is an automated system for tracking and filming Unmanned Aerial Vehicle (UAV) flights trials, however this technique could be applied to many other areas such as estimating the 6 DOF pose of a camera-equipped sensor network node.

When conducting flight trials with an UAV it is useful to film the flight as the video footage can help to refine performance or establish the cause of any failures. A continuous video stream of the UAV is also useful for the UAV operator when there is no direct line of sight from the GCS (Ground Control Station) to the UAV. Manual filming with a tripod-mounted camcorder for long periods is a laborious task and automating this task frees up one person. It can also produce more reliable tracking results since the camera operator could lose sight of the UAV at crucial moments during the flight test due to fatigue or distraction. We have developed a system which can perform this task based on known GPS information of the UAV, or via vision-based tracking of the vehicle. For GPS-based tracking, it is necessary to know the pose (3D location and orientation) of the tracking PTZ camera's center of rotation.

Since our camera is mounted to the roof of a GCS vehicle, its pose is different with each flight trial deployment,

Stefan Hrabar and Volker Hilsenstein are with the Australian Research Centre for Aerospace Automation (ARCAA), 22-24 Boronia Road, Eagle Farm, Queensland, Australia {Stefan.Hrabar,Volker.Hilsenstein}@csiro.au

Peter Corke is with the Queensland University of Technology, 2 George St, Brisbane, Queensland, Australia Peter.Corke@qut.edu.au

Fig. 1. PTZ camera in the background while tracking the UAV.

and measuring this pose directly with the required accuracy is not practical. Before conducting long range flights, a short camera localization flight is conducted within visual range. The UAV is tracked in the image while the image coordinates, UAV GPS position and PTZ pan and tilt angles are recorded. These are then used to estimate the camera pose using a non-linear least squares optimization procedure. This procedure also reduces the setup time for the tracking system compared to hand measuring the pose.

This optimization technique requires an initial estimate for the PTZ pose. We show that for the present application an initial estimate based on GPS position reading and a rough compass alignment are sufficient to ensure convergence. We also present a method for obtaining an initial estimate by transforming the problem into the standard pose estimation problem for a fixed camera and then applying EPnP [1], a non-iterative, globally convergent pose estimation algorithm.

We present experimental results and a study of the sensitivity of the iterative approach to the number of points used, their spatial distribution around the PTZ, the accuracy of the initial pose estimate and measurement noise. We show that data from a single orbit of the UAV around the PTZ is sufficient for localization and that the estimated pose is accurate enough for tracking subsequent flights.

The remainder of this paper is organized as follows: In Section II we present related work and in Section III we present the localization technique. Experimental results are presented in Section IV while in Section V we analyze how sensitive the proposed technique is to various factors. In Section VI we discuss the use of EPnP to seed the optimization and draw conclusions in Section VII.

II. Related Work

Estimating the 6 DOF pose (or 'Exterior Orientation') of a camera using *n* corresponding 3D points and their 2D projections is a widely studied problem in robotics, photogrammetry [2], computer graphics and computer vision. This is often referred to as the Perspective-*n*-Point problem (P*n*P). A number of iterative [3] [4] [5] and non-iterative approaches [1] have been presented to date. In general the iterative approaches produce more accurate results however can be sensitive to the initial estimate and are more computationally expensive.

Our problem differs from the typical P*n*P problem as we are solving for the pose of a PTZ camera (we consider only the case of constant zoom) and not a rigidly fixed camera. We therefore introduce an additional set of rotations (pan and tilt) in the relationship between 3D and 2D coordinates. Our problem also differs in the nature of the 3D points. Typically a number of points are taken from a static scene, with known relative positions between them (such as from a calibration target). The camera pose relative to the coordinate frame of these points can then be determined. In our case a single point is used per frame and the point's location in a global reference frame is known, as is the pan/tilt angle of the camera. Once a number of measurements have been collected over time and pan and tilt angles have been used to transform the points into the PTZ coordinate frame, the problem becomes equivalent to the standard P*n*P one. A variety of techniques could then be used to solve for the PTZ pose.

