PLANAR HOMOGRAPHY: ACCURACY ANALYSIS AND APPLICATIONS

Shahriar Negahdaripour[*], *Ricard Prados, Rafael Garcia*

Computer Vision and Robotics Group
Dept. of Electronics, Informatics and Automation
University of Girona, Spain

ABSTRACT

Projective homography sits at the heart of many problems in image registration. In addition to many methods for estimating the homography parameters [5], analytical expressions to assess the accuracy of the transformation parameters have been proposed [4]. We show that these expressions provide less accurate bounds than those based on the earlier results of Weng et al. [7]. The discrepancy becomes more critical in applications involving the integration of frame-to-frame homographies and their uncertainties, as in the reconstruction of terrain mosaics and the camera trajectory from flyover imagery. We demonstrate these issues through selected examples.

1. INTRODUCTION

The registration of various frames in a video mosaic has numerous applications, including generation of terrain mosaics from flyover image transects in underwater and airborne systems [3]. The abilities to accurately compute the view-to-view transformation parameters and to assess their accuracies are equivalently important issues. For example, reprojection error bounds can be used to establish search windows in processing of the data recursively to improve the image matching and registration.

Let I and I' be the images of a scene from two camera viewpoints. Under many conditions and for a number of applications, it may be assumed that the scene is (approximately) planar, and thus the image transformation $I \rightarrow I'$ can be described by a planar projective homography. Given a pair of image correspondences $\{\mathbf{p} = (x, y, 1)^T$ and $\mathbf{p}' = (x', y', 1)^T$ in I and I', respectively, the projective homography $\mathbf{H}$ establishes the constraint between each correspondence according to $\rho' \mathbf{p}' = \mathbf{Hp}$, where ρ' is a constant scaling. A number of methods, including closed-form linear methods, have been given to estimate $\mathbf{H}$ (up to a scale) from $N \geq$ correspondences $\{\mathbf{p}_i, \mathbf{p}'_i\}$, $i = 1 : N$. Assuming that the image position measurements are corrupted by Gaussian noise, Criminisi et al. [4] give closed-form expressions to estimate the variance of the 8 independent parameters[1]. The derivation is

claimed to be based on the earlier results of Weng et al. [7], who reported a comprehensive error analysis of motion and structure parameters from image correspondences. Criminisi et al. also stated that their results, being better conditioned than solutions given in [2] if only $N = 4$ correspondences are used, also provide better estimates when the $N > 4$ matchings have relatively small measurement noise levels. We will show that these expressions provide less accurate bounds than those derived from a singular value perturbation analysis, as first proposed by Weng et al. [7]. We also explore the behavior of the uncertainty bounds in applications involving the integration of frame-to-frame homographies, as in the reconstruction of mosaics and the camera trajectory from flyover imagery.

2. TERMINOLOGY AND BACKGROUND

Let $\mathbf{H}$ denote the planar homography that maps points $\mathbf{p}$ in I onto $\mathbf{p}'$ in I' based on up-to-scale transformation $\mathbf{p}' = \mathbf{Hp}$. For simplicity, we express the up-to-scale homography as[2]

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix}$$

If $\mathbf{h}$ denotes the vector form of $\mathbf{H}$, the above homograhy can be expressed by two linear constraint equations on $\mathbf{h}$. If we have N correspondences $\{\mathbf{p}_i, \mathbf{p}'_i\}$ $(i = 1 : N)$, $2N$ constraint equations can be written $\mathbf{Ah} = \mathbf{0}$, where

$$\mathbf{A} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1 x_1 & -x'_1 y_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1 x_1 & -y'_1 y_1 & -y'_1 \\ & & & & \vdots & & & & \\ x_N & y_N & 1 & 0 & 0 & 0 & -x'_N x_N & -x'_N y_N & -x'_N \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_N x_N & -y'_N y_N & -y'_N \end{bmatrix}$$

Let $\mathbf{A} = \mathbf{U\Sigma V}^T$ denote the singular value decomposition of $\mathbf{A}$. Equivalently, we consider the eigenvalue decomposition $\mathbf{Q} = \mathbf{A}^T \mathbf{A} = \mathbf{VD}^2 \mathbf{V}^T$, where $\Sigma = \mathbf{D}^2$. The linear solution of the 9×1 unit homography vector $\hat{\mathbf{h}}$ is the eigenvector associated with the smallest singular value: $\hat{\mathbf{h}} = v_9$ (diagonal elements of Σ are arranged in descending order, and v_i is the i-th column of $\mathbf{V}$). Subsequently, $\mathbf{h}$ can be determined by scaling: $\mathbf{h} = \hat{\mathbf{h}}/\hat{h}_9$. We are interested in the variations of

---

[1]Nine elements determined up to a constant scale factor.

