Exploiting Homologies in Global Calibration of a Multi-imager Array

H. Harlyn Baker
Hewlett-Packard Laboratories
Palo Alto, CA 94024

Zeyu Li
University of California
Berkeley, CA 94720

Constantin Papadas
ISD, SA
Athens, Greece

ABSTRACT

Advances in building high-performance camera arrays have opened the opportunity—and challenge—of using these devices for synchronized 3D and multi-viewpoint capture [2]. A requirement of using camera arrays for metric work is that their relative poses be known. With a structured array—that is, one whose imagers' intrinsics and relative positions do not change—it is feasible to perform this calibration once, before any use. We present progress in developing a new approach to this calibration that capitalizes on high quality homographies between pairs of imagers to develop a global optimal solution delivering epipoles and fundamental matrices simultaneously for the entire system. The method exploits what we identify as the *Rank-One-Perturbation-of-Identity* (ROPI) structure of homologies in posing a unified SVD estimator for the parameters. We summarize the theory, and present both qualitative depictions and quantitative assessments of our results.

Index Terms—multi-imager calibration, camera array, homography, ROPI, epipole, fundamental matrix

1. INTRODUCTION

1.1. Multi-imager Camera System

In earlier work, we presented a PC-based multi-imager camera array based on cell-phone-type inexpensive imagers, employed for the task of creating megapixel video mosaicked panoramas [2]. This paper reports further research in exploiting this camera system.

Named after an amphitheater at the base of the Parthenon, the Herodion camera is a high performance capture system built around a direct memory access (DMA) PCI interface. It streams Bayer-format video to PC memory through a three-layer tree structure: an imager layer, a concentrator layer, and a frame grabber layer. Up to 24 imagers, grouped in 6's, are attached to leaf concentrators. Two leaf concentrators connect to a middle concentrator, up to two of which can connect to the PCI-bus frame grabber. Different configurations are supported, down to a single imager. Details of the camera system can be found in that earlier publication. In distinction to others [8], these data are uncompressed, synchronized at the pixel level, and running into a single PC. Redesign for PCI-X (for greater than 8Gbps bandwidth) is underway, which will give us up to 96 synchronized VGA streams (or about 14 HD streams). To support community developments in these areas, we have licensed the imaging system for commercial sale through our system developers [12].

Several camera arrays have been designed to application requirements (Figure 1), with pixel distribution set to meet imaging needs.

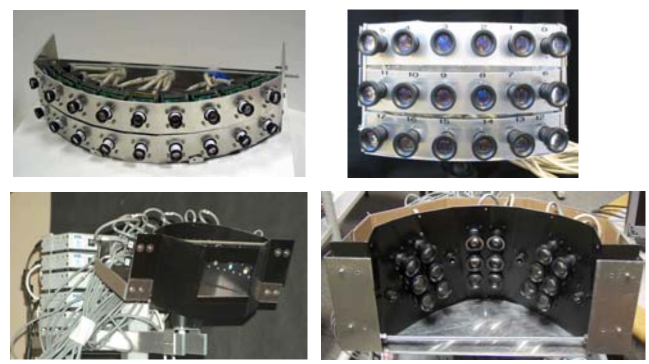

Figure 1: Multi-imager camera arrays.

As mentioned, initial developments with this camera system have centered on building real-time megapixel video mosaics for videoconferencing. Some of these imagers may have fields of view that have considerable overlap with others' (see Figure 3). In these regions, we wish to exploit the multi-perspective imaging by using stereo processing to extract range.

Working in tasks of videoconferencing and document/ artifact imaging, we use planar homographies to control image composition. Our interests have now moved to using similar multi-imager configurations to provide, in addition, scene range, and this has turned our attention to upgrading our quite-accurate homographies to derive fundamental matrices.

1.2. Line-based Plane Homographies

Traditional 3D calibration has used point correspondence to solve for the fundamental matrix between a pair of cameras. We were interested, instead, in building upon our good homographies (refined through a global bundle adjustment) and achieving a calibration that incorporated the information from all imagers simultaneously.

Planar calibration targets provide an effective mechanism for determining projective relationships between a pair of cameras [3,6,7,10,11]. We used patterns of lines projected to a facing wall as the calibration features in our mosaicking work, and determined a globally optimal set of homographies that exploited all observations of all imagers. This ensured, for example, that lines straight in the scene would be straight in the mosaic—even if the observing imagers had no overlap. Homographies are related to the fundamental matrix through imager epipoles [4]:

$$F \cong [e']_\times H$$

If we were able to determine the epipoles among all of these imagers, then we could provide their 3D calibration, and this would permit us to compute range using, for example, epipolar-line stereo matching. Our homographies can provide us with these epipoles.

Consider $\mathbf{H}_1$ in Figure 2, formed between imager I and imager I' using the calibration plane π_1. If we reposition the imagers (or, equivalently, reposition the plane, as shown), then we can determine $\mathbf{H}_2$, relating I and I' using the calibration plane π_2. We can form a new transform ${}^{2,1}\mathbf{H} = \mathbf{H}_2^{-1}\mathbf{H}_1$, by observing that $\mathbf{H}_1$ maps points in I to I' through π_1 and $\mathbf{H}_2^{-1}$ maps points in I' back to I through π_2. This combination transform has some interes-ting properties. Formed from two homographies between two cameras induced by two planes, it is termed an *homology* [4]. Certain features of an homology are invariant. The epipole e, determined by the imager relationships, is unaffected by the planes, with ${}^{2,1}\mathbf{H}e = e$. Similarly, the intersection line of the planes, determined by the plane positions, is unaffected by the relative positions of the imagers. These two invariances provide us with a means for combining multi-plane and multi-imager observations into an integrated linear epipole estimator.

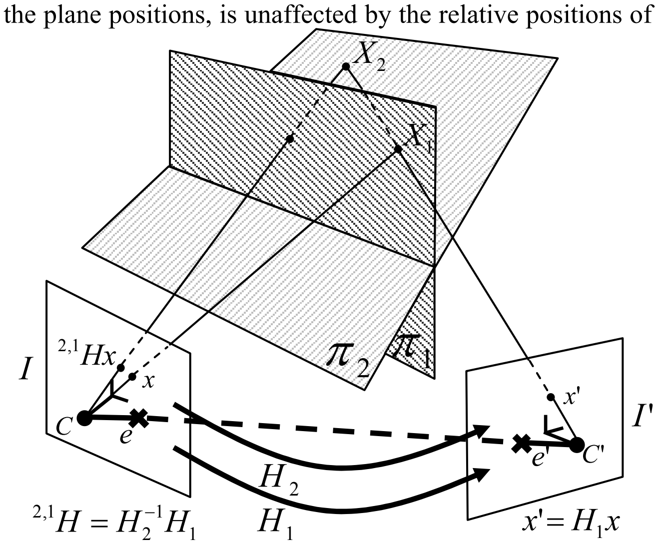

Figure 2: Homology from plane homographies.

Ashdown et al. [1] present a related calibration notion also using multiple planes. While each plane can provide one homography matrix, the H's at the intersection of two planes (the fixed line invariant) are unlikely to be consistent, that is, $H_1 x \neq H_2 x$. To resolve this, they determine point correspondences along the intersection and refine the homographies to force the invariant using a DLT algorithm. Since we compute homographies from separate images—each with a view filled with the plane—our intersection is consistent and forms part of the constraint set.

Given two homographies between two cameras induced by two planes, $\mathbf{H}_2^{-1}\mathbf{H}_1$ can be identified as a "*rank-one perturbation of identity* (ROPI) *matrix*" [5]. This matrix can provide us more information. Zelnik-Manor and Irani [9] observed that each homology matrix has two degenerate unit eigenvalues, and used this information to get ``true" or ``corrected-scale'' versions of the homology. However, in a real scenario with noise, it is problematic to assess such equality. We prove elsewhere [5] that this matrix must have second eigenvalue equal to 1, and this provides us a normalization mechanism for establishing a set of simul-taneous equations for solving in a least-squares sense for the invariants.

2. MULTI-IMAGER CALIBRATION

Development of the mathematics may proceed from the case of two imagers and two planes; instead we leave that development to our earlier paper and present the formulation for multiple imagers observing multiple planes.

2.1. Estimation of the Epipoles

Consider multiple cameras and multiple planes, where one of the cameras provides the frame of reference ($k = 0$). Denoting homology matrices ${}^{i,j}H_k = ({}^iH_k)^{-1}({}^jH_k)$, where i and j are planes indices ($i = 1,2,\ldots,n-1$, $j = i+1,\ldots,n$), and $k = 1,2,\ldots,K$ is the camera index, we have [9]:

$$\begin{cases} {}^{i,j}H_1 = I_{3x3} + e_1\, {}^{i,j}v^T \\ \quad\vdots \\ {}^{i,j}H_k = I_{3x3} + e_k\, {}^{i,j}v^T \\ \quad\vdots \\ {}^{i,j}H_K = I_{3x3} + e_K\, {}^{i,j}v^T \end{cases}$$

Together, $n(n-1)K/2$ equations can be acquired, and the final objective function can be written:

$$J=\sum_{k=1}^{K}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left\| {}^{i,j}H_k - I - e_k\, {}^{i,j}v^T \right\|_F^2$$

where e_k is the epipole of camera *k* corresponding to the reference camera, and ${}^{i,j}v$ is the intersection of planes *i* and *j*. e_k is only related to cameras *k* and the reference camera, and is independent of the planes. On the other hand, ${}^{i,j}v$ is the intersection of the plane *i* and *j* and will be independent of the cameras. We can present this formula in matrix form:

$$\begin{pmatrix} \cdots {}^{i,j}H_1 - I \cdots \\ \cdots {}^{i,j}H_2 - I \cdots \\ \cdots \quad \vdots \quad \cdots \\ \cdots {}^{i,j}H_K - I \cdots \end{pmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_K \end{bmatrix} \begin{bmatrix} {}^{1,2}v^T & {}^{1,3}v^T & \cdots & {}^{n-1,n}v^T \end{bmatrix}$$

Defining the left side of the above equation as *A,* each row of *A* corresponds to one camera. This is also an MLE or least-squares problem and different methods can be used to solve it (Levenberg-Marquard, etc.). Using SVD with $A=U\Sigma V^T$, the optimal values for e_k and ${}^{i,j}v$ arise from the first singular value[1] of the matrix *A*.

While we could estimate e' as $e'=H_i e$, we prefer to derive it directly, using the method used to obtain e. This we do by reversing the roles of H_1 and H_2 in the homology.

2.2. Estimation of the Fundamental Matrix

Since $F \cong [e']_\times H$, given the epipole-pair e and e', and a set of homographies, $H_i, i=1,2,\cdots,n$, the optimal fundamental matrix, in the Frobenius norm sense, is $F^*=\sum_{i=1}^{n}[e']_\times H_i$, having rank 2. We could weight this mean using estimate confidences.

3. EVALUATION

We demonstrate the performance of our global calibration system using the 2x9 multi-imager camera at the top left of Figure 1. In Figure 3 we sketch sample epipolar lines across adjacent (horizontal and vertical) imagers.

Lacking ground truth for our experimental tests, we have adopted what seems to be a reasonable method (similar to

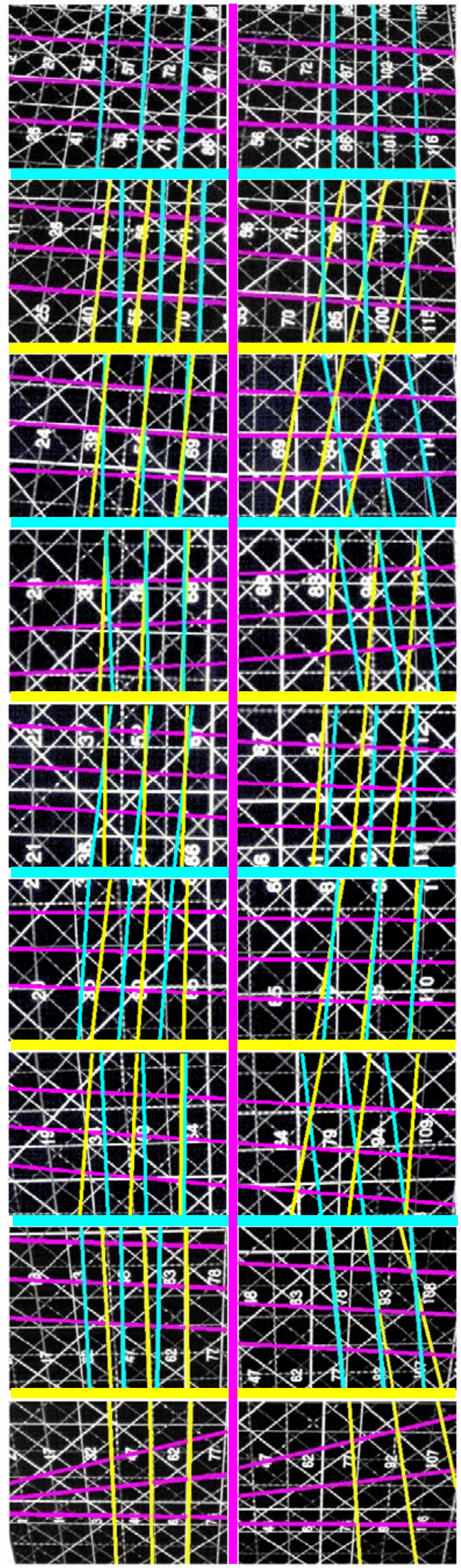

Figure 3: Corresponding epipolar lines for 2x9 multi-imager configuration: yellow-cyan alternating horizontally, magenta vertically

1 http://voteview.com/ideal_point_Eckart_Young_Theorem.htm

Luong [6] but using RMS error) for evaluating relative estimation quality. We presume our homographies are, in some sense, ground truth, and determine how well the derived fundamental matrices act with respect to them. We justify this assumption by noting that the homographies provide a correspondence across imagers for points in their plane—mapping x to x'. The fundamental matrix ensures $x'^T Fx = x'^T l_x = 0$, that is, x' lies on the epipolar line l_x corresponding to x. We take about 1% of the points in one image of a pair, and determine where the homography takes each in its paired image. We then map the same points through the fundamental matrix, determining their epipolar lines. Measuring the distance of each homography-transferred point from its epipolar line gives us a measure of the quality of the fundamental matrix for stereo reconstruction. Since we have demonstrated our homographies to be very accurate, this should provide a view on the quality of our fundamental matrix and epipole estimation.