Pose estimation for PTZ cameras has received much attention lately because of their widespread used in multi-camera surveillance systems. For these systems it is useful to know the relative camera poses, or their respective pose within a global reference frame. A variety of techniques have been presented such as in [6] were the relative camera poses can be estimated provided they have overlapping Field of Views (FOVs). These approaches generally assume the features lie on a horizontal plane. In [7] a sparse set of surveyed points is used to estimate the pose of PTZ cameras in an outdoor environment. These approaches assume the PTZ is stationary while taking a measurement, and that the 3D points are stationary.

Other examples of using a mobile robot to estimate the pose of fixed cameras in the environment include [8] and [9]. In [8] the robot is not tracked by the camera but it also carries a camera and uses its own localization information to estimate the location of features in the environment. The static cameras are then localized by matching these features in their own FOVs. In [9] visual markers are placed on a mobile robot which is tracked in an indoor environment. The image coordinates of the marker and localization information of the robot are used to localize static cameras in this environment.

III. Camera Localization

A. Problem statement

Our aim is to command pan (φ) and tilt (ϑ) angles for a PTZ camera such that the projection of a UAV is kept centered in the image. The geometric configuration is shown in Figure 2. The 3D position $\mathbf{P}$ of the UAV at a given instant is known from GPS telemetry data, but we need to estimate the exterior pose $W_{\text{cw}} = [R_{\text{cw}}\ \mathbf{t}_{\text{cw}}]$ of the PTZ unit as it is required to calculate φ, ϑ such that the projection of $\mathbf{P}$ is centered in the image. We determine this 6 DOF pose from point correspondences between image points $\mathbf{q}_i$ and world points $\mathbf{P}_i$ for different combinations of pan and tilt angles φ_i, ϑ_i. Such point correspondences are obtained during short localization flights where we track the UAV over time and record the GPS telemetry data, the image points, and the values of the angular encoders.

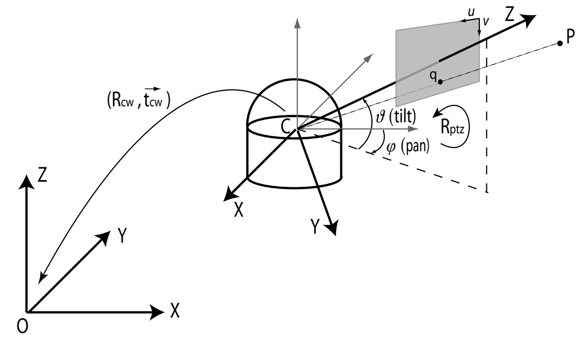

Fig. 2. Camera coordinate system.

B. Determining the exterior pose

We determine the exterior pose by searching for the pose which minimizes the sum of squared reprojection errors (in pixels) between a set of world points and their corresponding image points.

For a pinhole camera, a world point $\mathbf{P}$ in the camera coordinate system projects to a point $\mathbf{q}$ in the image plane according to

$$\mathbf{q} = sM\mathbf{P} \tag{1}$$

where s is an arbitrary scale factor, $\mathbf{q} = (u, v, 1)^{\mathrm{T}}$ and $\mathbf{P}= (x, y, z, 1)^{\mathrm{T}}$ are the image and world points in homogeneous notation and

$$M = \begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

is the camera intrinsic matrix, where $(x_0, y_0)^{\mathrm{T}}$ is the principal point and f_x, f_y are the horizontal and vertical focal length in pixels; for square pixels $f_x = f_y$. We assume the camera's intrinsic parameters have been determined and the images have been corrected for lens distortions. For a camera positioned in a global coordinate system, the 3D point coordinates need to be transformed from the global to the camera coordinate frame before applying the above projection. For a PTZ camera, the exterior pose can be decomposed into two parts, a fixed exterior pose of the unit plus a variable rotation describing the relative pose of the camera with respect to its zero position φ_0, ϑ_0. This leads to the following projection equation:

$$\mathbf{q} = sMR_{\text{ptz}}W_{\text{cw}}\mathbf{P}, \tag{3}$$

where $R_{\text{ptz}} = R_{\vartheta}R_{\varphi}$ is a 3×3 rotation matrix representing pan and tilt of the sensor with respect to the PTZ zero position (φ_0, ϑ_0), and $W_{\text{cw}} = [R_{\text{cw}}\ \mathbf{t}_{\text{cw}}]$ is the fixed exterior pose of the unit consisting of a rotation and a translation of the PTZ unit with respect to the world coordinate system. These coordinate transformations are illustrated in Figure 2.

