PAPER *Special Issue on Machine Vision Applications*

Optimal Homography Computation with a Reliability Measure*

Kenichi KANATANI†, **Naoya OHTA**†, *and* **Yasushi KANAZAWA**††, *Members*

SUMMARY We describe a theoretically optimal algorithm for computing the homography between two images. First, we derive a theoretical accuracy bound based on a mathematical model of image noise and do simulation to confirm that our renormalization technique effectively attains that bound. Then, we apply our technique to mosaicing of images with small overlaps. By using real images, we show how our algorithm reduces the instability of the image mapping.

key words: homography, image mosaicing, reliability evaluation, renormalization, theoretical accuracy bound

1. Introduction

A *homography* is a mapping that typically occurs between two perspective images of a planar surface in the scene. Computing homography plays an essential role in aerial image registration and analysis of runway and road images viewed from airplanes and vehicles. Since far-away scenes can effectively be regarded as planar surfaces, we can integrate multiple images into one continuous image by computing homographies; this process is known as *image mosaicing* [10], [11], [16].

If we know the homography between two images of a planar surface, we can determine the position of the surface and the relative motion of the camera up to scale [8], [9], [15]. This fact can be used to build a scene understanding system based on objects with planar surfaces [14]. If the scene consists mostly of a single planar surface (e.g., the ground) on which there are non-planar objects, we can detect them by mapping one image onto the other via a homography and searching for non-overlapping regions. We can also reconstruct 3-D from the degree of non-overlapping [1], [3], [4]. Homography computation also plays an essential part in calibrating cameras and structured-light projection systems using planar patterns [2], [13].

However, there are some issues which have not received much attention:

- In the past, the least-squares method (described later) has been used widely because of computational simplicity. However, the least-squares method produces statistical bias [6], [8], because *homogeneous coordinates* are treated as if they were physical quantities without distinguishing measurement data from artificial constants. Since noise occurs in "image coordinates" (*inhomogeneous coordinates*) subject to *Euclidean geometry*, we must work in the Euclidean framework [12].
- Traditionally, the performance of a new algorithm is tested by random noise simulation but usually compared only with "previously proposed methods." However, an algorithm is *optimal* only when it attains a *theoretical accuracy bound*.

Recently, Kanatani [6] proposed a theory of statistical optimization, generalizing various types of geometric estimation in an abstract framework and deriving a scheme for optimal estimation and reliability evaluation. In this paper, we apply his theory to homography computation. First, we derive a theoretical accuracy bound based on a mathematical model of image noise. Then, we describe a technique called *renormalization* that theoretically attains the accuracy bound in the first order. By random noise simulation, we confirm that our renormalization method effectively attains that bound.

Finally, we apply our algorithm to mosaicing of images with *small overlaps*. Since matching points are restricted in very narrow strips in the images, their small perturbation would cause large image deformations. Using real images, we demonstrate this vulnerability and show how our algorithm reduces it.

2. Representation of Homography

A *homography* is an image mapping in the form

$$x' = \frac{Ax + By + C}{Px + Qy + R}, \quad y' = \frac{Dx + Ey + F}{Px + Qy + R}. \tag{1}$$

If we define vectors $\boldsymbol{x}$ and $\boldsymbol{x}'$ and matrix $\boldsymbol{H}$ by

$$\boldsymbol{x} = \begin{pmatrix} x/f \\ y/f \\ 1 \end{pmatrix}, \quad \boldsymbol{x}' = \begin{pmatrix} x'/f \\ y'/f \\ 1 \end{pmatrix}, \tag{2}$$

$$\boldsymbol{H} = \begin{pmatrix} A & B & C/f \\ D & E & F/f \\ P/f & Q/f & R/f^2 \end{pmatrix}, \tag{3}$$

Manuscript received September 22, 1999.
Manuscript revised March 2, 2000.

†The authors are with the Department of Computer Science, Gunma University, Kiryu-shi, 376-8515 Japan.
††The author is with the Department of Knowledge-based Information Engineering, Toyohashi University of Technology, Toyohashi-shi, 441-8580 Japan.

*This work was in part supported by the Ministry of Education, Science, Sports and Culture, Japan under the Grant-in-Aid for Scientific Research C(2) (No. 11680377).

