Problem summary

• Human pose estimation in images via tree models
• Attempt to answer the following critical questions:
  ➢ Are simple tree models sufficient?
  ➢ How to use tree models in human pose estimation?
  ➢ How shall we use combined parts with single parts?
• Latent tree models for discovering graphical model structure
  ➢ Exact inference
  ➢ Visual categorization for combined parts
  ➢ Better performance

Latent tree models for human pose

• Learn a tree structure directly from our observations without making many assumptions of the physical constraints
• Information distance: \( d_{ij} = -\log\left(\frac{Cov(X_i, X_j)}{\sqrt{Var(X_i)Var(X_j)}}\right)\)

Our Approach

A framework for integrating primitive parts and combined parts [1]
➢ Primitive parts (non-oriented): geometric clustering [4]
➢ Combined parts: Visual Categorization SVM+HOG [3]
➢ Tree structured models Learned directly from data
➢ Textbook example of exact inference

Visual categorization for combined parts

Results

Dataset:
➢ LSP: 2000 images, subject-centric
➢ PARSE: 305 images, image-centric
➢ Pascal Dog dataset: subset

LSP

Dog pose

Conclusion

• Tree models for human pose estimation are efficient
• Latent tree is an effective tool for recovering model structure
• Learning visual category of combined part becomes important.

References


NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council