Dynamic facial expression recognition using a discrete choice model

Thomas Robin **Michel Bierlaire** **Javier Cruz**

9th november 2009

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

The context

- Recent interest for emotion recognition in transportation

  → Driving assistance

  - Safety
  - Mobility

  → Well-being measuring of users

  - Improve public transportation offers
  - Improve car comfort

The context

- Emotion: **mental** and **physiological** state associated with a wide variety of feelings, thoughts and behavior.

- Emotions signs easy to measure with non-intrusive techniques for transportation users:

  - Behavior
  - Facial expression
  - Voice intonation

The context

Driving assistance

→ Adapt car behavior to a danger

Well being measuring

→ Improve level of service

Objectives

- Model the facial expression recognition made by a person looking at a face video sequence
- Model explicitely the **dynamic process**
- **No classification**
- Estimate the model on **behavioural** data (relax ground truth assumptions)

Outline

- **Estimation data:** - Video: Database and feature extraction
  - Internet survey
- **Model:** - Framework
  - Specification
  - Estimation
  - Predictions

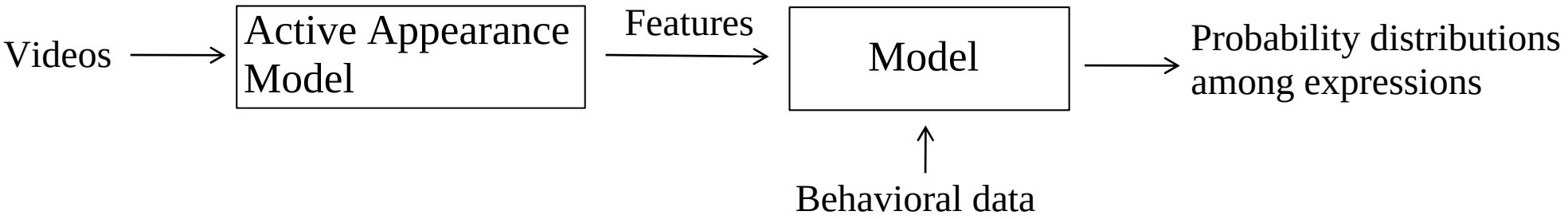
Presented at STRC 2009

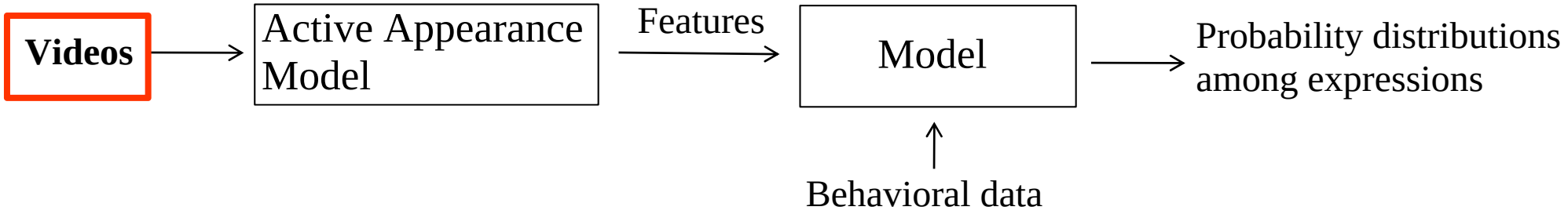
- **Validation:** - Procedure
  - Prediction on estimation data
  - Cross-validation
  - New video database
- **New Models:** - Model 2
  - Model 3
  - Model 4
  - New validation

NEW

Introduction

- Model overview:




Videos
Active Appearance Model
Features
Model
Probability distributions among expressions
Behavioral data

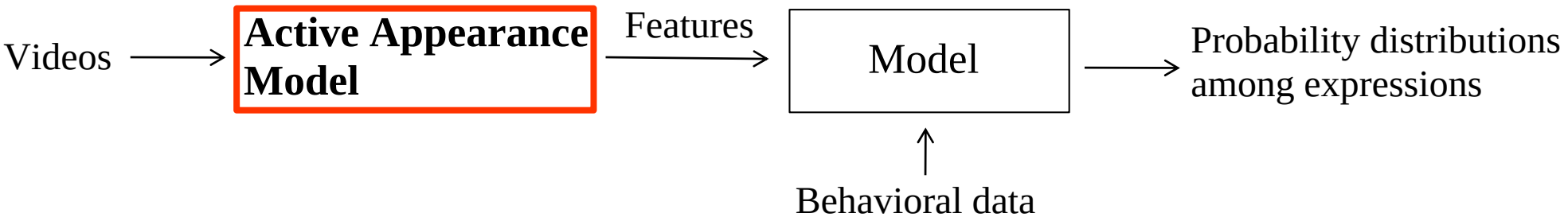
Data: video database

- The Technical University Munich database (TUM)
  Facial Expression and Emotion Database (FEED)

➡ Students faced to a video, natural expressions recorded

138 sequences, 18 subjects

Data: video database


Videos
Active Appearance Model
Features
Model
Probability distributions among expressions
Behavioral data

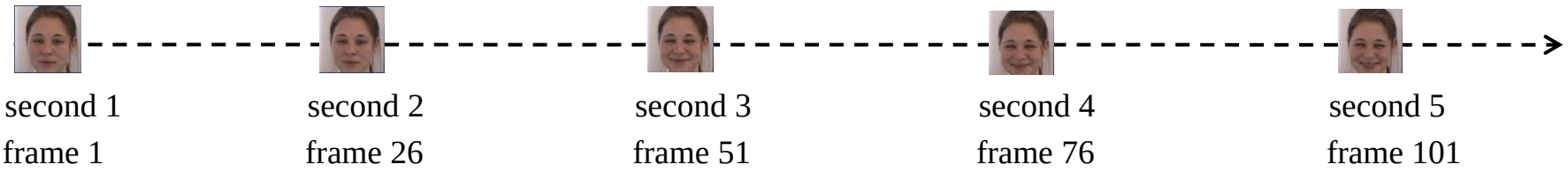
Features extraction: Active Appearance Model

- Video = succession of images, called frames

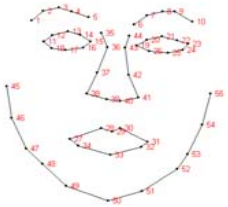
  ⇒ information extracted on each frame

- Hypothesis: individual perception evolves at regular time step (1 s)

  ⇒ - a video contains 25 frames per second
  - first frame of each second retained



Features extraction: Active Appearance Model



AAM

Shape, x = $(x_1, y_1, \ldots, x_n, y_n)^T$

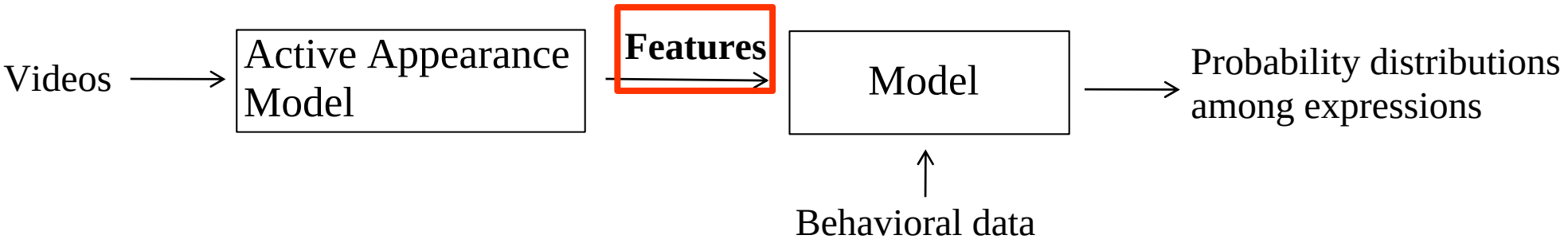
Texture

(C vectors)

FACS

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

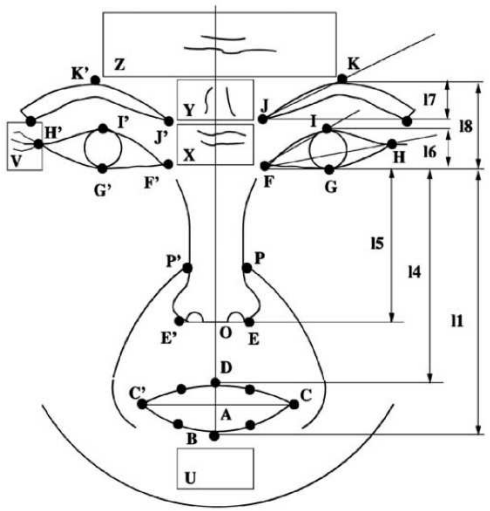

Videos
Active Appearance Model
Features
Model
Probability distributions among expressions
Behavioral data

Features: Facial Action Coding System

- FACS developped by Ekman and Friesen (1978)
- Mesurement units, called “Action Units” (AUs) associated to expressions

➡ **leading standard for measuring facial expressions**

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



TRANSP-OR

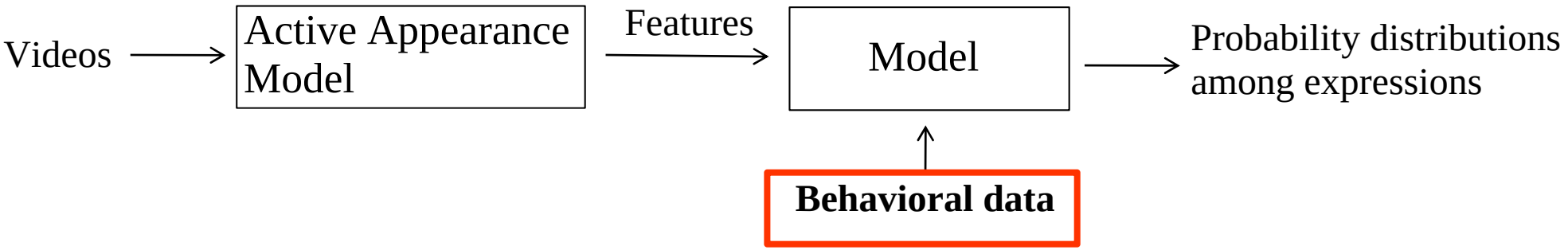
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Features: C vectors

- Direct output of the Principal Component Analysis (PCA) conducted in the AAM

➡ It characterises both **face shape** and **face shadows**

⟶ C vector: 100 elements


Videos
Active Appearance Model
Features
Model
Probability distributions among expressions
Behavioral data

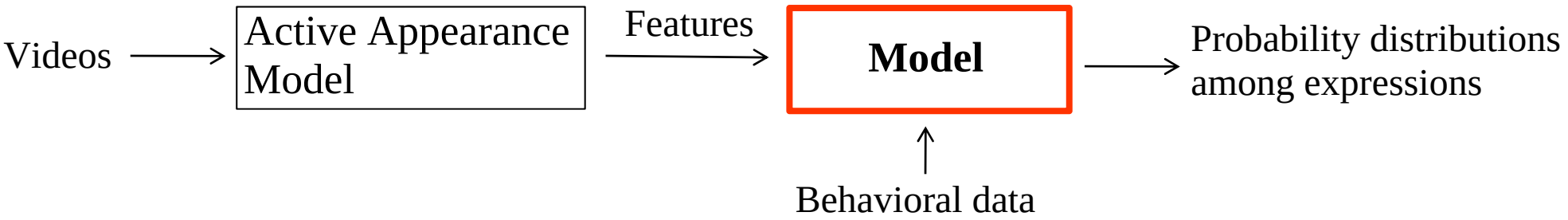
Data: internet survey

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

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

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

- 1 database of video is used:
  - Facial Expression and Emotion Database (FEED)


Videos
Active Appearance Model
Features
Model
Probability distributions among expressions
Behavioral data

Model: framework

- Combination of 2 DCM

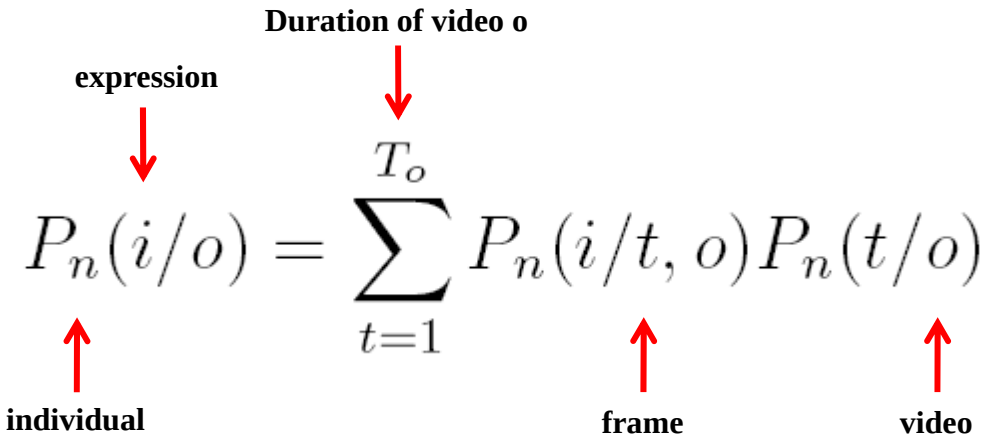
→ **Instantaneous expression perception sub-model**

- Similar to static model
- Choice set: expressions
- Attributes: facial features

→ **Video frames weighting sub-model**

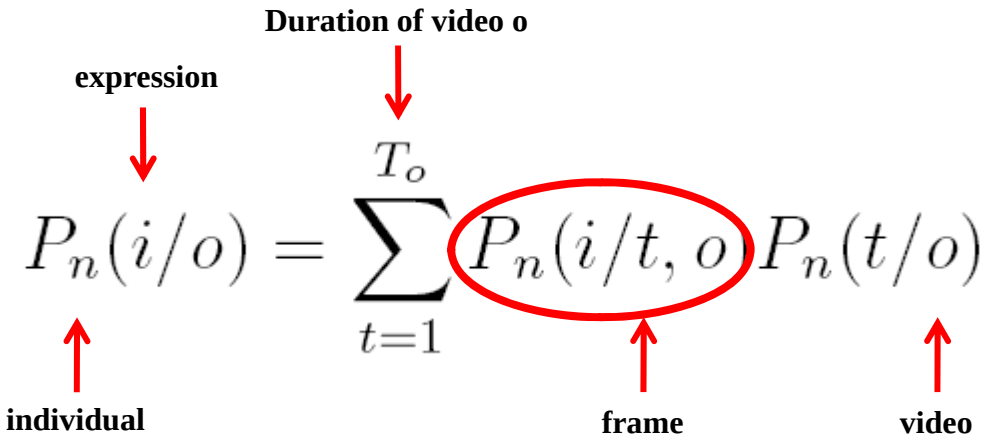
- Capture influence of each frame on final expression perception
- Choice set: frames (depend on the video)
- Attributes: dynamic features, such as facial features derivatives

Model: framework



- $P_n(i/t,o)$: Instantaneous expression perception sub-model.
- $P_n(t/o)$: Video frames weighing sub-model.
- $P_n(t/o)$: Model.

Model: expression perception sub-model



- Choice set: 9 expressions (Happiness, Surprise, Fear, Disgust, Sadness, Anger, Neutral, Other, Don't know)
- Logit model
- Memory effect : captured in expression utilities

Model: expression perception sub-model

- Memory effect : captured in expression utilities

Directly:

$$V_n(i|t,o) = Vgeneric_n(i|t,o) + a_i Vgeneric_n(i|t-1,o)$$

Estimated parameter

Convex combination:

$$V_n(i|t,o) = (1-a_i)Vgeneric_n(i|t,o) + a_i Vgeneric_n(i|t-1,o)$$

$a_i \in [0,1]$

$$V_n(i|t,o) = Vgeneric_n(i|t,o) + a_i(Vgeneric_n(i|t-1,o) - Vgeneric_n(i|t,o))$$

a_i unbounded

What is the best?

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model: expression perception sub-model

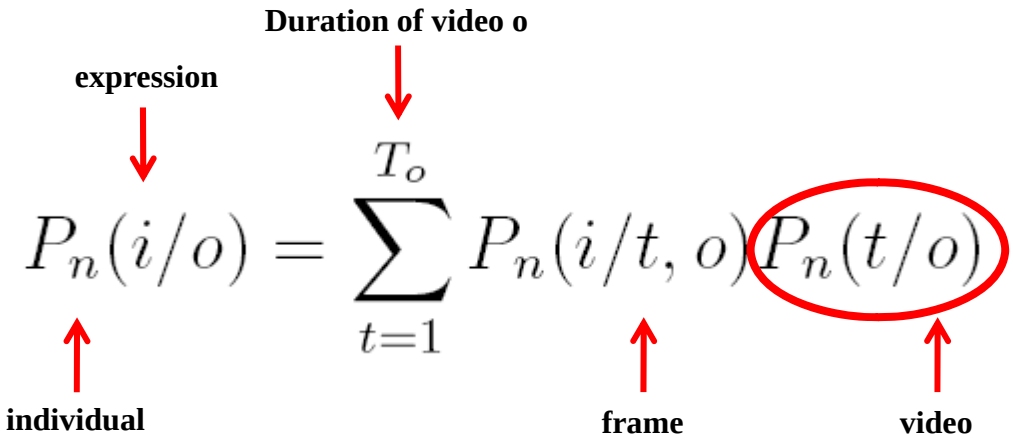
Specification of $Vgeneric_n(i|t,o)$

- Inspiration:

Sorci, M., Antonini, G., Cruz, J., Robin, T., Bierlaire, M. and Thiran, J.-P. (n.d.). Modelling human perception of static facial expressions, *Image and Vision Computing* . Accepted for publication.

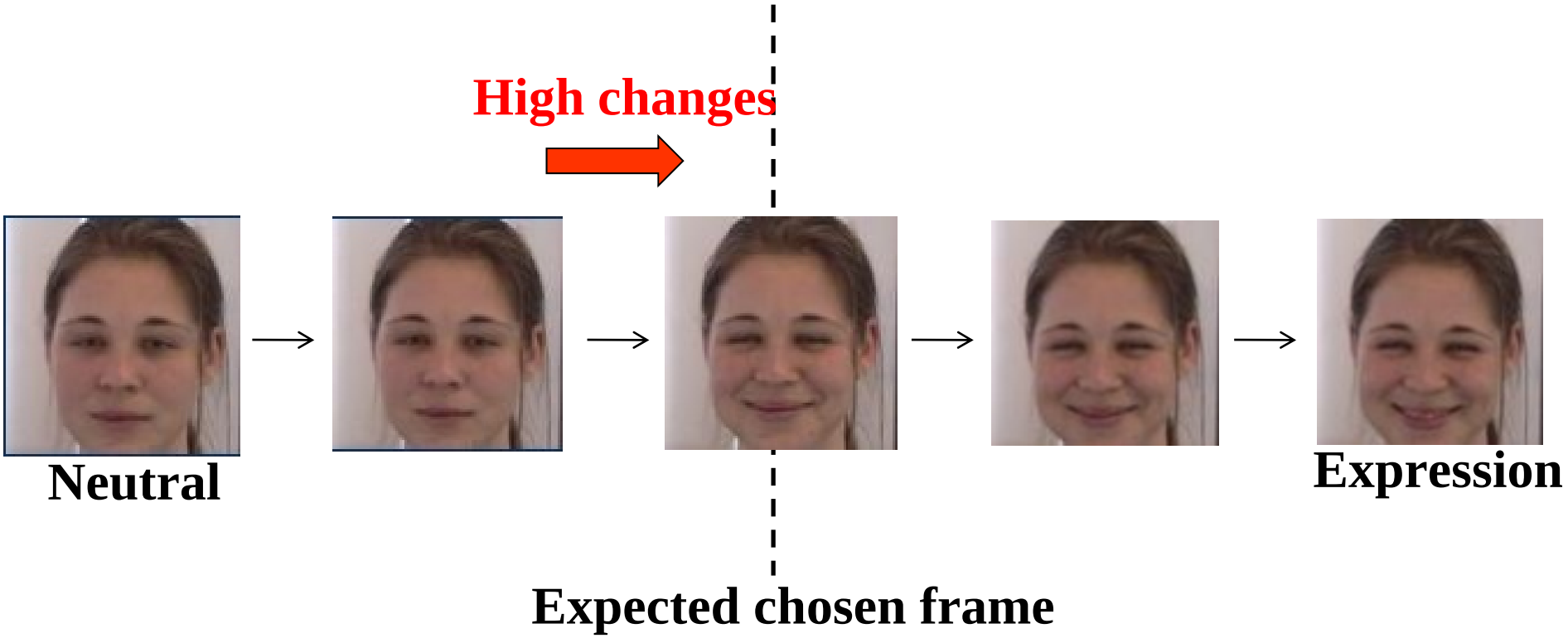
- Linear in parameters and attributes specification:
  - Alternative Specific Constants (**ASC**)
  - Measures corresponding to **AU** (FACS)
  - **No** Expression Descriptive Units (EDU)
  - Elements of **C vectors** (outputs of AAM)

Model: frames weighting sub-model



- Choice set: Frames of the videos (it depends on *o*)
- Logit model
- Utility specification: Derivatives of facial features

Model: frames weighting sub-model



What characterize the frame choice?

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model: expression perception sub-model

- Specification of $V_n(t|o)$:

  - Linear in attributes and parameters (for simplicity)

  - Generic parameters

  - Frame Attributes:

    1) Facial features of frame *t*
    2) Difference between Facial features of frames *t* and *t-1*
    3) Difference between Facial features of frames *t* and *0*

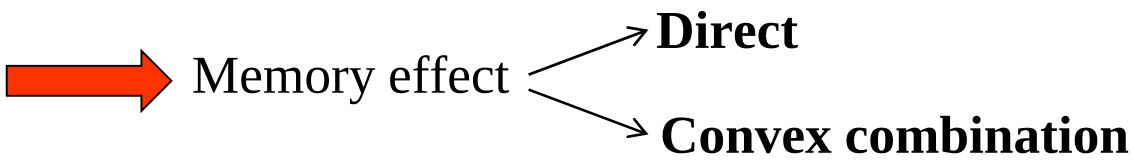
Hard to specify

Model 1: specification

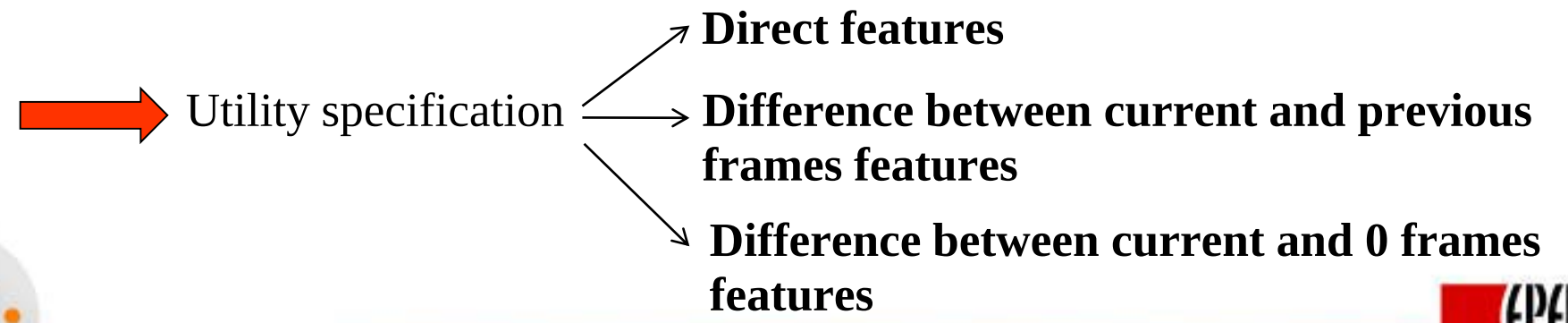
- **Aim**: find the best model in terms of **fit**, **interpretability** and **prediction**

- What can we play with?

  - Sub-model of expression perception

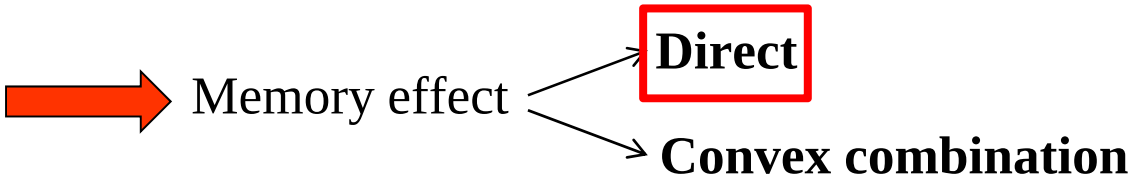


  - Sub-model of frame influence

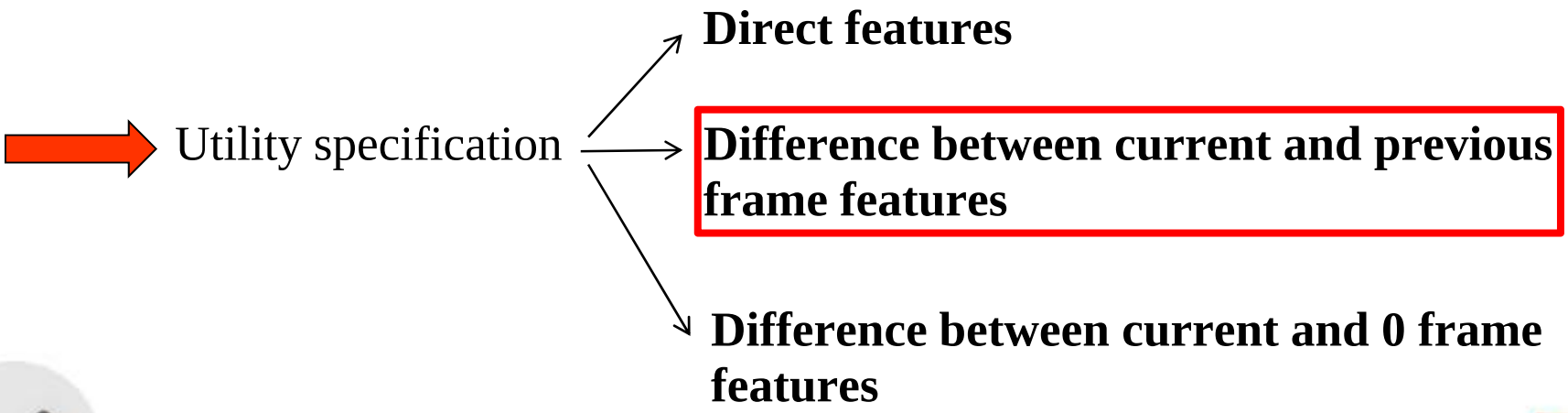


Model 1: specification

- Sub-model of expression perception



- Sub-model of frame influence



Model 1: estimation

- Likelihood maximization
- Estimation program based on the BIOGEME software
- General model fit:

Nb of observations:	418
Nb of parameters:	44
Null log-likelihood:	−918.43
Cste log-likelihood:	−826.31
Final log-likelihood:	−565.71
$\bar{\rho}^2$:	0.336

Model 1: estimation

Sub-model of expression perception

	Id	Parameter name	Value	Std-error	t-test 0
ASCs	1	*ASC_A*	−5.19	4.60	−1.13
	2	*ASC_D*	7.73	3.59	2.15
	3	*ASC_DK*	−0.94	0.33	−2.81
	4	*ASC_F*	−13.51	5.30	−2.55
	5	*ASC_H*	17.45	3.83	4.56
	6	*ASC_O*	6.40	2.61	2.45
	7	*ASC_SA*	7.94	5.12	1.55
	8	*ASC_SU*	−13.67	2.35	−5.82
AUs	9	*b_broweye_l2_SA*	422.93	90.55	4.67
	10	*b_broweye_l3_SU*	95.63	17.77	5.38
	11	*b_broweye_r2_A_D_F_SA_SU*	−109.60	22.01	−4.98
	12	*b_eye_angle_below_l_F*	1.95	1.41	1.38
	13	*b_eye_angle_l_F_SA*	20.79	4.34	4.79
	14	*b_eye_angle_r_F_SA*	−7.39	2.39	−3.10
	15	*b_eye_brow_angle_l_SA*	−13.12	3.19	−4.11
	16	*b_eye_mouth_dist_l2_D*	−45.37	21.68	−2.09
	17	*b_eye_mouth_dist_l_H_O_SA*	−47.20	27.93	−1.69
	18	*b_eye_nose_dist_l_A*	215.62	29.71	7.26
	19	*b_eye_nose_dist_l_D_F_O_SA*	85.47	17.46	4.89
	20	*b_eye_nose_dist_r_D_F_O_SA_A*	−122.63	15.29	−8.02
	21	*b_leye_h_F*	270.06	87.16	3.10
	22	*b_leye_h_SU*	244.97	42.60	5.75
	23	*b_mouth_h_A_D_H_SA_F_SU*	95.50	20.91	4.57
	24	*b_mouth_nose_dist2_A_SA*	−253.08	47.10	−5.37
	25	*b_mouth_nose_dist_H*	−267.43	37.67	−7.10
	26	*b_mouth_w_A_D_F_H_O*	28.29	12.26	2.31
C vector	27	*b_XC_1_SU*	65.54	14.97	4.38
	28	*b_XC_1_F*	40.46	27.28	1.48
	29	*b_XC_1_D*	104.21	14.75	7.07
	30	*b_XC_1_A*	137.69	21.17	6.50
	31	*b_XC_2_H*	16.06	8.47	1.90
	32	*b_XC_2_SU*	31.30	10.49	2.98
Memory effect	33	*A_H*	−0.71	0.14	−5.16
	34	*A_D*	−0.24	0.08	−2.91
	35	*A_SA*	−0.42	0.11	−3.86
	36	*A_A*	−0.21	0.09	−2.48

Sub-model of frame influence

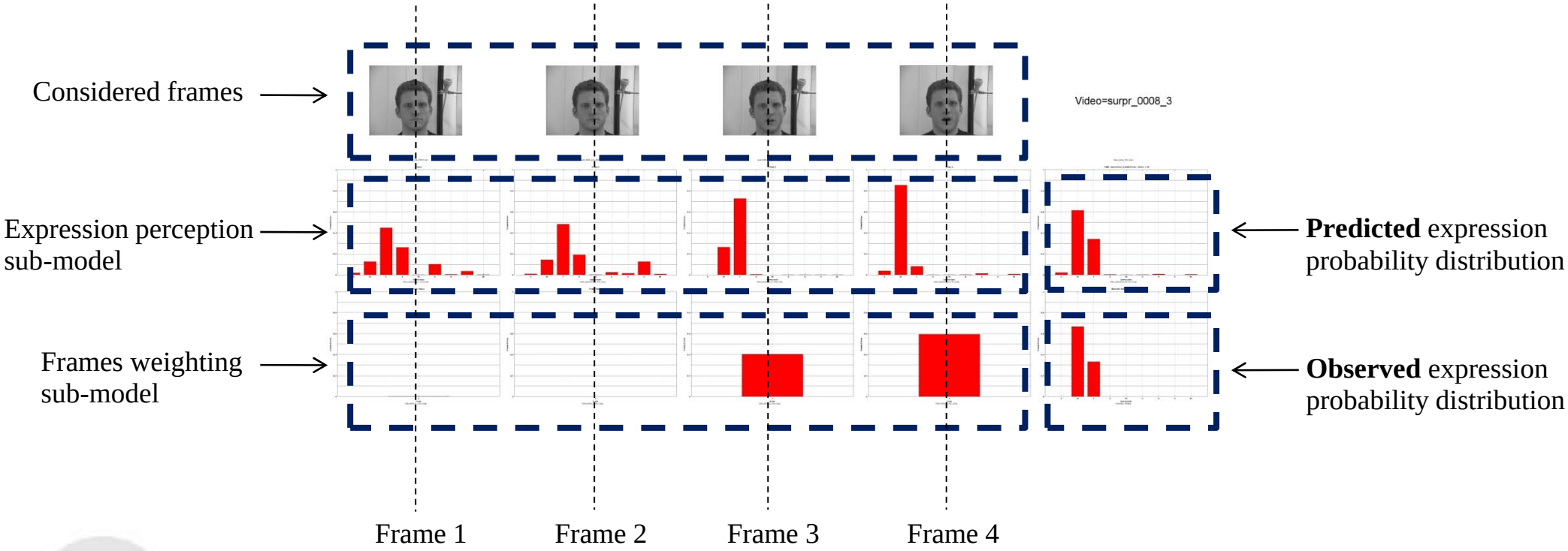
	Id	Parameter name	Value	Std-error	t-test 0
C vector deriv	1	*b_XC_1_deriv_FRAME*	−5.67	1.96	−2.90
	2	*b_FRAME_eye_h_deriv*	−59.69	14.75	−4.05
	3	*b_FRAME_eye_brow_angle_deriv*	3.12	0.80	3.88
	4	*b_FRAME_Xmouth_h_deriv*	112.20	30.50	3.68
	5	*b_FRAME_Xmouth_w_deriv*	−15.13	3.91	−3.87
AUs deriv	6	*b_XC_2_deriv_FRAME*	−17.61	4.28	−4.12
	7	*b_XC_3_deriv_FRAME*	16.19	3.75	4.31
	8	*b_XC_5_deriv_FRAME*	−7.04	2.01	−3.50

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model 1: prediction

- Check model validity
- Prediction display example:

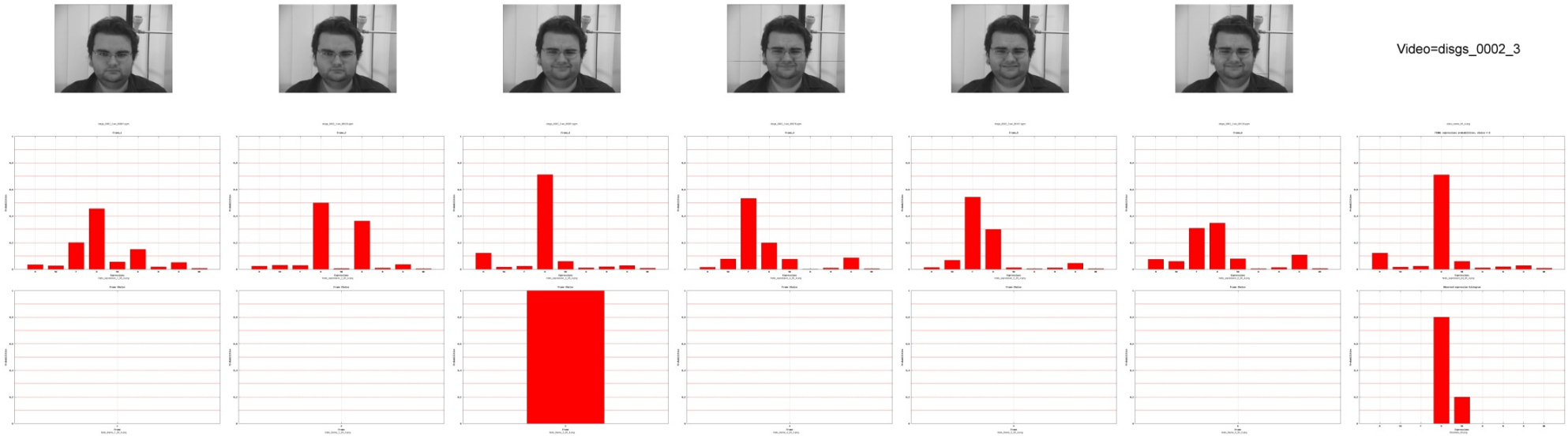


TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model 1: prediction 1

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



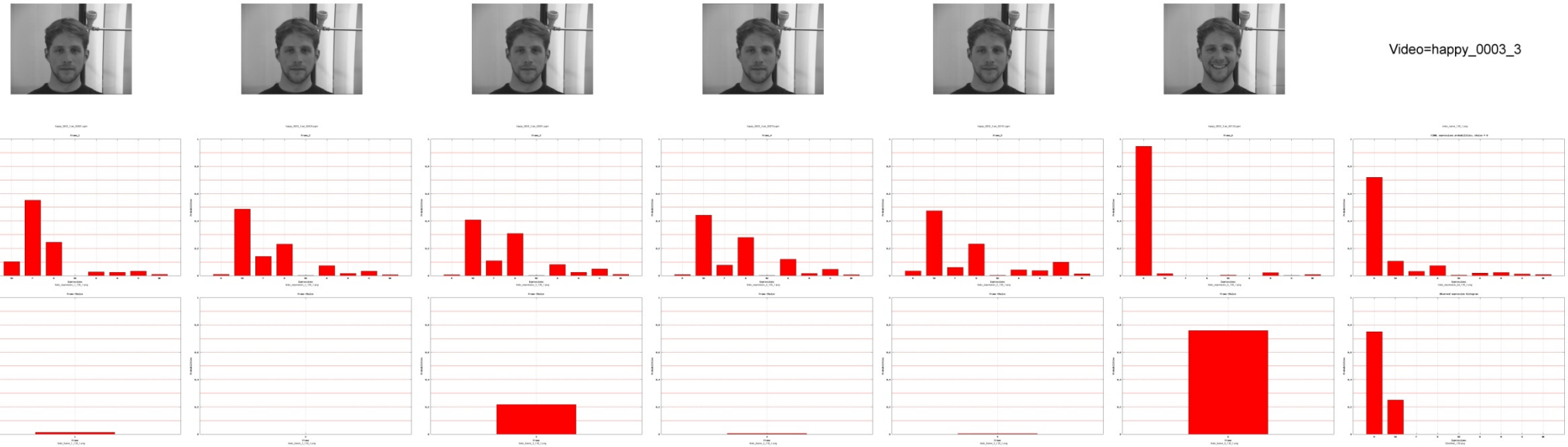
GOOD

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model 1: prediction 2

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



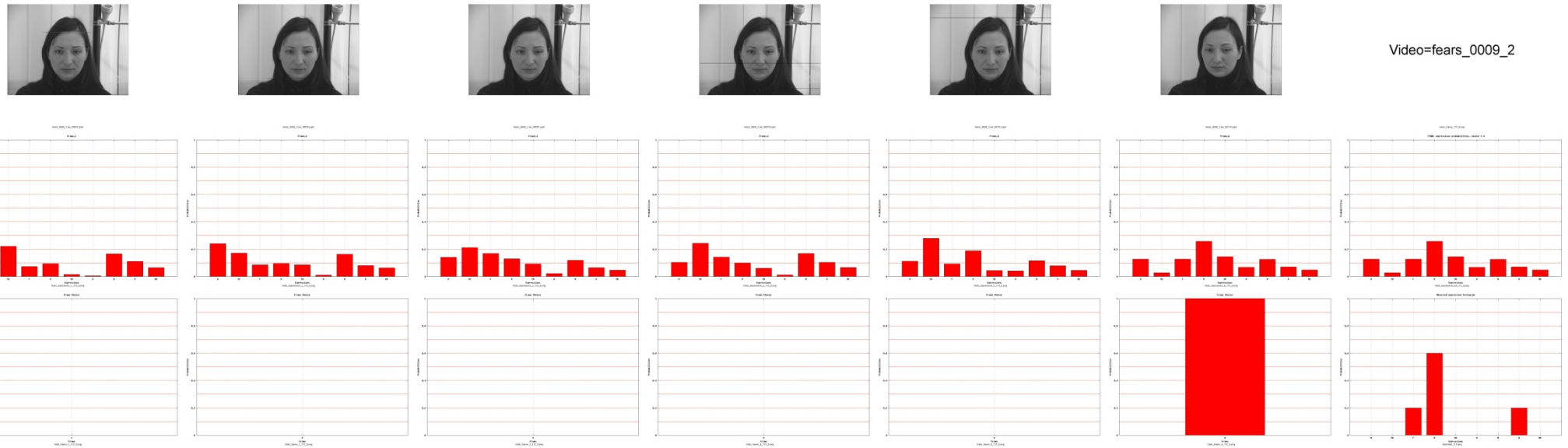
GOOD

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model 1: prediction 3

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



NOT GOOD

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Validation: introduction

- Prove model applicability
- Inspired from:

Robin, T., Antonini, G., Bierlaire, M. and Cruz, J. (2009). Specification, estimation and validation of a pedestrian walking behavior model, *Transportation Research Part B: Methodological* **43**(1): 36–56.

- Procedure:

  - Prediction on estimation data
  - Cross-validation on estimation data
  - Prediction on a new video database (not used for model estimation)

Validation: prediction on estimation data

- Outlier: observation with predicted choice probability less than 1/9

→ Outlier percentage

ASC model	model 1
35.89%	18.90%

Validation: prediction on estimation data

- Distance between observed and predicted expression distribution

→ Sum over all videos $$T = \sum_{o} (\sum_{i=1}^{9} \frac{(N_{io} - n_o p_{io})^2}{n_o p_{io}})$$

N_{io} : number of labels for expression *i* and video *o*

n_o : number of labels for video *o*

p_{io} : predicted probability of expression *i* for video *o*

Null model	ASC model	model 1
1831.65	1472.38	1512.20

Validation: cross-validation on estimation data

Save data for validation

- Data are divided **randomly** in 5 subsets, each one containing 20% of the facial videos

- Model **estimation** on 4 subsets, **simulation** on the remaining one

- 5 experiments

- Comparison between observations and predictions, with the outlier percentage

Validation: cross-validation on estimation data

- Cross-validation:

  - Data statistics:

	data 1	data 2	data 3	data 4	data 5
Number of videos	14	14	14	14	17
Number of observations	65	102	74	79	98

  - Outlier percentage :

	simulation 1	simulation 2	simulation 3	simulation 4	simulation 5
ASC model	23.08%	50.98%	35.13%	27.85%	35.71%
model 1	23.07%	28.43%	25.67%	31.64%	Not converged

Validation: new video database

- Apply the model on a new video database not involved in the estimation process

→ Database collected during summer 2009 by Clément Denis

- Natural expressions (65)

But, poor resolution and lack of luminosity

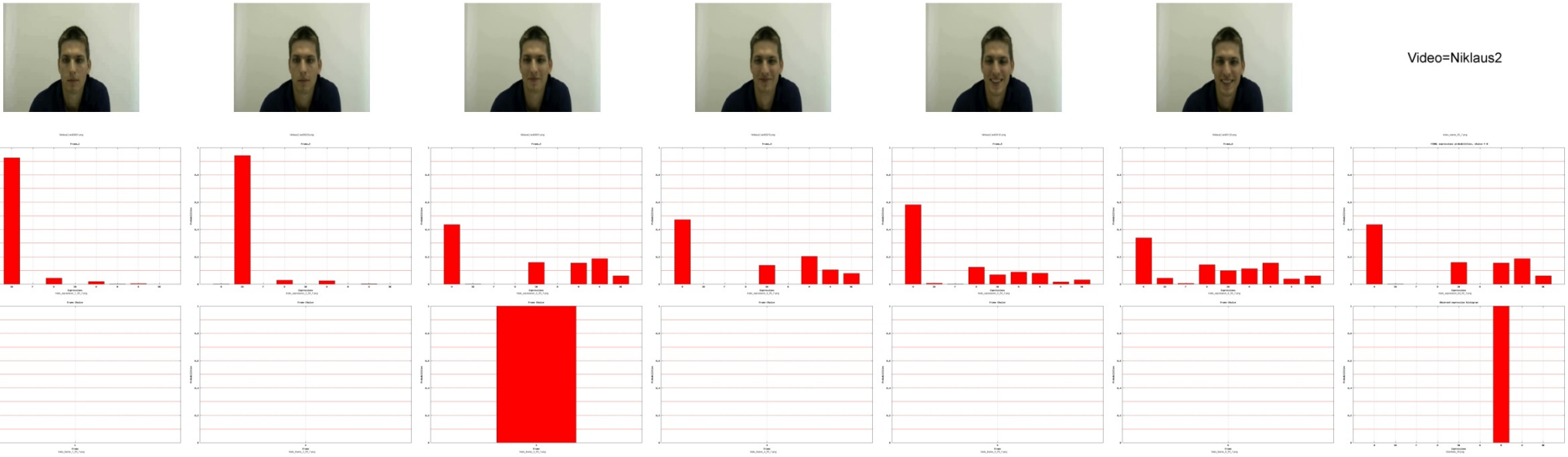
→ Man Machine Interface (MMI) by Maya Pantik

- Played expressions using FACS (1395)
- Natural expressions (167)
- Benchmark for computer vision algorithm