[2]This representation assumes that $h_9 \neq 0$.

the solution with noisy observations. Define the observation vector $\mathbf{m}$ of N left-and-right matching pixel coordinates:

$$\mathbf{m} = [x_1, y_1, x_1', y_1', x_2, y_2, x_2', y_2', \ldots, x_N, y_N, x_N', y_N']_{4N \times 1}$$

The noisy measurement vector $\hat{\mathbf{m}} = \mathbf{m} + \delta\mathbf{m}$ comprises zero-mean Gaussian noise vector $\delta\mathbf{m}$ with covariance $\mathbf{C}_m$. A simplified assumption is a normal distribution $(0, \sigma)$; that is, $\mathbf{C}_m = \sigma^2 \mathbf{I}_{4N \times 4N}$, which is reasonable when the errors are primarily due to quantization effects and the localization inaccuracies of the feature detector, e.g. Harris interest point operator [1]. Outliers violate the assumed normal noise distribution model. However, it is typically the case that some robust estimation method, e.g. RANSAC or LMedS, can be used as a first step to identify the outliers, before the proposed linear solution is applied to the inliers (which satisfy the assumed noise model). Here, we assume knowledge of σ, and are interested to estimate the variances of h_i ($i = 1 : 8$).

3. COVARIANCE OF PROJECTIVE HOMOGRAPHY

Due to space limitation, we refer the reader to section 5 in [4], given for the estimation of the covariance of the homography parameters. This is denoted $\mathbf{C}_{hc}$ here. We also derive $\mathbf{C}_{ho}$, the covariance derived here as the new estimate.

From $\delta\mathbf{m}$, we can determine the variation in the measurement matrix $\mathbf{A}_{2N \times 9}$, or $\mathbf{Q}_{9 \times 9}$. Assume that $\mathbf{Q}$ is perturbed by $\Delta\mathbf{Q}$, where $\mathbf{q}_{81 \times 1}$ and $\delta\mathbf{q}_{81 \times 1}$ are the corresponding vector forms. For small variations $-\max\{\delta q_i\} << 1$, where q_i denotes i-th element of $\mathbf{q}$ –it can be shown [6, 7] that up to first-order, the eigenvalues and eigenvectors of $\mathbf{Q}$ vary according to $\delta\lambda_i = \boldsymbol{v}_i^T \Delta\mathbf{Q}\, \boldsymbol{v}_i$ and $\delta\boldsymbol{v}_i = \mathbf{V}\Psi_i\mathbf{V}^T\Pi_i\delta\mathbf{q}$ where

$$\Psi_i = \text{diag}\{(\lambda_i - \lambda_1)^{-1}, \ldots, 0, \ldots, (\lambda_i - \lambda_9)^{-1}\}$$

$$\Pi_i = \begin{bmatrix} v_{i1}\mathbf{I}_{9\times9} & v_{i2}\mathbf{I}_{9\times9} & \ldots & v_{i9}\mathbf{I}_{9\times9} \end{bmatrix}$$

The covariances of the eigenvectors $\mathbf{C}_{v_i} = \mathbf{V}\Psi_i\mathbf{V}^T\Pi_i\mathbf{C}_q\Pi_i^T\mathbf{V}\Psi_i^T\mathbf{V}^T$ allows us to write

$$\mathbf{C}_{\hat{h}} = \mathbf{C}_{v_9} = \mathbf{V}\Psi_9\mathbf{V}^T\Pi_9\mathbf{C}_q\Pi_9^T\mathbf{V}\Psi_9^T\mathbf{V}^T$$