For a set of points $\mathbf{P}_i = (x_i, y_i, z_i)^{\text{T}}$ and their pan and tilt angles $\varphi_i, \vartheta_i, i = 1, \ldots, n$, with $n \geq 3$, we can compare the predicted reprojection $\hat{\mathbf{q}}_i = (u_i, v_i, 1)^{\text{T}}$ given an estimate of the pose to the actual coordinates extracted from the image $\mathbf{q}_i = (u_i, v_i, 1)^{\text{T}}$. This leads to the following objective function:

$$\epsilon_{W_{\text{ptz}}} = \sum_{i=1}^{n} \|\hat{\mathbf{q}}_i - \mathbf{q}_i\|^2, \tag{4}$$

the sum of squared reprojection errors. Starting with an initial estimate for the pose, we obtain a refined pose estimate by minimizing $\epsilon_{W_{\text{ptz}}}$ subject to the six free parameters of the pose.

IV. Experiments

A. Implementation

Our UAV tracking system utilizes a UAV Vision GV170 PTZ dome (shown in the background of Figure 1) which supports continuous panning, and a tilt range of $\pm 90°$ from the vertical. Optical encoders provide pan and tilt information at a resolution of $0.02°$. Telemetry data is received from the UAV at a rate of 50 Hz via a wireless link. The transmitted UAV pose data is estimated onboard by fusing data from an IMU, GPS, magnetometer and barometer with an Extended Kalman Filter.

The UAV is tracked in the image using an OpenCV [10] implementation of the CAMSHIFT algorithm [11]. This provides the image coordinates (u, v) of the UAV in flight. A Proportional-Derivative (PD) controller uses the measured image coordinates as input, and produces pan and tilt rate commands to keep the UAV centered in the image. To collect data for the pose estimation procedure, the helicopter is flown within visual range of the camera while P, u, v, φ and ϑ values are logged. The PTZ zoom is fixed at the shortest focal length setting, as the intrinsic calibration procedure was performed for this setting.

A Matlab implementation of the Levenberg-Marquardt algorithm [12] takes the logged data as input, and minimizes the objective function (4).

B. Evaluation of Proposed Technique

To evaluate the proposed technique we conducted a number of pose estimation flight trials, and report the results from one trial below. Once the pose had been estimated, it was used by the tracking system during a second flight to track the UAV. Figure 3 shows an image taken by the tracking camera during this flight.

Two performance metrics were used to evaluate the estimated pose; the PTZ pose offset from a manually measured 'ground truth' pose, as well as the mean reprojection error (in pixels) of 3D points onto the 2D image plane. Since manual measurements of the ground truth pose have limited precision, the mean reprojection error is a better metric for the quality of the pose estimation and it is directly related to system performance.

Fig. 3. A sample image from the tracking PTZ camera as it tracked the UAV.

The approximate ground truth position and yaw angle were measured by co-locating and aligning the UAV and PTZ, and taking the mean GPS and compass readings over two minutes. This gave a position estimate within the error bounds of the GPS. The height of the PTZ above ground was measured by hand, while roll and pitch angles were measured with a digital inclinometer. To facilitate measuring the the ground truth pose for this validation experiment, the PTZ was mounted to a tripod which was placed on the ground. For normal flight operations the tripod is placed on the roof of a vehicle, making it impractical to take precise measurements in this way.

The optimization process requires an initial estimate of the pose, and when the PTZ is mounted to the roof this initial estimate is obtained by placing the UAV near the vehicle and taking a GPS position reading from telemetry. For these experiments, although we had a more accurate pose measurement than usual, we did not use this to seed the optimization. Instead we used a roughly measured estimate, making it more realistic in terms of the normal procedure. As shown in Section VI, the initial pose estimate can also be obtained by using the EPnP algorithm, which we plan to use in future deployments.

TABLE I

REPROJECTION ERROR AND POSE OFFSET RELATIVE TO MEASURED GROUND TRUTH POSE FOR A LOCALIZATION FLIGHT AND SUBSEQUENT TRACKING FLIGHT.

