Automatic facial expression recognition

A discrete choice approach

Michel Bierlaire

transp-or.epfl.ch

Transport and Mobility Laboratory, EPFL

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Outline

- Facial expressions
- Data
- Variables
- Static expressions
- Dynamic expressions
- Conclusion

Facial expressions

"The face is the most extraordinary communicator, capable of accurately signaling emotion in a bare blink of a second, capable of concealing emotion equally well"

Deborah Blum

Facial expressions

1872 Darwin: universality of facial expressions

1971 Ekman: six primary emotions, that possess each a distinctive content together with a unique facial expression:

- Happiness
- Sadness
- Fear
- Anger
- Disgust
- Surprise

1978 Ekman's Facial Action Unit Coding System (FACS)

Facial expressions

- Active field of research
- Mostly in the machine learning and computer vision communities
- Some difficulties:
  - Context dependency (time, gestures, verbal reaction, etc.)
  - Ambiguity of expressions
  - Need for a ground truth

Data: facial expressions

Cohn-Kanade database

- Testbed for research in automatic facial image analysis
- About 500 image sequences from 100 subjects
- Subjects were instructed by an experimenter to perform a series of 23 facial displays
- Expressions based on FACS action units (see later)

Data: facial expressions

Example of recorded video:

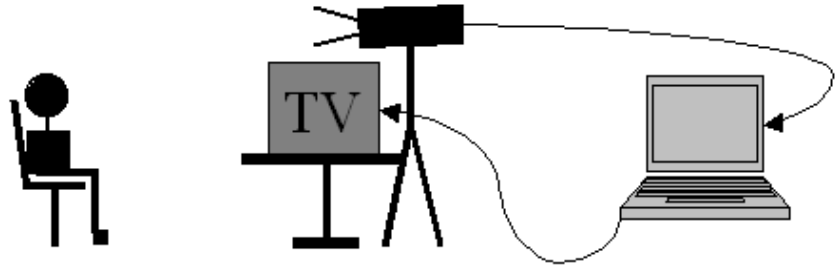
Data: facial expressions

Example of recorded video:

Data: facial expressions

Facial Expressions and Emotion Database, TU Munich

Experimental setup:



Data: facial expressions

Example of video shown:

Data: facial expressions

Example of recorded video:

Data: facial expressions

Another example of recorded video:

Data: facial expressions

- Database: 18 different individuals
- Each individual performed all six desired actions three times.
- Additionally, three sequences doing no expressions at all are recorded.
- Total: 399 sequences.

For more information:

www.mmk.ei.tum.de/ waf/fgnet

Data: choice

Choice experiment:

- Present an image or a video sequence
- Ask the respondent to select the most appropriate expression among
  - `Happiness`
  - `Surprise`
  - `Fear`
  - `Disgust`
  - `Sadness`
  - `Anger`
  - `Neutral`
  - `Other`
  - `I don’t know`

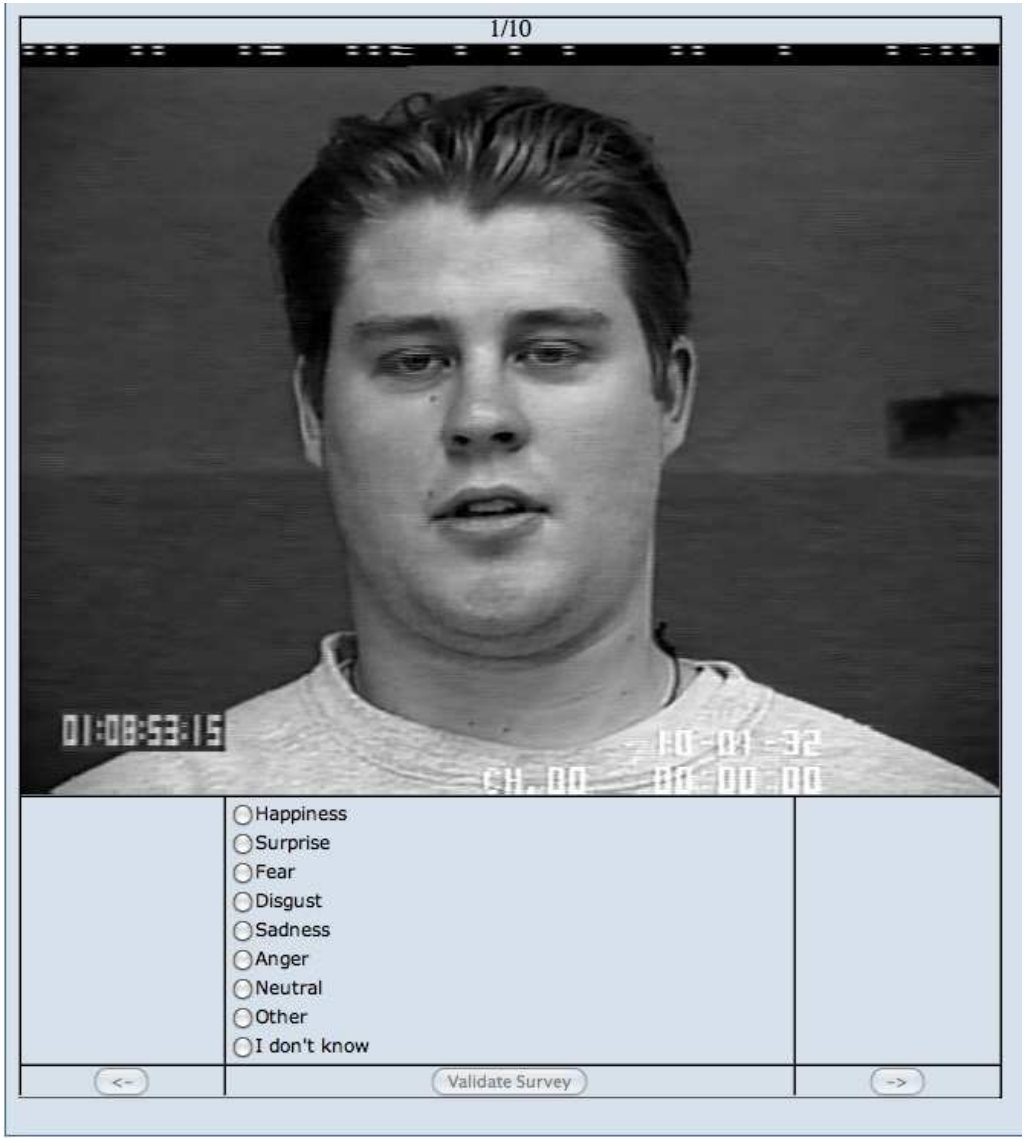
Data: choice

Survey for static images:

`lts5www.epfl.ch/face`

Designed by Matteo Sorci

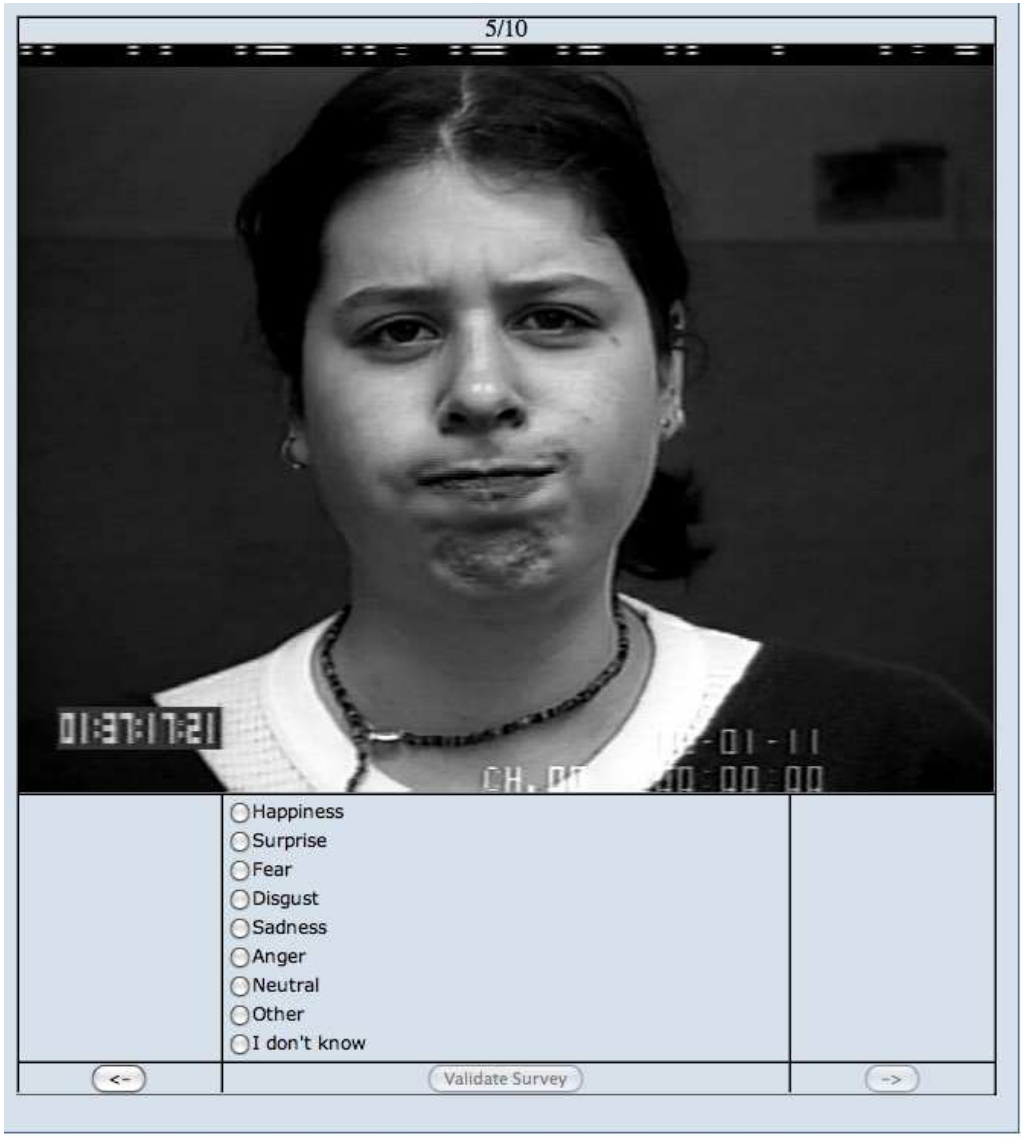
Data: choice



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Data: choice


5/10
Happiness
Surprise
Fear
Disgust
Sadness
Anger
Neutral
Other
I don't know
<-
Validate Survey
->