But account not yet activated

Validation: new video database

- Model 1
- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



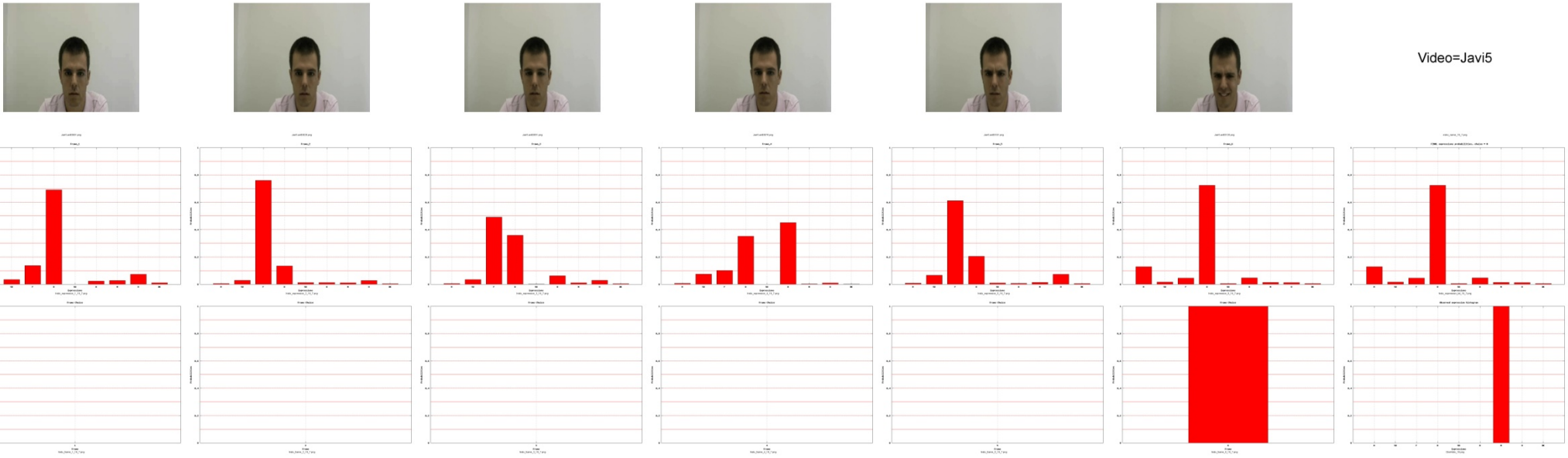
GOOD

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Validation: new video database

- Model 1
- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



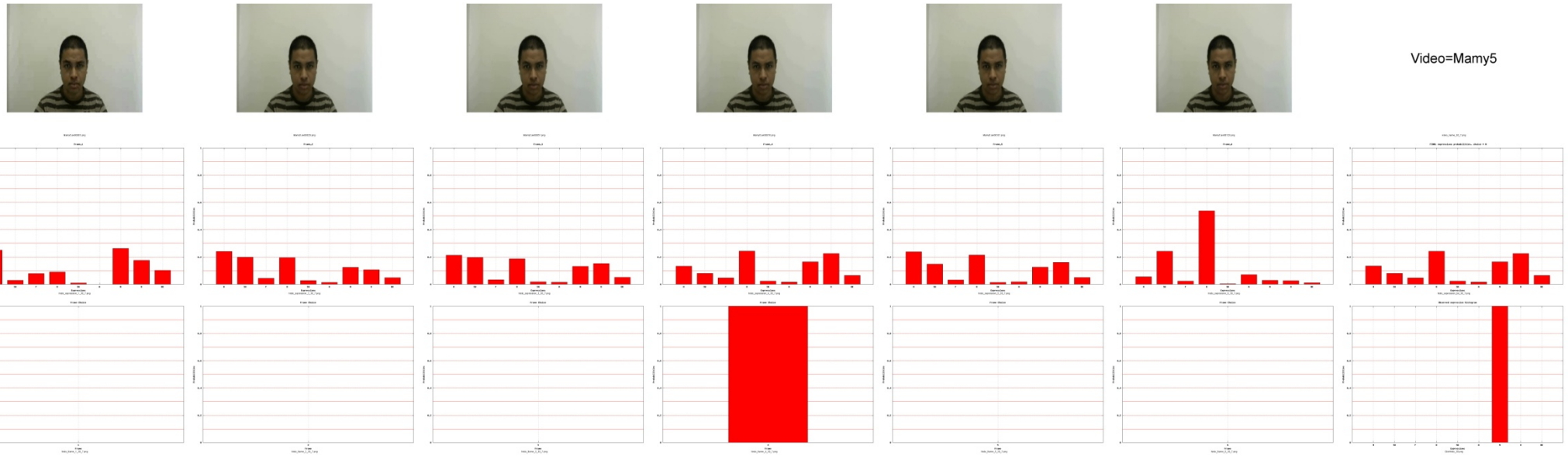
GOOD

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Validation: new video database

- Model 1
- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



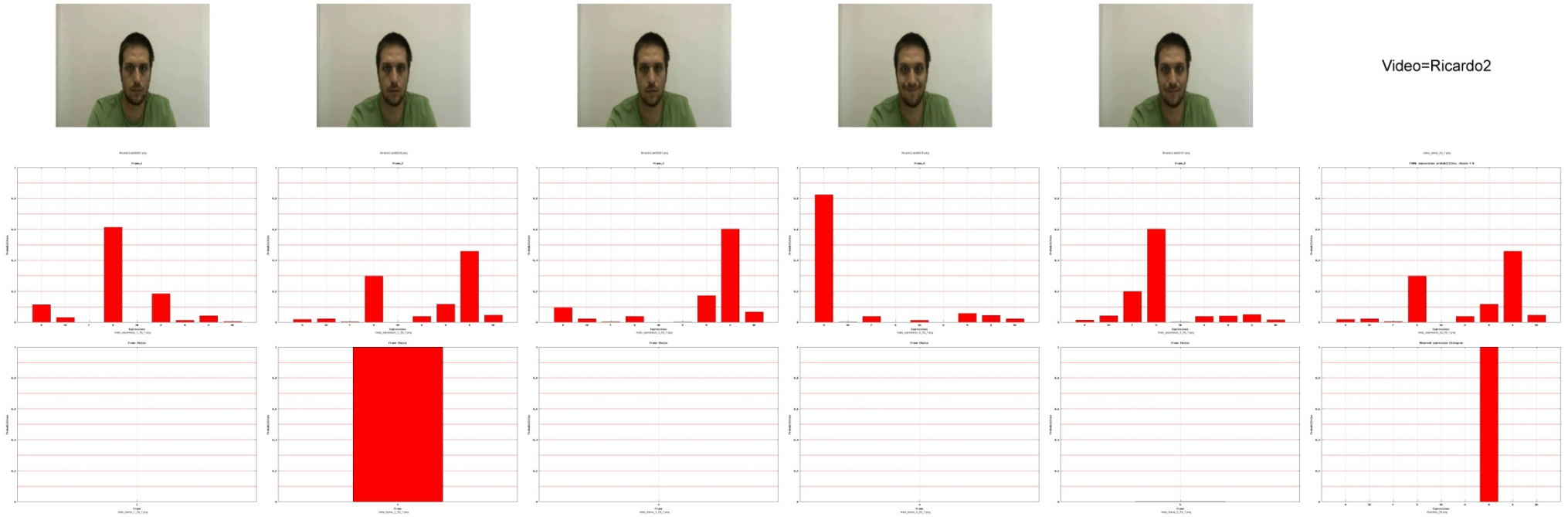
NOT GOOD

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Validation: new video database

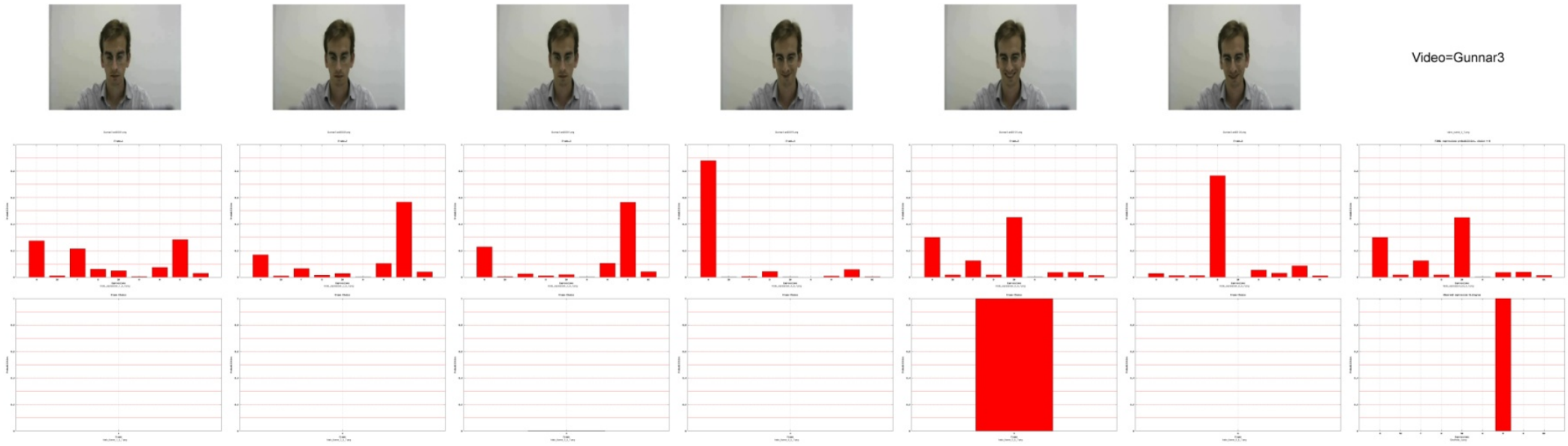
- Model 1
- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



BAD

Validation: new video database

- Model 1
- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



BAD

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

New models: introduction

- What are the problems of model 1?

  - Model of frame influence too sensitive

  - Validation results not so good

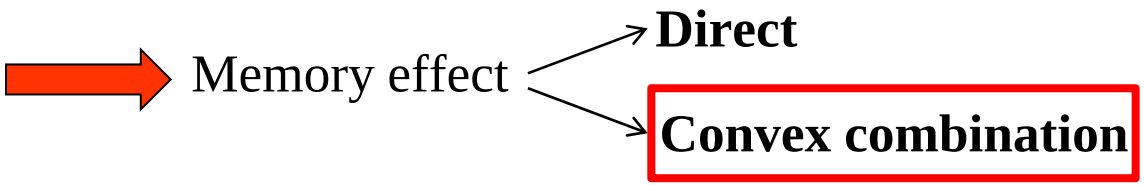
- How results can be improved?

  - Specification

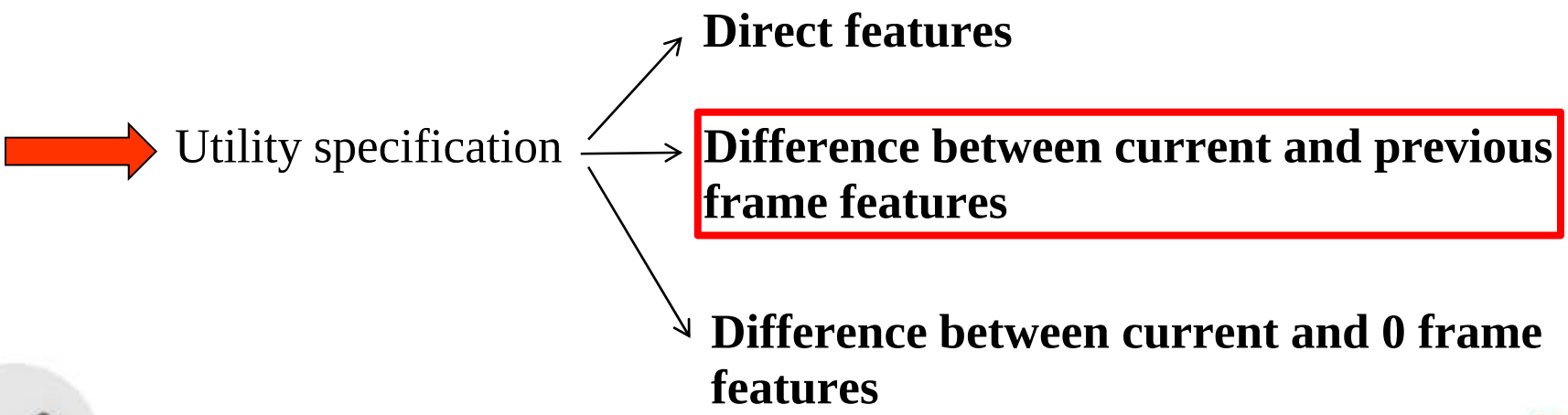
  - Better adaptation of the validation procedure

Model 2: specification

- Sub-model of expression perception



- Sub-model of frame influence



Model 2: estimation

- General model fit:

Model 1:

Nb of observations:	418
Nb of parameters:	44
Null log-likelihood:	−918.43
Cste log-likelihood:	−826.31
Final log-likelihood:	−565.71
$\bar{\rho}^2$:	0.336

Model 2:

Nb of observations:	418
Nb of parameters:	44
Null log-likelihood:	−918.43
Cste log-likelihood:	−826.31
Final log-likelihood:	−561.72
$\bar{\rho}^2$:	0.340

Model 2: estimation

Sub-model of expression perception

	Id	Parameter name	Value	Std-error	t-test 0
ASCs	1	*ASC_A*	−0.91	3.47	−0.26
	2	*ASC_D*	8.71	2.45	3.56
	3	*ASC_DK*	−0.93	0.34	−2.74
	4	*ASC_F*	−9.44	5.25	−1.80
	5	*ASC_H*	8.16	2.31	3.54
	6	*ASC_O*	7.73	1.98	3.90
	7	*ASC_SA*	6.37	3.53	1.81
	8	*ASC_SU*	−10.50	2.01	−5.23
AUs	9	*b_broweye_l2_SA*	296.57	80.98	3.66
	10	*b_broweye_l3_SU*	81.81	16.07	5.09
	11	*b_broweye_r2_A_D_F_SA_SU*	−93.22	16.70	−5.58
	12	*b_eye_angle_below_l_F*	2.00	1.53	1.31
	13	*b_eye_angle_l_F_SA*	15.43	3.46	4.46
	14	*b_eye_angle_r_F_SA*	−5.61	1.66	−3.39
	15	*b_eye_brow_angle_l_SA*	−8.35	4.12	−2.03
	16	*b_eye_mouth_dist_l2_D*	−27.08	15.01	−1.80
	17	*b_eye_mouth_dist_l_H_O_SA*	−29.26	16.50	−1.77
	18	*b_eye_nose_dist_l_A*	141.24	23.74	5.95
	19	*b_eye_nose_dist_l_D_F_O_SA*	52.04	13.65	3.81
	20	*b_eye_nose_dist_r_D_F_O_SA_A*	−92.32	12.03	−7.67
	21	*b_leye_h_F*	285.91	80.42	3.56
	22	*b_leye_h_SU*	208.40	42.05	4.96
	23	*b_mouth_h_A_D_H_SA_F_SU*	69.76	17.46	4.00
	24	*b_mouth_nose_dist2_A_SA*	−156.57	33.99	−4.61
	25	*b_mouth_nose_dist_H*	−146.64	29.14	−5.03
	26	*b_mouth_w_A_D_F_H_O*	16.23	8.98	1.81
C vector	27	*b_XC_1_SU*	66.36	15.04	4.41
	28	*b_XC_1_F*	33.83	24.02	1.41
	29	*b_XC_1_D*	71.89	11.19	6.42
	30	*b_XC_1_A*	96.12	19.43	4.95
	31	*b_XC_2_H*	8.66	4.94	1.75
	32	*b_XC_2_SU*	31.38	10.36	3.03
Memory effect	33	*A_H*	−1.01	0.34	−2.99
	34	*A_D*	−0.69	0.28	−2.51
	35	*A_SA*	−0.96	0.30	−3.19
	36	*A_A*	−0.50	0.22	−2.24

Sub-model of frame influence

	Id	Parameter name	Value	Std-error	t-test 0
C vector deriv	37	*b_XC_1_deriv_FRAME*	−5.21	1.81	−2.88
	38	*b_FRAME_eye_h_deriv*	−63.60	14.65	−4.34
	39	*b_FRAME_eye_brow_angle_deriv*	3.14	0.73	4.28
	40	*b_FRAME_Xmouth_h_deriv*	93.42	20.42	4.57
	41	*b_FRAME_Xmouth_w_deriv*	−17.72	4.12	−4.31
AUs deriv	42	*b_XC_2_deriv_FRAME*	−18.80	4.36	−4.32
	43	*b_XC_3_deriv_FRAME*	16.74	3.59	4.66
	44	*b_XC_5_deriv_FRAME*	−6.71	1.81	−3.70