Flight	Mean Re-pro. Error (pix)	Position Offset (x, y, z) (m)	Angular Offset (r, p y) (deg)
Localization	17.94	(-0.49, 3.55, -1.13)	(0.96, 1.16, 4.40)
Tracking (estimated pose)	50.38	(0.89, 3.21, -1.23)	(-0.98, -1.62, 4.86)
Tracking (measured pose)	137	(0, 0, 0)	(0, 0, 0)

The results for both flights are tabulated in Table I, while Figures 4 - 7 plot these results. Figure 4 shows the 3D points used for pose estimation, as well as the 'ground truth' and estimated poses. Figure 5 shows the reprojection errors for this flight. Figures 6 and 7 show the same information for the second flight where the UAV was actively tracked. We see the reprojected points are all well within the 720×576 image, and from the grouping of the measured points we see the UAV was kept in the center region of the image for the entire flight. This indicates that the estimated pose was sufficiently accurate for tracking purposes. Also in Table I we show what the mean reprojection error would have been for the tracking flight if the hand-measured 'ground truth' pose had been used. These reprojection errors are shown in Figure 8. We see that with the carefully measured pose the tracking system would have kept the UAV within the camera FOV, however it would not have been as well centered as when using the optimized pose.

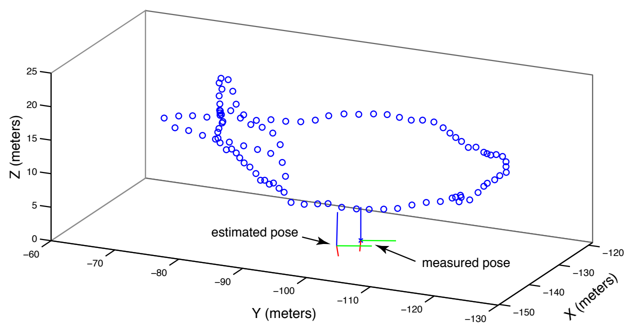

Fig. 4. 3D points of the localization flight with the measured and estimated PTZ poses.

V. SENSITIVITY ANALYSIS

As shown by the results in Section IV, the camera localization is effective for points covering a large segment of the sky, and with a reasonable initial estimate of the pose. In order to streamline the localization process, it is desirable to fly a path which is compact and select a

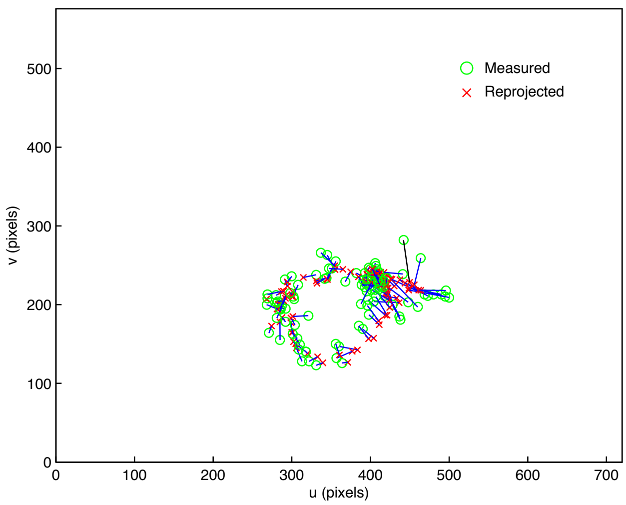

Fig. 5. Measured and reprojected image coordinates for the localization flight, based on the pose estimated from this data. These are shown within the bounds of the 720×576 image.

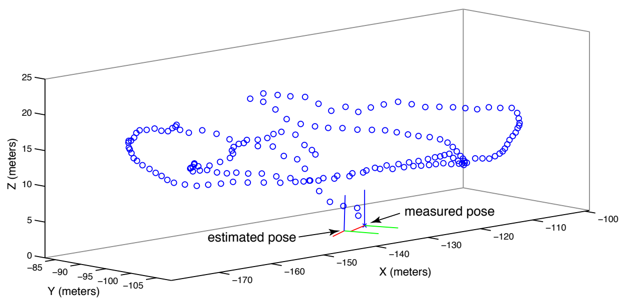

Fig. 6. 3D points of the tracking flight with the measured and estimated PTZ poses.

minimum set of 3D points which still give an adequate localization result. We are therefore interested in evaluating what type of flight pattern would be suitable, and how many points from such a flight would be needed.