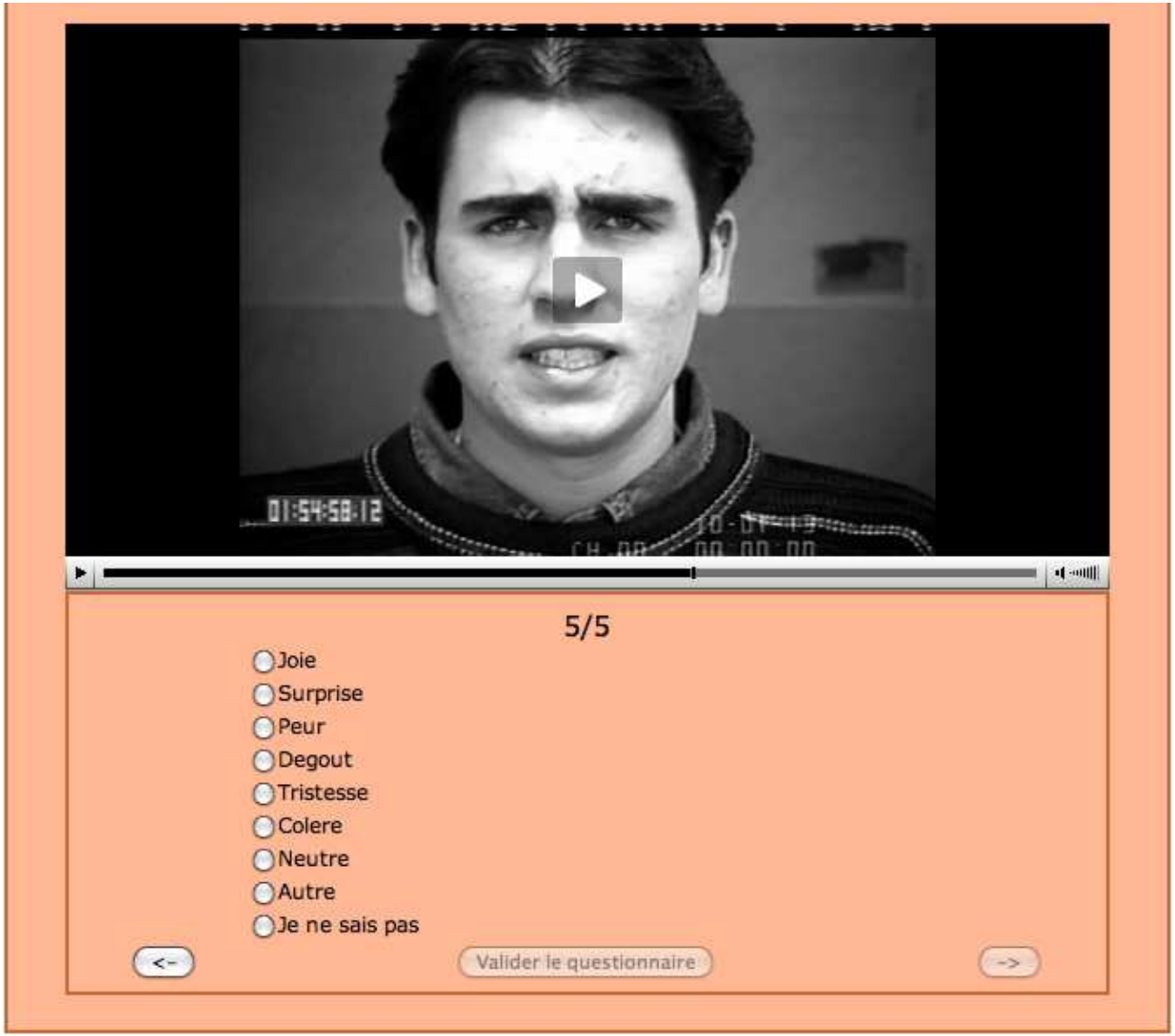
Data: choice

Survey for video sequences:

`transp-or2.epfl.ch/videosurvey`

Designed by Thomas Robin and Javier Cruz

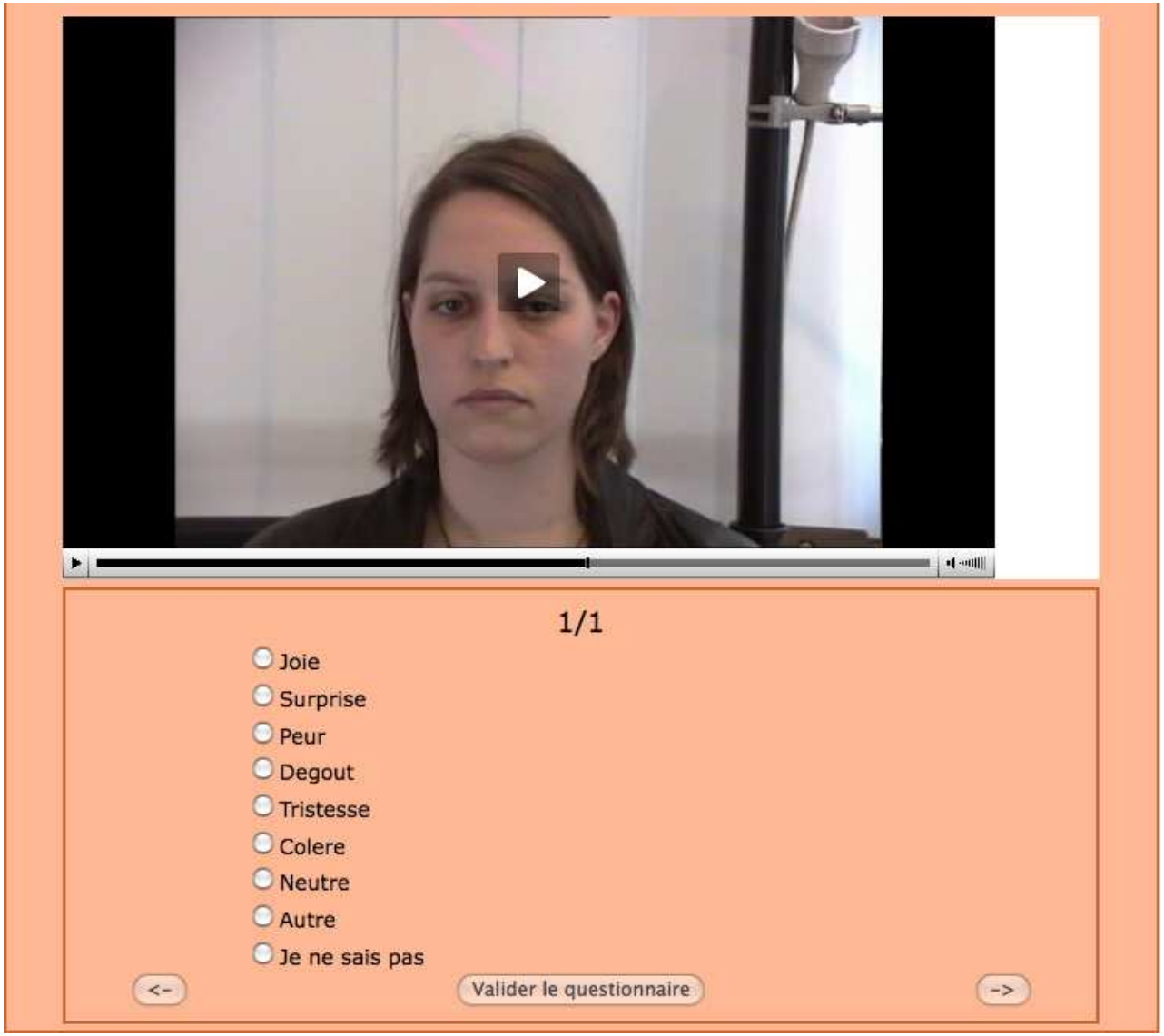
Data: choice



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Data: choice



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Variables

- Objective: explain the choice of an expression using variables describing the face
- Require image analysis techniques
- Cootes, Edwards and Taylor (2001) *Active Appearance Models* PAMI, 23, 681-685.
  - Statistical models of shape and texture
  - Shape and texture are often correlated
  - Correlation learned from PCA
  - Another PCA is used combining the previous two

Data extraction

Shape:

- Automatic extraction of a list of points

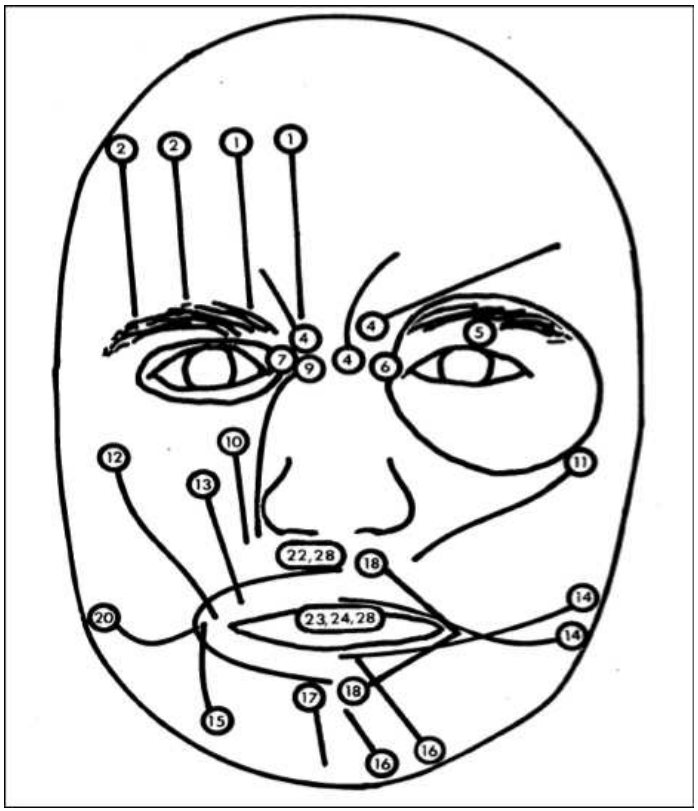
- $x = (x_1, y_1, \ldots, x_n, y_n)'$

Explanatory variables

- FACS Action Units
- Expression Descriptive Units
- Active Appearance Model

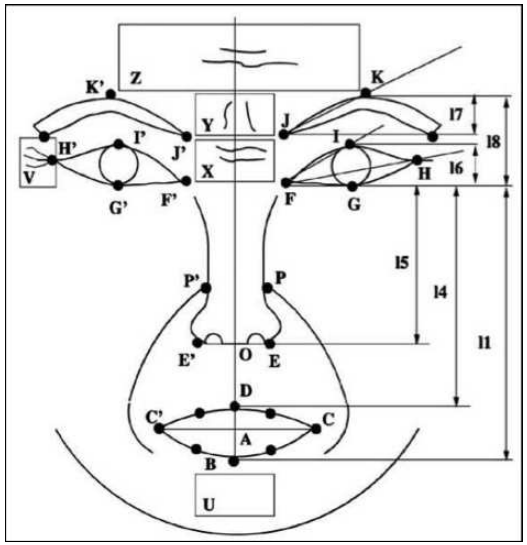
FACS Action Units

- 1978: Ekman & Friesen proposed the Facial Action Coding System
- Measurement units called *Action Units* (AUs)
- AUs are contraction or relaxations of one or more muscles
  - 46 AUs account for changes in facial expression
  - 12 AUs describe changes in gaze direction and head orientation



FACS: leading standard for measuring facial expressions

FACS Action Units



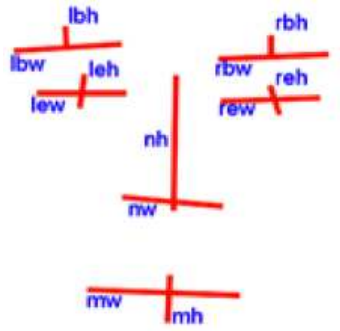
AU1	AU2	AU4	AU5	AU6	AU7
					
Inner Brow Raiser	Outer Brow Raiser	Brow Lowerer	Upper Lid Raiser	Cheek Raiser	Lid Tightener
AU9	AU10	AU12	AU15	AU16	AU17
					
Nose Wrinkler	Upper Lip Raiser	Lip Corner Puller	Lip Corner Depressor	Lower Lip Depressor	Chin Raiser
AU20	AU23	AU24	AU25	AU26	AU27
					
Lip Stretcher	Lip Tightener	Lip Pressor	Lips part	Jaw Drop	Mouth Stretch

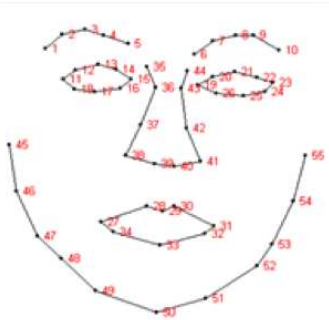
www.bk.isy.liu.se/candide

Expression Descriptive Units

EDU Measures	Measures definition
EDU1	$\frac{1}{2}(\frac{leh}{lew} + \frac{reh}{rew})$
EDU2	$\frac{lbh}{lbw}$
EDU3	$\frac{rbh}{rbw}$
EDU4	$\frac{mh}{mw}$
EDU5	$\frac{nh}{nw}$
EDU6	$\frac{lew}{mw}$
EDU7	$\frac{leh}{mh}$
EDU8	$\frac{leh+reh}{lbh+rbh}$
EDU9	$\frac{lew}{nw}$
EDU10	$\frac{nw}{mw}$
EDU11	$\frac{EDU2}{EDU4}$
EDU12	$\frac{EDU3}{EDU4}$
EDU13	$\frac{EDU2}{EDU10}$
EDU14	$\frac{EDU3}{EDU14}$

Proposed by Antonini, Sorci, Bierlaire and Thiran (2006)



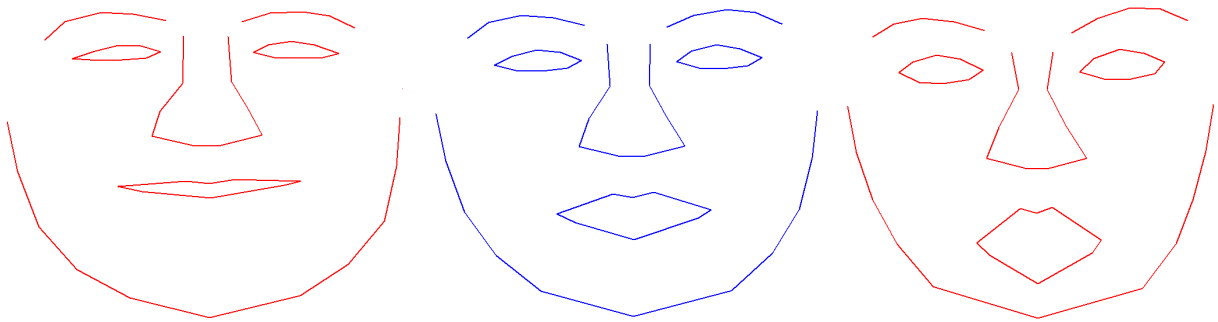


Active Appearance Model

Shape:

- $x = (x_1, y_1, \ldots, x_n, y_n)'$
- Apply PCA

$$x = \bar{x} + P_s b_s$$



Active Appearance Model

Texture:

$\Longrightarrow$ warp to mean shape $\Longrightarrow$

Texture is represented by a vector g

Active Appearance Model

Texture: Apply PCA

$$g = \bar{g} + P_g b_g$$

Finally, apply PCA on b_s and b_g

$$b = \begin{pmatrix} W b_s \\ b_g \end{pmatrix} = Qc = \begin{pmatrix} Q_s \\ Q_g \end{pmatrix} c$$

Varying c changes both shape and texture (demo using AAM lab...)

Static expressions

Discrete choice model:

Decision maker : a person evaluating an expression

Choice set : happiness, surprise, etc. (see surveys)

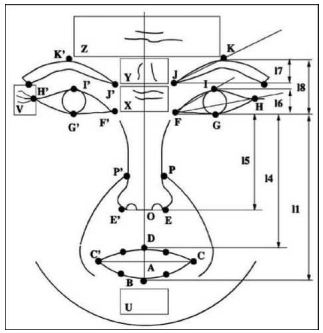
Attributes : FACS, EDU, AAM

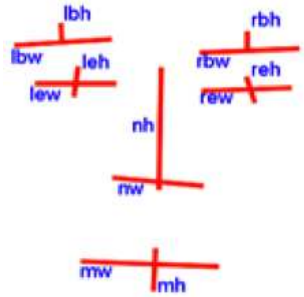
Characteristics : possibility to account for heterogeneity of perceptions and interpretations (not done here)

Model : up to now, logit model

The model

$$V_i = \mathrm{ASC}_i + \underbrace{\sum_k I_{ik}\beta_{ik}^{\mathrm{FACS}}\mathrm{AU}_k}_{\text{FACS}} + \underbrace{\sum_h I_{ih}\beta_{ih}^{\mathrm{EDU}}\mathrm{EDU}_h}_{\text{EDU}} + \underbrace{\sum_\ell I_{i\ell}\beta_{i\ell}^{\mathrm{AAM}}\mathrm{AAM}_\ell}_{\text{AAM}}$$





Data: 39000 observations from 1718 respondents

Model	Nbr of parameters	LL
AU only	93	-57121
AU + EDU	120	-55027
AU + EDU + AAM	145	-54657

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Some parameters

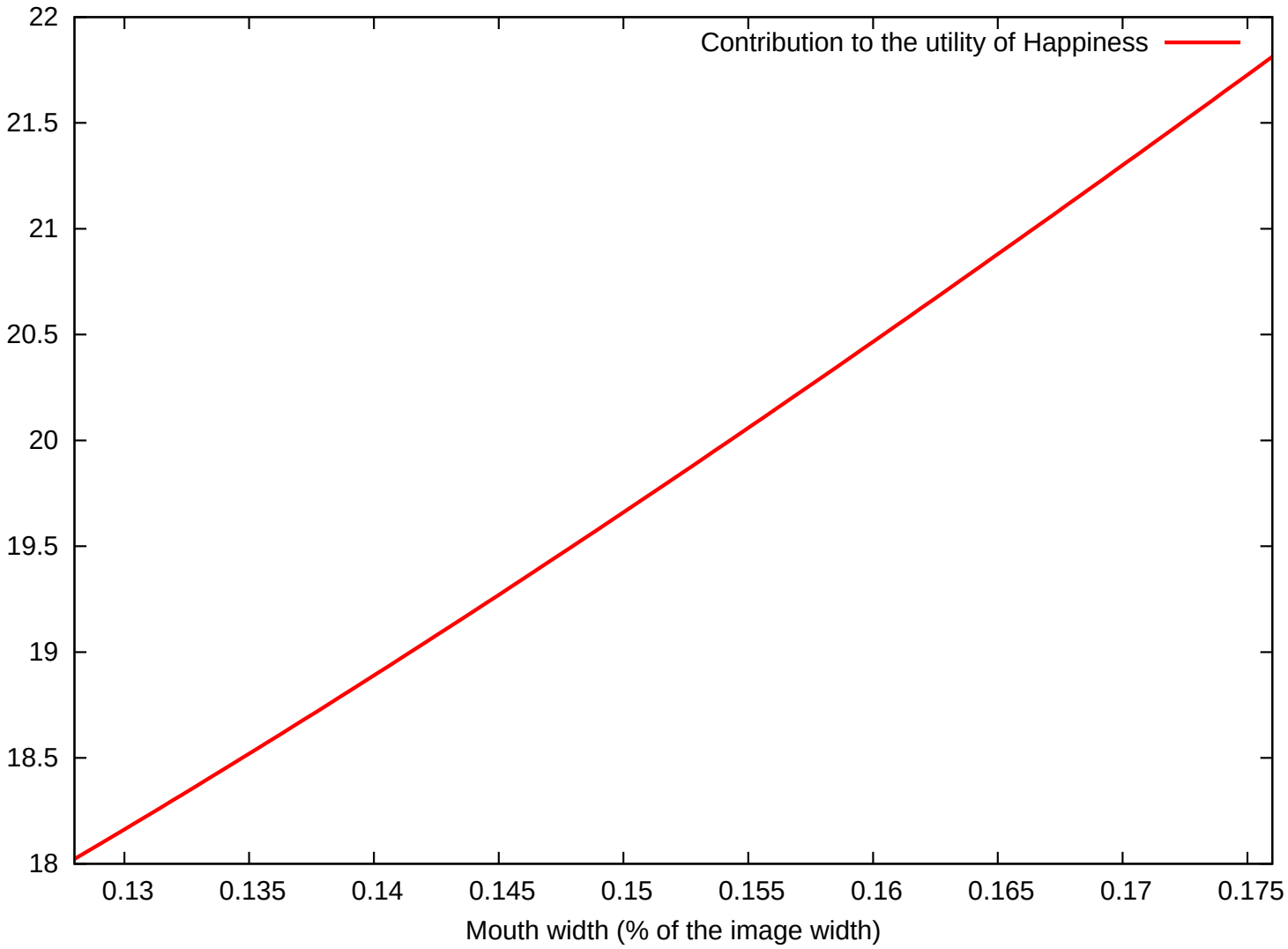
Effects of mouth width on utility of “happiness”

$$U_{\text{happiness}} = \cdots + \beta_1 \frac{\text{mouth height}}{\text{mouth width}} + \beta_2 \text{mouth width} + \cdots$$

	Estimate	t-test
β_1	8.38	8.25
β_2	105	37.64

The wider the mouth, the happier...