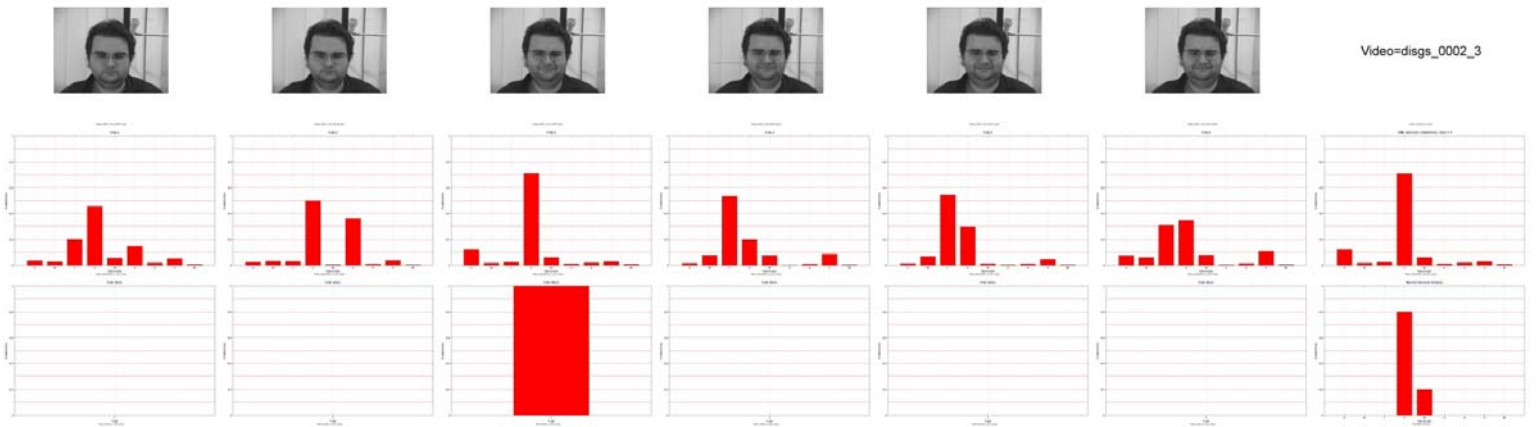
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

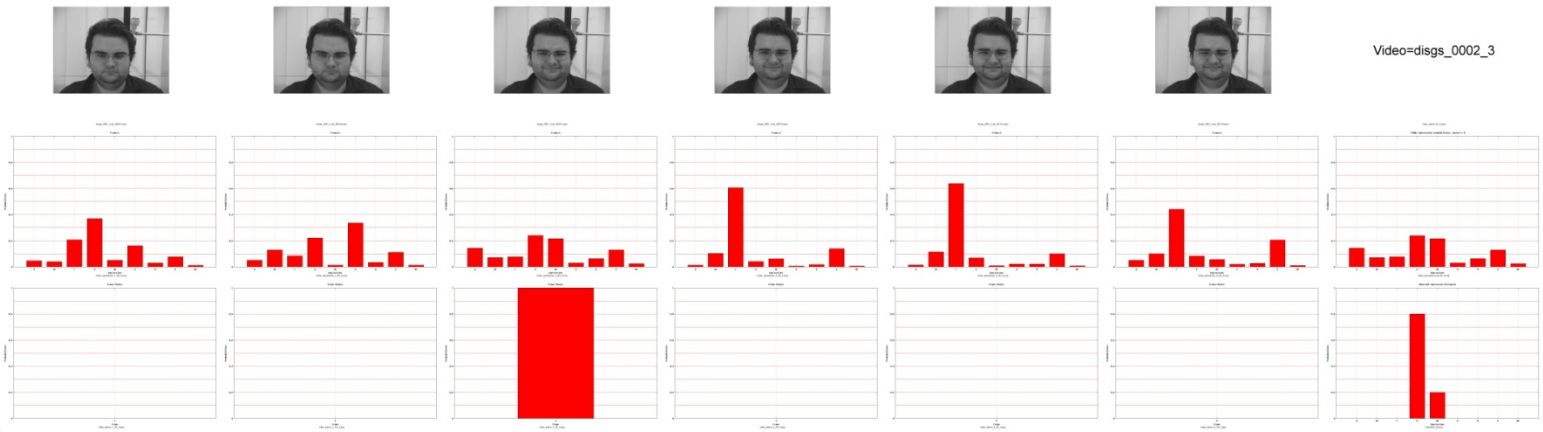
Model 2: prediction 1

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

Model 1:



Model 2:

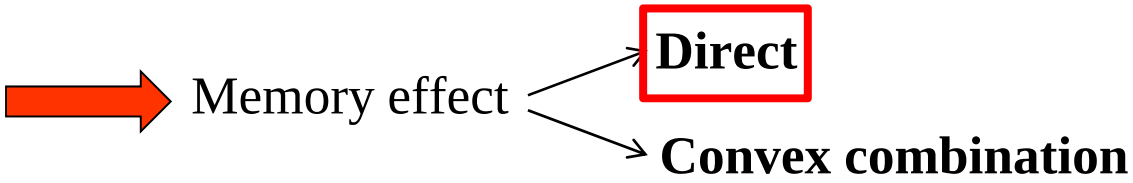


TRANSP-OR

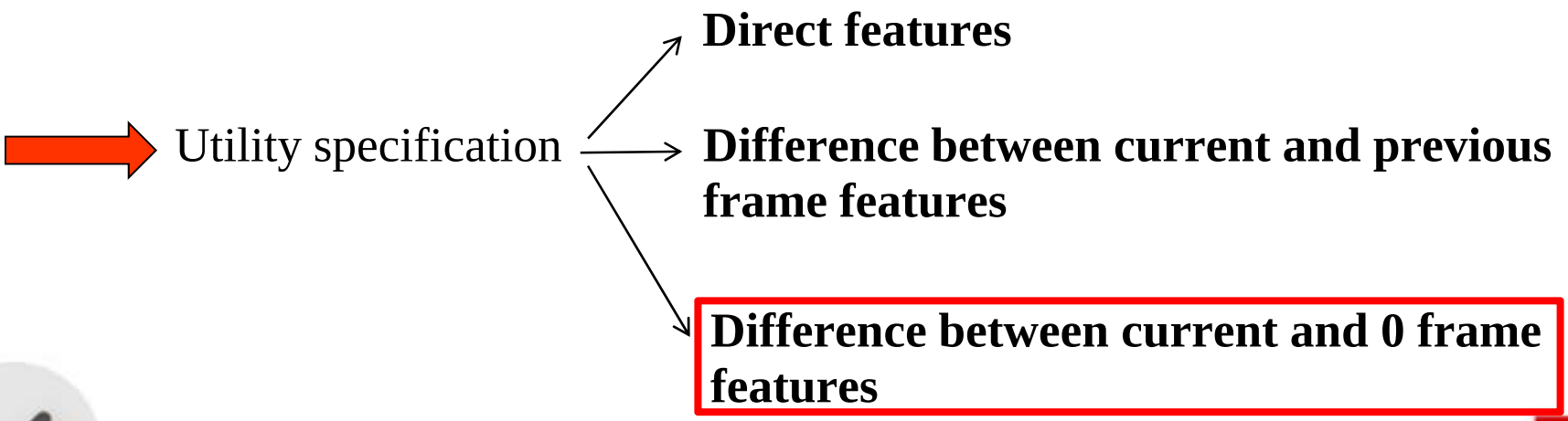
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model 3: specification

- Sub-model of expression perception



- Sub-model of frame influence



Model 3: estimation

- General model fit:

Model 1:

Nb of observations:	418
Nb of parameters:	44
Null log-likelihood:	−918.43
Cste log-likelihood:	−826.31
Final log-likelihood:	−565.71
$\bar{\rho}^2$:	0.336

Model 3:

Nb of observations:	418
Nb of parameters:	40
Null log-likelihood:	−918.43
Cste log-likelihood:	−826.31
Final log-likelihood:	−628.54
$\bar{\rho}^2$:	0.272

Model 3: estimation

Sub-model of expression perception

	Id	Parameter name	Value	Std-error	t-test 0
ASCs	1	*ASC_A*	−17.15	6.20	−2.76
	2	*ASC_D*	0.05	3.17	0.01
	3	*ASC_DK*	−1.03	0.34	−3.04
	4	*ASC_F*	−21.86	6.74	−3.24
	5	*ASC_H*	10.62	5.87	1.81
	6	*ASC_O*	2.29	2.57	0.89
	7	*ASC_SA*	3.58	4.60	0.78
	8	*ASC_SU*	−16.26	2.90	−5.60
AUs	9	*b_broweye_l2_SA*	310.61	96.36	3.22
	10	*b_broweye_l3_SU*	86.50	18.61	4.65
	11	*b_broweye_r2_A_D_F_SA_SU*	−70.48	18.56	−3.80
	12	*b_eye_angle_below_l_F*	1.69	1.69	1.00
	13	*b_eye_angle_l_F_SA*	12.06	4.36	2.76
	14	*b_eye_angle_r_F_SA*	−4.89	2.73	−1.79
	15	*b_eye_brow_angle_l_SA*	−9.67	4.88	−1.98
	16	*b_eye_mouth_dist_l2_D*	−45.00	16.31	−2.76
	17	*b_eye_mouth_dist_l_H_O_SA*	−35.27	32.97	−1.07
	18	*b_eye_nose_dist_l_A*	192.37	39.23	4.90
	19	*b_eye_nose_dist_l_D_F_O_SA*	65.19	20.85	3.13
	20	*b_eye_nose_dist_r_D_F_O_SA_A*	−97.52	17.98	−5.42
	21	*b_leye_h_F*	435.32	99.89	4.36
	22	*b_leye_h_SU*	225.31	49.87	4.52
	23	*b_mouth_h_A_D_H_SA_F_SU*	121.56	26.54	4.58
	24	*b_mouth_nose_dist2_A_SA*	−181.29	44.03	−4.12
	25	*b_mouth_nose_dist_H*	−255.18	55.01	−4.64
	26	*b_mouth_w_A_D_F_H_O*	45.05	13.15	3.43
C vector	27	*b_XC_1_SU*	48.06	19.96	2.41
	28	*b_XC_1_F*	82.65	26.61	3.11
	29	*b_XC_1_D*	91.42	19.10	4.79
	30	*b_XC_1_A*	174.43	36.07	4.84
	31	*b_XC_2_H*	91.34	17.31	5.28
	32	*b_XC_2_SU*	51.60	14.46	3.57
Memory effect	33	*A_H*	−0.33	0.16	−2.11
	34	*A_D*	0.13	0.13	0.99
	35	*A_SA*	−0.22	0.22	−0.97
	36	*A_A*	−0.12	0.10	−1.19

Sub-model of frame influence

	Id	Parameter name	Value	Std-error	t-test 0
C vector deriv	37	*b_XC_0_deriv_1stFRAME*	−7.66	1.45	−5.28
	38	*b_XC_1_deriv_1stFRAME*	3.13	1.32	2.36
	39	*b_XC_2_deriv_1stFRAME*	2.53	0.71	3.55
	40	*b_XC_4_deriv_1stFRAME*	−1.95	0.87	−2.24

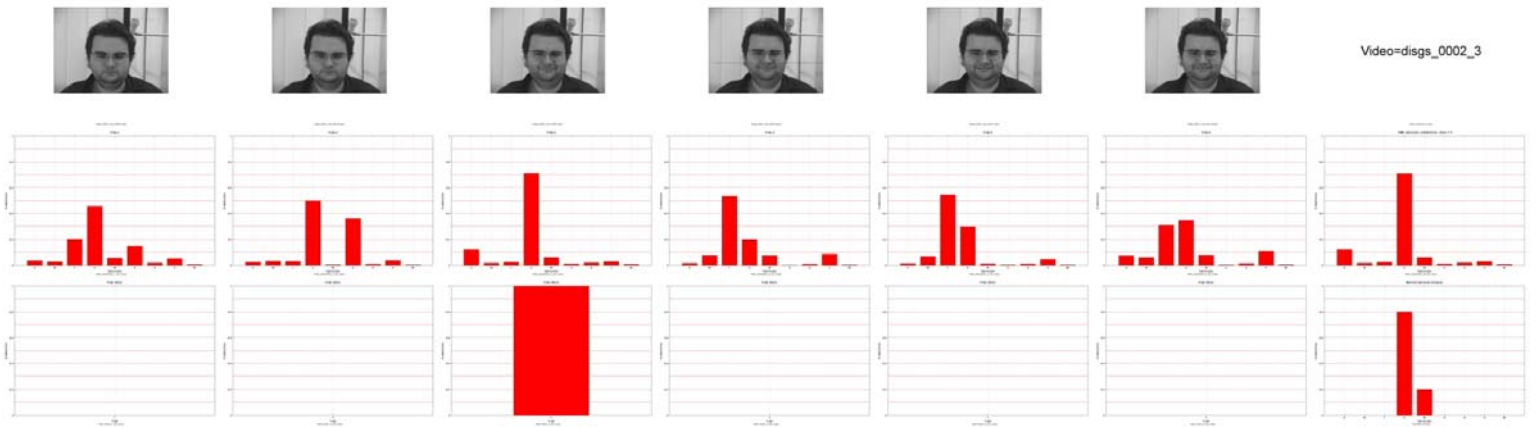
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

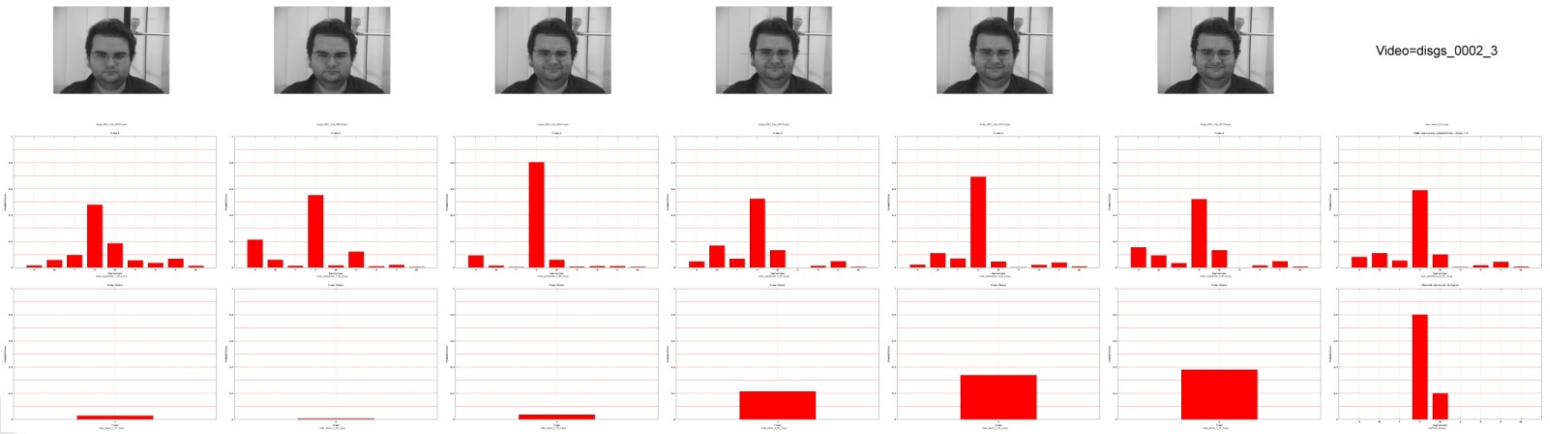
Model 3: prediction 1

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

Model 1:



Model 3:

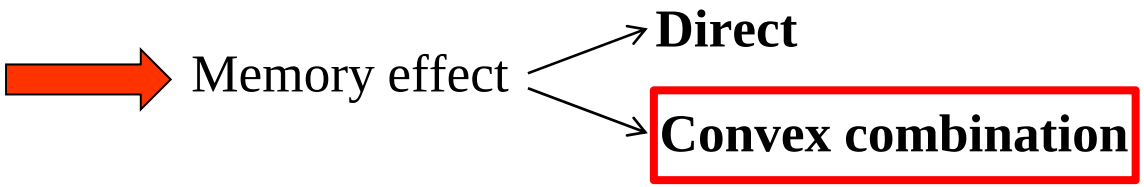


TRANSP-OR

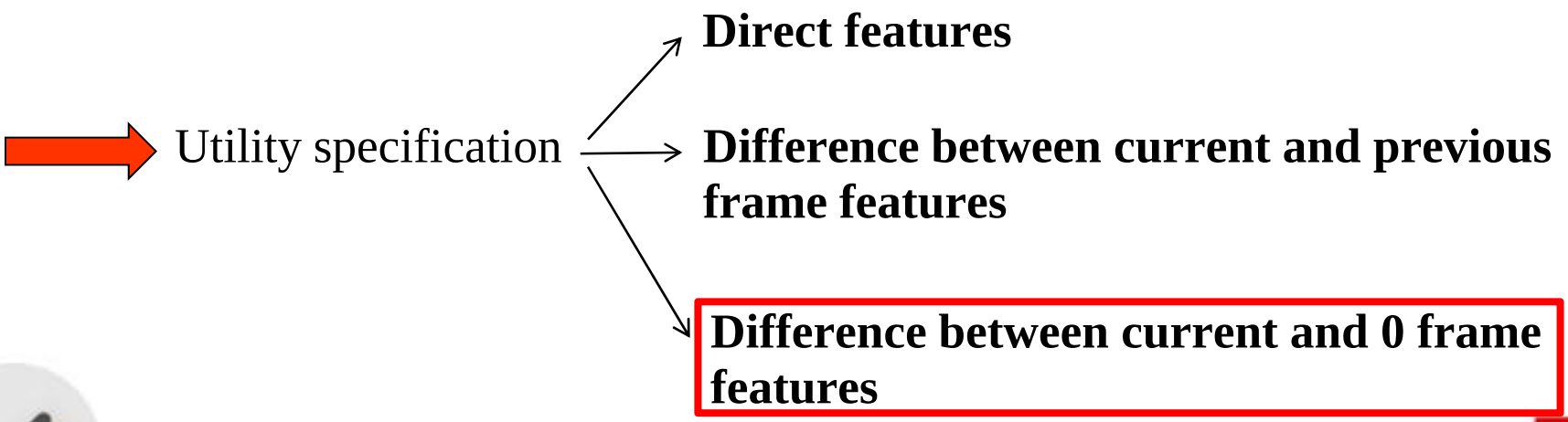
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model 4: specification

- Sub-model of expression perception



- Sub-model of frame influence



Model 4: estimation

- General model fit:

Model 1:

Nb of observations:	418
Nb of parameters:	44
Null log-likelihood:	−918.43
Cste log-likelihood:	−826.31
Final log-likelihood:	−565.71
$\bar{\rho}^2$:	0.336

Model 4:

Nb of observations:	418
Nb of parameters:	40
Null log-likelihood:	−918.43
Cste log-likelihood:	−826.31
Final log-likelihood:	−607.776
$\bar{\rho}^2$:	0.295

Model 4: estimation

Sub-model of expression perception

ASCs

Id	Parameter name	Value	Std-error	t-test 0
1	*ASC_A*	−10.41	4.49	−2.32
2	*ASC_D*	4.69	3.15	1.49
3	*ASC_DK*	−0.94	0.34	−2.76
4	*ASC_F*	−23.54	6.99	−3.37
5	*ASC_H*	9.95	3.77	2.64
6	*ASC_O*	6.93	2.36	2.94
7	*ASC_SA*	6.00	5.37	1.12
8	*ASC_SU*	−16.59	2.76	−6.00

AUs

Id	Parameter name	Value	Std-error	t-test 0
9	*b_broweye_l2_SA*	253.16	166.60	1.52
10	*b_broweye_l3_SU*	104.75	17.89	5.86
11	*b_broweye_r2_A_D_F_SA_SU*	−98.45	22.53	−4.37
12	*b_eye_angle_below_l_F*	2.97	1.88	1.58
13	*b_eye_angle_l_F_SA*	14.84	4.28	3.47
14	*b_eye_angle_r_F_SA*	−9.42	3.36	−2.80
15	*b_eye_brow_angle_l_SA*	−3.98	7.38	−0.54
16	*b_eye_mouth_dist_l2_D*	−47.69	18.85	−2.53
17	*b_eye_mouth_dist_l_H_O_SA*	−80.09	24.58	−3.26
18	*b_eye_nose_dist_l_A*	161.33	30.25	5.33
19	*b_eye_nose_dist_l_D_F_O_SA*	49.55	19.03	2.60
20	*b_eye_nose_dist_r_D_F_O_SA_A*	−87.41	15.60	−5.60
21	*b_leye_h_F*	584.81	125.88	4.65
22	*b_leye_h_SU*	225.82	51.64	4.37
23	*b_mouth_h_A_D_H_SA_F_SU*	126.76	23.34	5.43
24	*b_mouth_nose_dist2_A_SA*	−178.82	42.13	−4.24
25	*b_mouth_nose_dist_H*	−224.37	42.53	−5.28
26	*b_mouth_w_A_D_F_H_O*	25.68	9.27	2.77

C vector

Id	Parameter name	Value	Std-error	t-test 0
27	*b_XC_1_SU*	56.25	17.79	3.16
28	*b_XC_1_F*	71.84	26.79	2.68
29	*b_XC_1_D*	101.44	18.48	5.49
30	*b_XC_1_A*	125.85	26.40	4.77
31	*b_XC_2_H*	32.84	9.12	3.60
32	*b_XC_2_SU*	64.65	15.01	4.31

Memory effect

Id	Parameter name	Value	Std-error	t-test 0
33	*A_H*	−2.20	0.58	−3.82
34	*A_D*	0.14	0.13	1.12
35	*A_SA*	1.04	0.22	4.81
36	*A_A*	−0.40	0.20	−2.04

Sub-model of frame influence

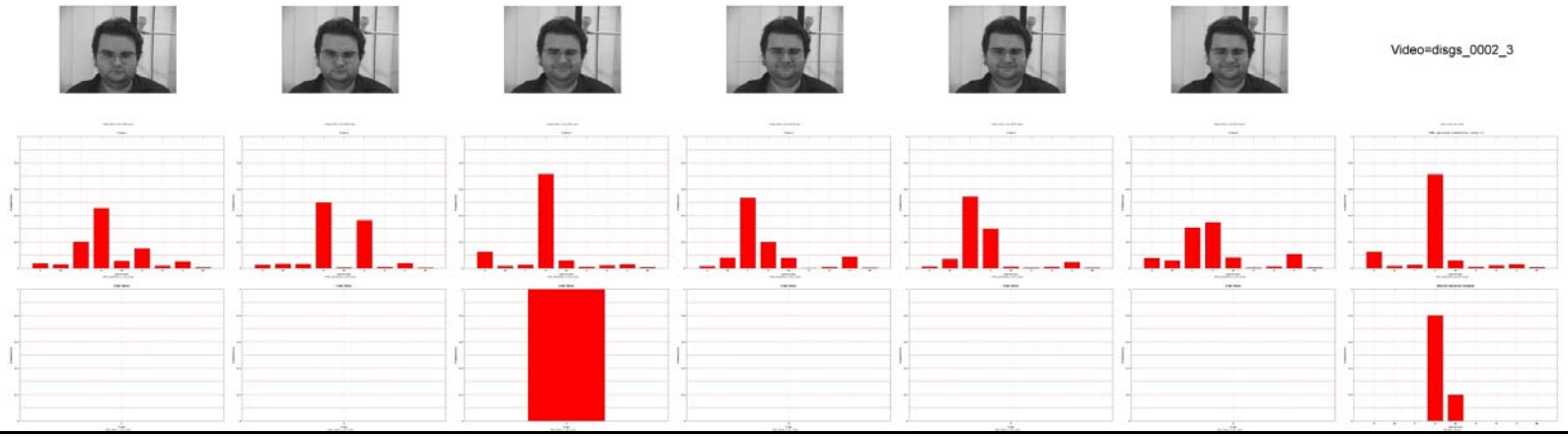
C vector deriv

Id	Parameter name	Value	Std-error	t-test 0
37	*b_XC_0_deriv_1stFRAME*	−5.95	1.02	−5.82
38	*b_XC_1_deriv_1stFRAME*	3.02	1.01	2.98
39	*b_XC_2_deriv_1stFRAME*	2.39	0.58	4.10
40	*b_XC_4_deriv_1stFRAME*	−0.96	0.70	−1.37

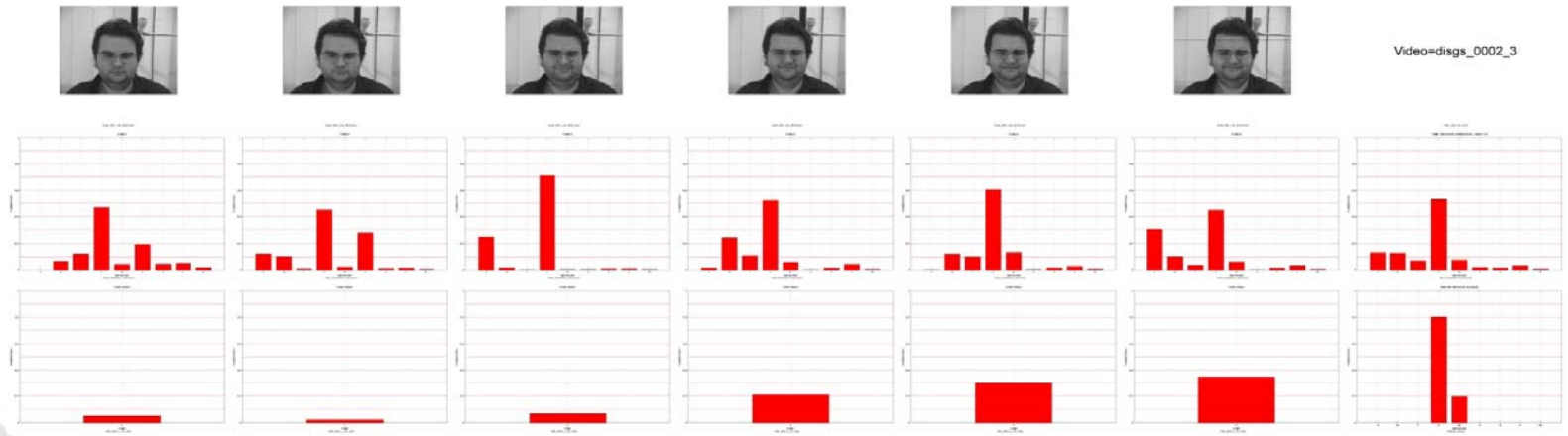
Model 4: prediction 1

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

Model 1:



Model 4:



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

New validation: prediction on estimation data

- Outlier: observation with predicted choice probability less than 1/9

→ Outlier percentage

ASC model	model 1	model 2	model 3	model 4
35.89%	18.90%	17.94%	17.93%	18.18%

New validation: prediction on estimation data

- Distance between observed and predicted expression distribution

→ Sum over all videos $$T = \sum_{o} \left(\sum_{i=1}^{9} \frac{(N_{io} - n_o p_{io})^2}{n_o p_{io}} \right)$$

N_{io} : number of labels for expression *i* and video *o*

n_o : number of labels for video *o*

p_{io} : predicted probability of expression *i* for video *o*

Null model	ASC model	model 1	model 2	model 3	model 4
1831.65	1472.38	1512.20	1091.00	2251.63	1342.90

New validation: cross-validation on estimation data

- Cross-validation:

  - Data statistics:

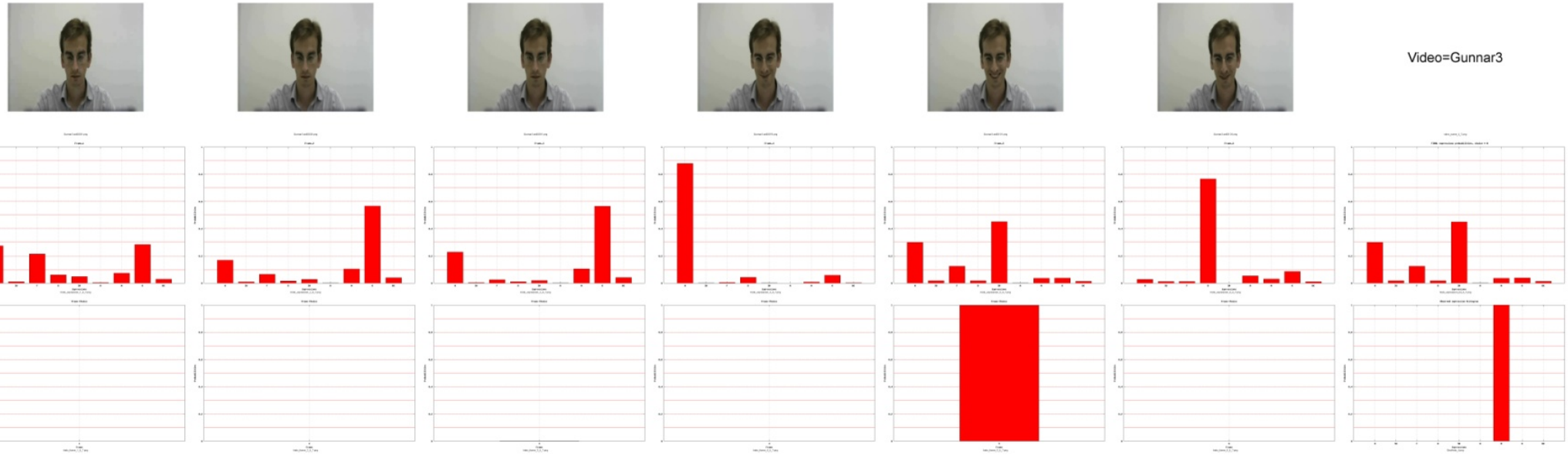
	data 1	data 2	data 3	data 4	data 5
Number of videos	14	14	14	14	17
Number of observations	65	102	74	79	98

  - Outlier percentage :

	simulation 1	simulation 2	simulation 3	simulation 4	simulation 5
ASC model	23.08%	50.98%	35.13%	27.85%	35.71%
model 1	23.07%	28.43%	25.67%	31.64%	Not converged
model 2	35.38%	35.29%	21.62%	27.84%	Not converged
model 3	36.92%	40.20%	22.97%	32.91%	50.00%
model 4	33.85%	34.31%	22.97%	50.63%	42.86%

New validation: new video database

- Model 1
- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

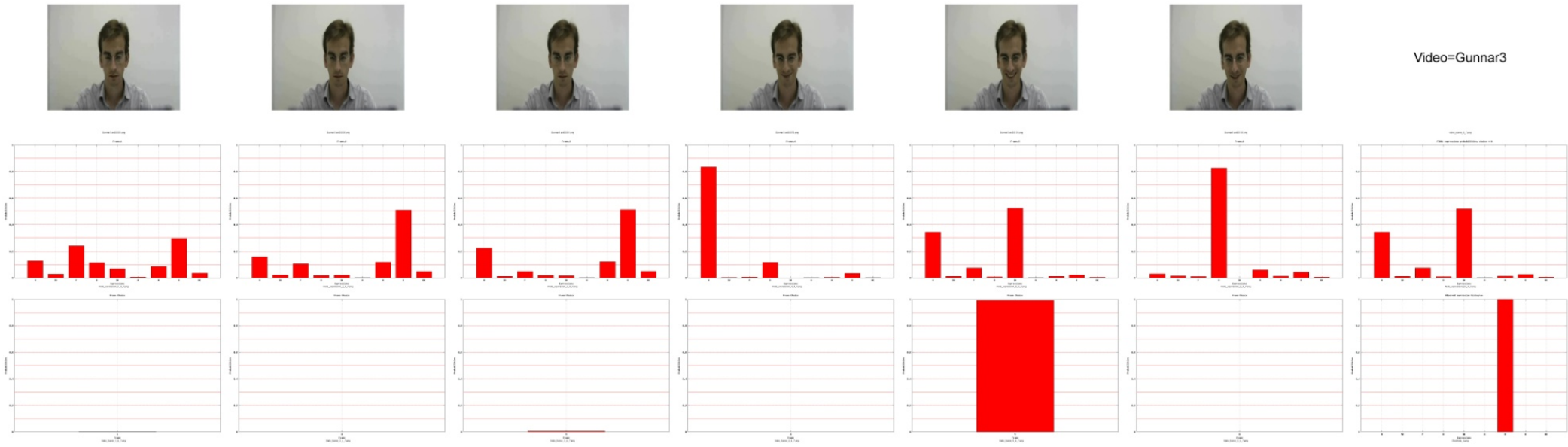


TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

New validation: new video database

- Model 2
- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

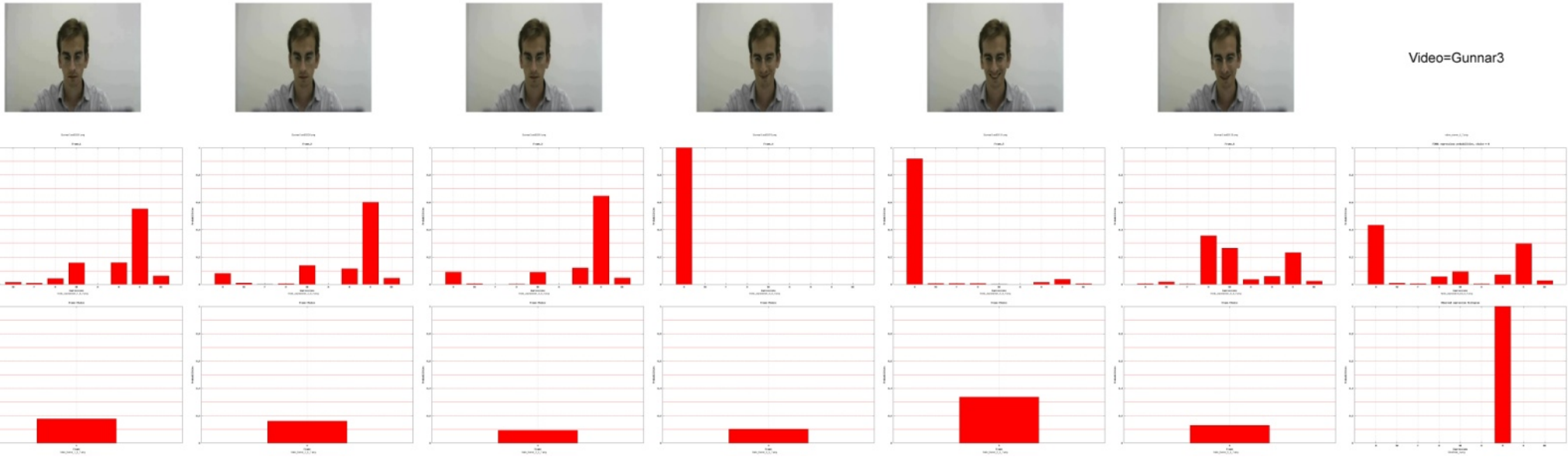


TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

New validation: new video database

- Model 3
- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

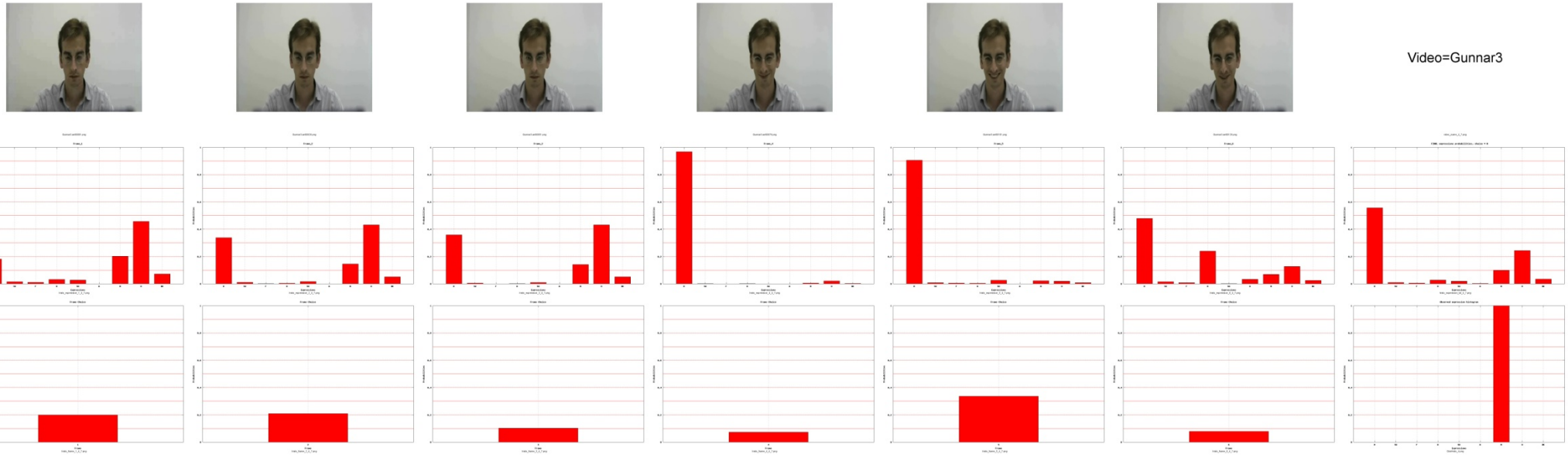


TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

New validation: new video database

- Model 4
- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



TRANSP-OR

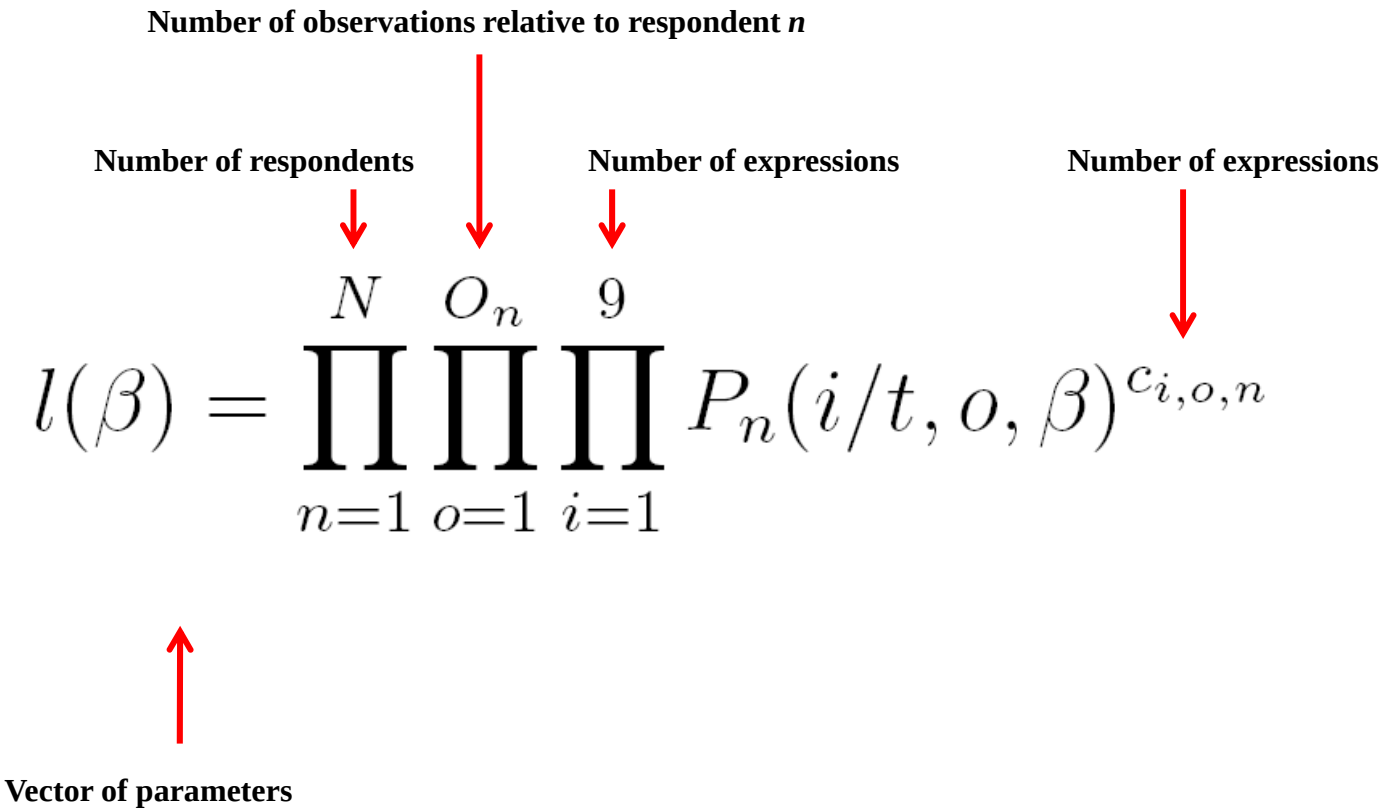
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Conclusions and Perspectives

- Conclusion:
  - Database of face video annotations
  - Model estimated using behavioral data
  - Validation procedure
- Perspectives:
  - Improve the specification
  - Validation
    - ➡ Use of the MMI database

