Cross-Domain 3D Equivariant Image Embeddings

Carlos Esteves      Avneesh Sud      Zhengyi Luo     Kostas Daniilidis     Ameesh Makadia
Universal image embeddings

Images: Wikimedia commons (cc)
Why is this hard?

Images: Wikimedia commons (cc)
Why is this hard?
Conventional approaches

Images: Wikimedia commons (cc)
Conventional approaches
Our approach

Images: Wikimedia commons (cc)
3D equivariant embeddings

Images: Wikimedia commons (cc)
Embeddings on the sphere!

\[ \Lambda_R \]

Images: Wikimedia commons (cc)
How to learn cross-domain embeddings?
What is the supervision?

No task-specific supervision!
Relative pose estimation

No pose regression/supervision!
Relative pose estimation

Inputs

Embeddings

A to pose of B
B to pose of A
Results on real images from ObjectNet3D

Ours: 13.75 deg, Regression: 36.52 deg (median error).

Dataset: Xiang et al. ObjectNet3D: A Large Scale Database for 3D Object Recognition. ECCV'16.
Novel view synthesis

Training time: reconstruct the input
Novel view synthesis

Test time: generate any view from any other view
A novel approach to learning geometric image embeddings

- equivariant to 3D rotations
- needs only unaligned meshes as training inputs
- generalizes to geometric tasks without typical task-specific supervision
  - e.g. no pose or view synthesis supervision
- avoids the difficulties of traditional approaches
  - e.g. no pose regression or full 3D structure prediction
Cross Domain 3D Equivariant Image Embeddings

06:30 -- 09:00 PM @ Pacific Ballroom #25