Fundamental Matrices from Moving Objects Using Line Motion Barcodes

Yoni Kasten   Gil Ben Artzi   Shmuel Peleg   Michael Werman
The Hebrew University of Jerusalem

Introduction

The Goal: Calibration of multi-camera systems from significantly different viewpoints, when the scene has multiple moving objects.

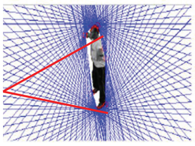
Prior Work: When finding corresponding points is difficult, use corresponding epipolar lines from dynamic silhouettes. Motion Barcode are used in order to accelerate the search.

Our Contribution: When there are multiple objects in the scene, previous methods fail. Our method can handle such cases by sampling lines.

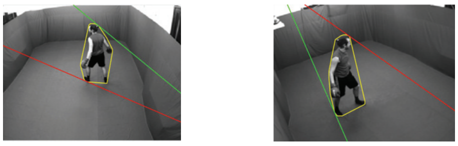
Prior Work

Ben-Artzi et al. *Camera Calibration from Dynamic Silhouettes Using Motion Barcodes,* CVPR'16.

- Uniformly sampling tangent lines around the convex hull of the silhouette in each frame.



- In each frame time, tangent lines from one camera view are compared to tangent lines from second camera view.
- The pair of lines having highest motion barcode similarity is selected.
- Number of RANSAC iterations is much lower than previous papers (Sinha & Pollefeys, IJCV 2010).



- Limitation: With multiple objects in the scene – convex hulls do not include the same objects.

Camera A   Camera B

Motion Barcodes for Lines (Ben-Artzi et al)

For a segmented video (moving objects vs. static background),

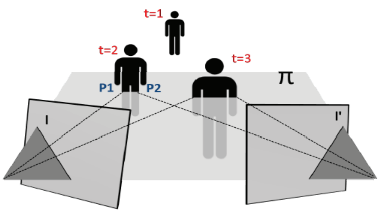
For line l at frame t:

$$b_l(t) = \begin{cases} 1 & l \text{ intersects object} \\ 0 & \text{Otherwise} \end{cases}$$

Barcode correlation defines the temporal similarity between two lines:

$$d_t(l, l') = corr(b_l(t), b_{l'}(t))$$

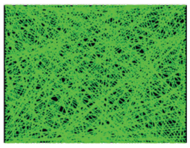
Corresponding epipolar lines have correlated Motion Barcodes.



Three time instances from two videos, having different viewpoints, showing a single moving person. Motion barcodes for the two corresponding lines (*l* and *l'*) in the two videos are 011.

Method

- The Goal: Find the Fundamental Matrix between two cameras viewing a dynamic scene.
- Lines are sampled in both cameras by connecting two uniformly sampled points on the border of the frame (see picture).
- Motion Barcodes are computed for all sampled lines.



Uniformly sampled lines

- Only informative lines are used in next steps, i.e. lines whose motion barcodes have enough zeros and ones. This leaves n_1 line Motion Barcodes for Camera A, and n_2 line Motion Barcodes for Camera B.
- The correlations of all line motion barcodes from Camera A with all those from Camera B are computed.

$$\begin{bmatrix} corr(b_{A,1}b_{B,1}) & \cdots & corr(b_{A,1}b_{B,n_2}) \\ \vdots & \ddots & \vdots \\ corr(b_{A,n_1}b_{B,1}) & \cdots & corr(b_{A,n_1}b_{B,n_2}) \end{bmatrix}$$

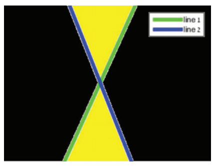
correlation matrix of size $n_1 \times n_2$

- If the correlation of a pair of lines is in the mutual top 3 of each other, i.e. top 3 in both row and column, it is considered a candidate.
- The 1,000 candidate pairs having the highest correlations are taken as candidates for corresponding epipolar lines.
- As a result we get 1000 pairs of corresponding epipolar line candidates: $\{(l_1, l_1'), \dots, (l_{1000}, l_{1000}')\}$
- Experimentally the probability that a candidate pair is true, i.e. both lines are correct epipolar lines, is 0.7 for synthetic data, and 0.37 for real data.

RANSAC samples

In each RANSAC iteration:

- two candidate pairs of lines are randomly sampled: (l_1, l_1') *and* (l_2, l_2')
- The intersection of each two lines in each picture gives two candidate epipoles: $e = l_1 \times l_2, \quad e' = l_1' \times l_2'$
- Another pair of lines passing through the epipoles is chosen: $(l_3, l_3') = \arg\min_{(l_i, l_i') \in \{candidates\} \setminus \{(l_1, l_1'), (l_2, l_2')\}} d(l_i, e) + d(l_i', e')$
- The epipolar line homography *H* is calculated using the DLT algorithm, and its consistency with all other candidates is measured



For a candidate pair (l_i, l_i'): The image area between Hl_i and l_i' is used as a similarity measure.

The pair (l_i, l_i') is considered an inlier if the area is small enough.

The Fundamental Matrix F is computed from the most consistent epipolar line homography:

$$F = [e']_x H^{-T}$$

Results

- Running 10,000 RANSAC iterations for each pair of cameras in each dataset resulted in a Fundamental Matrix for each camera pair.
- The Symmetric Epipolar error of the resulting F was measured on ground truth points.

Dataset	Average error	Number of good pairs	
Cubes	0.31	10/10	
Thin Cubes	0.79	21/21	
Pets 2009	1.69	5/6