Omni-Rig Sensors: What Can be Done With a Non-Rigid Vision Platform?

Amnon Shashua
Institute of Computer Science
Hebrew University of Jerusalem
Jerusalem 91904 ISRAEL
http://www.cs.huji.ac.il/∼ shashua/

Abstract

We describe the principles of building a moving vision platform (a Rig) that once calibrated can thereon self-adjust to changes in its internal configuration and maintain an Euclidean representation of the 3D world using only projective measurements. Formally, we address the question of how to obtain an invariant 3D projective representation from a dynamic collection of cameras.

We show that the maximal generality is reached when the rig consists of 5 cameras whose center of projection remain fixed relative to each other during the motion. In other words, the non-rigid component motion may consist of change of internal parameters and relative camera orientations. The configuration reduces to 3 views when the Rig is built using a single physical camera with half-mirrors (beam-splitters) for creating 3 distinct views. We also briefly discuss 2-view configurations using half-mirrors and the principle behind adapting the configuration to allow for zoom lenses in the system.

The new research paradigm on non-rigid rigs (we term "Omni-Rig") is applicable to the design of Vision-based sensors that after calibration can move in space while changing critical elements of their configuration — such as changing focus on the fly, zoom, relative camera orientation and inclination of focal plane to object's surface orientation — without the need for recalibration, i.e., using only projective calculations throughout its motion.

1 Introduction

The projective framework of Structure from Motion (SFM), also known as "3D-from-2D" problem, is supported by a relatively large body of literature on the techniques for taking matching image features (points and lines) across multiple views produce a projective representation of the three-dimensional (3D) positions of the features in space (the special cases of 2,3 views are theoretically interesting, for a review see [6, 18]). There are tasks, such as visual recognition by alignment, image transfer in Photogrammetry and image-based rendering in Computer Graphics in which a projective representation is sufficient (for example, [17, 2, 12]). However, in mensuration, navigation and robot guidance tasks an Euclidean representation is necessary — and to that end a growing body of literature is dedicated to the problem of obtaining an Euclidean representation [7, 4, 11, 22, 1]. We divide the various approaches into the following classes:

- **Euclidean Input:** In the most straightforward realization of this approach one requires a minimal set of 5 control points whose positions in the desired Euclidean space are known. This approach is obvious due to the fact that a projective basis is defined by 5 points and that the projective group includes the Euclidean group as a subgroup. A more advanced variation on this theme is to recover the projective to Euclidean ("Proj-2-Euc") transformation from Euclidean cues, such as from known angles between lines (such as orthogonal lines) and known distances between points [4].
- **Fixed Internal Parameters Assumption:** Here the goal is to recover the internal parameters of the camera *explicitly* by making the assumption of a single moving camera whose internal parameters are unknown but remain fixed throughout. These methods are based on recovering invariants to Euclidean transformations, notably the absolute conic and absolute quadric [7, 15, 19] (and [1, 13, 9] in the case of restricted camera motion). More recent approaches [11, 22] assume a stereo rig (of two cameras) moving rigidly (i.e., the relative orientation between the two cameras remains fixed) with fixed (yet unknown) internal parameters.
- **Varying Subset of Internal Parameters:** Using the machinery developed for recovering the absolute quadric [19, 10], one can tolerate variation in

a subset of the internal parameters as the number of views increases beyond three. For example, by counting arguments one can show that in the absence of skew 8 views are required for solving for the remaining 4 internal parameters per camera. This idea has been so far described fully and implemented successfully for purely varying focal length [16]. Due to the highly non-linear equations involved, the implementation of these approaches are most suited for relatively large sequences of views, as reported in [16].

In this paper, we pose the question of Projective-to-Euclidean relationship differently. Instead of asking how to calibrate (i.e. recover the internal parameters) a rig of cameras, we ask under what conditions can we *preserve* a fixed projective representation of the 3D space under a non-rigid motion of the rig of cameras?

The relationship to Euclidean reconstruction is clear: if we are able to preserve a fixed projective representation of the 3D space under non-rigid motion, then we need to recover the Projective-to-Euclidean transformation only once and it would remain valid from there on even though the rig has gone out of calibration (i.e. it's internal configuration has changed — whether voluntarily or due to instabilities of the rig's internal structure).