Recalling that the homography $\mathbf{h}$ is determined from $\hat{\mathbf{h}}$ by scaling (such that $h_9 = 1$) the covariances of $\mathbf{h}$ and $\hat{\mathbf{h}}$ are related by the transformation $\mathbf{C}_{ho} = \mathbf{J}\,\mathbf{C}_{\hat{h}}\,\mathbf{J}^T$ where

$$\mathbf{J} = \left(\frac{\partial\mathbf{h}}{\partial\hat{\mathbf{h}}}\right) = \begin{bmatrix} 1/h_9 & 0 & \ldots & 0 & -h_1/h_9^2 \\ 0 & 1/h_9 & \ldots & 0 & -h_1/h_9^2 \\ \vdots & & & & \\ 0 & 0 & \ldots & 1/h_9 & -h_8/h_9^2 \end{bmatrix}_{8 \times 9}$$

In the above equations, $\mathbf{C}_q$ is determined from

$$\mathbf{C}_q = \left(\frac{\partial\mathbf{q}}{\partial\mathbf{a}}\right)\left[\left(\frac{\partial\mathbf{a}}{\partial\mathbf{m}}\right)\mathbf{C}_m\left(\frac{\partial\mathbf{a}}{\partial\mathbf{m}}\right)^T\right]\left(\frac{\partial\mathbf{q}}{\partial\mathbf{a}}\right)^T$$

where $\mathbf{a}$ is the vector form of $\mathbf{A}$:

$$\mathbf{a} = [\; x_1, y_1, 1, 0, 0, 0, -x_1'x_1, -x_1'y_1, -x_1', \\ 0, 0, 0, x_1, y_1, 1, -y_1'x_1, -y_1'y_1, -y_1', \ldots \\ 0, 0, 0, x_N, y_N, 1, -y_N'x_N, -y_N'y_N, -y_N']_{18N \times 1}$$

4. SIMULATIONS

We use computer simulations to compare the closed-form expressions for estimating the variances of the projective homography parameters, given in [4] and derived here (denoted $\mathbf{C}_{hc}$ and $\mathbf{C}_{ho}$, respectively). In all but two tests, we start with $N = 20$ points $\mathbf{p} = [x(i, j), y(i, j), f]$ on a regular grid $i = 64 : 64 : 320$ and $j = 64 : 64 : 256$ ($f = 320$ is assumed). In these two our simulations, only a minimum 4 points near the image corners are used. We construct matching pairs $\{\mathbf{p}, \mathbf{p}'\}$ based on a pre-specified homography $\mathbf{H}$; we use the well-know interpretation $\mathbf{H} = \mathbf{R} + \mathbf{t}\mathbf{n}^T$ in terms of the motion $\{\mathbf{R}, \mathbf{t}\}$ of a camera relative to a planar scene with surface normal $\mathbf{n} = [-P, -Q, 1]/Z_o$, where P and Q control the surface slant and tilt angles, and Z_o its distance from the camera. Measurements noise from a normal distribution with standard deviation σ_m is then added to corresponding pairs to subsequently estimate the homography, and calculate its error. The process is repeated 1000 times with different noise samples to compute statistical measures. In case 1, the experimental variance of each parameter in h_i ($i = 1 : 8$) is compared with the analytical estimates $\sigma_{hk} = \sqrt{diag\{\mathbf{C}_{hk}\}}$ ($k = c, o$). In the second case, the estimated noisy homographies map certain corners of a real image into 1000 noisy matches in the 2nd view. The final experiment deals with the generation of a 10-frame sequence, and the integration of frame-to-frame homographies and the corresponding variances.

4.1. Case 1

In the fist simulation, we have used $P = 0.2$, $Q = -0.2$, and $Z_o = 10$, with translational motion $\mathbf{t} = [0.1, 0.2, 0.05]$ and rotational matrix $\mathbf{R} = \mathbf{R}_x\mathbf{R}_y\mathbf{R}_z$, where $\mathbf{R}_a$ denotes a rotation about axis a with angle θ_a ($\theta_x = 0.05$ [rad], $\theta_y = 0.05$ and $\theta_y = 0$).

