Modeling Mutual Context of Object and Human Pose

HELLO!

Yash Patel, 201301134
Shrenik Jain, 201403006

1.

Human pose estimation

HUMAN POSE ESTIMATION

Human-Object Interaction

Holistic image based classification

Detailed understanding and reasoning

- **Human pose estimation**

2.

OBJECT DETECTION

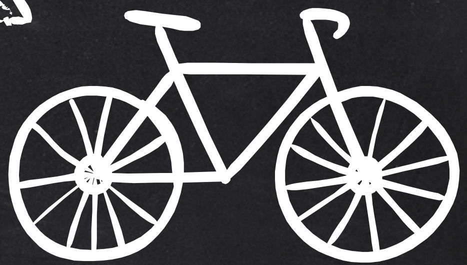
OBJECT DETECTION

Human-Object Interaction

Holistic image based classification

Detailed understanding and reasoning

- Human pose estimation
- **Object detection**



Big concept

Mutual context of Object and human pose in HOI.

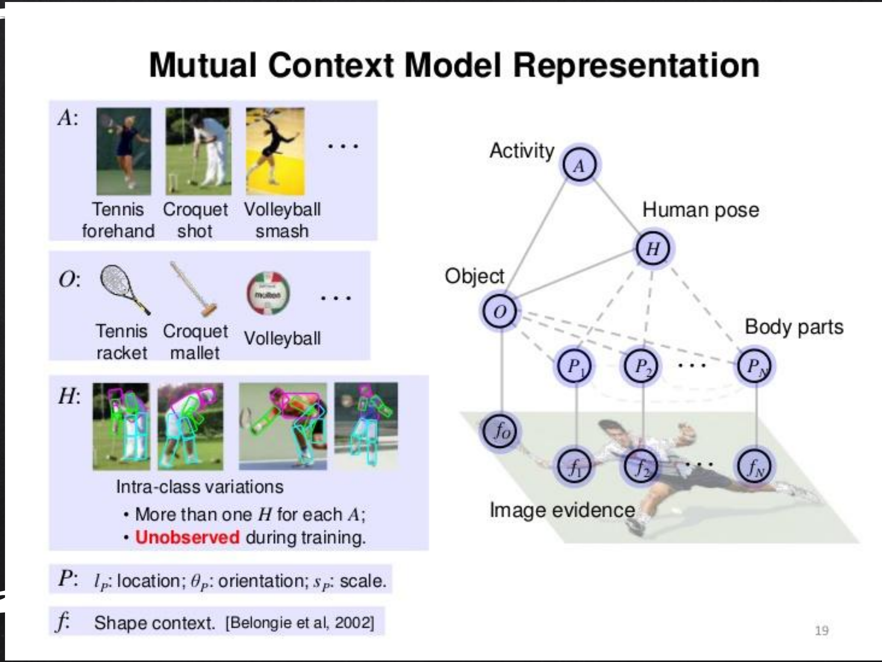
Mutual context of object and human pose

Model Representati on

Model learning

Model Inference

Mutual context model representation

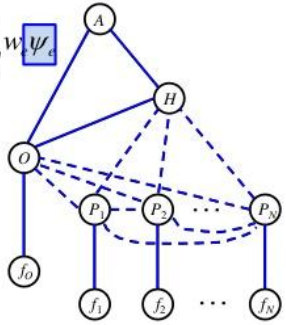



Pattern
recognition
: model
learning

MODEL LEARNING

Model Learning

$$\Psi = \sum_{e \in E} w_e \boxed{\psi_e}$$



Goals:

Hidden human poses

Structural connectivity

Potential parameters

Potential weights

Approach:

- Maximum likelihood

  $\psi_e(A,O)$ $\psi_e(A,H)$ $\psi_e(O,H)$

  $\psi_e(H,P_n)$ $\psi_e(O,P_n)$ $\psi_e(P_m,P_n)$

- Standard AdaBoost

  $\psi_e(O,f_O)$ $\psi_e(P_n,f_{P_n})$

WORK MODULES

FEATURE EXTRACTION : EVIDENCE RETRIEVAL

MODEL GENERATION : STRUCTURED LEARNING

TESTING INTERFACE : FIND CLOSEST MATCHING MODEL

1.

FEATURE EXTRACTION / EVIDENCE RETRIEVAL

FEATURE EXTRACTION / EVIDENCE RETRIEVAL

1).

TRAIN A DETECTOR USING SHAPE CONTEXT AS FEATURE DESCRIPTOR : AS SPECIFIED BY AUTHOR, USES CONCEPT OF ADAPTIVE BOOSTING .

FOR BOTH OBJECTS AND HUMAN BODY PARTS

2).

USING HOG AND SVM CLASSIFIER.

3).

Fast R-CNN: Fast Region-based Convolutional Networks for object detection

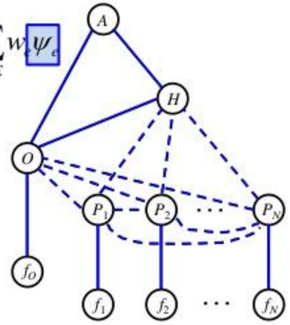
2.

MODEL GENERATION : STRUCTURED LEARNING

MODEL GENERATION : STRUCTURED LEARNING MODEL USING MARKOV RANDOM FIELD

Model Learning

$$\Psi = \sum_{e \in E} w_e \boxed{\psi_e}$$



Approach:

- Maximum likelihood

  $\psi_e(A,O)$ $\psi_e(A,H)$ $\psi_e(O,H)$

  $\psi_e(H,P_n)$ $\psi_e(O,P_n)$ $\psi_e(P_m,P_n)$

- Standard AdaBoost

  $\psi_e(O,f_O)$ $\psi_e(P_n,f_{P_n})$

Goals:

Hidden human poses

Structural connectivity

Potential parameters

Potential weights

MODEL GENERATION : STRUCTURED LEARNING

1). Co-occurrence context for the activity class, object, and human pose.

2). $\psi e(O, Pn)$ models the spatial relationship between the object O and the body part Pn, which is computed by bin($lO - lPn$) $\cdot$ bin($\theta O - \theta Pn$) $\cdot$ $\mathcal{N}(sO/sPn)$

3). $\cdot\, \psi e(Pm, Pn)$ models the spatial relationship between different body parts.

4). $\psi e(H, Pn)$ models the compatibility between the pose class H and a body part Pn.

5). $\psi e(O, fO)$ and $\psi e(Pn, fPn)$ model the dependence of the object and a body part with their corresponding image evidence. (as discussed in previous section).

MODEL GENERATION : STRUCTURED LEARNING

1). Hill-climbing structure learning :

Hill climbing with tabu list acts as structure space.

2). Max-margin parameter estimation :

Parameter estimation using max-margin so as to ensure generalized learning.

3). Analysis of our learning algorithm

3.

TESTING INTERFACE : FIND CLOSEST MATCHING MODEL

REFERENCED MATERIAL

Research Paper :

http://vision.stanford.edu/documents/YaoFei-Fei_CVPR2010b.pdf

DataSet :

Sports images. HOI referred are sports related poses.

DataSet details :

Cricket Cricket Croquet defensive shot bowling shot Tennis Tennis Volleyball forehand serve smash

THANKS!

Any questions?

yash.patel@students.iiit.ac.in
shrenik.jain@research.iiit.ac.in