Eq. (1) can be written as

$$\boldsymbol{x}' = k\boldsymbol{H}\boldsymbol{x}, \tag{4}$$

where k is an arbitrary nonzero constant. To avoid numerical instability, the scale factor f is chosen in such a way that x/f and y/f have order of 1. Since the scale of the matrix $\boldsymbol{H}$ is irrelevant, we adopt normalization $\|\boldsymbol{H}\| = 1$, where the norm of a matrix $\boldsymbol{A} = (A_{ij})$ is defined by $\|\boldsymbol{A}\| = \sqrt{\sum_{i,j=1}^{3} A_{ij}^2}$. We hereafter call both the matrix $\boldsymbol{H}$ and the mapping it defines the "homography $\boldsymbol{H}$."

3. Description of Uncertainty

Regarding data as random variables, we describe the uncertainty of data points (x_α, y_α) and (x'_α, y'_α) by their covariance matrices $\boldsymbol{\Sigma}_\alpha$ and $\boldsymbol{\Sigma}'_\alpha$. It follows that the vectors $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ have the following singular covariance matrices (the superscript $\top$ denotes transpose):

$$V[\boldsymbol{x}_\alpha] = \frac{1}{f^2}\begin{pmatrix} \boldsymbol{\Sigma}_\alpha & \boldsymbol{0} \\ \boldsymbol{0}^\top & 0 \end{pmatrix},$$

$$V[\boldsymbol{x}'_\alpha] = \frac{1}{f^2}\begin{pmatrix} \boldsymbol{\Sigma}'_\alpha & \boldsymbol{0} \\ \boldsymbol{0}^\top & 0 \end{pmatrix}. \tag{5}$$

Since it is difficult to predict the uncertainty of each data point in advance, we assume that only the relative tendency of noise occurrences is known. In other words, we assume that the covariance matrices $V[\boldsymbol{x}_\alpha]$ and $V[\boldsymbol{x}'_\alpha]$ are known *up to scale* and write

$$V[\boldsymbol{x}_\alpha] = \epsilon^2 V_0[\boldsymbol{x}_\alpha], \quad V[\boldsymbol{x}'_\alpha] = \epsilon^2 V_0[\boldsymbol{x}'_\alpha]. \tag{6}$$

We call the unknown magnitude ϵ the *noise level*. The matrices $V_0[\boldsymbol{x}_\alpha]$ and $V_0[\boldsymbol{x}'_\alpha]$, which we call the *normalized covariance matrices*, specify the relative dependence of noise occurrence on positions and orientations. If no prior knowledge is available for them, we simply assume isotropy and homogeneity and use the default values $V_0[\boldsymbol{x}_\alpha] = V_0[\boldsymbol{x}'_\alpha] = \mathrm{diag}(1, 1, 0)$ (the diagonal matrix with diagonal elements 1, 1, and 0 in that order).

Let $\hat{\boldsymbol{H}} = (\hat{H}_{ij})$ be an estimate of the homography, and $\bar{\boldsymbol{H}} = (\bar{H}_{ij})$ its true value. The uncertainty of the estimate $\hat{\boldsymbol{H}}$ is measured by its *covariance tensor*

$$\mathcal{V}[\hat{\boldsymbol{H}}] = E[\mathcal{P}\Big((\hat{\boldsymbol{H}} - \bar{\boldsymbol{H}}) \otimes (\hat{\boldsymbol{H}} - \bar{\boldsymbol{H}})\Big)\mathcal{P}^\top], \tag{7}$$

where $E[\,\cdot\,]$ denotes expectation. The symbol $\otimes$ denotes tensor product: for matrices $\boldsymbol{A} = (A_{ij})$ and $\boldsymbol{B} = (B_{ij})$, the $(ijkl)$ element of their tensor product is $A_{ij}B_{kl}$. For tensors $\mathcal{P} = (P_{ijkl})$ and $\mathcal{T} = (T_{ijkl})$, the product $\mathcal{P}\mathcal{T}\mathcal{P}^\top$ is a tensor whose $(ijkl)$ element is $\sum_{m,n,p,q=1}^{3} P_{ijmn}P_{klpq}T_{mnpq}$. The $(ijkl)$ element of the tensor $\mathcal{P}$ in Eq. (7) is given by

