But What About Their Hands?

Daphne Tsatsoulis, David Forsyth, Derek Hoiem

{tsatsou2, daf, dhoiem}@illinois.edu

Department of Computer Science, University of Illinois at Urbana-Champaign

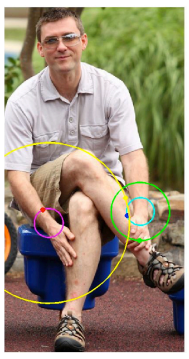
Introduction

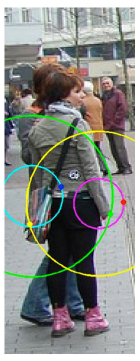
Wrist detection and localization is important for understanding human activities from images. We describe a method to predict wrist locations from a single 2D image. Our method predicts the location of a person's wrists in space relative to their body, for example, detecting a wrist to be a number of units above the right shoulder and behind the torso. We extract features from 2D images and learn to predict a set of attributes that encode a 3D body configuration. We then use these predicted attributes to predict wrist locations. This method extends naturally to video.

Ours vs. Baseline

Our Right Hand Our Left Hand

Baseline Right Hand Baseline Left Hand

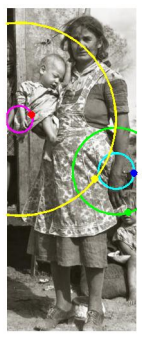




- We can find wrists in challenging images from multiple datasets.
- We predict the location of both right and left wrists
- We can predict occluded wrist locations
- Improved image resolution leads to better predictions

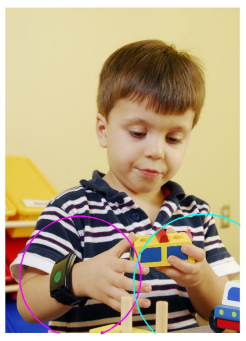
Coming soon:

- Our method can employ 3D body prior information
- Our method can predict 3D configurations

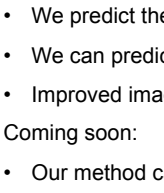


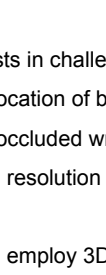
Examples

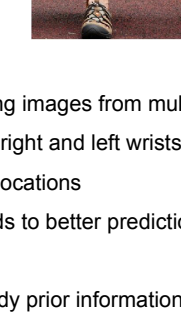
Georgia Tech Dataset:

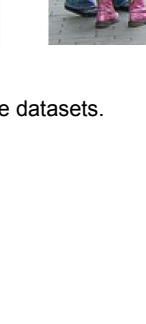


H3D Dataset:

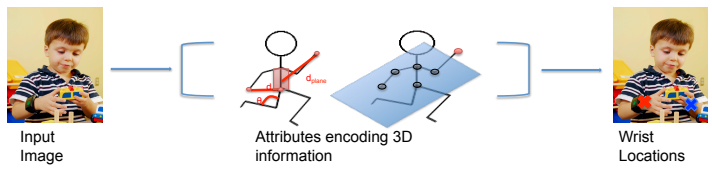




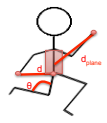




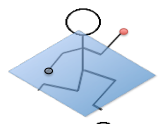
Method



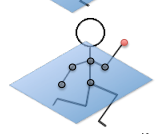
We extract features from our 2D training images. These are used to predict long (~10^3) attribute vectors that encode the 3D geometric correlations of the body's joints. Such correlations include the angle between joints and the distance of all joints to a hyperplane. We have tested a number of attributes as described below:



Geometric Attributes: Encode the joint distances, angles and distances to planes created by other joints.

