FAST 6DOF POSE ESTIMATION WITH SYNTHETIC TEXTURELESS CAD
MODEL FOR MOBILE APPLICATIONS
Bowen Chen, Juhan Bae, Dibyendu Mukherjee
Epson Canada Ltd.

PROBLEM
Conventional methods for 6 Degrees-of-Freedom (DoF) pose estimation require one or more of the following:
• Large-scale real training data
• Synthetic textured data
• Use of RGB-D for better inference if trained from only synthetic
• Multiple stages of regression or classification for accuracy when extended to multiple objects and large view-ranges

CONTRIBUTIONS
Propose two algorithms (VIEWMOD, BBOX9) that
• Only require synthetic textureless CAD model for training.
• Use of only RGB information during inference
• Real-time inference for mobile CPUs.

SYNTHETIC-TO-REAL DOMAIN ADAPTATION
1. Project 3D model to 2D image with random textures and lighting
2. Apply Random scaling, in-plane rotation, and background with noise (Gaussian blur, motion blur, and additive noise)
3. Apply Laplacian Filter

VIEWMOD
• Use 2D bounding box detection with view-classification followed by a LINEMOD [1] based pose estimation
• Fast and accurate two-stage inference with improved interpretability to detect failures.

BBOX9
Perform a one-stage direct regression of a 3D bounding box surrounding the object, followed by a PnP routine to estimate the object’s 6 DoF pose:

\[ L(x, c, l, g) = \frac{1}{N} \left( L_{\text{corr}}(x, c) + \alpha L_{\text{loc}}(x, l, g) \right) \]

\[ L_{\text{loc}}(x, l, g) = \sum_{i=0}^{N} \sum_{m=1}^{M} x_{ij} \text{smooth}_{L1}(l_{im}^m - \hat{g}_{im}^m) \]

\[ \hat{g}_{im}^m = \begin{cases} (g_{im}^m - d_{im}^m)/d_{im}^m, & \text{if } m \text{ is odd} \\ (g_{im}^m - d_{im}^m)/d_{im}^m, & \text{otherwise} \end{cases} \]

RESULTS
BB8 uses real RGB images while VIEWMOD and BBOX9 only use synthetic textureless CAD models.

MOBILE INFERENCE
VIEWMOD and BBOX9 take ~200ms per frame using CPU on Google Pixel 2, using a Tensorflow API.

CONCLUSION
We introduced an efficient and user-friendly 6DoF pose estimation framework for mobile applications:
• Effective domain adaptation strategy to use synthetic textureless CAD.
• Real-time inference for mobile CPUs.

Reference: