Deciphering the Face

Aleix M. Martinez
Computational Biology
and Cognitive Science Lab
aleix@ece.osu.edu
Models of Face Perception

• **Features**: Shape vs. texture.

• 2D vs. 3D

• **Form** of the computational space:

  Continuous vs. Categorical
What we are going to show

• What is the form of the computational space in human face perception? **Hybrid approach**: Linear combination of continuous representations of categories.

\[ c_1 + c_2 + \ldots + c_n \]

• What are the dimensions? Mostly **configural**.

• In computer vision we need **precise detailed detection** of faces and facial features.
Identity

Same or different?
Identity

Same or different?
Identity

Same or different?
Identity, expression, gender, etc.
Dimensions of the Face Space

Same or different?

Configural processing
Form of the Computational Face Space

Exemplar-based model

Exemplar cells

Mid-level vision

Low-level vision

Norm-based model
Facial Expressions of Emotion
Muscle Positions Model
• Global shape (bone structure) determines identity – configural.

• But ONLY muscles are responsible for expression, interaction...
Configural Processing

Emotion perception in emotionless faces

Neutral

Sad

Angry

Neth & Martinez, JOV, 2009.
Stimuli

Neth & Martinez, JOV, 2009.
Experiment

Cursor = 600 ms

Image1 = 600 ms

Mask = 500 ms

Cursor = 600 ms

Image2 = 600 ms

Mask until response

Less, same, more.
Configurational Processing

Neth & Martinez, JOV, 2009.
Configural Processing

Neth & Martinez, JOV, 2009.
Norm-based Face Space

Multidimensional Face Space

Sadness

Anger

More difficult

Easier

Neth & Martinez, JOV, 2009.
Configural Processing

Neth & Martinez, JOV, 2009.
Computational Space

Neth & Martínez, Vision Research, 2010
Computational Space

Neth & Martinez, Vision Research, 2010
American Gothic Illusion

Neth & Martinez, Vision Research, 2010
Why Configural Features?
Why Configural cues?

sad  neutral  angry

Neth & Martinez, Vision Research, 2010; Du & Martinez, 2011
Proposed Hybrid Model:
Recognizing other emotion labels

\[ c_1 + c_2 + \ldots + c_n \]

Happily surprised

Martinez, CVPR, 2011

Angrily surprised
Configural Processing = Precise detection of facial features

3,930 images

4.2 pixels error (1.5%)

Ding & Martinez, PAMI, 2010
Face Detection
**Features VS context**

**Observation:** Most detections are near the correct location – they are not incorrect, they are *imprecise*.

**Key idea:** Use context information to train where *not* to detect faces and facial features.

Ding & Martinez, CVPR, 2008; PAMI, 2010
Observation: Most detections are near the correct location – they are not incorrect, they are imprecise.

Key idea: Use context information to train where not to detect faces and facial features.

Ding & Martinez, CVPR, 2008; PAMI, 2010
Features VS context

**Observation:** Most detections are near the correct location – they are not incorrect, they are *imprecise*.

**Key idea:** Use context information to train where *not* to detect faces and facial features.

Ding & Martinez, CVPR, 2008; PAMI, 2010
Subclass Discriminant Analysis

Between-subclass scatter matrix:

$$
\Sigma_B = \sum_{i=1}^{C} \sum_{j=1}^{H_i} p_{ij} (\mu_{ij} - \mu)(\mu_{ij} - \mu)^T.
$$

Basis vectors:

$$
\Sigma_B V = \Sigma_X V \Lambda.
$$

How many subclasses (H):

Minimize the conflict, $K$.

Zhu & Martinez, PAMI, 2006
Precise Detailed Detection

Error: 6.2 pixels (2%) vs Manual: 4.2 (1.5%)

Ding & Martinez, CVPR, 2008; PAMI, 2010
Detection + non-rigid SfM

Gotardo & Martinez, PAMI, 2011; Gotardo & Martinez, CVPR, 2011.
Take Home Messages

• What is the form of the computational space in human face perception? Linear combination of known categories.

\[ c_1 + c_2 + \ldots + c_n \]

• What are the dimensions? Mostly configural.

• Precise detection of facial features.
CBCSL


National Institutes of Health
National Science Foundation