$$P_{ijkl} = \delta_{ik}\delta_{jl} - \bar{H}_{ij}\bar{H}_{kl}, \tag{8}$$

where δ_{ij} is the Kronecker delta, taking 1 for $i = j$ and 0 otherwise. Since a homography is normalized to have unit norm, it is a point on an 8-dimensional sphere S^8 in the 9-dimensional parameter space $\mathcal{R}^9$. The tensor $\mathcal{P}$ projects the deviation $\hat{\boldsymbol{H}} - \bar{\boldsymbol{H}}$ onto the tangent space $T_{\bar{\mathbf{H}}}(S^8)$ at $\bar{\boldsymbol{H}}$; the uncertainty of computation is measured in the plane orthogonal to $\bar{\boldsymbol{H}}$ in $\mathcal{R}^9$. The *root-mean-square error* is given by

$$\mathrm{rms}(\hat{\boldsymbol{H}}) = \sqrt{E[\|\mathcal{P}(\hat{\boldsymbol{H}} - \bar{\boldsymbol{H}})\|^2]}. \tag{9}$$

Its domain is $0 \leqq \mathrm{rms}(\hat{\boldsymbol{H}}) \leqq 1$.

4. Theoretical Accuracy Bound

If we use bars to denote the true values, the problem is formally stated as follows:

Problem 1: Estimate a matrix $\boldsymbol{H}$ such that

$$\bar{\boldsymbol{x}}'_\alpha \times \boldsymbol{H}\bar{\boldsymbol{x}}_\alpha = \boldsymbol{0} \tag{10}$$

from noisy data $\{\boldsymbol{x}_\alpha\}$ and $\{\boldsymbol{x}'_\alpha\}$.

Applying the general theory of Kanatani [6], we obtain the theoretical accuracy bound as follows:

$$\mathcal{V}[\hat{\boldsymbol{H}}] \succ \epsilon^2\Big(\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3} \bar{W}_\alpha^{(kl)}\Big(\boldsymbol{e}^{(k)} \times \bar{\boldsymbol{x}}'_\alpha\Big) \otimes \bar{\boldsymbol{x}}_\alpha \otimes \Big(\boldsymbol{e}^{(l)} \times \bar{\boldsymbol{x}}'_\alpha\Big) \otimes \bar{\boldsymbol{x}}_\alpha\Big)_8^-, \tag{11}$$

$$\bar{\boldsymbol{W}}_\alpha = \Big(\bar{\boldsymbol{x}}'_\alpha \times \bar{\boldsymbol{H}}V_0[\boldsymbol{x}_\alpha]\bar{\boldsymbol{H}}^\top \times \bar{\boldsymbol{x}}'_\alpha + (\bar{\boldsymbol{H}}\bar{\boldsymbol{x}}_\alpha) \times V_0[\boldsymbol{x}_\alpha] \times (\bar{\boldsymbol{H}}\bar{\boldsymbol{x}}_\alpha)\Big)_2^-. \tag{12}$$

Consequently, the root-mean-square error is bounded in the form

$$\mathrm{rms}(\hat{\boldsymbol{H}}) \geqq \sqrt{\mathrm{tr}\mathcal{V}[\hat{\boldsymbol{H}}]}, \tag{13}$$

where the *trace* of a tensor $\mathcal{T} = (T_{ijkl})$ is defined by $\mathrm{tr}\mathcal{T} = \sum_{k,l=1}^{3} T_{klkl}$.

In Eq. (11), $\mathcal{T} \succ \mathcal{S}$ means that $\mathcal{T} - \mathcal{S}$ is a positive semi-definite symmetric tensor, and the operation $(\,\cdot\,)_r^-$ denotes the (Moore-Penrose) generalized inverse of rank r (discussed later). The product $\boldsymbol{a} \times \boldsymbol{T} \times \boldsymbol{a}$ of a vector $\boldsymbol{a} = (a_i)$ and a matrix $\boldsymbol{T} = (T_{ij})$ is a matrix whose (ij) element equals $\sum_{k,l,m,n=1}^{3} \varepsilon_{ikl}\varepsilon_{jmn}a_k a_m T_{ln}$, where ε_{ijk} denotes the Eddington epsilon, taking 1 and -1 if (ijk) is an even and odd permutation of (123), respectively, and 0 otherwise. We define $\boldsymbol{e}^{(1)} = (1, 0, 0)^\top$, $\boldsymbol{e}^{(2)} = (0, 1, 0)^\top$, and $\boldsymbol{e}^{(3)} = (0, 0, 1)^\top$.

