Estimating Smile Intensity: A better way

Jeffrey M. Girard
Jeffrey F. Cohn
Fernando De La Torre

Pattern Recognition Letters, In Press
The Whole World Smiles

- Smiles are salient **socio-emotional** signals
  - Valence, Dominance, Affiliation

- Smiling is linked to **psychological phenomena**
  - Gender, personality, and culture
  - Health and interpersonal outcomes
The “Power” of a Smile

- Most researchers have focused on detecting smile occurrence.
- The signal value of a smile is highly dependent on its intensity.
  - Also important for modeling dynamics.
- How can we most efficiently estimate a smile’s intensity?
  - Can existing binary classifiers provide such estimates, ...
  - ...or do we need to explicitly train models on intensity labels?

A B C D E
Is there a shortcut?

- SVM trained on binary labels
Is there a shortcut?

- SVM trained on binary labels
- Finds the best hyperplane to separate smiles and non-smiles
Is there a shortcut?

- SVM trained on binary labels
- Finds the best hyperplane to separate smiles and non-smiles
- **May estimate smile intensity**... assuming higher intensity frames are further from the hyperplane
Is there a shortcut?

- SVM trained on binary labels
- Finds the best hyperplane to separate smiles and non-smiles
- May estimate smile intensity... assuming higher intensity frames are further from the hyperplane
Is there a shortcut?

- SVM trained on binary labels
- Finds the best hyperplane to separate smiles and non-smiles
- May estimate smile intensity... assuming higher intensity frames are further from the hyperplane
- However, nothing in the SVM requires this to be the case, and many factors may influence the distance to the hyperplane
Or is the only road direct?

**Multiclass SVM (one-against-one)**

**Support Vector Regression**

- Minimize:
  \[
  \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*)
  \]

- Constraints:
  \begin{align*}
  y_i - wx_i - b &\leq \varepsilon + \xi_i \\
  wx_i + b - y_i &\leq \varepsilon + \xi_i^*
  \end{align*}

\[\xi_i, \xi_i^* \geq 0\]
Spontaneous Expression Databases

BP4D-Spontaneous Database
- 120,000 frames from 30 participants
- High quality video, high expressiveness
- Expert Coding ($F_1 = 0.96; ICC = 0.92$)

Spectrum Depression Database
- 200,000 frames from 33 participants
- Highly challenging, psychiatric context
- Expert Coding ($F_1 = 0.71; ICC = 0.92$)
Methods for Expression Analysis

Image Metrics
- Mathew’s AAM

Localized Gabor

Laplacian Eigenmap

Two-Class (Binary) SVM
- Multiclass (Intensity) SVM

Regression (Intensity) SVR
Average Performance Across Methods

- BP4D-SFE Spectrum

- ICC

- Two-Class
- Multiclass
- Regression

Bar chart showing comparison of ICC for Two-Class, Multiclass, and Regression methods across BP4D-SFE and Spectrum.
Intensity Level Separation in BP4D

**Two-Class (Binary)**

ICC = 0.533

**Regression (Intensity)**

ICC = 0.893

*Blue Box* = 25th to 75th percentile, *Red Line* = Median, *Blue Lines* = 1.5 x IQR
Intensity Level Separation in Spectrum

**Two-Class (Binary)**

- ICC = 0.398

**Regression (Intensity)**

- ICC = 0.558

Blue Box = 25th to 75th percentile, Red Line = Median, Blue Lines = 1.5 x IQR
Intensity Estimation Demo
Intensity Estimation Demo
Conclusions

- Distance to the hyperplane did not yield competitive performance
- Multiclass and regression models far outperformed this shortcut
- There is no substitute for training on intensity ground truth labels
- Research would benefit from moving beyond binary models of smiling

Estimating smile intensity: A better way
*Pattern Recognition Letters, In Press*
10.1016/j.patrec.2014.10.004

http://tinyurl.com/smileintensity