Motivation & Contribution
- Real-time matching, mobile devices require efficient descriptors.
- Many existing fast descriptors sacrifice invariance for speed.
- BAFT is a fast, binary, quasi affine-invariant local image feature
- Based on ideas from ORB descriptor [6], Harris/Hessian Affine detectors [10]; adapt sampling pattern to normalization matrix.
- Higher degree of invariance to skew vs. rotation
- Higher AUC; works well especially for large perspective changes

BAFT Algorithm
1. Detecting Keypoints
   - Detect keypoints using FAST, compute second moment matrix
   \[ M(x_p) = \sum_{p \neq q} w(p, q) \begin{bmatrix} I_x^2(x_q) & I_x I_y(x_q) \\ I_x I_y(x_q) & I_y^2(x_q) \end{bmatrix} \]
   - Eigenvalues of \( M \) measure ‘cornerness’, pick \( N \) best keypoints
   - Matrix stored in memory with keypoint

2. Building the Descriptor
   - Sample image around keypoint, adjusted for affine transformations
   - Winner-take-all hashing from samples (16-128 bytes descriptors)

Experimental Validation & Results

Database
- ASIFT dataset [14]
- Perspective changes
- 5 sets of 10 images

Conclusions
- BAFT: novel affine-invariant local image descriptor; computationally efficient (similar to ORB, much faster than SIFT).
- Achieves robustness to perspective changes without sacrificing speed (outperforms ORB, SIFT, AKAZE, and others).

References