For a tensor $\mathcal{T} = (T_{ijkl})$ and a matrix $\boldsymbol{A} = (A_{ij})$, we say that $\boldsymbol{A}$ is an *eigenmatrix* of $\mathcal{T}$ for *eigenvalue* λ if $\mathcal{T}\boldsymbol{A} = \lambda\boldsymbol{A}$, where the product $\mathcal{T}\boldsymbol{A}$ is a matrix whose (ij) element is $\sum_{k,l=1}^{3} T_{ijkl}A_{kl}$. The eigenmatrices and eigenvalues of a tensor can be computed by identifying a tensor and a matrix with a 9×9 matrix and a 9-dimensional vector, respectively [6].

A tensor $\mathcal{T} = (T_{ijkl})$ is said to be *symmetric* if $T_{ijkl} = T_{klij}$. A symmetric tensor has nine real eigenvalues $\{\lambda_i\}$. The corresponding eigenmatrices $\{\boldsymbol{U}_i\}$ can be chosen to be an orthogonal system of matrices of unit norm, where the inner product of matrices $\boldsymbol{A} = (A_{ij})$ and $\boldsymbol{B} = (B_{ij})$ is defined by $(\boldsymbol{A}; \boldsymbol{B}) = \sum_{i,j=1}^{3} A_{ij} B_{ij}$. A symmetric tensor is *positive semi-definite* if its eigenvalues are all nonnegative.

Let $\lambda_1 \geqq \cdots \geqq \lambda_9$ ($\geqq 0$) be the eigenvalues of a positive semi-definite symmetric tensor $\mathcal{T}$, and $\{\boldsymbol{U}_1, ..., \boldsymbol{U}_9\}$ the corresponding orthonormal set of eigenmatrices of unit norm. If $\lambda_r > 0$, the (*Moore-Penrose*) *generalized inverse* of $\mathcal{T}$ of rank r is computed as follows [6]:

$$\Big(\mathcal{T}\Big)_r^- = \sum_{i=1}^{r} \frac{\boldsymbol{U}_i \otimes \boldsymbol{U}_i}{\lambda_i}. \tag{14}$$

5. Maximum Likelihood Estimation

It can be shown [6] that the theoretical accuracy bound can be attained in the first order (i.e., ignoring terms of $O(\epsilon^4)$) by minimizing the squared *Mahalanobis distance*

$$\begin{aligned} J = &\sum_{\alpha=1}^{N} (\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}}_\alpha, V_0[\boldsymbol{x}_\alpha]_2^-(\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}}_\alpha)) \\ &+ \sum_{\alpha=1}^{N} (\boldsymbol{x}'_\alpha - \bar{\boldsymbol{x}}'_\alpha, V_0[\boldsymbol{x}'_\alpha]_2^-(\boldsymbol{x}'_\alpha - \bar{\boldsymbol{x}}'_\alpha)) \end{aligned} \tag{15}$$

subject to the constraint (10). We denote the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ by $(\boldsymbol{a}, \boldsymbol{b})$. Using Lagrange multipliers and introducing first order approximation, we can eliminate the constraint (10) and express Eq. (15) in the form

$$J = \sum_{\alpha=1}^{N} (\boldsymbol{x}'_\alpha \times \boldsymbol{H}\boldsymbol{x}_\alpha, \boldsymbol{W}_\alpha(\boldsymbol{x}'_\alpha \times \boldsymbol{H}\boldsymbol{x}_\alpha)), \tag{16}$$

where $\boldsymbol{W}_\alpha$ is the matrix obtained by replacing, in Eq. (12), the true values $\bar{\boldsymbol{x}}_\alpha$, $\bar{\boldsymbol{x}}'_\alpha$, and $\bar{\boldsymbol{H}}$ by the data $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ and the variable $\boldsymbol{H}$, respectively. Let $\hat{J}$ be the residual, i.e., the minimum of the function J. It can be shown that $\hat{J}/\epsilon^2$ is subject to a χ^2 distribution with $2(N-4)$ degrees of freedom to a first approximation [6]. Hence, an unbiased estimator of ϵ^2 is obtained in the form

$$\hat{\epsilon}^2 = \frac{\hat{J}}{2(N-4)}. \tag{17}$$

Since the resulting solution $\hat{\boldsymbol{H}}$ is optimal, we can evaluate its covariance tensor to a first approximation by replacing, in the right-hand side of Eq. (11), the true values $\bar{\boldsymbol{x}}_\alpha$, $\bar{\boldsymbol{x}}'_\alpha$, and $\bar{\boldsymbol{H}}$ by the data $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ and the solution $\hat{\boldsymbol{H}}$, respectively. The square noise level ϵ^2 is replaced by the estimate (17).

