Cascaded Pose Regression

Overview

- Principled approach to pose estimation
- Built using pose-indexed features
- Learning from training examples
- Natural coarse to fine estimation

Pose-Indexed Features

- Key property: weak invariance

Training Algorithm

Input: Data \((I_i, \theta_i)\) for \(i = 1 \ldots N\)
1: \(\theta^0 = \text{arg min}_\theta \sum \phi(\theta, \theta_i)\)
2: \(\theta^t = \theta^0 \text{ for } i = 1 \ldots N\)
3: for \(t = 1\) to \(T\) do
4: \(x_i = h^t(\theta^{t-1}, I_i)\)
5: \(\hat{\theta}_t = \theta_t^{t-1} \cdot \hat{\theta}_i\)
6: \(R^t = \text{arg min}_R \sum \phi(R(x_i), \hat{\theta}_i)\)
7: \(\theta_t = \theta^{t-1} \cdot \theta_t\)
8: \(\epsilon_t = \sum \phi(\theta_t, \theta_i) / \sum \phi(\theta^{t-1}, \theta_i)\)
9: If \(\epsilon_t \geq 1\) stop
10: end for
11: Output \(R = (R^2, \ldots, R^T)\)

Human Consistency Analysis

- Human vs human performance.

Booted-Reg: boosted regression [17] using same features as CPR.
Rand-16 Best: oracle selects the best of 16 random poses.
CPR-1: CPR with a single \((K = 1)\) starting pose.
CPR-16-Clust: CPR with 16 starting poses followed by clustering.
CPR-16-Best: CPR with 16 starting poses, oracle selects best.

Performance Breakdown

Cascaded Pose Regression (CPR)

- Number of Stages
- Data Augmentation

Conclusion

Advantages:
- Principled learning driven approach
- General, robust and accurate
- Data augmentation
- Fast (2-3ms per image)

Disadvantages:
- Pose-indexed feature design
- Supervision (~250 labels)

Future Work:
- Integrate with detection
- Integrate with tracking