In evaluating our solution, we show measures of RMS error over all adjacent (i.e., stereo-suitable) imagers at all subsets of plane combinations (from 3 to 7), deriving a set of measures that reveal decreasing RMS error with increasing number of planes. Figure 4 shows typical results.

4. SUMMARY AND CONCLUSIONS

We have presented results in using a global optimization method for estimating the epipoles and fundamental matrices of a multi-imager camera system. Preliminary homographies are determined using lines projected to a planar field, and used in constructing a mosaicked video stream. Homologies are formed from these homographies, which can be decomposed as ROPI matrices whose second singular values are unity. This gives us robust matrix normalization, which permits us to consider all planes and multiple cameras simultaneously without regard for the usual "up-to-a-scale'' issue. By stacking all homologies together, the epipole estimation can be solved by SVD. Since we exploit all the constraints between calibration planes and cameras, this method is robust to noise.

Our approach does not depend upon camera intrinsics, so it applies to uncalibrated cameras and in situations where we don't have metric information about the planes.

We expect camera systems such as ours will provide the mechanisms that enable multi-viewpoint and 3D imaging, and optimization techniques such as the one presented here will help us exploit these image streams by delivering high quality globally optimal calibration models. Our next studies will include using these results for 3D reconstruction from camera arrays.

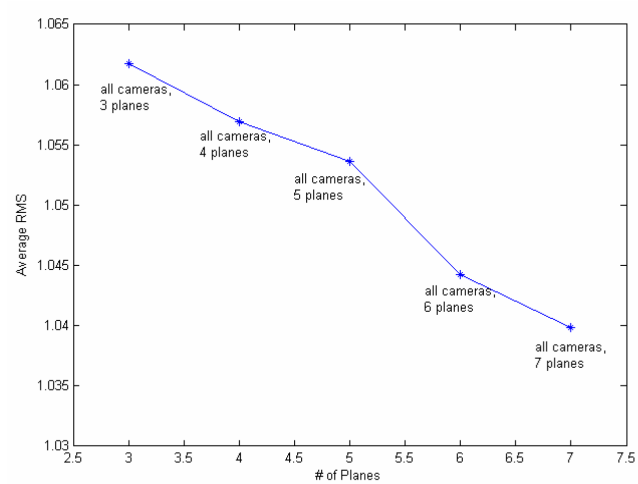

Figure 4: Reduction in RMS error with increasing number of homology constraints.

5. REFERENCES

[1] M. Ashdown, M. Flagg, R. Sukthankar, J. Rehg, ``A flexible projector-camera system for multi-planar displays,'' IEEE Press, CVPR, 165-172, 2004.

[2] H. Baker, D. Tanguay, ``A Multi-Imager Camera for Variable-Definition Video (XDTV),'' Springer-Verlag, MRCS 2006, 594–601, 2006.

[3] A.Criminisi, I.Reid, A.Zisserman, ``Duality, Rigidity and Planar Parallax,'' Springer-Verlag, ECCV, 846-861, 1998.

[4] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge Press, 2000.

[5] Z. Li, H. Baker, R. Schreiber, ``Estimation of Epipoles and Fundamental Matrices in Multi-camera Multi-plane Systems,'' Hewlett-Packard Laboratories Tech Report HPL-2007-31.

[6] Q.-T. Luong, O.D. Faugeras, ``Determining the Fundamental Matrix with Planes: Unstability and New Algorithms.'' IEEE Press, CVPR, 489-494, 1993.

[7] B. Triggs, ``Autocalibration from Planar Scenes,'' Springer-Verlag, ECCV, 89-105, 1998.

[8] Wilburn, B., N. Joshi, V. Vaish, M. Levoy, M. Horowitz, ``High speed videography using a dense camera array,'' IEEE Press, CVPR, 294-302, 2004.

[9] L. Zelnik-Manor, M. Irani, ``Multiview Constraints on Homographies,'' IEEE PAMI, Vol. 24, No. 2, 214-223, 2002.

[10] Z. Zhang, ``Flexible Camera Calibration by Viewing a plane from unknown orientations,'' IEEE Press, ICCV, 666–673, 1999.

[11] H. Zollner, R. Sablatnig, ``Comparison of Methods for Geometric Camera Calibration using Planar Calibration Targets,'' Vienna University of Technology, 2004.

[12] http://www.isd.gr Products/Herodion