The eigenmatrix $\boldsymbol{U}_{\max}$ of the resulting covariance tensor $\mathcal{V}[\hat{\boldsymbol{H}}]$ for the maximum eigenvalue $\lambda_{\max}$ indicates the orientation in the 9-dimensional parameter space $\mathcal{R}^9$ along which error is most likely to occur; $\lambda_{\max}$ is the variance along it. We perturb the solution $\hat{\boldsymbol{H}}$ along that orientation in both directions by the standard deviation $\sqrt{\lambda_{\max}}$ and normalize the resulting values, which we call *primary deviation pair* and view as typical instances of perturbation:

$$\begin{aligned} \boldsymbol{H}^{(+)} &= N[\hat{\boldsymbol{H}} + \sqrt{\lambda_{\max}}\boldsymbol{U}_{\max}], \\ \boldsymbol{H}^{(-)} &= N[\hat{\boldsymbol{H}} - \sqrt{\lambda_{\max}}\boldsymbol{U}_{\max}]. \end{aligned} \tag{18}$$

The operator $N[\,\cdot\,]$ denotes normalization to unit norm.

6. Renormalization

We apply a technique called *renormalization* [6], which is described as follows:

1. Let $c = 0$ and $\boldsymbol{W}_\alpha = \boldsymbol{I}$, $\alpha = 1, ..., N$.
2. Define the following tensor $\mathcal{M}$:
$$\begin{aligned} \mathcal{M} = &\frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3} W_\alpha^{(kl)} (\boldsymbol{e}^{(k)} \times \boldsymbol{x}'_\alpha) \otimes \boldsymbol{x}_\alpha \\ &\otimes (\boldsymbol{e}^{(l)} \times \boldsymbol{x}'_\alpha) \otimes \boldsymbol{x}_\alpha. \end{aligned} \tag{19}$$
3. Compute the following tensor $\mathcal{N} = (N_{ijkl})$:
$$\begin{aligned} N_{ijkl} = &\frac{1}{N}\sum_{\alpha=1}^{N}\sum_{m,n,p,q=1}^{3} \varepsilon_{imp}\varepsilon_{knq} W_\alpha^{(mn)} \\ &\Big(V_0[\boldsymbol{x}_\alpha]_{jl} x'_{\alpha(p)} x'_{\alpha(q)} + V_0[\boldsymbol{x}'_\alpha]_{pq} x_{\alpha(j)} x_{\alpha(l)}\Big). \end{aligned} \tag{20}$$
4. Compute the nine eigenvalues $\lambda_1 \geqq \cdots \geqq \lambda_9$ of tensor
$$\hat{\mathcal{M}} = \mathcal{M} - c\mathcal{N} \tag{21}$$
and the corresponding orthonormal system of eigenmatrices $\{\boldsymbol{H}_1, ..., \boldsymbol{H}_9\}$ of unit norm.
5. If $\lambda_9 \approx 0$, stop. Else, update c and $\boldsymbol{W}_\alpha$ in the following way and go back to Step 2:
$$c \leftarrow c + \frac{\lambda_9}{(\boldsymbol{H}_9; \mathcal{N}\boldsymbol{H}_9)}, \tag{22}$$
$$\begin{aligned} \boldsymbol{W}_\alpha \leftarrow &\Big(\boldsymbol{x}'_\alpha \times \boldsymbol{H}_9 V_0[\boldsymbol{x}_\alpha]\boldsymbol{H}_9^\top \times \boldsymbol{x}'_\alpha \\ &+ (\boldsymbol{H}_9\boldsymbol{x}_\alpha) \times V_0[\boldsymbol{x}'_\alpha] \times (\boldsymbol{H}_9\boldsymbol{x}_\alpha)\Big)_2^-. \end{aligned} \tag{23}$$

In the above procedure, $\boldsymbol{I}$ denotes the unit matrix, and $x_{\alpha(i)}$ and $x'_{\alpha(i)}$ denote the ith components of vectors $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$, respectively. The symbol $W_\alpha^{(kl)}$ denotes the (kl) element of the matrix $\boldsymbol{W}_\alpha$, and $V_0[\boldsymbol{x}_\alpha]_{ij}$ and $V_0[\boldsymbol{x}'_\alpha]_{ij}$ denote the (ij) elements of the normalized covariance matrices $V_0[\boldsymbol{x}_\alpha]$ and $V_0[\boldsymbol{x}'_\alpha]$, respectively.