Some parameters



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

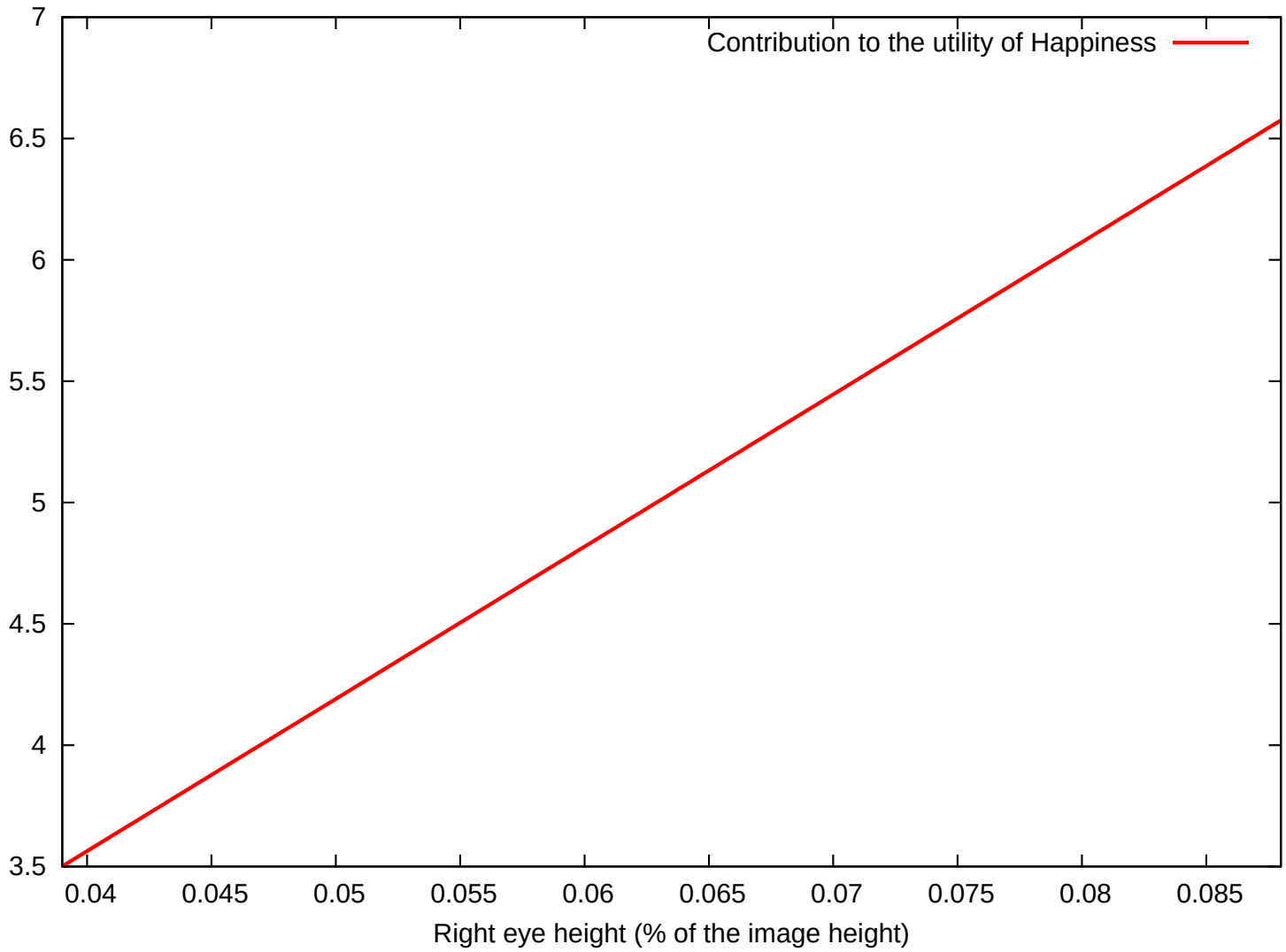
Some parameters

Effects of the height of the right eye on happiness

$$U_{\text{happiness}} = \cdots + \beta_1 \frac{1}{2}\left(\frac{\text{left eye height}}{\text{left eye width}} + \frac{\text{right eye height}}{\text{right eye width}}\right) + \beta_2 \frac{\text{left eye height}+\text{right eye height}}{\text{left eyebrow height}+\text{right eyebrow height}} + \beta_3 \text{right eye height} + \cdots$$

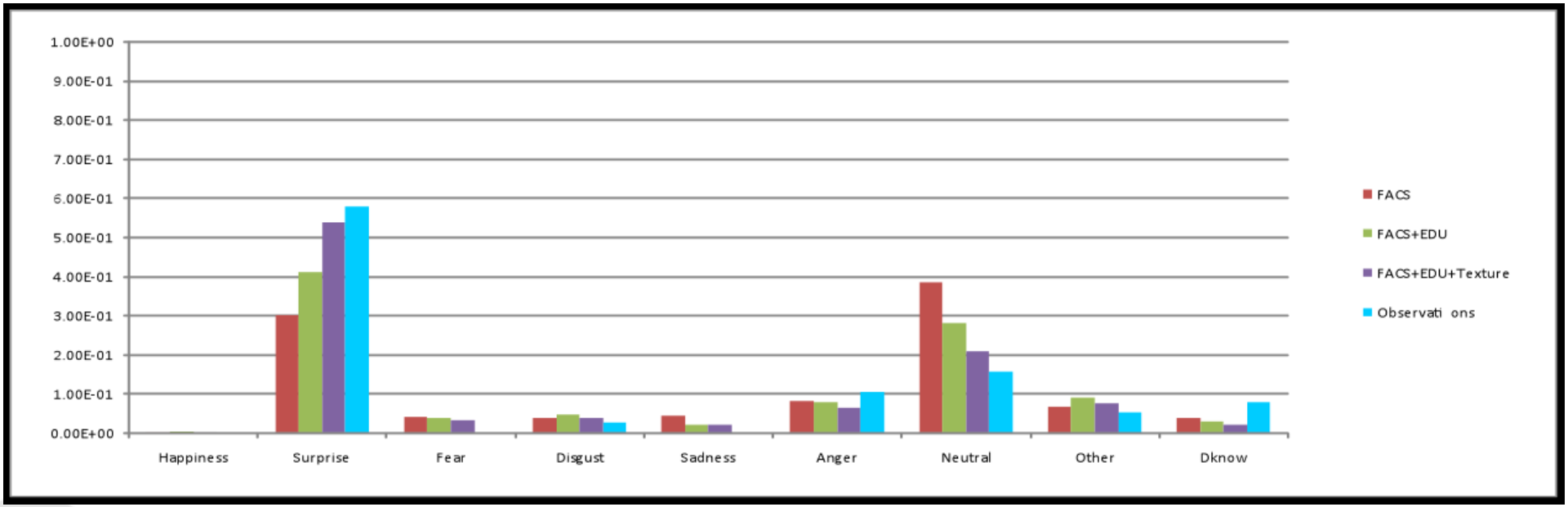
	Estimate	t-test
β_1	-4.61	-5.54
β_2	6.15	8.89
β_3	36	3.95

Some parameters



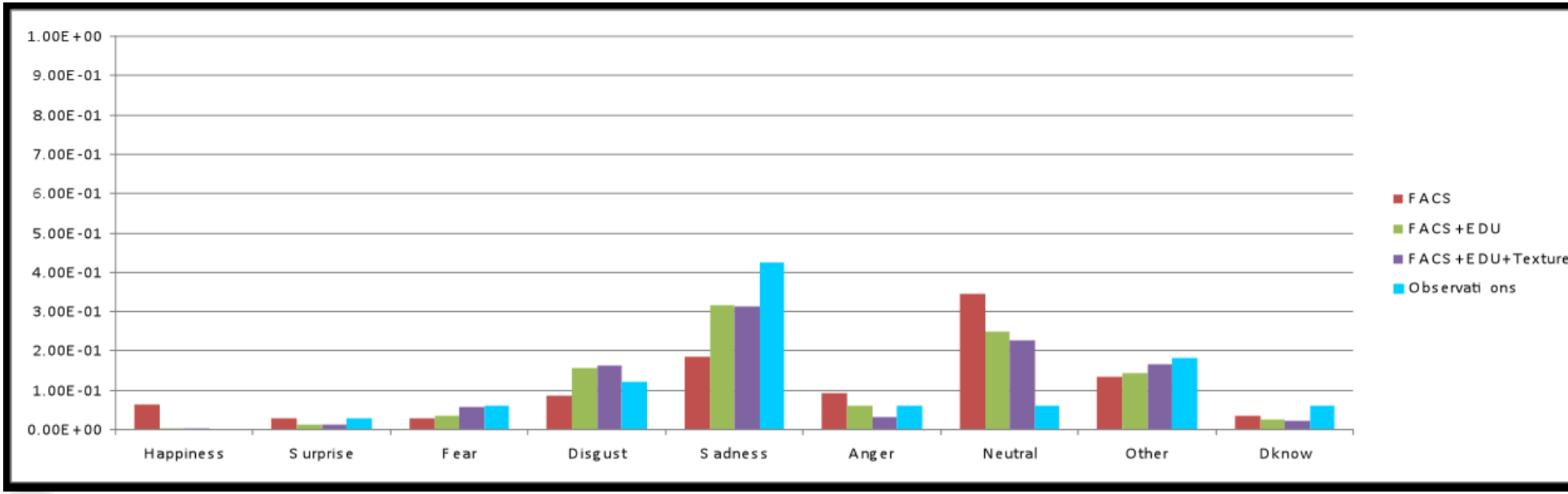
Simulation: good

38 observations

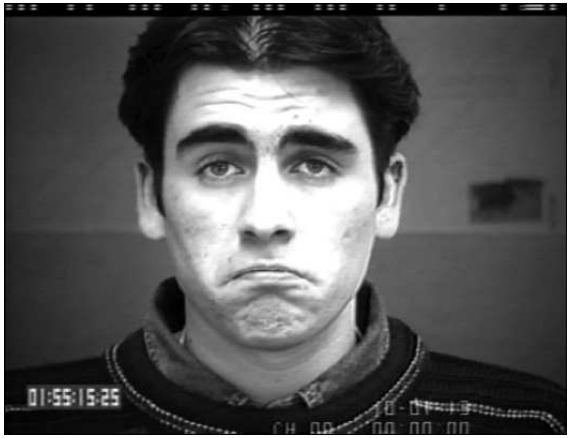


Simulation: not too bad

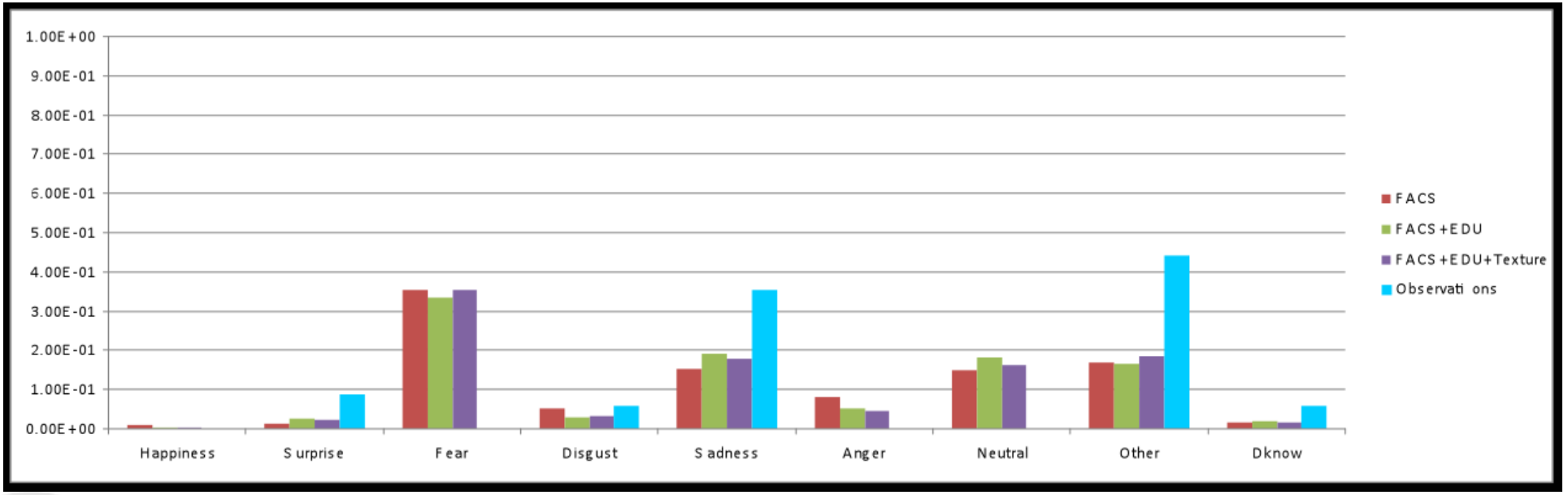
38 observations



Simulation: poor



38 observations



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Dynamic Expressions

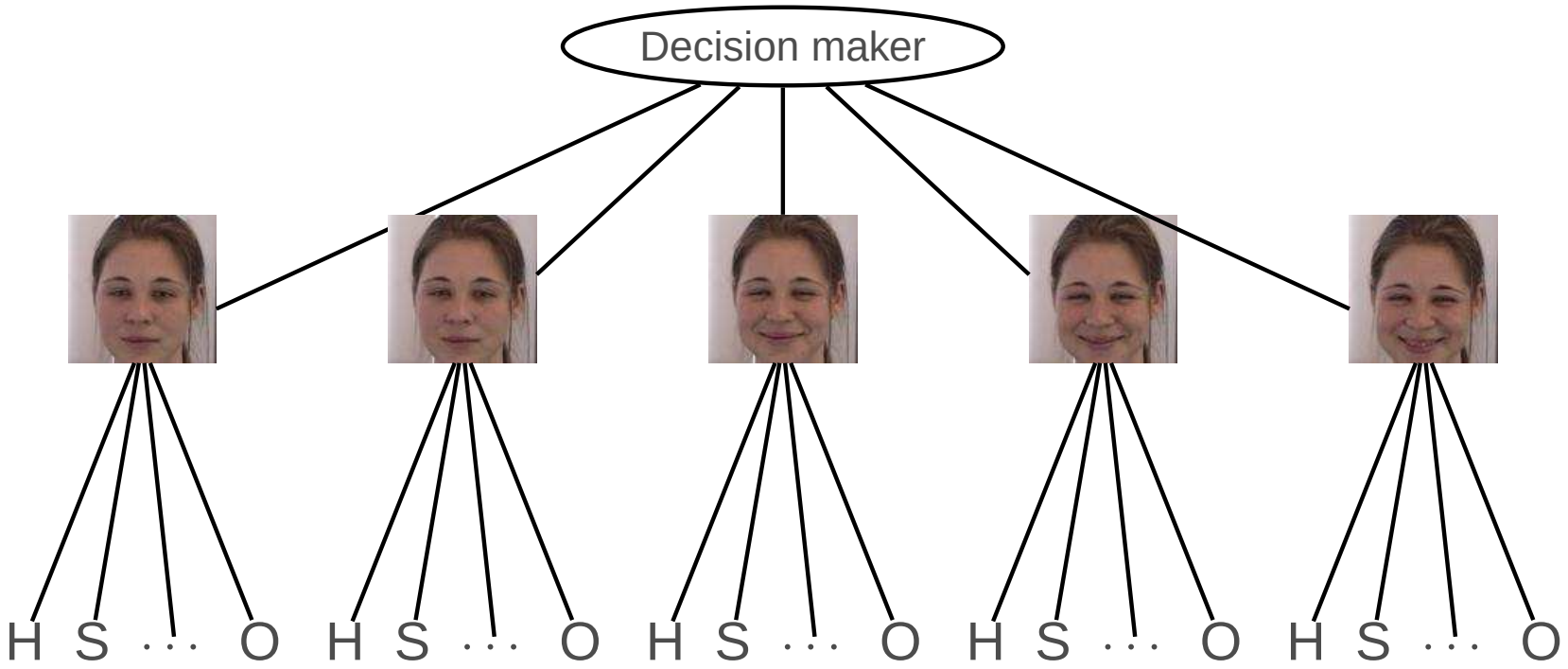
Video = sequence of images

Dynamic Expressions

Modeling assumptions :

- Time discretization (frames) : one second
- One frame will trigger the choice
- The choice of this frame is not observed
- For a given frame, we use the same model as for static expressions

Dynamic Expressions



Models and variables

- Video sequence o
- Decision maker n
- Prob. that frame t is selected: $P_n(t|o)$ (logit). Variables:
  - image attributes, e.g. size of the mouth
  - duration since beginning
- Prob. that expression i is selected in frame t: $P_n(i|t,o)$. Static model (logit)
- Prob. that expression i is selected

$$P_n(i|o) = \sum_t P_n(i|t,o)P_n(t|o)$$

Models and variables

- Panel effect: random parameter ξ_n

$$P_n(i_1, ..., i_{O_n}|\xi_n) = \prod_{o=1}^{O_n} P_n(i_o|o, \xi_n) = \prod_{o=1}^{O_n} \sum_t P_n(i_o|t, o) P_n(t|o, \xi_n)$$

- Integrate

$$P_n(i_1, ..., i_{O_n}) = \int_\xi \prod_{o=1}^{O_n} \sum_t P_n(i_o|t, o) P_n(t|o, \xi) f(\xi) d\xi$$

- Log likelihood

$$\mathcal{L} = \sum_{n=1}^{N} \log \int_\xi \prod_{o=1}^{O_n} \sum_t P_n(i_o|t, o) P_n(t|o, \xi) f(\xi) d\xi$$

Dynamic expression

- Ongoing project
- Estimation procedure under development
- Model specification must be investigated further

Summary

- Main idea: consider the identification of an expression as a choice
- Explanatory variables: features of the face
- Static case:
  - combine three types of variables
  - obtain meaningful model with significant parameters
- Dynamic case:
  - extend the static model
  - assume that one frame is representative

The faces of the team

Thomas Robin

Matteo Sorci

Jean-Philippe Thiran

Javier Cruz

Gianluca Antonini

Michel Bierlaire

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE