Organizer/Lecturer
Marc Pollefeys
Katholieke Universiteit Leuven

54

Obtaining 3D Models With a Hand-Held Camera

Course Notes

The 29th International Conference on
Computer Graphics and Interactive Techniques

Henry B. Gonzalez Convention Center
San Antonio, Texas USA

Conference: 21–26 July 2002
Exhibition: 23–25 July 2002

SAN ANTONIO
SIGGRAPH
2002

Obtaining 3D Models With a Hand-Held Camera

Marc Pollefeys
Center for Processing of Speech and Images
University of Leuven, Belgium

Lecture Notes SIGGRAPH Course
Wednesday 24 July, 2002

Acknowledgments

At this point I would like to thank the people who contributed directly or indirectly to the text for this tutorial. First of all I would like to express my gratitude towards professor Luc Van Gool, head of the Industrial Image Processing group (VISICS/PSI/ESAT) of the K.U.Leuven. I am very grateful to my colleagues Maarten Vergauwen and Kurt Cornelis who are not only doing great research work (of which you'll find many examples in this text), but also proofread the whole text. Next I would also like to thank Frank Verbiest, Jan Tops, Joris Schouteden, Reinhard Koch, Tinne Tuytelaars and Benno Heigl who have contributed to the work presented in this text. I would also like to acknowledge the Fund for Scientific Research - Flanders (Belgium) for granting me a Postdoctoral Fellowship. The financial support of the FWO project G.0223.01 and the European IST projects InViews, Attest, Vibes and Murale are also gratefully acknowledged.

Notations

To enhance the readability the notations used throughout the text are summarized here.

For matrices bold face fonts are used (i.e. **A**). 4-vectors are represented by $\mathtt{A}$ and 3-vectors by $\mathtt{a}$. Scalar values will be represented as a.

Unless stated differently the indices i, j and k are used for views, while l and m are used for indexing points, lines or planes. The notation $\mathbf{A}_{ij}$ indicates the entity **A** which relates view i to view j (or going from view i to view j). The indices i, j and k will also be used to indicate the entries of vectors, matrices and tensors. The subscripts P, A, M and E will refer to projective, affine, metric and Euclidean entities respectively

$\mathbf{P}$	camera projection matrix (3×4 matrix)
$\mathtt{M}$	world point (4-vector)
Π	world plane (4-vector)
$\mathtt{m}$	image point (3-vector)
$\mathtt{l}$	image line (3-vector)
$\mathbf{H}^{\Pi}_{ij}$	homography for plane Π from view i to view j (3×3 matrix)
$\mathbf{H}_{\Pi i}$	homography from plane Π to image i (3×3 matrix)
$\mathbf{F}$	fundamental matrix (3×3 rank 2 matrix)
$\mathtt{e}_{ij}$	epipole (projection of projection center of viewpoint i into image j)
$\mathbf{T}$	trifocal tensor ($3 \times 3 \times 3$ tensor)
$\mathbf{K}$	calibration matrix (3×3 upper triangular matrix)
$\mathbf{R}$	rotation matrix
Π_∞	plane at infinity (canonical representation: $W = 0$)
Ω	absolute conic
	(canonical representation: $X^2 + Y^2 + Z^2 = 0$ and $W = 0$)
Ω^*	absolute dual quadric (4×4 rank 3 matrix)
ω_∞	absolute conic embedded in the plane at infinity (3×3 matrix)
ω^*_∞	dual absolute conic embedded in the plane at infinity (3×3 matrix)
ω	image of the absolute conic (3×3 matrices)
ω^*	dual image of the absolute conic (3×3 matrices)
$\sim$	equivalence up to scale ($A \sim B \Leftrightarrow \exists\lambda \neq 0 : A = \lambda B$)
$\|\mathbf{A}\|_F$	indicates the Frobenius norm of **A** (i.e. $\sum_{ij} a_{ij}^2$)
$\mathcal{F}(\mathbf{A})$	indicates the matrix **A** scaled to have unit Frobenius norm
	(i.e. $\frac{\mathbf{A}}{\|\mathbf{A}\|_F}$)
$\mathbf{A}^\top$	is the transpose of **A**
$\mathbf{A}^{-1}$	is the inverse of **A** (i.e. $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$)
$\mathbf{A}^\dagger$	is the Moore-Penrose pseudo inverse of **A**

Contents