7. Accuracy and Efficiency

Figure 1 shows two synthetic images of a grid pattern. We evaluated the accuracy and efficiency of our algorithm by adding random Gaussian noise of standard deviation σ (pixels) to the coordinates of the grid points independently and computing the homography between them by using the default noise model.

Figure 2(a) plots the root-mean-square error computed for each σ by repeating the computation 50 times with different noise. The symbol ● is for our method, and the dotted lines indicate the theoretical lower bound given by Eq. (13). The symbol ◇ is for the least-squares method (also called the *algebraic-distance minimization*) minimizing

$$\sum_{\alpha=1}^{N} \|\boldsymbol{x}'_\alpha \times \boldsymbol{H}\boldsymbol{x}_\alpha\|^2. \tag{24}$$

Since this is a quadratic form in $\boldsymbol{H}$, the solution is immediately obtained by eigenvalue analysis. For this simplicity, this method has been widely used. However, it can be proved that the least-squares solution has statistical bias while the renormalization solution has the same order of accuracy as the maximum likelihood solution [6]. Indeed, we can confirm from Fig. 2(a) that not only does our method always perform better than the least-squares method but that it also attains the theoretical lower bound.

Figure 2(b) plots the corresponding average computation time in seconds on a Sun Ultra-1 workstation (Solaris 2.5.1): the solid lines are for our algorithm and the dashed lines are for the least-squares method. The least-squares method has the advantage if speed has priority at the cost of accuracy.

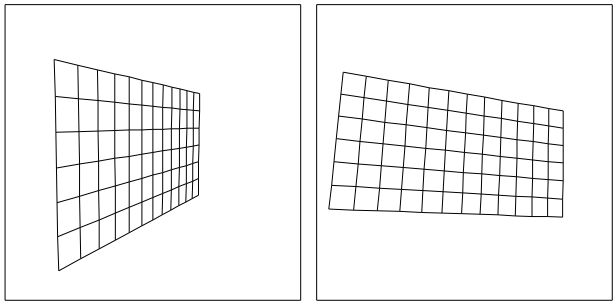


Fig. 1 Synthetic images of a grid pattern.

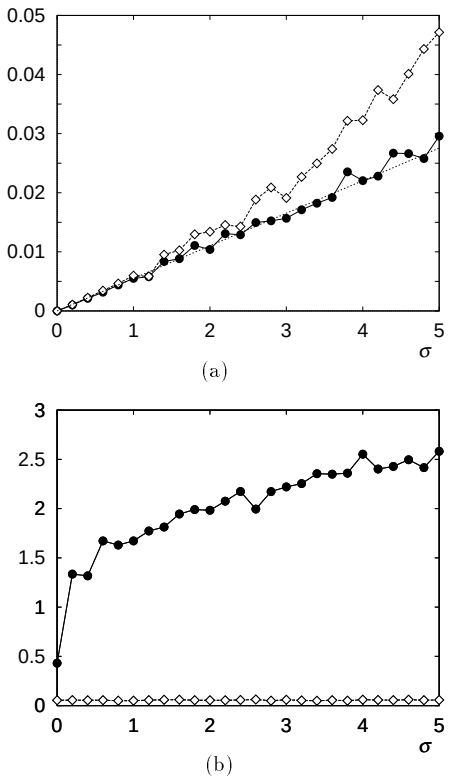
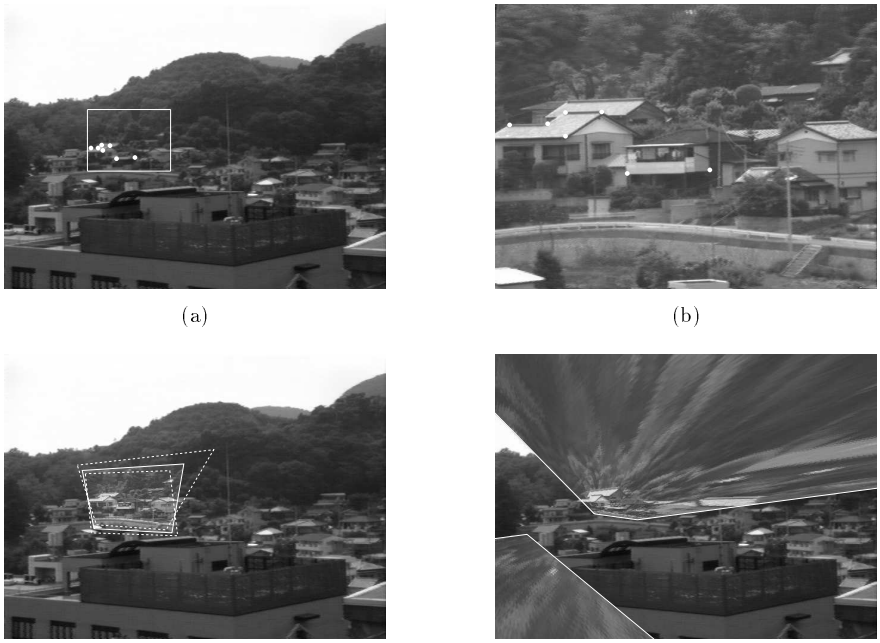


