Behavioural modeling of facial expression recognition

Thomas Robin, Prof. Michel Bierlaire, Javier Cruz

20th january 2009

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

The context

Human

Face image OR
Face video sequence

Face expression decision

Applications:

- **Driver's attention state**
- **Smart meeting rooms**
- **Human-Machine interfaces**

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Objectives

- Model the facial expression recognition made by a person looking at a
  1. face **image**
  2. face **video sequence**

- Model explicitely the **decision process**

- Estimate the model on **behavioural** data (not classification)

Outline

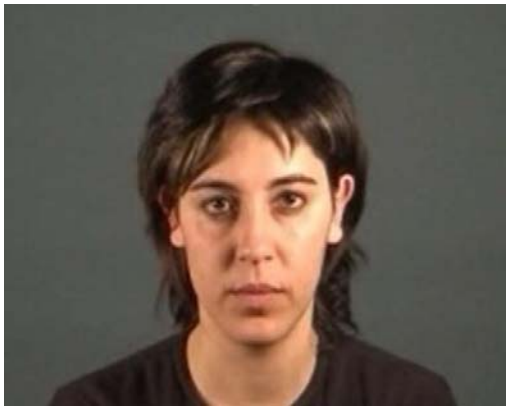
- **Introduction**
- **Features extraction**
- **Static facial expression recognition** (images)
- **Dynamic facial expression recognition** (video sequences)
- **Conclusion and Perspectives**

Introduction

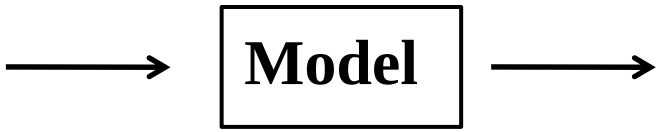
Introduction: facial expression recognition

- **Static:**

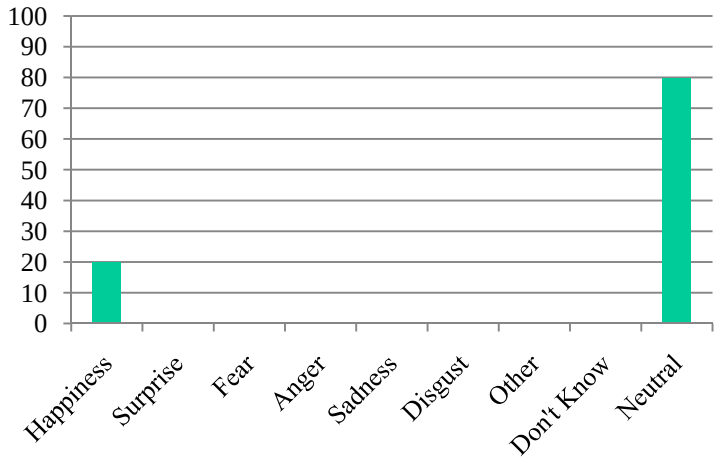
Input:



image



Output:



- **Dynamic:** | **Input** = face video

Output = sequence of probability distribution

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Introduction: Modeling

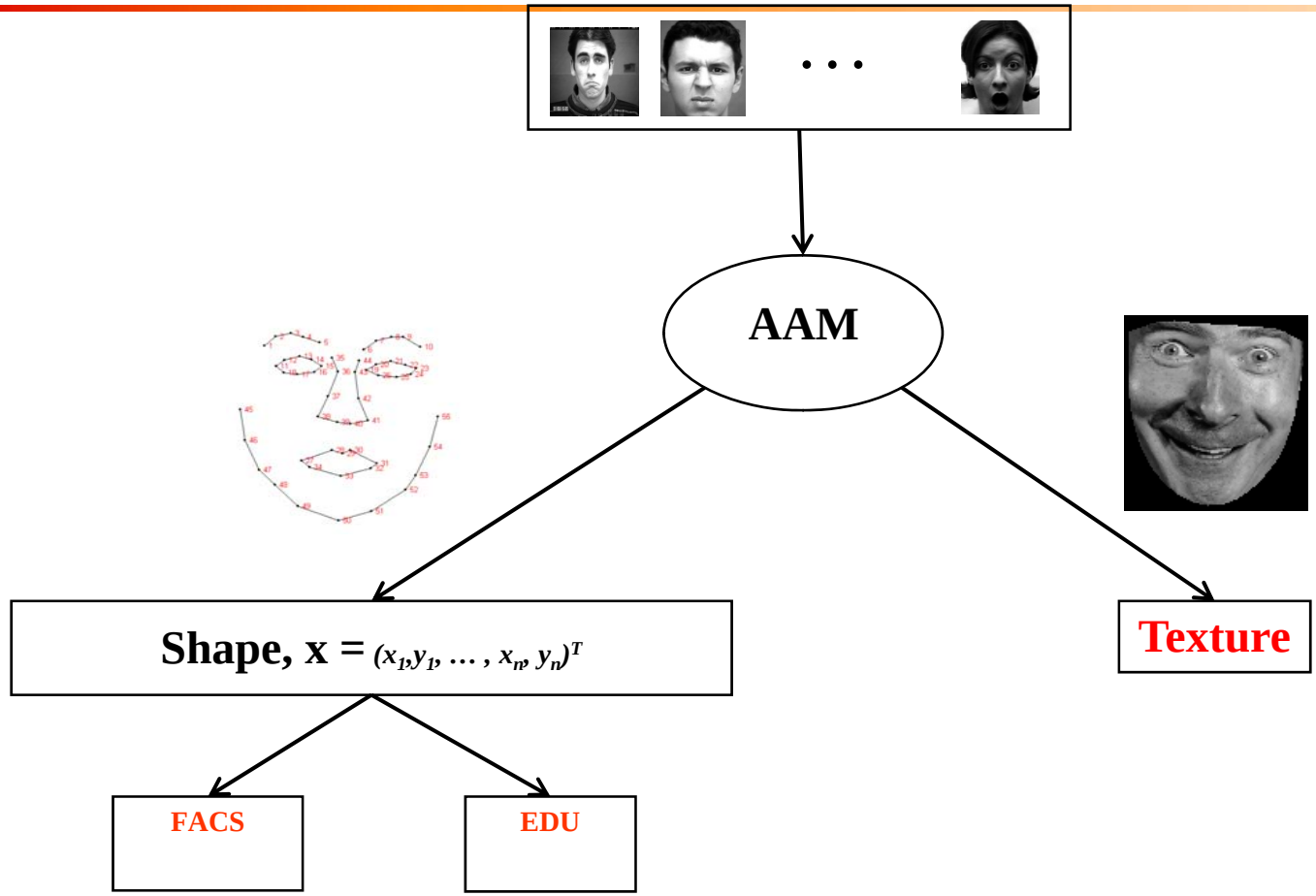
- **Extraction of images/videos informations (features)**
  - ➡ Active Appearence Model (AAM)
- **STATIC face expression recognition**
  - ➡ Discrete Choice Models (DCM)
- **DYNAMIC face expression recognition**
  - ➡ Extension of Discrete Choice Models
  - ➡ Inspired from dynamic models

Introduction: Discrete Choice Models

- **Econometric** models developped since the 50's (transportation, marketing)
- **Disagregate** model
- No need of ground truth → **difference with classification**
- Finite set of choice alternatives (choice set)
- Based on the utility maximisation theory
- Estimated by **likelihood maximization** → **Need of behavioural data**

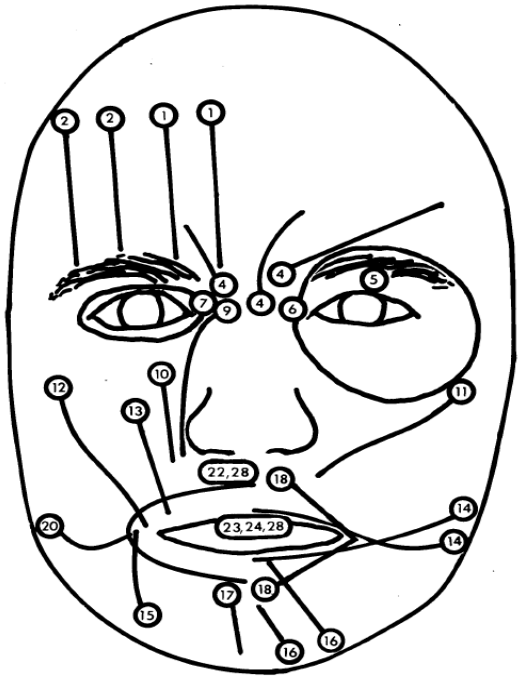
Features extraction

Features extraction: Active Appearance Model



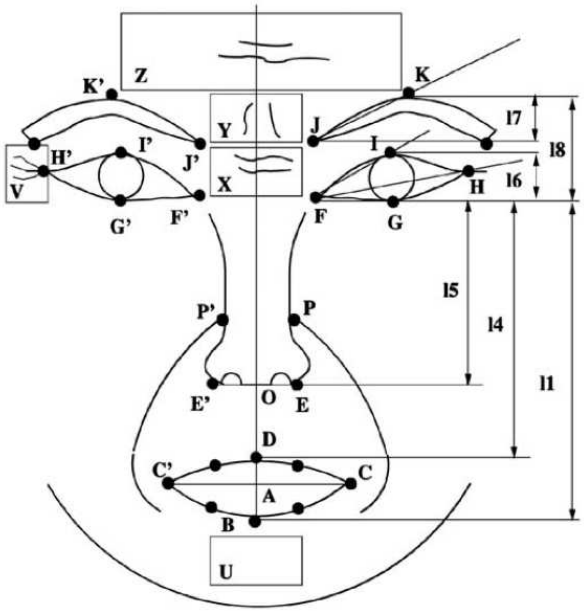
Features extraction: Active Appearance Model FACS

- In 1978 Ekman and Friesen developed the Facial Action Coding System
- Mesurement units: “Action Units” (Aus)
  - AUs are contractions 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



The FACS has become the leading standard for measuring facial expressions

Features extraction: Active Appearance Model FACS



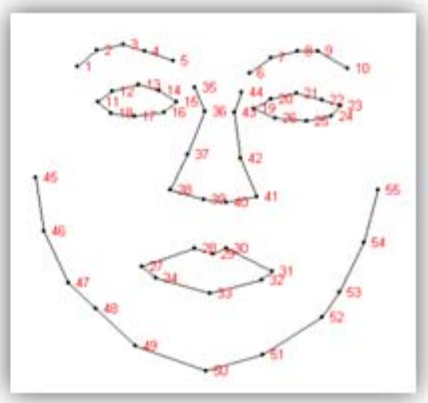
AU1	AU2	AU4	AU5	AU6	AU7
					
Inner Brow Raiser	Outer Brow Raiser	Brow Lowerer	Upper Lid Raiser	Cheek Raiser	Lid Tightener
AU9	AU10	AU12	AU15	AU16	AU17
					
Nose Wrinkler	Upper Lip Raiser	Lip Corner Puller	Lip Corner Depressor	Lower Lip Depressor	Chin Raiser
AU20	AU23	AU24	AU25	AU26	AU27
					
Lip Stretcher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

TRANSP-OR

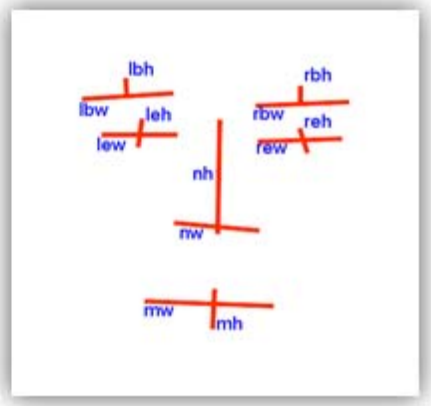
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Features extraction: Active Appearance Model EDU

- Expression Descriptive Units by Antonini, Sorci, Bierlaire and Thiran in « Discrete Choice Models for Static Facial Expression Recognition »



EDU1	$\frac{lew+rew}{leh+reh}$	**EDU8**	$\frac{leh+reh}{lbh+rbh}$
EDU2	$\frac{lbw}{lbh}$	**EDU9**	$\frac{lew}{nw}$
EDU3	$\frac{rbw}{rbh}$	**EDU10**	$\frac{nw}{mw}$
EDU4	$\frac{mw}{mh}$	**EDU11**	EDU2 / EDU4
EDU5	$\frac{nh}{nw}$	**EDU12**	EDU3 / EDU4
EDU6	$\frac{lew}{mw}$	**EDU13**	EDU2 / EDU10
EDU7	$\frac{leh}{mh}$	**EDU14**	EDU3 / EDU10



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Features extraction: Active Appearance Model C parameters

- Results of the **Principal Component Analysis (PCA)** of the images

Levels of grey on the faces

Static face expression recognition

Static case: model presentation

→ Use of a Discrete Choice Model

- A choice theory defines :
  - **a decision maker** : a human
  - **alternatives** : possible expressions
  - **attributes of alternatives** : characteristics of the images (FACS, EDU, C parameters)
  - **decision rule :** utility maximisation theory
- Choice set :

  Happiness, Surprise, Anger, Fear, Disgust, Sadness, Neutral, Other, Don't know

Static case: DCM, decision rule

- Utility maximisation theory

  - Association of a function, called **utility** to each alternative
  - It depends on the **alternative *i*** , and on the **decision maker *n***

  ➡ The decision maker ***n*** will choose the alternative ***i*** which has the **higher utility**

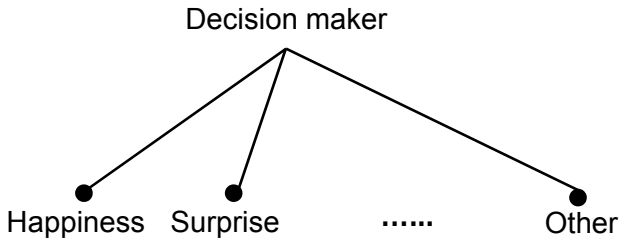
$$U_{in} = V_{in} + \epsilon_{in}$$

V_{in} **: deterministic part of the utility of alternative *i* for individual *n***

ϵ_{in} **: error term, different assumptions can be made on its distribution**

Static case: DCM, specification of the error term

- We suppose ϵ_{in} independent and identically distributed (iid) with an extreme value distribution



Multinomial Logit model

$P_n(i|C_n)$: probability for the individual ***n*** to choose the alternative ***i***

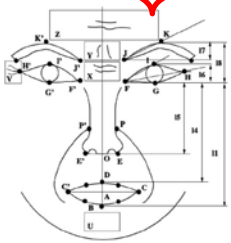
$$P_n(i|C_n) = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}}$$

C_n : Choice set, depends on the individual ***n***

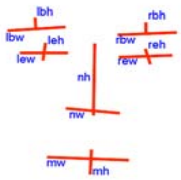
Static case: DCM, utility specification

- 3 different specifications

$$V_j = ASC_j + \sum_{k=1}^{K_1} I_{kj}\,\beta_{kj}^{FACS}\,FACS_k + \sum_{h=1}^{K_2} I_{hj}\,\beta_{hj}^{EDU}\,EDU_h + \sum_{l=1}^{K_3} I_{lj}\,\beta_{lj}^{C}\,C_l$$



+



+

model 1

model 2

model 3

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Static case: behavioural data

- Estimated by **likelihood maximization** → **Need of behavioural data**

- **People have to label randomly chosen pictures**
- **People report their socio-economics characteristics (age, formation, job…)**
- **1718 participants for more than 39000 labeled images**

http://lts5www.epfl.ch/face

Static case: behavioural data

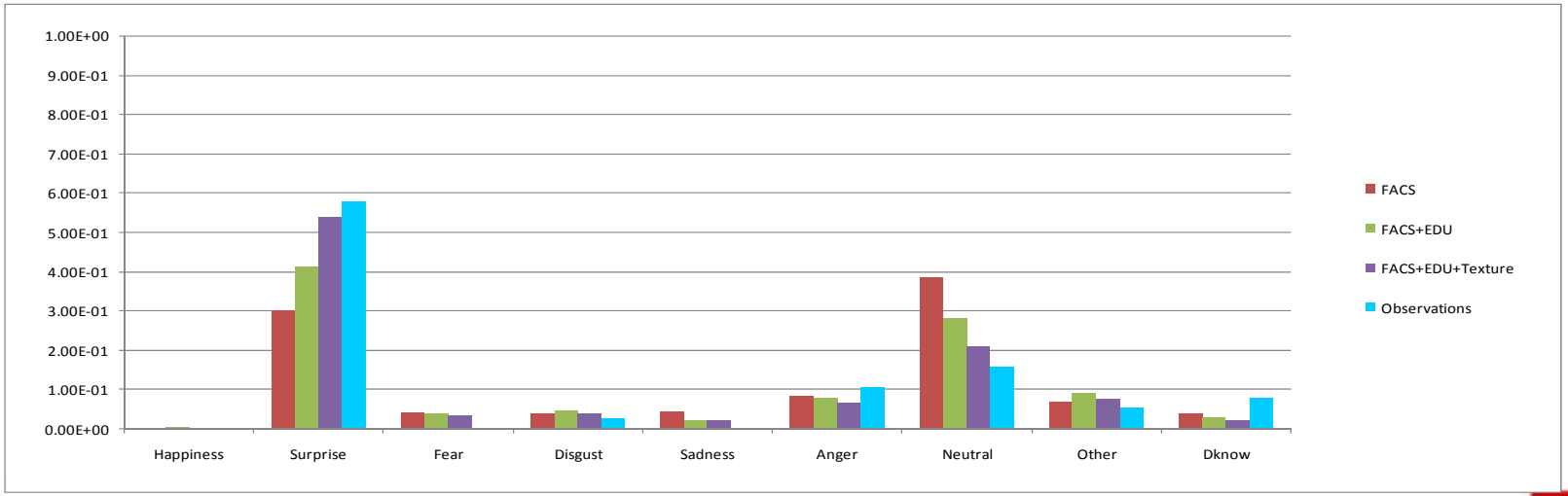
- Estimation of the models with Biogeme (biogeme.epfl.ch)

$$L = \sum_{n=1}^{N} \left(\sum_{j=1}^{9} y_{jn} \log(P_{jn}) \right)$$

- Model 1: “FACS” (Primary AU + Secondary AU + Transient Features)
  - **93 parameters , LL = - 57121**
- Model 2: “FACS + EDU”
  - **120 parameters , LL = - 55027**
- Model 3: “FACS + EDU + C COEFFICIENTS”
  - **145 parameters , LL = - 54657**

Static case: simulation

38 observations

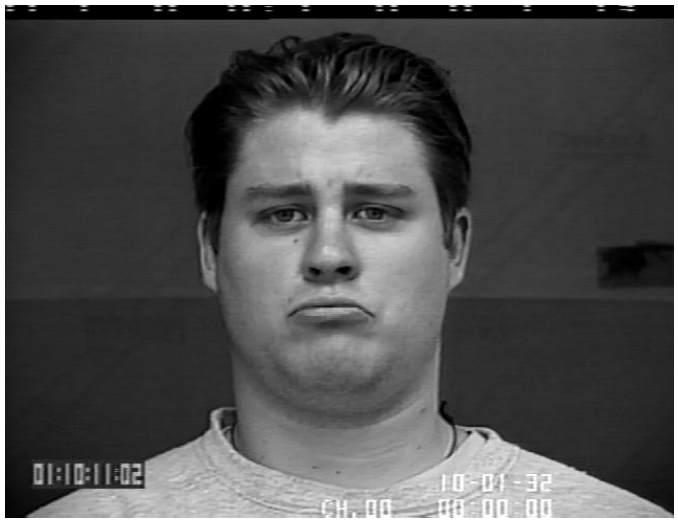


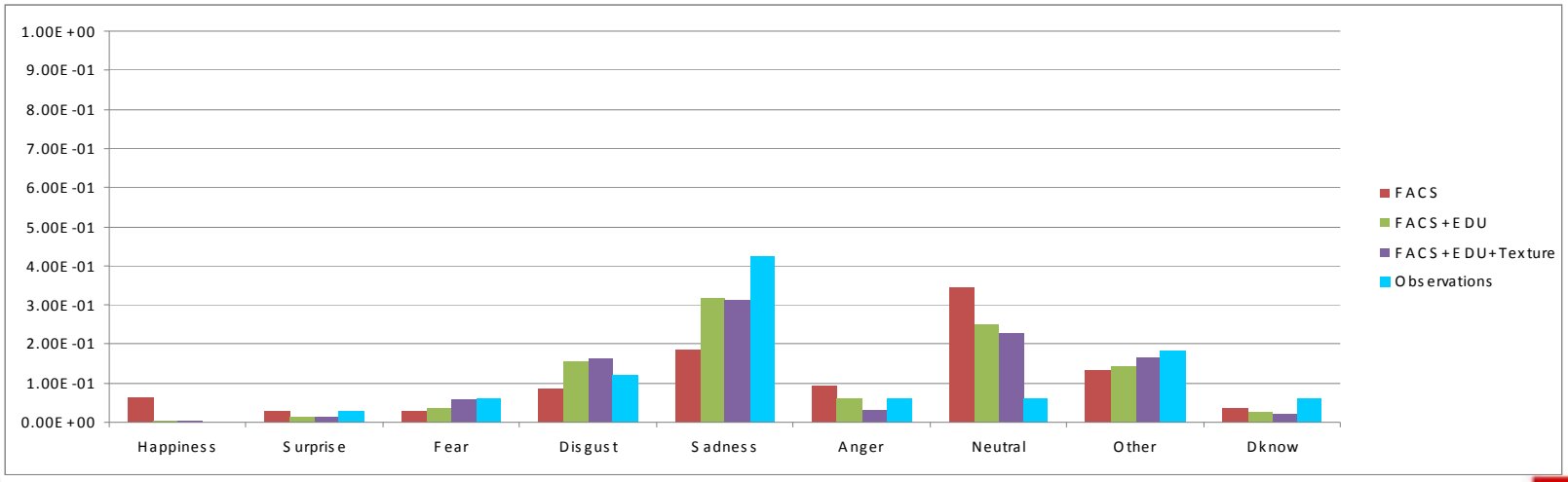
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Static case: simulation

33 observations



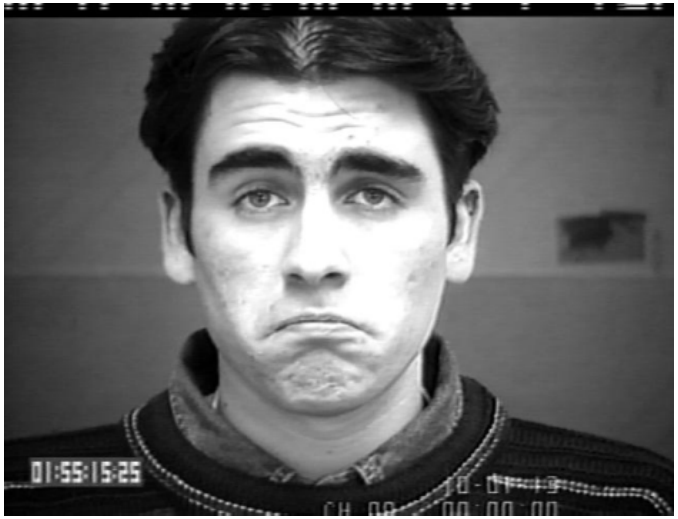


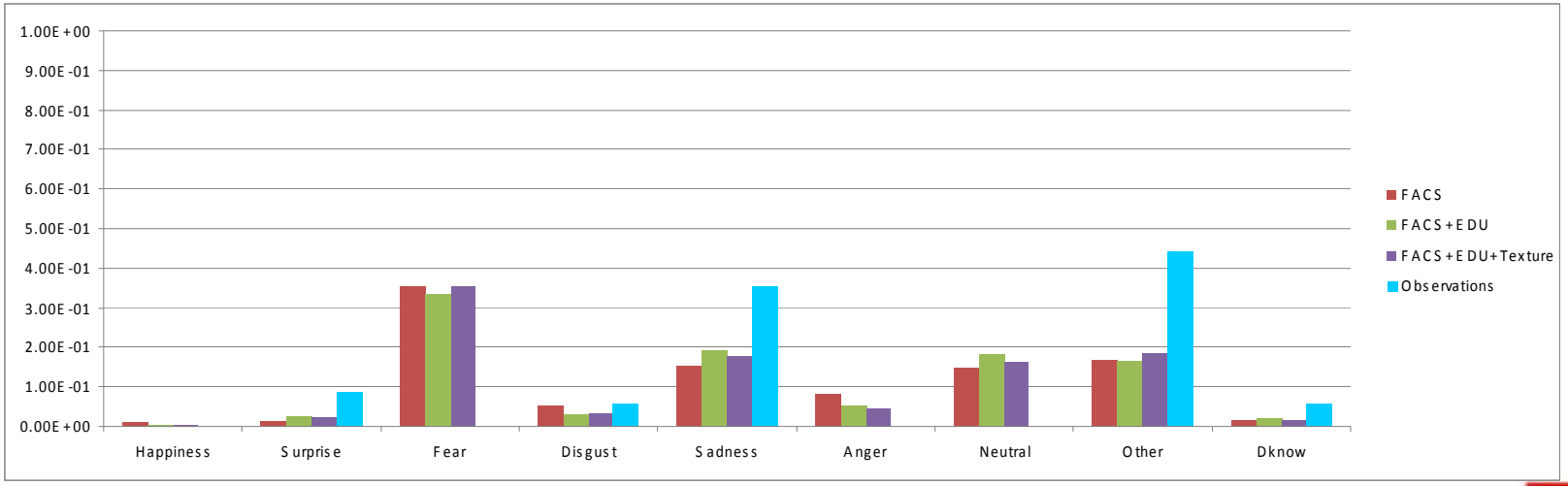
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Static case: simulation

34 observations





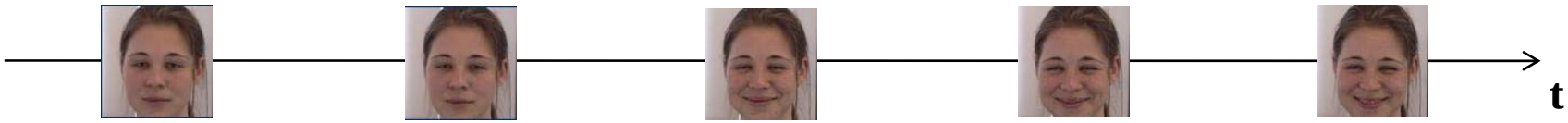
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Dynamic face expression recognition

Dynamic case: introduction

- Video sequence = succession of images



- Expression evaluation made at each frame (or time step, 1s)

  ➡ **Latent pre-choices**

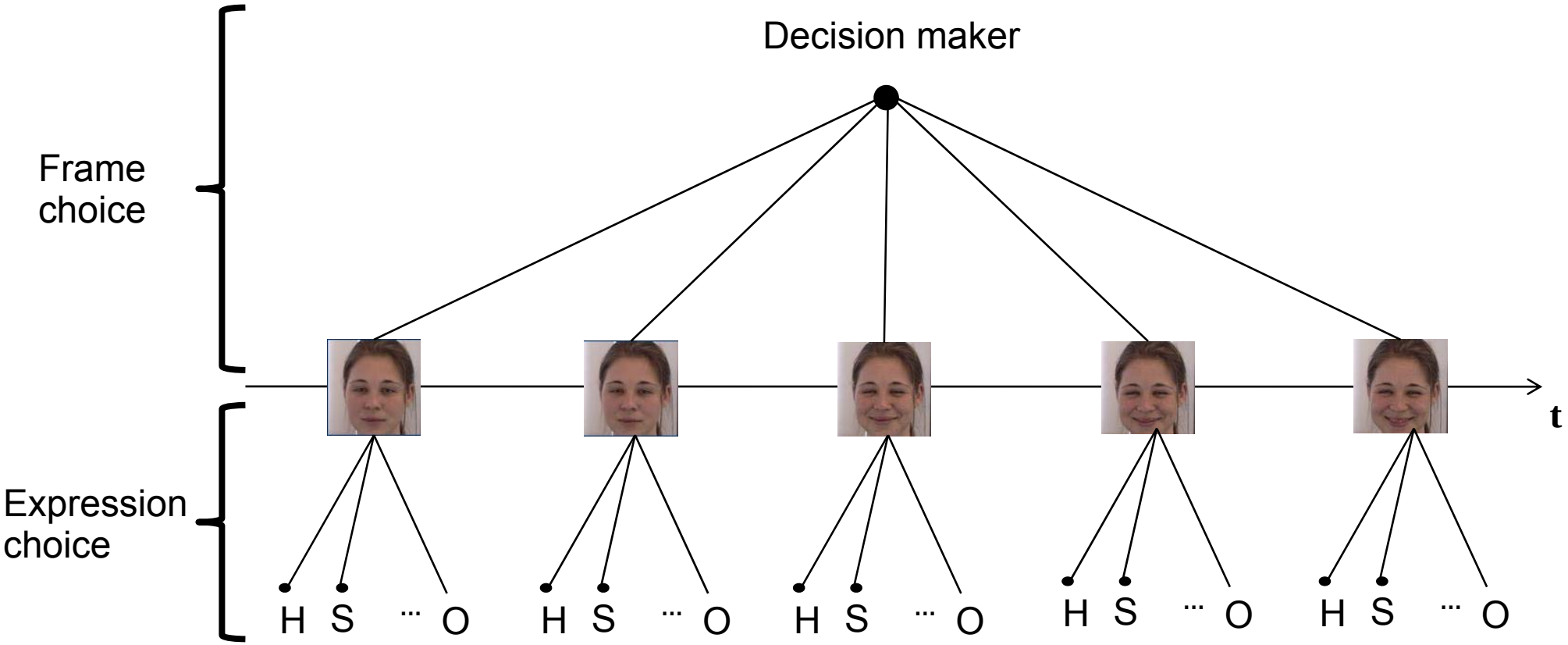
  ➡ **Discrete choice model**

- Need of weight the frame evaluation in the choice expression process

  ➡ **Discrete choice model**

- Estimation by likelihood maximization

Dynamic case: introduction



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Dynamic case: the model

Probability of choosing the expression ***i***, for the individual ***n***, watching the video sequence ***o***, in the frame ***t***

multivariate density function of ξ_n

$$P_{o,n}(i) = \int \sum_{t=1}^{T_o} P_{o,n}(i/t, \xi_n) P_{o,n}(t) f(\xi_n) d\xi_n$$

Probability of choosing the expression ***i***, for the individual ***n***, watching the video sequence ***o***

Probability for the individual ***n***, when watching at the video ***o*** to make his choice when faced to the frame ***t***

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Dynamic case: behavioral data

- Survey conducted at the address below(English, French, Italian, Spanish):

  http://transp-or2.epfl.ch/videosurvey/

- Respondents have to:
  - create an account
    → Socioeconomics characteristics
  - label some video sequences with expressions
    → observations

- 2 databases of video are used:
  - Cohn-Kanade
  - Technical University Munich (TUM)

Dynamic case: behavioral data: video database

- The Cohn-Kanade database

  ➡ Actors **playing** expressions, according to the Facial Action Coding System (**FACS**)

55 sequences, 11 subjects

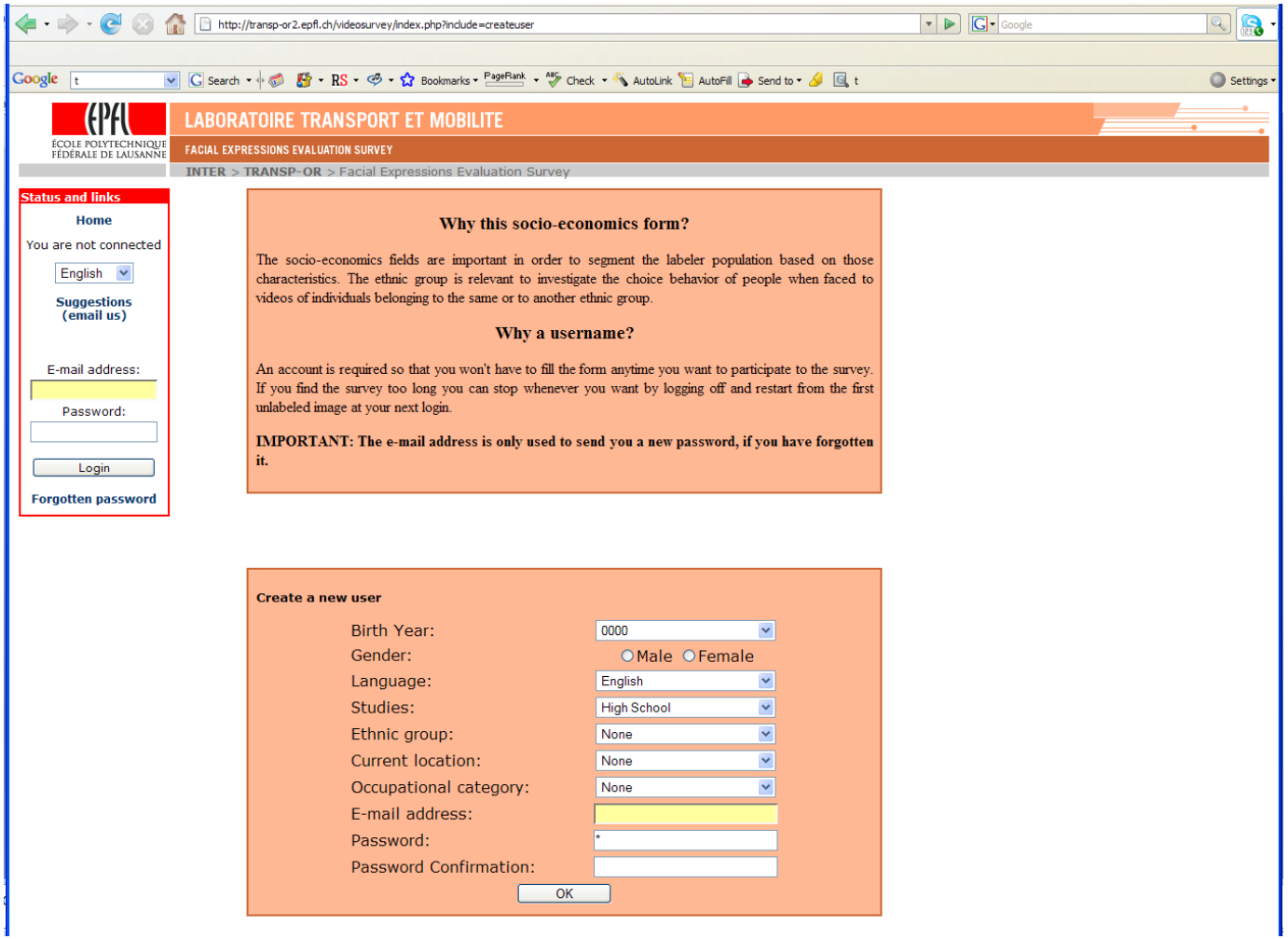
Dynamic case: behavioral data: video database

- The Technical University Munich database (TUM)

  ➡ Students faced to a video, natural expressions recorded

399 sequences, 18 subjects

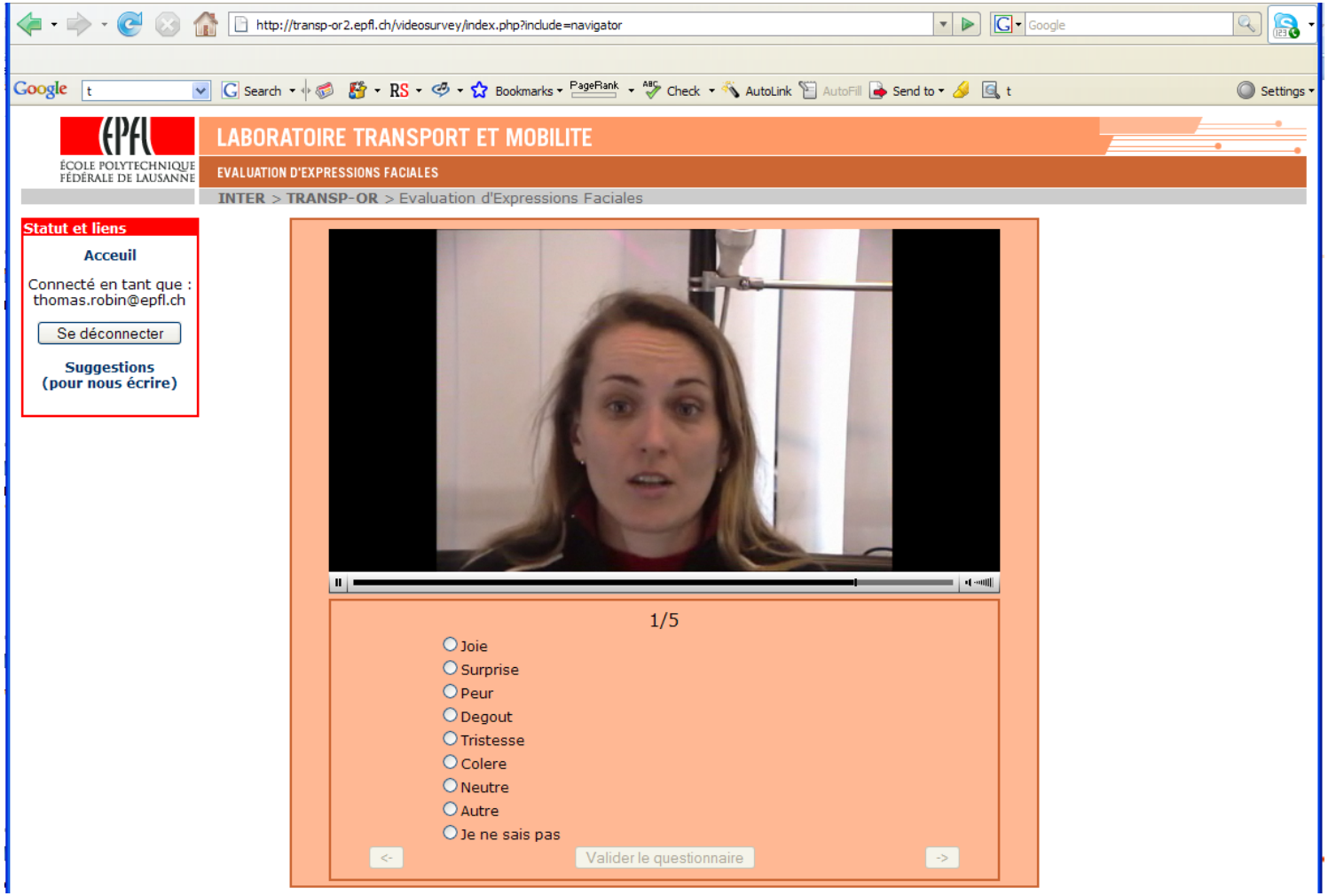
Dynamic case: behavioral data: internet survey



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Dynamic case: behavioral data: internet survey



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Dynamic case: likelihood function

$$l = \prod_{n=1}^{N} \prod_{o=1}^{O_n} P_{o,n}(i)$$

$$l = \prod_{n=1}^{N} \prod_{o=1}^{O_n} \left(\prod_{i=1}^{E} \int \sum_{t=1}^{T_o} P_{o,n}(i/t, \xi_n) P_{o,n}(t) f(\xi_n) d\xi_n \right)^{c_{i,o,n}}$$

$$l = \prod_{n=1}^{N} \prod_{o=1}^{O_n} \left(\prod_{i=1}^{E} \int \sum_{t=1}^{T_o} \frac{\exp(v_{i,t,o,n}(\xi_n))}{\sum_{j=1}^{E} \exp(v_{j,t,o,n}(\xi_n))} \frac{\exp(\overline{v}_{t,o,n})}{\sum_{a=1}^{T_o} \exp(\overline{v}_{a,o,n})} f(\xi_n) d\xi_n \right)^{c_{i,o,n}}$$

Conclusions and Perspectives

- Conclusion:
  - **static** case: Discrete Choice Model
  - **dynamic** case: new model framework

- Perspectives for the **dynamic** case:
  - implementation of the likelihood maximization
  - model estimation: find a satisfactory specification
  - model validation: measure the prediction power

Thank you for your attention