To appreciate the difference between the two questions, consider the following thought experiment. We are given a rig of $x \geq 2$ cameras placed in room $\mathcal{A}$. Assume that we have obtained somehow the Proj-2-Euc transformation, either through direct input of Euclidean information coming from the features of the room, or through one of the self-calibration methods (by moving the rig rigidly while assuming that the internal parameters remain fixed). Clearly, as long as the rig moves rigidly in space we can obtain an Euclidean representation of space because the projective representation is fixed (or can be made fixed), thus the Proj-2-Euc transformation remains valid. Also clearly, if we apply a non-rigid motion, i.e., change the relative positioning (and internal parameters) among the cameras, but remain in the same room, then by using the overlap between the features of $\mathcal{A}$ seen before and after the non-rigid change one can chain the projective-to-projective transformations together (since they form a group) to obtain back the Euclidean representation.

However, consider placing the rig in a different room $\mathcal{B}$, such that there is no overlap between the features of $\mathcal{A}$ and the features of $\mathcal{B}$. Once the rig is in $\mathcal{B}$, apply a non-rigid change to the rig and take a *single* snap-shot of the room. *Can we recover an Euclidean representation of $\mathcal{B}$ using only projective calculations, i.e., without using any Euclidean input from $\mathcal{B}$?*

Note that existing approaches for self-calibration are not designed to handle such a situation. In room $\mathcal{B}$ we do not have the freedom to take multiple snapshots while applying a rigid motion to the rig, nor can we assume that the internal parameters remain fixed (because we have a multiple camera situation), and since there is no overlap with the previous room $\mathcal{A}$ we cannot chain together projective transformations. Yet, the question we have posed is *less ambitious* than the self-calibration paradigm because we do allow for the possibility of using Euclidean input from room $\mathcal{A}$ in order to calibrate once (and for all) the Proj-2-Euc transformation.

The question of restoring the Proj-2-Euc transformation from dynamically changing projective representations (due to non-rigid changes of the rig) has an interesting practical source. A constructive answer to the question above provides means for designing a rig of cameras that can change its internal configuration, such as focus, zoom, and relative rotation, during mensuration. The only requirement is that the rig be calibrated once (i.e., by obtaining the Proj-2-Euc transformation), then during mensuration the field of view and the working distance can change considerably without the need for Euclidean input or overlap between image input taken prior to the change of camera configuration.

Our working assumption is that from the application point of view it is relatively manageable to start with a known Proj-2-Euc (by simply taking a snapshot of a calibration object prior to the mensuration process). Yet, it is unrealistic to expect for Euclidean input each time the camera configuration has changed, and it may be unrealistic (or very restrictive) to ask for matching features across images before and after the change of rig configuration (for example, change of field of view is large, or the rig is positioned far a way from its previous location, etc.).

1.1 Formal Statement of the Problem

Our problem formulation is strictly projective: Let $\mathcal{R}$ be the projective representation of 3D space projected onto $x \geq 2$ 2D views modeled by a set of $x \geq 2$ uncalibrated pin-hole cameras. Let the set of x cameras undergo some non-rigid (projective) transformation $\mathcal{G}$. What should be $x, \mathcal{G}$ such that the projective representation $\mathcal{R}$ of 3D space is preserved?

We will show that $\mathcal{G}$ can be at most an arbitrary 2D projective transformation *per* image plane — i.e., change of internal parameters and camera orientation per camera — and that for x views we have the measurements for recovering at most $2(x-1)$ parameters per image plane. Thus, a rig of 5 cameras is sufficient for tolerating an arbitrary 2D collineation per image plane (more than 5 cameras would not add anything fundamental), and 3 cameras are sufficient for tolerating an arbitrary change in the internal parameters (focal length, aspect ratio and principle point) per camera.

We will then consider Rig designs that allow for a single physical camera with half-mirrors for creating multiple views. We show that we have $2x(x-1)$ measurements for x views, thus the general case reduces to 3 views (instead of 5) and the restricted case of variable internal parameters reduces to 2 views (instead of 3). We will also briefly discuss the principle behind adapting the configuration to allow for zoom lenses in the system.

We will start below with a couple of "toy" examples to illustrate the basic idea and then address the problem in its full generality.

1.2 Few "Toy" Examples

Let $x = 2$, i.e., we have a rig of two cameras, with views $p_1, p_2 \in \mathcal{P}^2$ and let F be the fundamental matrix describing the relative orientation of the two cameras, i.e., $p_2^\top F p_1 = 0$. Assume we change the internal parameters of the rig by changing the focal length and aspect ratio of each camera. Thus, $p_1' = M_1 p_1$ and $p_2' = M_2 p_2$, where M_1, M_2 are diagonal matrices $M_i = diag(f_i, g_i, 1)$, $i = 1, 2$ (note that f_i is the change of focal length and f_i/g_i is the change in aspect ratio). By substitution we obtain:

$$p_2'^\top F' p_1' = p_2^\top (M_2^\top F' M_1) p_1 = p_2^\top F p_1 = 0$$

hence, $F \cong M_2^\top F' M_1$. We are given the fundamental matrices F, F' (before and after the change in rig configuration), and we are interested in undoing the effect of changing configuration, i.e., maintain the same projective representation of $\mathcal{P}^3$ produced by the rig with fundamental matrix F. Clearly, the goal is achieved if we can recover M_1, M_2 from the equation $F \cong M_2^\top F' M_1$. It can be readily verified that the parameters $f_i, g_i, i = 1, 2$ are uniquely determined from the above equation.

Consider next an example with $x = 3$ cameras with views p_1, p_2, p_3 and Fundamental matrices F_{12}, F_{23} and F_{13}, i.e., $p_3^\top F_{23} p_2 = 0$, etc. Let the change of camera configuration be represented by three orthogonal image-plane transformations $p_i' = R_i p_i$, $i = 1, 2, 3$, i.e., $R_i^\top R_i = R_i R_i^\top = I$. Let the three new Fundamental matrices be F_{12}', F_{23}' and F_{13}' and we have the following relations: $F_{ij} \cong R_j^\top F_{ij}' R_i$, $i < j = 1, 2, 3$. Again, we are given F_{ij} and F_{ij}' and wish to preserve the projective representation of $\mathcal{P}^3$ produced by the rig represented by F_{ij} while using the rig represented by F_{ij}'. As before, the goal is clearly achieved if one can recover R_i. To see how this can be done we do the following. Note that,

$$F_{12} F_{13}^\top F_{23} \cong R_2^\top (F_{12}' F_{13}'^\top F_{23}') R_2$$

or equivalently,

$$R_2 (F_{12} F_{13}^\top F_{23}) \cong (F_{12}' F_{13}'^\top F_{23}') R_2$$

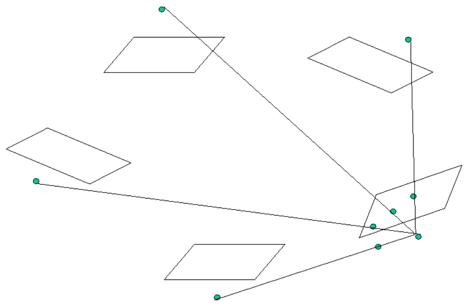

Figure 1. A sketch of the proof of Theorem 1. The centers of projection are fixed, the image planes are allowed to move. Each image plane "sees" the other 4 COPs which are its epipoles. The change in the epipole positions provide the measurements for recovering the motion of each image plane, and thus undoing the non-rigid change in the rig's configuration.

which is of the well known form of matrix equation "$XA = BX$" whose solution can be found in ([8], pp. 215). The scale in the equation can also be determined as follows[1]: $\lambda A = X^\top B X$, thus $det(\lambda A) = det(B)$ from which we obtain λ^3 and thus λ. We can set up similar matrix equations for R_1, R_2, and thus solve for R_i, $i = 1, 2, 3$.

The two examples given above[2] indicate that some information exists to undo the effect of changing camera configuration on the output representation of 3D space. We see also that the more cameras we have the more we can undo. *Our challenge, therefore, is to address the question in its generality, i.e., what transformations can be undone and with how many cameras.*

2 Preserving Euclidean Representation of a Non-Rigid Rig

The main result is summarized in the theorem below. The basic idea, sketched in Fig. 1, is that each image "sees" the camera centers (the epipoles) of the remaining $x - 1$ cameras (the epipoles can be computed from the pairwise fundamental matrices). If the center of projection (COP) of all the x cameras move rigidly in space (their relative position remains fixed with respect to each other), then

[1] Thanks to Radu Horaud.

[2] Note that these were "toy" examples since focal length never changes alone (principle point changes as well), and orthogonal assumption holds only for calibrated cameras.

the change in position of the epipolar points in a particular image is due entirely to the change of position of the image plane with respect to its camera COP. Since a general transformation of the image plane is a homography (collineation) described by 8 parameters, and since each epipolar point contributes 2 measurements, we have the total of $2(x-1)$ measurements per image plane that must satisfy $2(x-1) \geq 8$. In other words, with a rig of 5 cameras whose COP are fixed with respect to each other, the non-rigid motion that can be tolerated (i.e., by undone using only projective calculations) is an arbitrary 2D homography per image plane (i.e., each image plane can arbitrarily change its position with respect to its COP).

Theorem 1 *A set of (no less than) 5 cameras in general position undergoing a non-rigid motion that maintains the relative locations of the centers of projection of the cameras fixed with respect to each other, can maintain a fixed projective representation of the 3D scene viewed by the cameras.*

Proof: let $p_i \in \mathcal{P}^2$, $i = 1, ..., 5$ be matching points across 5 views of some 3D point, and let F_{ij} denote the fundamental matrix between views i and j, i.e., $p_j^\top F_{ij} p_i = 0$. Let F'_{ij} be the fundamental matrices of the rig following the non-rigid motion. From the conditions of the problem, the relationship between the old and new Fundamental matrices is captured by:

$$F_{ij} \cong A_j^\top F'_{ij} A_i$$

for some general 2D collineations $A_1, ..., A_5$. We wish to show that A_i, $i = 1, ..., 5$ are recoverable from the relationship above, but cannot be recovered with fewer cameras.

Let e_{ij}, also known as "epipoles", be the projection of the center of projection (COP) of camera i onto the image plane of camera j in the old configuration and likewise e'_{ij} in the new configuration (rig following non-rigid motion). It is well known that $F_{ij}^\top e_{ij} = 0$ and $F'^\top_{ij} e'_{ij} = 0$. Clearly, for a fixed j_o, the relationship between e_{ij_o} and e'_{ij_o}, $\forall i \neq j_o$, depends *solely* on A_{j_o}, i.e., $A_{j_o} e_{ij_o} \cong e'_{ij_o}$. To see why this is so, note that

$$0 = F_{ij_o}^\top e_{ij_o} = A_i^\top F'^\top_{ij_o} A_{j_o} e_{ij_o}$$

holds for all $i \neq j_o$, and because $F'^\top e'_{ij} = 0$ it follows that $A_{j_o} e_{ij_o} \cong e'_{ij_o}$.

Therefore, A_{j_o} is determined by the four matching epipoles e_{ij_o} and e'_{ij_o}, $\forall i \neq j_o$. Since a 2D collineation is uniquely determined by 4 matching points in the 2D plane, it follows that we need at least 5 cameras in order to recover A_i, $i = 1, ..., 5$ which represents the non-rigid change in rig configuration. Since the 4 matching points in each image should be in general position (no 3 are on a line), this could happen if and only if the 5 COPs of the cameras are in general position (no 3 on a line, no 4 on a plane). □

It is worthwhile noting that the number 5 being both the number of cameras we need and the number of points for a projective basis is not a coincident. In fact it is known that in a projective sense a camera is dual to a point, i.e., all what matters for reconstruction up to a projectivity can be captured solely by the COP of the camera [21, 5].

A general homography transformation includes a change in camera orientation and a change in the internal parameters (focus, aspect-ratio, skew and principle point). If only a change in internal parameters is desired, then in the absent of skew (which holds for CCD image planes) we must satisfy $2(x-1) \geq 4$, thus 3 cameras would suffice:

Corollary 1 *One can maintain a fixed projective representation of the 3D scene under changing internal parameters, in the absence of skew, with a rig of 3 cameras.*

3 Catadioptric Rigs

Consider the situation in which all the image-plane homographies associated with the non-rigid motion of the x cameras coincide. What should x be? Clearly, all the epipolar points combined contribute measurements for the single unknown homography matrix. Thus, for a general homography transformation $2x(x-1) \geq 8$ should hold, i.e., $x = 3$ cameras would suffice (modulo degenerate configurations discussed later). Likewise, for the reduced group of changing internal parameters only, we have $2x(x-1) \geq 4$, i.e., $x = 2$ cameras would suffice.

Such a concept could be easily implemented by having a single physical camera (lens system + CCD array) and mirrors (or half mirrors) for creating multiple views. Nayar [14] refers to a system with refracting and reflecting elements with a single COP as a *catadioptric camera*, thus in the same spirit we will refer to such a system with multiple COPs as a *catadioptric rig*. Fig 2 shows an example of a 2-view catadioptric rig. In this design a single image captures two (distinct) views. Alternative designs in which the multiple views are coordinated with multiple snapshots are also possible.

Unlike the regular rig (of 5 distinct cameras) for which the only requirement was on the position (general) of the COPs, with a catadioptric rig the degenerate configurations involve the position of the image planes as well. Clearly, if the three COPs are collinear then there are only 3 distinct epipoles (instead of 6) which is not sufficient for recovering a general homography (we need 4). Thus, an obvious constraint on the design is to have the COPs far from being on a line. The remaining degenerate configurations involve at least distinct 4 epipoles whose positions are not general (3 are on a line). It can be shown that these situations can

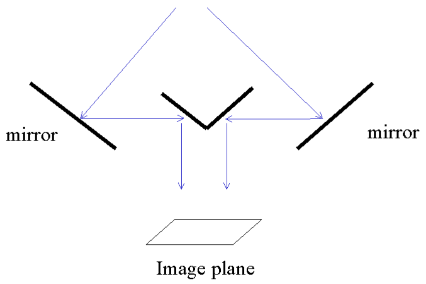

Figure 2. A sketch of a catadioptric rig creating two views (with distinct centers of projection) in a single image. The thick lines represent mirrors.

be avoided if all three image planes are not pairwise parallel (i.e., must have a pairwise vergence between cameras). Likewise, in the 2-view catadioptric rig the system is degenerate if and only if the translation between the COPs is along the rotation axis between the image planes (this includes in particular the case were the image plane are parallel).

Finally, in order to incorporate a Zoom lens in the design of a catadioptric rig, yet still maintain a fixed COP, the simplest idea is to turn the system into an orthographic projection using a telecentric aperture [20] in front of the lens, or by mounting a relay lens onto the perspective lens [3]. In an orthographic system there are further degenerate configurations to be aware of, but will not be addressed here.

4 Summary and Applications

One can view the immediate application of Theorem 1 in terms of designing a multi-camera moving platform (a "rig") whose internal configuration is variable yet produces a fixed projective representation $\mathcal{R}$ of the 3D scene. If at any time the mapping g from the representation $\mathcal{R}$ to the Euclidean representation $\mathcal{E} = g\mathcal{R}$ is given (by using control points, for example), then since $\mathcal{R}$ is fixed throughout, then so is g. Thus the rig would continue to produce an Euclidean representation of the scenes it encounters even though the internal configuration of the rig undergoes changes.

The design principles involved in the mechano-optical implementation require a variable moving CCD plane. This is in contrast to typical designs in which the CCD plane is fixed and the lens elements are mobile (for changing focus for example). In a multiple lens system (the regular design of 5 cameras) the lens elements should be fixed and the entire variability is placed on the CCD planes. In a single lens (catadioptric rig) system there is only one CCD plane whose mobility determines the range of non-rigid transformations. For general mobility a 3-view catadioptric rig is required and for purposes of changing focus alone (CCD plane moves only along the vertical axis thereby changing 4 internal parameters) a 2-view catadioptric rig is sufficient.

The variability of the rig can be viewed either as a *passive* or an *active* property. The passive aspect of a variable rig has to do with stability. Once the rig is mounted (and calibrated) the working assumption is then that the calibration holds for a certain period, say throughout a measurement cycle. This assumption requires a that relative positions between the lenses and the CCD arrays remain fixed. The requirement of rigidity in turn adds weight and cost to the rig. The application of Theorem 1 in this respect reduces the material rigidity requirement as the CCD plane(s) are decoupled from the optical elements.

The active aspect of a variable rig is that one can design the optical system such that each camera could re-focus, change the field of view, change the camera orientation or change the tilt and pan of the focal plane in order to measure slanted surfaces. The only requirement is that the variability be placed on the CCD array rather than on the lens position. In the case of 3 cameras (or 2-view catadioptric rig), each CCD array can move forward and backward (to re-focus) and also lateral movements (this could be a passive property). Thus, the design of such an optical system should fix the lenses and allow for CCD motion for purposes of re-focusing. The optical system can then tolerate lateral perturbations (as a passive property) without affecting the accuracy of the Euclidean reconstruction.

In the general case of 5 cameras (or 3-view catadioptric rig) one can allow for tilt and pan of the CCD array as well. This would allow re-focusing on slanted surfaces. Alternatively, if the lens is allowed to rotate around its optical center (not an easy design objective), then the working distance (through re-focusing and re-orientation) of the vision system can be increased due to control of overlap between scene points across the image planes.