Fig. 2 (a) Root-mean-square errors for our algorithm (●), the least-squares method (◇), and the theoretical lower bound (dotted lines). (b) Computation time (seconds) for our algorithm (●) and the least-squares method (◇).

8. Instability of Homography

Figure 3(a) is a real image of an outdoor scene. Figure 3(b) is a zoomed image of the same scene. We manually selected seven matching points marked in the images and computed the homography by using the default noise model. Figure 3(c) shows the mapping of Fig. 3(b) into Fig. 3(a) via the computed homography (white solid lines) and its primary deviation pair (white dashed lines). Figure 3(d) shows the corresponding result by least squares; its primary deviation pair is not defined. Figure 3(c) tells us that the upper-right part of the mapped frame is most uncertain. In fact, the least-squares solution in Fig. 3(d) is such that the upper-right part infinitely extends to upper right and reappears from lower left. In contrast, our algorithm yields a reasonably accurate solution even in such a degenerate configuration.

This type of instability of homography often occurs in creating a mosaic of images with small overlaps between them, since matching points are restricted in very narrow strips in the images. Figure 4(a) shows images of an outdoor scene. Figure 5(a) shows a mosaiced image obtained by optimally computing the homography from the six matching points marked in the images. Figure 5(b) is the corresponding result obtained by least squares. This is an exaggerated example, since more matching points can be available for this particular image pair. In general, however, we can clearly see that a small error can result in a large image distortion if only a small number of matching points concentrated in a small region are available. In such a case, the accuracy of homography computation critically affects the quality of the mosaiced image, and our algorithm best serves such a purpose.



(a) (b)

(c) (d)

Fig. 3 (a) A real image of an outdoor scene and selected seven matching points. (b) A zoomed image of the same scene; it corresponds to the white frame in (a). (c) The image mapping (solid lines) and its primary deviation pair (dashed lines) computed by our algorithm. (d) The image mapping computed by least squares.

Fig. 4 Outdoor scene.

(a) (b)

Fig. 5 (a) Mosaicing by an optimal homography. (b) Mosaicing by least squares.

9. Concluding Remarks

This paper has described an algorithm for optimally computing the homography between two images from point matches. First, we derived a theoretical accuracy bound based on a mathematical model of image noise. Using the renormalization technique, we then presented an algorithm that theoretically attains the accuracy bound in the first order; its C++ program is avaliable via the Web[†]. By random noise simulation, we have confirmed that our algorithm effectively attains that bound. Applying our technique to mosaicing of images with small overlaps, we have shown how our algorithm reduces the instability of the image mapping.

This paper has focused on accuracy, but in practice mismatch detection is also an important issue to be studied separately [12].

Acknowledgments

The authors thank Yoshiyuki Shimizu of Sharp, Ltd. for helping them implementing a public-domain program of our algorithm. They also thank Long Quan and Bill Triggs of INRIA Rône Alpes for helpful discussions.

