Multi-view Tensors

Richard Hartley,
GE-CRD, Schenectady, NY.

The last few years have seen considerable interest in a the analysis and application of multilinear relations between corresponding features in images taken with a perspective camera. Given a point in space seen in two images, the image points satisfy an epipolar constraint expressed by the *fundamental matrix*. This is a 3×3 matrix $\mathtt{F}$ that satisfies the equation

$$\mathbf{x}'^{\top}\mathtt{F}\mathbf{x} = 0$$

where $\mathbf{x}$ and $\mathbf{x}'$ are any two corresponding points (expressed as homogeneous 3-vectors) in the two images. One way of expressing this is that the point $\mathbf{x}'$ matching a point $\mathbf{x}$ in the first image lies on a line $\mathtt{F}\mathbf{x}$, known as the epipolar line.

More recently, similar relationships have been discovered between corresponding points and lines in three and four views. The trifocal tensor relates points and lines in three views. The most basic relationship involves a point in one image and two corresponding lines in the other images.

$$\mathbf{x}_i \mathbf{l}'_j \mathbf{l}''_k T_i^{jk} = 0$$

in which repeated indices imply summation over the range $1 \ldots 3$. The geometric meaning of this relationship involves the common intersection of four planes, corresponding to the back-projection of lines in the image. The fundamental matrix, as well as a 4-view *quadrifocal* tensor may be similarly derived.

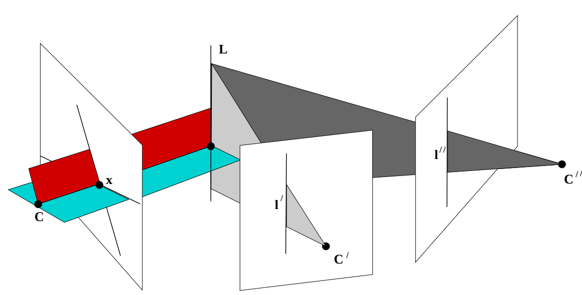

Figure 1: The intersection of 4-planes implies that the determinant constructed from the 4-vectors representing the planes must vanish. This gives rise to a linear relationship involving image coordinates, encapsulated in the trifocal tensor.

1 Applications of the multilinear tensors.

The multiview tensors, particularly the fundamental matrix and the trifocal tensor have found many applications in 3D scene reconstruction and construction of synthetic images. For instance, the trifocal tensor has found effective use in

- Robust point and line and curve matching between images.
- Reconstruction of projective and Euclidean representations of a viewed scene [2]
- Synthesis of images seen from different viewpoints [3]
- Self-calibration of a moving camera ([1])
- Independent motion detection ([4]).

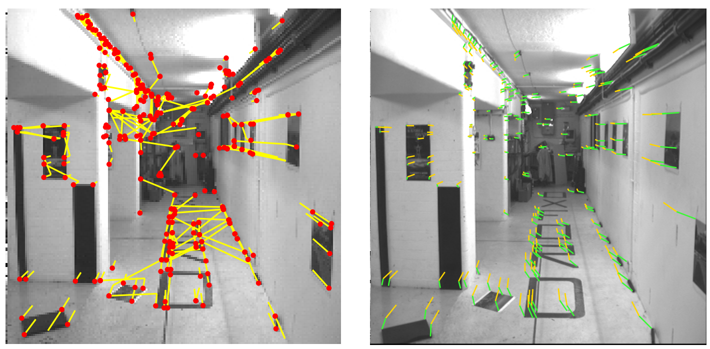

Figure 2: Correspondences between succeeding frames of a video sequence. Corner matching based on correlation alone gives many false matches (left). Retaining only matches satisfying the trifocal constraint across three views removes false matches (right).

References

[1] Martin Armstrong, Andrew Zisserman, and Richard I. Hartley. Self-calibration from image triplets. In *Computer Vision - ECCV '96, Volume I, LNCS-Series Vol. 1064, Springer-Verlag*, pages 3 – 16, 1996.

[2] Richard I. Hartley. Lines and points in three views and the trifocal tensor. *International Journal of Computer Vision*, 22(2):125–140, March 1997.

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

[4] Torr P.H.S., Zisserman A., and Murray D.W. Motion clustering using the trilinear constraint over three views. In R. Mohr and C. Wu, editors, *Europe-China Workshop on Geometrical Modelling and Invariants for Computer Vision*, pages 118–125. Springer–Verlag, 1995.