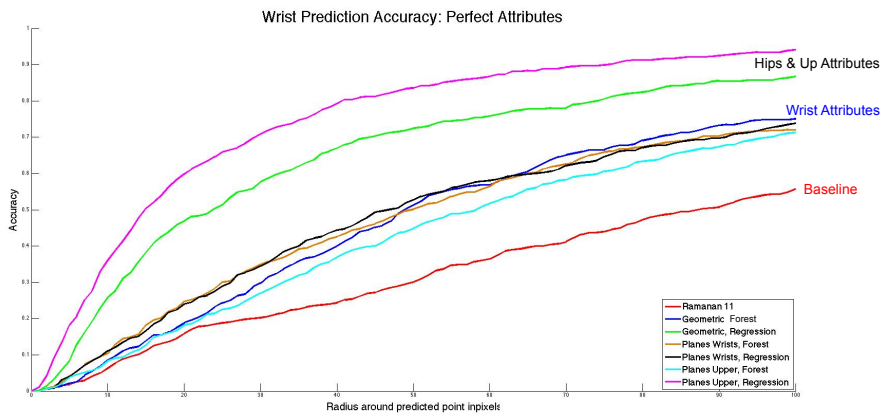


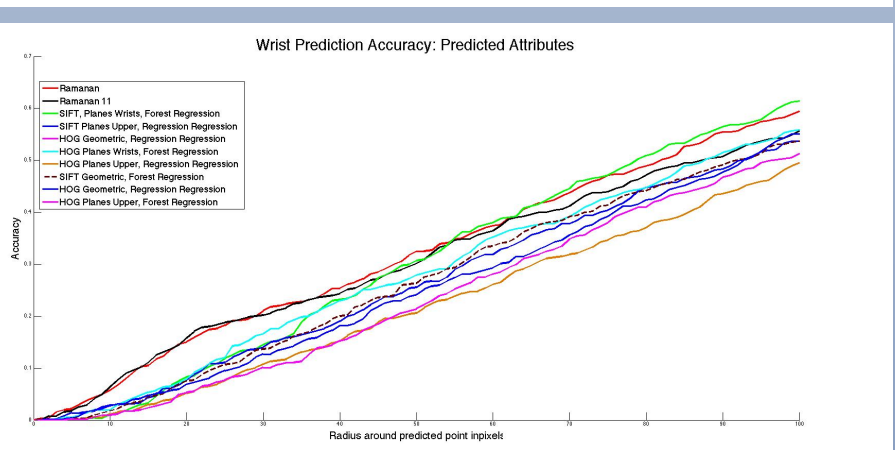
Planes Wrists Attributes: Encode the distance of the wrists to random hyperplanes.

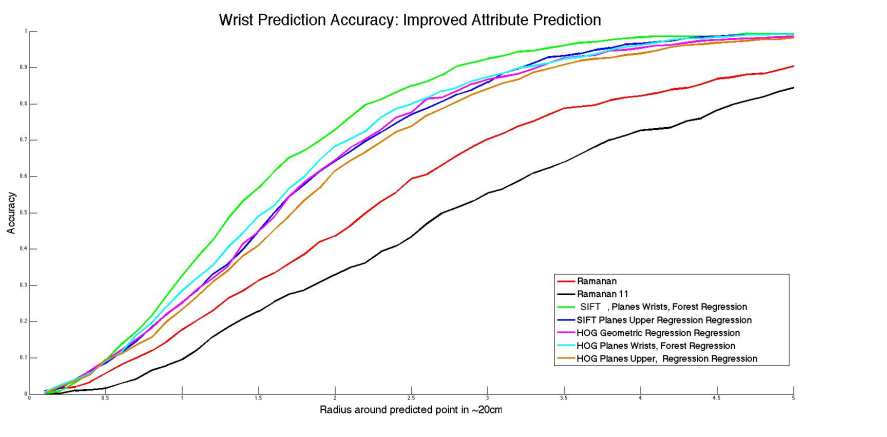


Planes Upper Attributes: Encode the distance of all upper body joints to random hyperplanes.

The predicted attributes are used to predict the wrist locations. Plot A shows our system's results assuming perfect attribute prediction for a number of attributes and wrist location predictions. We predict wrist locations by Regression or Random Forests. Plot B shows the results if we regress to predict the attribute vectors before predicting the locations. In Plot B we compare against Yang and Ramanan's 2008 and 2011 parsers. Plot C strongly suggests that higher image resolution results in improved wrist location predictions.







ILLINOIS NSF