References

[1] A.Criminisi, I.Reid and A.Zisserman, "Duality, rigidity and planar parallax," Proc. 5th Euro. Conf. Comput. Vision, vol.2, pp.846–861, Freiburg, Germany, June 1998,

[2] D.Q. Huynh, R.A. Owens and P.E. Hartman, "Calibrating a structured light stripe system: A novel approach," Proc. Int. J. Comput. Vision, vol.33, no.1, pp.73–86, 1999.

[3] M. Irani and P. Anandan, "Parallax geometry of pairs of points for 3D scene analysis," Proc. 4th Euro. Conf. Comput. Vision, vol.1, pp.17–30, Cambridge, U.K., April 1996.

[4] M. Irani, P. Anandan and D. Weinshall, "From reference frames to reference planes: Multi-view parallax geometry and applications," Proc. 5th Euro. Conf. Comput. Vision, vol.2, pp.829–845, Freiburg, Germany, June 1998.

[5] K. Kanatani, Geometric Computation for Machine Vision, Oxford University Press, Oxford, 1993.

[6] K. Kanatani, Statistical Optimization for Geometric Computation: Theory and Practice, Elsevier Science, Amsterdam, 1996.

[7] K. Kanatani, "Cramer-Rao lower bounds for curve fitting," Graphical Models Image Process., vol.60, no.2, pp.93–99, 1998.

[8] K. Kanatani and S. Takeda, "3-D motion analysis of a planar surface by renormalization," IEICE Trans. Inf. & Syst., vol.E78-D,no.8, pp.1074–1079, 1995.

[9] H.C. Longuet-Higgins, "The reconstruction of a plane surface from two perspective projections," Proc. Roy. Soc. Lond., ser.B, vol.227, pp.399–410, 1986.

[10] H.S. Sawhney, S.Hsu and R. Kumar, "Robust video mosaicing through topology inference and local to global alignment," Proc. 5th Euro. Conf. Comput. Vision, vol.2, pp.103–119, Freiburg, Germany, June 1998.

[11] R. Szeliski and H.-U. Shum, "Creating full view panoramic image mosaics and environment maps," Proc. SIGGRAPH'97, pp.251–258, Los Angeles, CA, U.S.A., August 1997.

[12] P. Torr and A. Zisserman, "Robust computation and parameterization of multiple view relations," Proc. 6th Int. Conf. Comput. Vision, pp.727–732, Bombay, India, January 1998.

[13] B. Triggs, "Autocalibration from planar surfaces," Proc. 5th Euro. Conf. Comput. Vision, vol.1, pp.89–105, Freiburg, Germany, June 1998

[14] T. Viéville, C. Zeller and L. Robert, "Using collineations to compute motion and structure in an uncalibrated image sequences," Int. J. Comput. Vision, vol. 20, no.3, pp.213–242, 1996.

[15] J. Weng, N. Ahuja and T.S. Huang, "Motion and structure from point correspondences with error estimation: Planar surfaces," IEEE Trans. Patt. Anal. Mach. Intell., vol.39, no.12, pp.2691–2717, 1991.

[16] I. Zoghlami, O. Faugeras and R. Deriche, "Using geometric corners to build a 2D mosaic from a set of images," Proc. IEEE Conf. Comput. Vision Patt. Recog., pp.420–425, Puerto Rico, June 1997.

[†] http://www.ail.cs.gunma-u.ac.jp/~kanatani/e/

Kenichi Kanatani received his Ph.D. in applied mathematics from the University of Tokyo in 1979. He is currently Professor of computer science at Gunma University. He is the author of *Group-Theoretical Methods in Image Understanding* (Springer, 1990), *Geometric Computation for Machine Vision* (Oxford, 1993) and *Statistical Optimization for Geometric Computation: Theory and Practice* (Elsevier, 1996).

Naoya Ohta received his Ph.D. from the University of Tokyo in 1998. He engaged in research and development of image processing systems at the Pattern Recognition Research Laboratories of NEC. He is currently Associate Professor of computer science at Gunma University. He was a research affiliate of the Media Laboratory in MIT from 1991 to 1992.

Yasuhi Kanazawa received his Ph.D. from Osaka Univeristy in 1997. He is currently Lecturer at the Department of Knowledge-based Information Engineering, Toyohashi University of Technology. His research interests include image processing and computer vision.