Also, measurement noise in u, v, $\mathbf{P}$, φ and ϑ will affect the result, so it is important to understand the effect of this noise and determine how robust the system is to it. Lastly, the optimization process requires an initial pose estimate, so it is useful to understand how accurate this initial estimate needs to be. We therefore analyze the sensitivity of the proposed technique with respect to:

- spatial distribution of 3D points,
- number of 3D points,
- measurement errors, and
- initial pose estimate.

Sensitivity to these factors was measured using data from a flight where the UAV was remotely piloted in various sections of the sky around the PTZ. Certain subsets of the data were then selected to either vary the number of points used, or the distribution of points. The PTZ pose was estimated using each subset, and the resulting pose was used to calculate the mean reprojection error for the

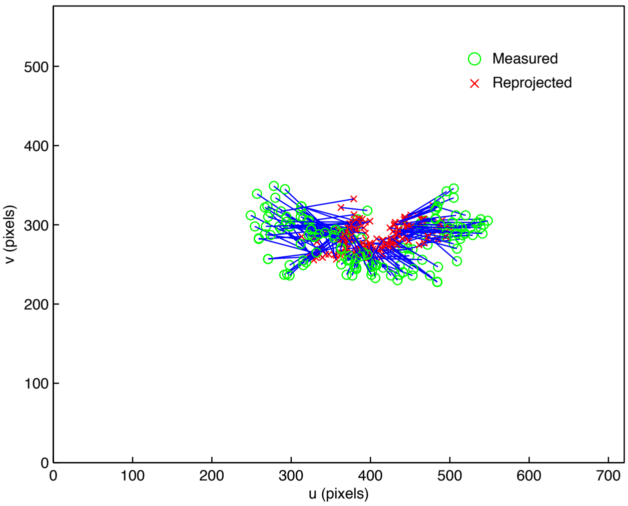

Fig. 7. Reprojection errors for the tracking flight based on pose estimated from the localization flight.

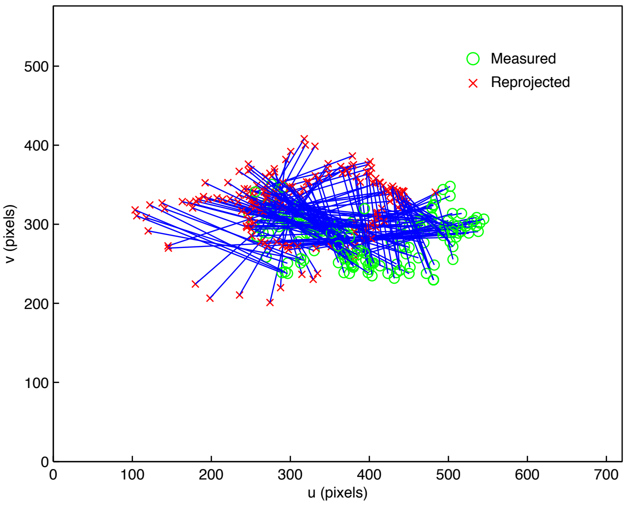

Fig. 8. Reprojection errors for the tracking flight based on the hand measured pose.

entire data set (not just for the subset).

A. Point Distribution

Eighteen different subsets of the data were chosen, varying from points clustered in a small section of the sky, through to points surrounding the PTZ on all sides. We categorized each subset as covering either 1, 2, 3 or 4 quadrants of the sky, and grouped them accordingly. Although the subsets varied in the number of points they contained, we see from Section V-B that varying the number of points above a certain amount has little impact on the results. The subsets all contained sufficient points to ensure point count variations would not significantly influence the results. Table II shows the mean and standard deviation of the mean reprojection errors for groups of subsets. As expected, we see the reprojection error improves as more quadrants are covered.

TABLE II

MEAN REPROJECTION ERRORS OF THE FULL DATASET BASED ON POSE ESTIMATES FROM SUBSETS COVERING DIFFERENT QUADRANTS OF THE SKY.

quadrants covered	Mean Repro. Error (pix)	std dev. (pix)
1	138.8	42.3
2	74.2	39.5
3	38.9	11.8
4	31.4	4.5

Figure 9 shows thumbnails of a few of the subsets used and their mean reprojection errors. The true and estimated poses are also shown in the thumbnails.

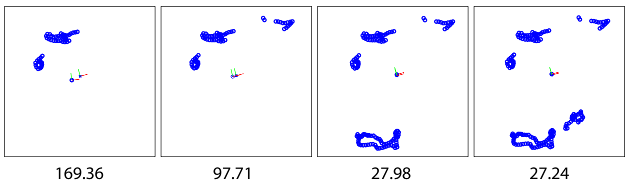

Fig. 9. Plan view of points used for localization, with resulting mean reprojection errors listed below.

B. Number of Points

To establish how sensitive the pose estimation is to the number of 3D points used, we subsampled the full dataset at different rates to produce subsets with varying point counts. In doing so we ensured the points in each subset were uniformly distributed around the PTZ so that point distribution did not affect the results. Table III lists the mean reprojection errors and the number of points used. Also listed in the 3rd column are the reprojection errors relative to the 'best' reprojection obtained using all 529 data points. We see little change in the relative reprojection error until less than 5 points are used in the pose estimation. Comparing the results to those of Section V-A, we see that as few as five points evenly distributed around the PTZ give better results than a large number of points that only cover three quadrants of the sky.

TABLE III

MEAN REPROJECTION ERRORS OF THE FULL DATASET BASED ON POSE ESTIMATES FROM SUBSETS WITH VARYING POINT COUNTS.

Num. Points	Mean Repro. Error	Rel. Error
529	24.4	1.0
265	24.4	1.0
106	24.4	1.0
53	24.6	1.0
11	26.1	1.1
7	26.9	1.1
6	28.5	1.2
5	33.3	1.4
4	67.7	2.8
3	73.5	3.0

C. Measurement Errors

Sensitivity to measurement errors was evaluated by adding zero-mean white noise to φ, ϑ, the image coordinates (u, v) and target position (x, y, z) of the dataset. The magnitude of these errors is representative of what could be expected in extreme cases with the tracking system. Position errors are due to the limited GPS position accuracy and from latencies in transmitting position data from the UAV to the tracking system. Pixel errors can result from the centroid of the CAMSHIFT tracker region not being centered on the UAV reference frame origin. Pan and tilt errors are due to latencies in reading these angles from the PTZ unit.

Table IV shows the errors added individually and applied together. Results are shown for using the entire dataset and for a small subset of six points. The effect of the errors on the estimated pose is indicated in the form of the mean reprojection error in each case. We see in all scenarios that using more points helps to reduce the mean reprojection error (indicating a better pose estimation). This is expected since the zero-mean errors average out overall for larger datasets.

TABLE IV
THE EFFECT OF ZERO-MEAN MEASUREMENT NOISE ON POSE ESTIMATION.

u,v noise	x,y,z noise	φ,ϑ noise	Num Points	Mean Repro. Error
0	0	$\pm 7.5°$	6	154.45
0	0	$\pm 7.5°$	529	105.39
± 20 pix	0	0	6	35.41
± 20 pix	0	0	529	25.4
0	$\pm$ 2.5m	0	6	59.6
0	$\pm$ 2.5m	0	529	49.93
± 20 pix	$\pm$ 2.5m	$\pm 7.5°$	6	146.66
± 20 pix	$\pm$ 2.5m	$\pm 7.5°$	529	112.4

Although the effect of zero-mean errors can be reduced by using a larger data set, systematic errors such as a bias in the pan or tilt readings will impact the estimated pose regardless of how many points are used.

D. Initial Pose Estimate

If the initial pose estimate used to seed the non-linear optimization is far from the true pose, the optimization can fail by falling into a local minimum. Sensitivity to the initial estimate is highly dependent on the configuration of the 3D points, as certain spatial configurations can lead to ambiguous pose solutions. For example if the points are co-planar in the horizontal plane, there is symmetry about this plane so the PTZ could be below looking up at the points, or above looking down. Another example would be a series of co-linear points, as there is symmetry around the line. In practice, such degenerate configurations are easily avoided by choosing an appropriate flight trajectory.