Fig. 1 (top) shows the results for $\sigma_m = 1$ with $N = 20$ correspondences (horizontal axis corresponds to homography parameters h_i ($i = 1 : 8$), and the vertical axis is the estimation error). Blue crosses depict the errors of the homography parameters for each of 1000 simulations, with red circles giving the (zero) mean error. The dashed blue envelop is the $\pm 3\sigma$ error bound computed experimentally, and the other two envelops in green and red are derived from analytical bounds $\pm 3\sigma_{hc}$ and $\pm 3\sigma_{ho}$, respectively. The latter nearly coincides with the experimental results. In [4], the authors claim that their solution provides better estimates in two cases: 1) Relatively small measurement noise levels, or 2) when minimum $N = 4$ image correspondences are utilized in the estimation of the homography. These cases are tested in the next three plots corresponding to $\sigma_m = 0.05$ with $N = 20$ $\sigma_m = 1/3$ with $N = 4$, and $\sigma_m = 1$ with $N = 4$. In all cases, the new results consistently provide a more accurate estimate of the $\pm 3\sigma_h$ error bounds.

4.2. Case 2

Fig. 2 shows the original and transformed images based on an assumed homography. In addition, selected corners have been mapped with 1000 homographies, estimated from noisy correspondences ($\sigma_m = 1$ is assumed). The green and red uncertainty ellipses of the transformed points have been determined from the homography variances σ_{hc} and σ_{ho}, respectively. The results for 4 selected points, $A - D$, demonstrate once again that σ_{ho} provides a more accurate and tighter uncertainty bound.

4.3. Case 3

A 10-frame sequence has been constructed based on a prescribed homography. Fig. 3 shows the sequence with an inter-frame motion of about 13-14 pixels. The blue stars depict the true positions of 3 sample pixels in various frames, at the center of uncertainty ellipses (computed from the two different techniques) that bound noisy positions of these points based on homographies estimated from 1000 simulations with noisy correspondences. For each of these three points, the noisy positions and bounding ellipses in frames 1, 4, 7 and 10 have been given in subsequent plots. As in previous 2 cases, our new results provide a tighter fit of the projected point distributions.

5. SUMMARY

Ability to not only estimate the transformation between frames but also to assess the confidence in these estimates is important in many applications involving motion estimation from video imagery. Computation of projective homography from frame-to-frame correspondences has been extensively studied in recent years [5], and analytical uncertainty bounds of the homography parameters and reprojection errors have been proposed [4]. Based on earlier results of Weng et al. [7], we have provided new expressions to estimate these bounds more accurately. This has been verified in a number of experiments based on the estimation of homography parameters, as well as the reprojection of image points based on the estimated homographies. These results are particularly important when dealing with long image sequences where frame-to-frame estimates need to be integrated to establish the camera position, to build an image mosaic, or generally to register various frames of a video sequence. Under investigation is the direct use of these uncertainty bounds in the construction of photo-mosaics.

6. REFERENCES

[1] D.P. Capel, "Image mosaicing and super-resolution," Ph.D. Thesis, Dept. of Engineering Science, Univ. of Oxford, 2001.
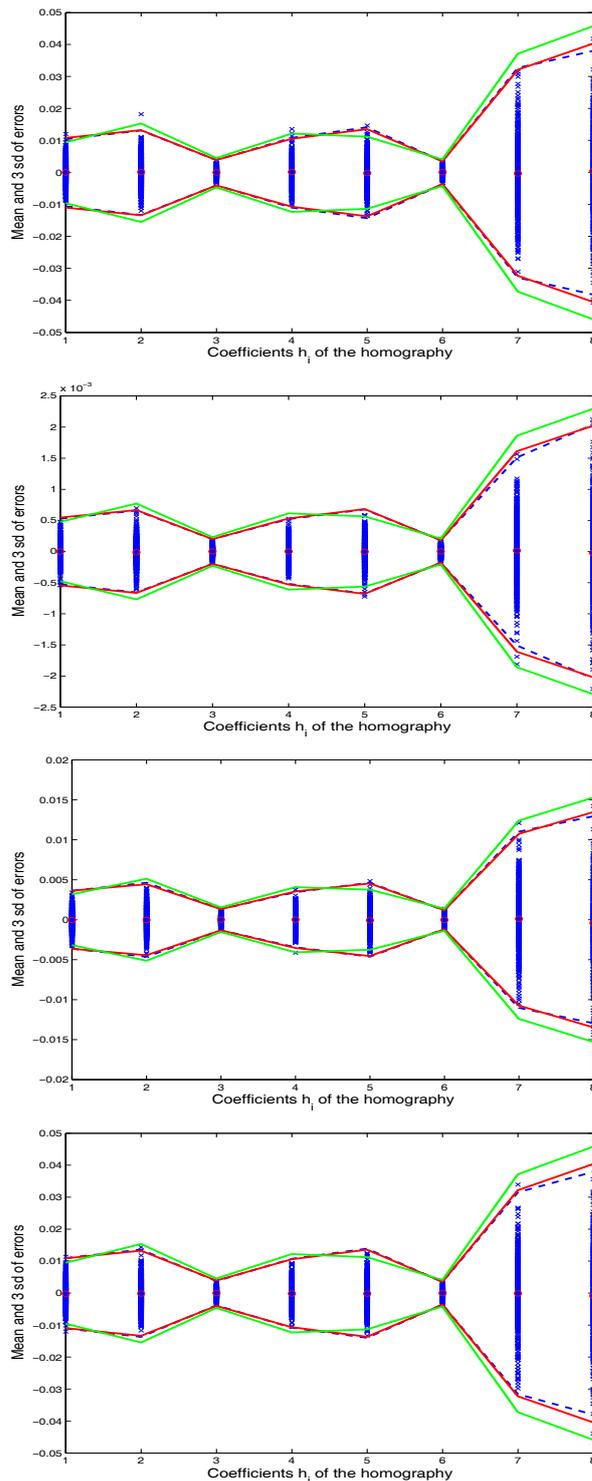

Fig. 1. Comparison between experimental (dashed-blue) and analytical uncertainty bounds of the plane homography coefficients for various noise levels and image correspondences ($\mathbf{C}_{hc}$ in green, and $\mathbf{C}_{ho}$ in red). In each case, the experimental homography represents 1000 simulations with different noise samples of variance σ_m and N point correspondences. From top to bottom, (1) $\sigma = 1$ [pix] with $N = 20$; (2) $\sigma = 0.05[pix]$ with $N = 20$; (3) $\sigma = 1/3[pix]$ with $N = 4$; (4) $\sigma = 1[pix]$ with $N = 4$.

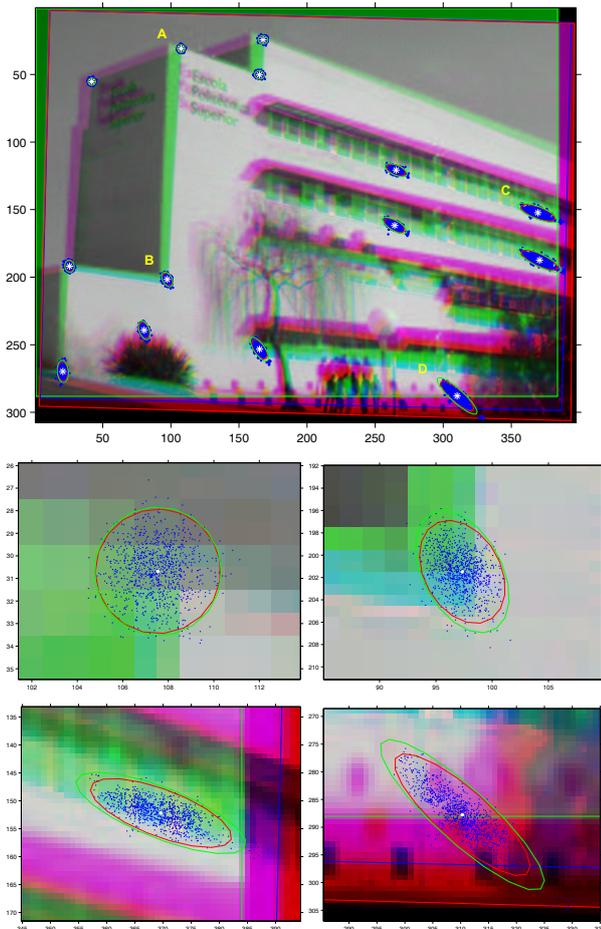**Fig. 2**. (top) Two views transformed according to homography $\mathbf{H}_3$ superimposed on two color channels. Selected corners are mapped based on homographies that are calculated from noisy image features (with variance $\sigma_m = 1$), and repeated 1000 times with different noise samples. For each corner, the analytical 3σ uncertainty ellipse is given based on $\mathbf{C}_{hc}$ (green) and $\mathbf{C}_{ho}$ (red). (bottom) Results have been magnified for 4 selected corners A, B, C and D.

