Incorporating Scene Context and Object Layout into Appearance Modeling

Hamid Izadinia*, Fereshteh Sadeghi*, Ali Farhadi
{izadinia, fsadeghi, ali}@cs.washington.edu

What is behind the black box?
- Joint reasoning over scene and objects
- Scene type
  - Expected objects and their style
- Nearby objects
  - Scene Type
  - Scene Layout
  - Object category

Our Approach
- Reasoning over interlacing components of a scene
  - Scene category
  - Scene specific appearance of objects (style, pose)
  - Objects layout in the scene

Learning
- Decoupling optimization of \( G \) and other parameters
- Optimize \( G \) by weighted maximum spanning tree
  - Nodes: Frequency of objects
  - Edges: Spatial consistency of object pairs
- Optimize \( W \) using large margin structured learning

Experimental results

Black Box Test:
- Predicting object pose
- Forced choice human study
- Ours wins in 74.7% of cases

Object Detection:
- Simultaneous scene recognition and object detection
- Scene structures improve localizing objects
- Intersection over union of object & ground truth

Scene Recognition:
- Use scene structures to generate features
  - Best scoring structures per scene category
  - Normalized locations of objects
  - Relative locations of objects

---

What is behind the black box?

- Joint reasoning over scene and objects
  - Scene type
    - Expected objects and their style
  - Nearby objects
    - Scene Type
    - Scene Layout
    - Object category

Our Approach
- Reasoning over interlacing components of a scene
  - Scene category
  - Scene specific appearance of objects (style, pose)
  - Objects layout in the scene

Learning
- Decoupling optimization of \( G \) and other parameters
- Optimize \( G \) by weighted maximum spanning tree
  - Nodes: Frequency of objects
  - Edges: Spatial consistency of object pairs
- Optimize \( W \) using large margin structured learning

Experimental results

Black Box Test:
- Predicting object pose
- Forced choice human study
- Ours wins in 74.7% of cases

Object Detection:
- Simultaneous scene recognition and object detection
- Scene structures improve localizing objects
- Intersection over union of object & ground truth

Scene Recognition:
- Use scene structures to generate features
  - Best scoring structures per scene category
  - Normalized locations of objects
  - Relative locations of objects