A reasonable initial estimate of many of the degrees of freedom is easily achieved in the field. For example the PTZ can be mounted horizontally (with the help of a bubble level built into the tripod), and aligned approximately using a compass. The height above ground is easily measured or guessed, and the remaining degrees of freedom (x, y) can be roughly measured by a GPS reading in the vicinity of the GCS.

For an initial estimate obtained in this manner, the degrees of freedom that are most likely to be incorrect are yaw and position. We have therefore tested the sensitivity to errors in these degrees of freedom. We applied a variety of offsets from the true values of x, y, z and yaw and used these as the initial estimate. The mean reprojection error was calculated for each initial estimate after running the optimization. We found that the optimization was very robust to errors in the position initial estimate when the yaw estimate error was small (position errors of up to ± 20m in each axis were handled well). In the reverse case with a good position initial estimate (± 2m), yaw errors of up to $\pm 70°$ still resulted in accurate pose estimates. Combined position errors of ± 10m and yaw errors of up to $\pm 45°$ also did not impact on the convergence of the optimization.

The practical implication of this analysis is that a suitable initial estimate is easily obtained for a typical deployment where the PTZ is levelled, aligned within $\pm 45°$ of a compass bearing, and placed within 10m of a known reference point.

VI. USING EPnP TO FIND AN INITIAL POSE ESTIMATE

We have also developed an EPnP-based [1] method to estimate the PTU pose without the need for an initial estimate. The pose estimates obtained using this method on data collected during the experiments are good enough to allow for tracking of the UAV, however, we have observed that in most cases the iterative approach performs better. The optimization can however be seeded with the output of this method instead of using a hand-measured initial pose estimate.

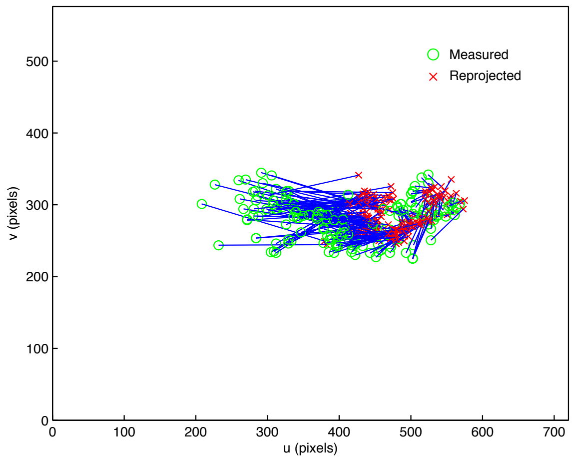

Fig. 10. Reprojection errors for the tracking flight based on pose estimated by the EPnP algorithm.

For example with the data shown in Figure 6, the mean reprojection error from the EPnP method is 75.42 and this is reduced to 50.96 with a subsequent iterative optimization of equation (4). This is the same result as when an initial estimate was used based on measurements taken in the field. Figure 10 shows the reprojection errors that result from using EPnP alone, and this can be compared to Figure 7 for the iterative approach.

This method, which we plan to use in future deployments, is based on transforming the pose estimation problem for the pan-tilt unit into the perspective n-point problem and applying the globally convergent EPnP algorithm, currently the best performing non-iterative pose-estimation algorithm.

The PTZ camera gimbal is designed such that the pan and tilt rotations are around the camera center. The image points in different views are therefore related by a homography [13]. Using the known pan and tilt angles φ_i, ϑ_i for a given image point $\mathbf{p}_i$ we can calculate the homography relating this image point to a corresponding image point in reference view φ_0, ϑ_0 as follows:

$$H_i = M R_{\varphi_i}^{-1} R_{\vartheta_i}^{-1} M^{-1},$$

where M is the camera projection matrix (2) and $R_{\varphi_i}^{-1} R_{\vartheta_i}^{-1}$ is the inverse rotation of the camera pan and tilt. When all image points are reprojected into the reference view using the homography, $\mathbf{q}_{i,0} = H_i \mathbf{q}_i$, we apply EP$n$P to determine the PTZ pose.

VII. Conclusions and Future Work

Our application is the automatic tracking and filming of UAV flight trials. This requires knowing the pose of the tracking PTZ camera sufficiently accurately to ensure the UAV is kept within the camera FOV. Measuring the pose directly to such a degree of accuracy is usually not practical, so we rely on estimating it indirectly with the proposed technique.

We have shown that the 6 DOF pose of a PTZ camera can be estimated by tracking a single 3D point over time. The proposed technique requires knowledge of the 3D positions of the tracked point, the corresponding 2D image coordinates as well as the pan and tilt angles of the camera. The camera's intrinsic parameters are also required. The 3D position of our target is available via telemetry downlink to the UAV ground station. Pixel coordinates of the UAV in the image are obtained by vision-based tracking of the UAV, while pan and tilt angles are read from the PTZ unit. Our experiments show that a short localization flight with the UAV orbiting the PTZ once provides sufficient data for localization. The obtained pose estimate is accurate enough for the system to subsequently track the UAV, keeping its projection near the image center. We include a detailed analysis of how sensitive the technique is to factors such as the number of points used, their distribution about the PTZ, measurement errors and the initial estimate for seeding the optimization. We show that it is more important to have points well distributed than to have many points, however more points do help to reduce the effect of measurement noise. Although the optimization is sensitive to the initial estimate, obtaining a suitable estimate is easily achieved for typical deployments of our tracking system. Alternatively, the non-iterative, globally convergent EPnP algorithm can be used to generate an initial estimate for the optimization, and we plan to employ this technique for future deployments.

We also plan to investigate determining the camera intrinsic parameters from the camera localization flight data instead of relying on pre-calibration of the camera. Because of potential GPS drift, it would be useful to have the pose estimate refined and updated in an online fashion during subsequent tracking flights, so this is an area of future work.

VIII. Acknowledgements

The authors would like to thank Torsten Merz for his help in obtaining the flight data. This work was funded by the Smart Skies project which is supported, in part, by the Queensland State Government Smart State funding scheme.

References

[1] F. Moreno-Noguer, V. Lepetit, and P. Fua, "Accurate non-iterative O(n) solution to the PnP problem," in *IEEE International Conference on Computer Vision*, Rio de Janeiro, Brazil, October 2007, pp. 1–8.

[2] P. Grussenmeyer and O. A. Khalil, "Solutions for exterior orientation in photogrammetry: A review," *The Photogrammetric Record*, vol. 17, pp. 615–634, 2002.

[3] A. Ansar and K. Daniilidis, "Linear pose estimation from points or lines," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 5, pp. 578–589, May 2003.

[4] P. Fiore, "Efficient linear solution of exterior orientation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 2, pp. 140–148, Feb. 2001.

[5] D. Lowe, "Fitting parameterized three-dimensional models to images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 13, no. 5, pp. 441–450, May 1991.

[6] I.-H. Chen and S.-J. Wang, "An efficient approach for the calibration of multiple PTZ cameras," *IEEE Transactions on Automation Science and Engineering*, vol. 4, no. 2, pp. 286 –293, April 2007.

[7] R. Collins and Y. Tsin, "Calibration of an outdoor active camera system," in *IEEE Computer Vision and Pattern Recognition (CVPR '99)*, vol. 1, June 1999, pp. 528–534.

[8] N. Leite, A. D. Bue, and J. Gaspar, "Calibrating a network of cameras based on visual odometry," in *in Proc. of IV Jornadas de Engenharia Electrónica e Telecomunicações e de Computadores*, Lisbon, Portugal, November 2008, pp. 174–179.

[9] T. Sasaki and H. Hashimoto, "Camera calibration using mobile robot in intelligent space," in *SICE-ICASE, 2006. International Joint Conference*, Oct. 2006, pp. 2657–2662.

[10] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*. Cambridge, MA: O'Reilly, 2008.

[11] G. Bradski, "Computer video face tracking for use in a perceptual user interface," *Intel Technology Journal,*, vol. Q2, 1998.

[12] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *The Quarterly of Applied Mathematics*, vol. 2, pp. 164–168, 1944.

[13] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.