[2] J.C. Clarke, "Modelling uncertainty: A primer", Tech. Report 2161/98, Univ. of Oxford, Dept. of Engineering Science, 1998.

[3] R. Garcia, X. Cufi, and V. Ila, "Recovering Camera Motion in a Sequence of Underwater Images through Mosaicking", *LNCS* no. 2652, pp. 255-262, Eds. Springer-Verlag, 2003.

[4] A. Criminisi, I. Reid, and A. Zisserman, "A Plane Measuring Device," *Image and Vis. Comp.* vol. 17(8) pp. 625–634, 1999.

[5] R.I Hartley, and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge Univ. Press, 2000.

[6] Z. Sun, V. Ramesh, and A.M. Tekalp, "Error characterization of the factorization method," CVIU, vol. 82(2), 2001.

[7] J. Weng, T. S. Huang, and N. Ahuja, "Motion and Structure from Two Perspective Views: Algorithms, Error Analysis, and Error Estimation," *IEEE Trans. on Patt. Anal. and Mach. Int.*, vol. 11(5), pp.451–476, 1989.
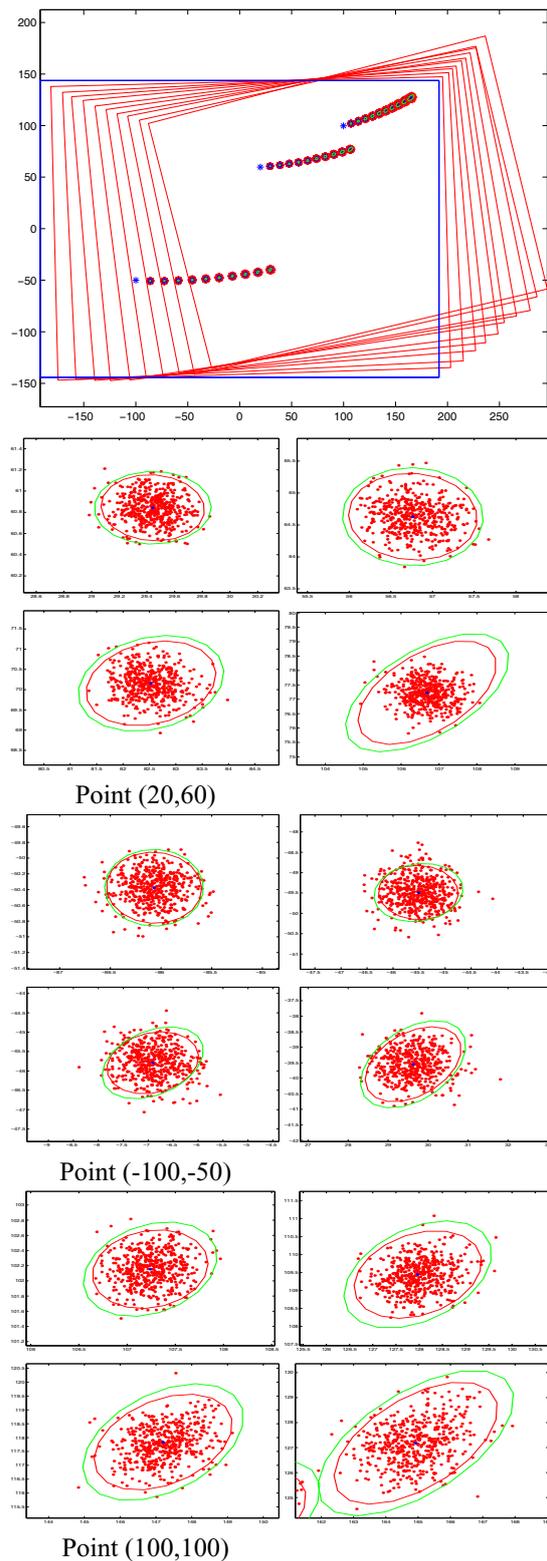
Fig. 3. Estimated positions of three sample points in a 10-frame sequence based on homographies estimated from noisy correspondences, repeating the experiment 1000 times. Analytical uncertainty ellipses have been given for $\mathbf{C}_{hc}$ (green) and $\mathbf{C}_{ho}$. For each sample point, the estimated positions and ellipses in frames 1, 4, 7, 10 are given in a 2×2 plot.