Taken together, it is possible to design a vision platform with a dynamic range of working distance and resolution (via re-focus and zoom) with lesser demands on rigidity that allow for light-weight designs, yet without the need for re-calibration.

Acknowledgments

I wish to thank Radu Horaud on pointing out how to resolve the scale issue in Section 1.2 and Danny Albeck for inspiring me to look into catadioptric designs.

References

[1] M. Armstrong, A. Zisserman, and R.I. Hartley. Self-calibration using image triplets. In *Proceedings of the European Conference on Computer Vision*, LNCS-1064, pages 3–16. Springer, April 1996.

[2] S. Avidan and A. Shashua. Novel view synthesis by cascading trilinear tensors. *IEEE Transactions on Visualization and Computer Graphics*, 4(3), 1998. Short version can be found in CVPR'97.

[3] M. Born and E. Wolf. *Principles of Optics*. Permagon Press, 1965.

[4] B. Boufma, R. Mohr, and F. Veillon. Euclidean constraints for uncalibrated reconstruction. In *Proceedings of the International Conference on Computer Vision*, pages 466–470, Berlin, Germany, May 1993.

[5] S. Carlsson. Duality of reconstruction and positioning from projective views. In *Proceedings of the workshop on Scene Representations*, Cambridge, MA., June 1995.

[6] O.D. Faugeras. Stratification of three-dimensional vision: projective, affine and metric representations. *Journal of the Optical Society of America*, 12(3):465–484, 1995.

[7] O.D. Faugeras, Q.T. Luong, and S.J. Maybank. Camera self calibration: Theory and experiments. In *Proceedings of the European Conference on Computer Vision*, pages 321–334, Santa Margherita Ligure, Italy, June 1992.

[8] F.R. Gantmacher. *The theory of matrices*, volume I. Chelsea publishing, NY, 1990.

[9] R. Hartley. Self calibration from multiple views with a rotating camera. In *Proceedings of the European Conference on Computer Vision*, pages 471–478, Stockholm, Sweden, May 1994.

[10] A. Heyden and K. Astrom. Euclidean reconstruction from constant intrinsic parameters. In *Proceedings of the International Conference on Pattern Recognition*, pages 339–343, Vienna, 1996.

[11] R. Horaud and G. Csurka. Self-calibration and Euclidean reconstruction using motions of a stereo rig. In *Proceedings of the International Conference on Computer Vision*, pages 96–103, Bombay, India, January 1998.

[12] S. Leveau and O.D. Faugeras. 3-D scene representation as a collection of images and fundamental matrices. Technical report, INRIA, Feb. 1994.

[13] T. Moons, L. Van Gool, M. Van Diest, and E. Pauwels. Affine reconstruction from perspective image pairs. In *The 2nd European Workshop on Invariants*, Ponta Delagada, Azores, October 1993.

[14] S.K. Nayar and S. Baker. Catadioptric image formation. In *Proceedings of the DARPA Image Understanding Workshop*, pages 1431–1437, New Orleans, LA, May 1997.

[15] M. Pollefeys and L. Van Gool. A stratified approach to metric self-calibration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 407–412, Puerto Rico, June 1997. IEEE Computer Society Press.

[16] M. Pollefeys, R. Koch, and L. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown camera parameters. In *Proceedings of the International Conference on Computer Vision*, Bombay, India, January 1998.

[17] A. Shashua. Algebraic functions for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):779–789, 1995.

[18] A. Shashua. Trilinear tensor: The fundamental construct of multiple-view geometry and its applications. In G. Sommer and J.J. Koenderink, editors, *Algebraic Frames For The Perception Action Cycle*, number 1315 in Lecture Notes in Computer Science. Springer, 1997. Proceedings of the workshop held in Kiel, Germany, Sep. 1997.

[19] B. Triggs. Autocalibration and the absolute quadric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 609–614, Puerto Rico, June 1997. IEEE Computer Society Press.

[20] M. Watanabe and S.K. Nayar. Telecentric optics for computational vision. In *Proceedings of the DARPA Image Understanding Workshop*, pages 781–785, New Orleans, LA, May 1997.

[21] D. Weinshall, M. Werman, and A. Shashua. Shape tensors for efficient and learnable indexing. In *Proceedings of the workshop on Scene Representations*, Cambridge, MA., June 1995.

[22] A. Zisserman, P.A. Beardsley, and I.D. Reid. Metric calibration of a stereo rig. In *Proc. of IEEE Workshop on Representation of Visual Scenes*, pages 93–100, Cambridge, MA, June 1995.