Joint Patch and Multi-label Learning for Facial Action Unit Detection

1 Beijing University of Posts and Telecommunications 2 Carnegie Mellon University 3 University of Pittsburgh Kaili Zhao* Wen-Sheng Chu* Fernando De la Torre† Jeffrey F. Cohn‡ Hongqiang Zhang∗

Problem

- Facial Action Unit (AU) detection
- Observations
  - AU relations are often co-occurring and competition between AUs.
- Objectives
  - Identify discriminative patch subset for each AU (patch learning)
  - Incorporate AU relations into model learning (multi-label learning)

Patch Learning (PL)

- Model
  - Minimize L(W) + αD(W), where D(W) = ∑_{i=1}^{n} ∑_{j=1}^{m} |w_{ij}|^α
- Patch importance learned for basic expressions
- The number of patches vs performance

Multi-label Learning (ML)

- Observed AU relations
- AU relations discovered from > 350,000 frames
- J Joint solution with ADMM: \( \min_{W_i} L(W_i) + \lambda_i D(W_i) \) where \( D(W_i) = \sum_{j=1}^{m} |w_{ij}|^\alpha \)

JPML Optimization

- Convergence of ADMM
- Convergence of patch learning

Experiments

- Settings
  - SIFT descriptor on pre-determined 49 facial landmarks
  - Leave-one-subject-out protocol
- Metrics
  - F1 Score
  - L2Boost
- Methods
  - L1-regularized logistic regression
  - Active Patch Learning
  - Multi-task Multi-kernel Learning
  - JPML: Our method
- GT6 [6]
- 60 2-minute spontaneous videos from 50 participants
- BLAUD [7]
- 45 spontaneous videos from 43 participants

Observations

1) AUs describe region-based muscle movements; there are local dependencies between features.
2) The combinations of AUs produce basic expressions; there are co-occurrence and competition between AUs.

Objectives

1) Identify discriminative patch subset for each AU (patch learning)
2) Incorporate AU relations into model learning (multi-label learning)