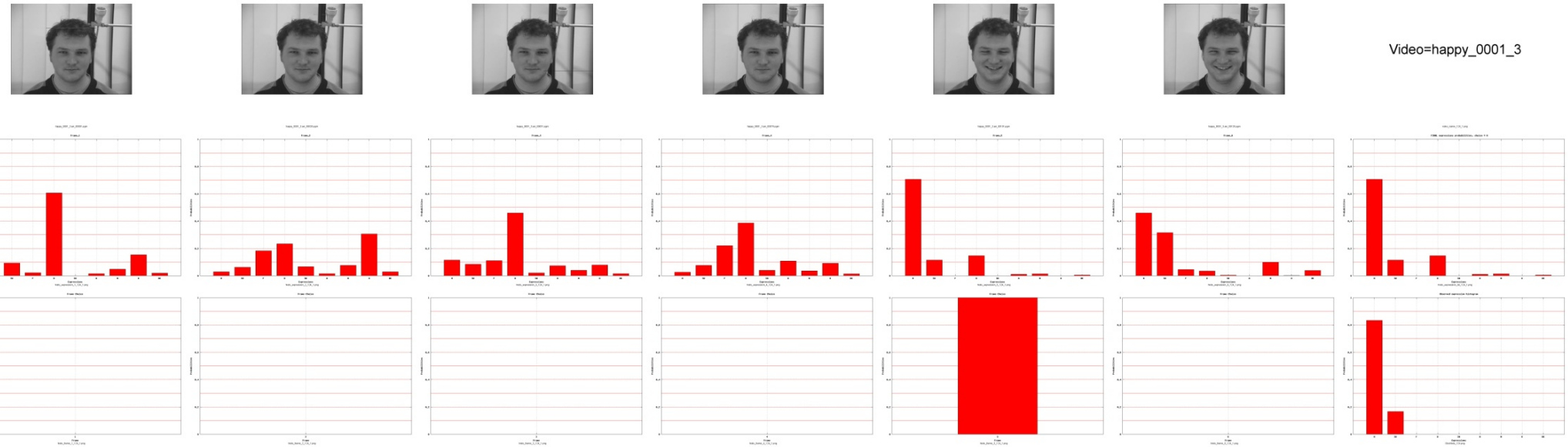
Thanks for your attention and help

Model: likelihood function



Model 1: prediction 4

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



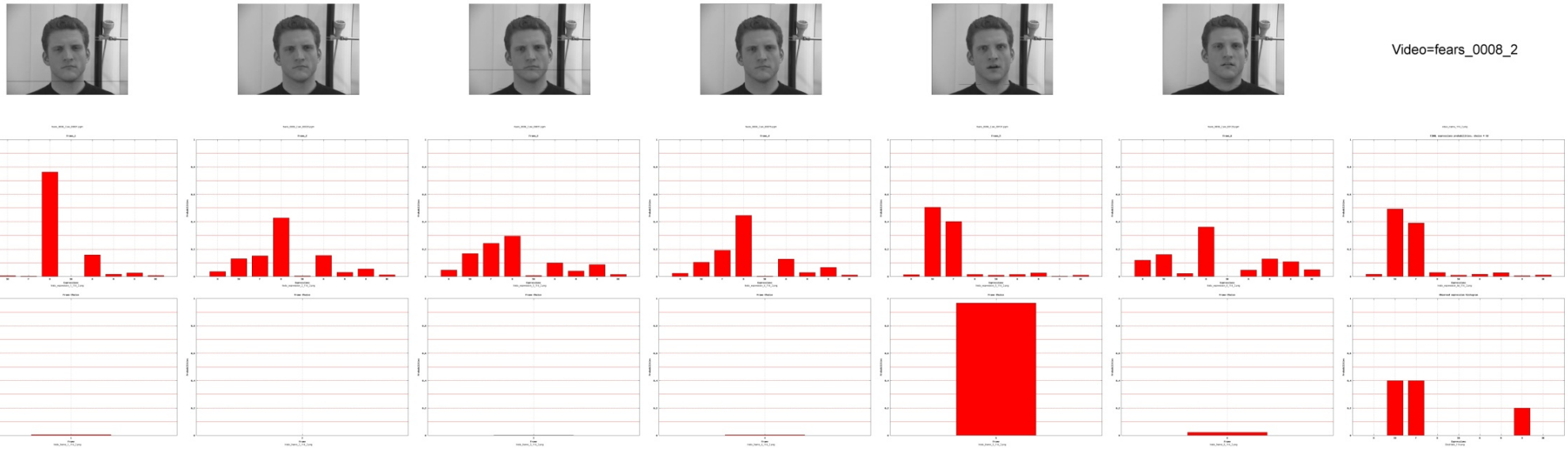
GOOD

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model 1: prediction 5

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**



GOOD

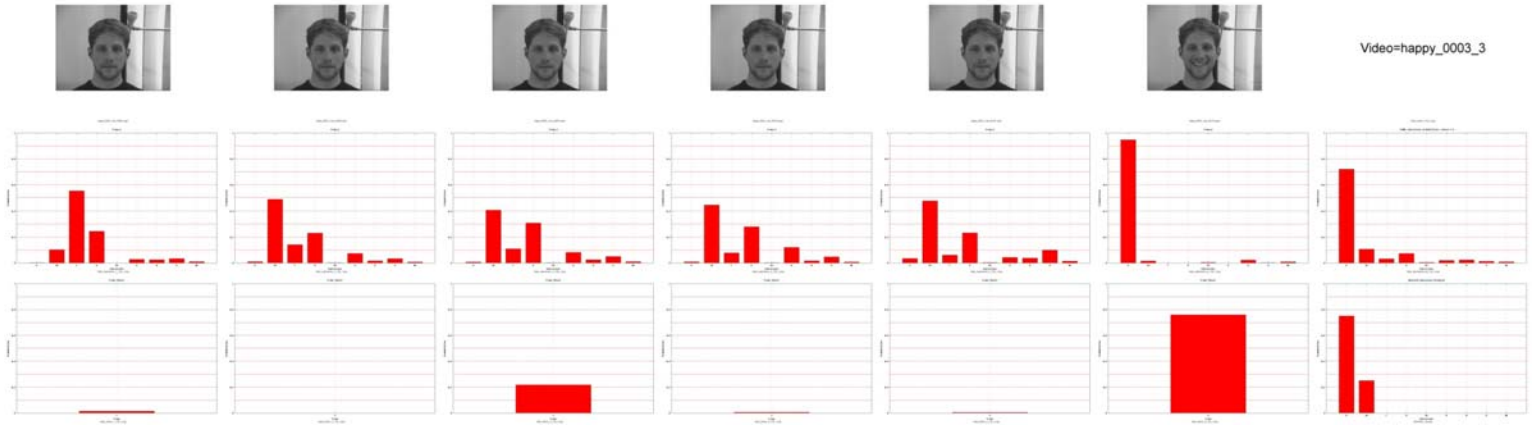
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

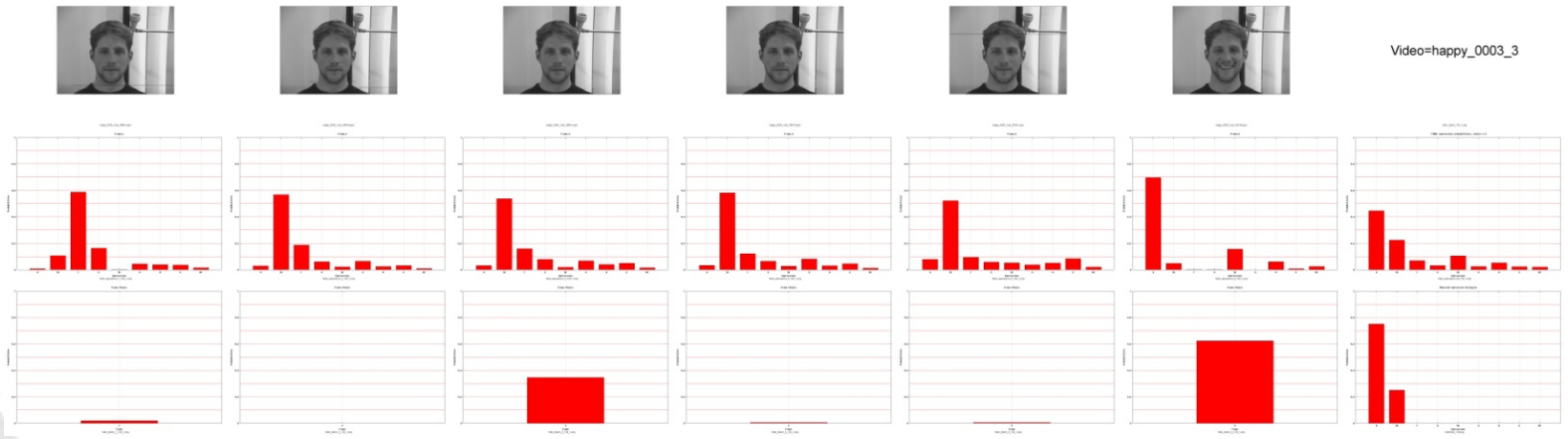
Model 2: prediction 2

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

Model 1:



Model 2:



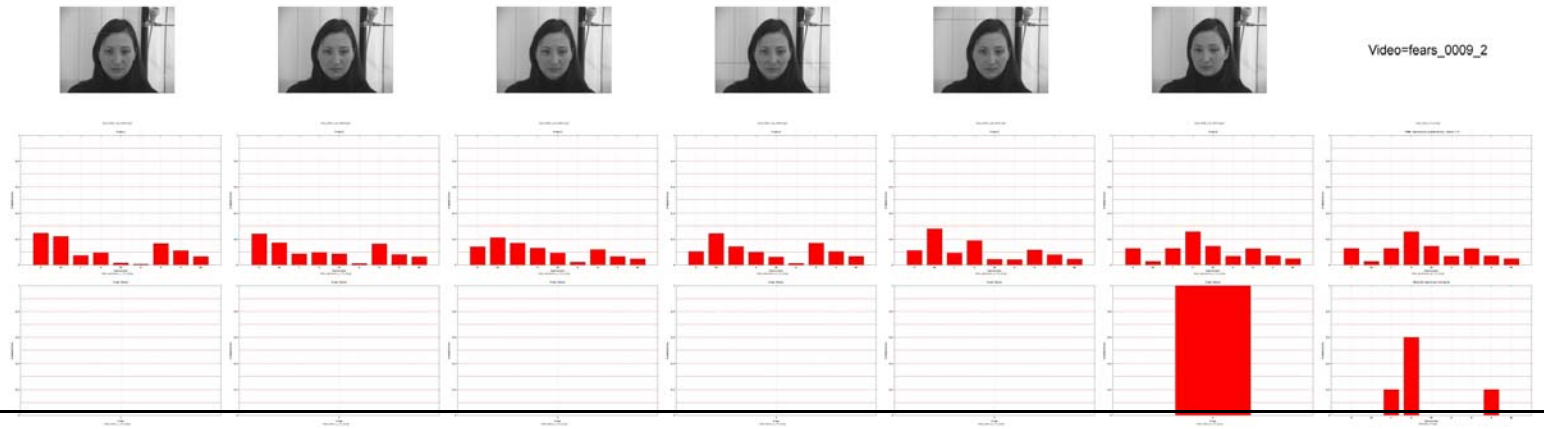
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

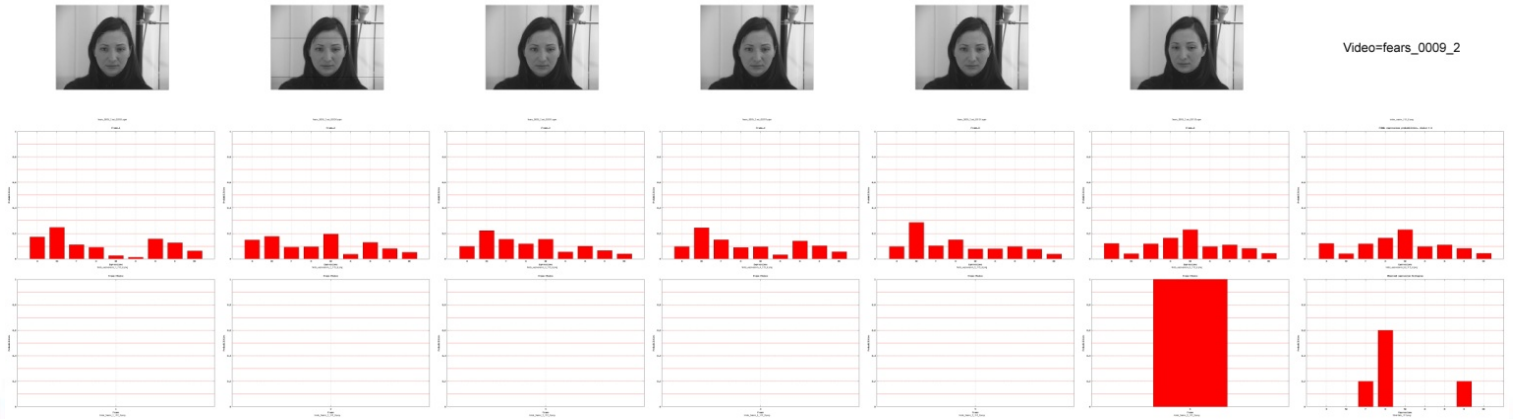
Model 2: prediction 3

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

Model 1:



Model 2:



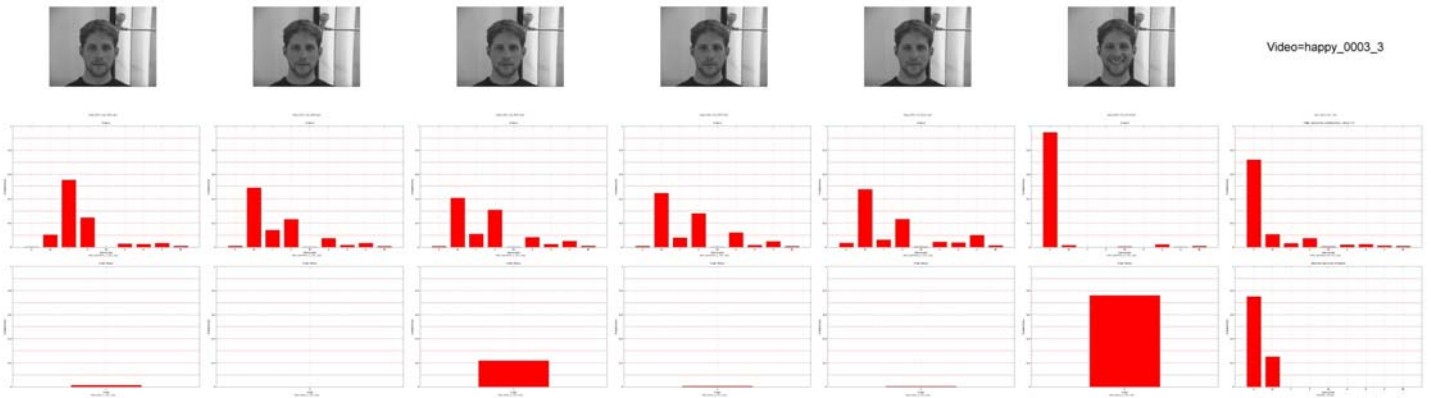
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

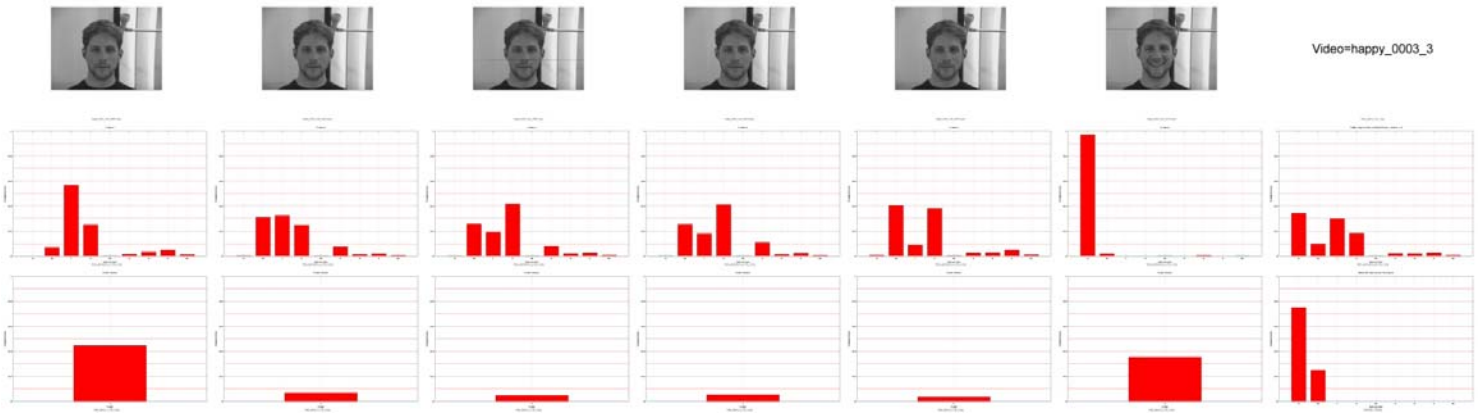
Model 3: prediction 2

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

Model 1:



Model 3:



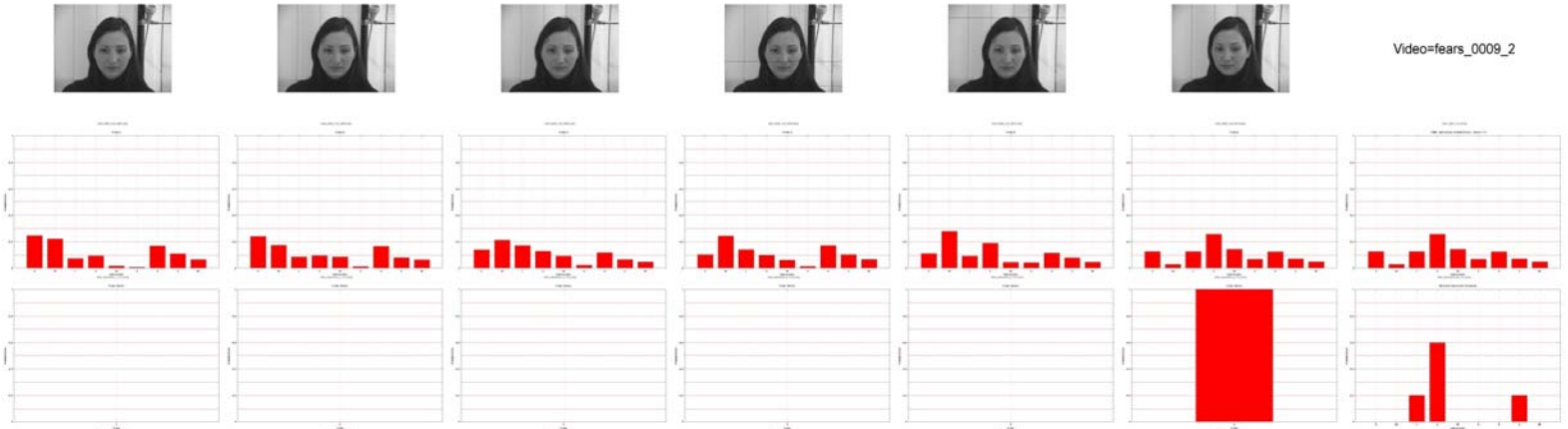
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

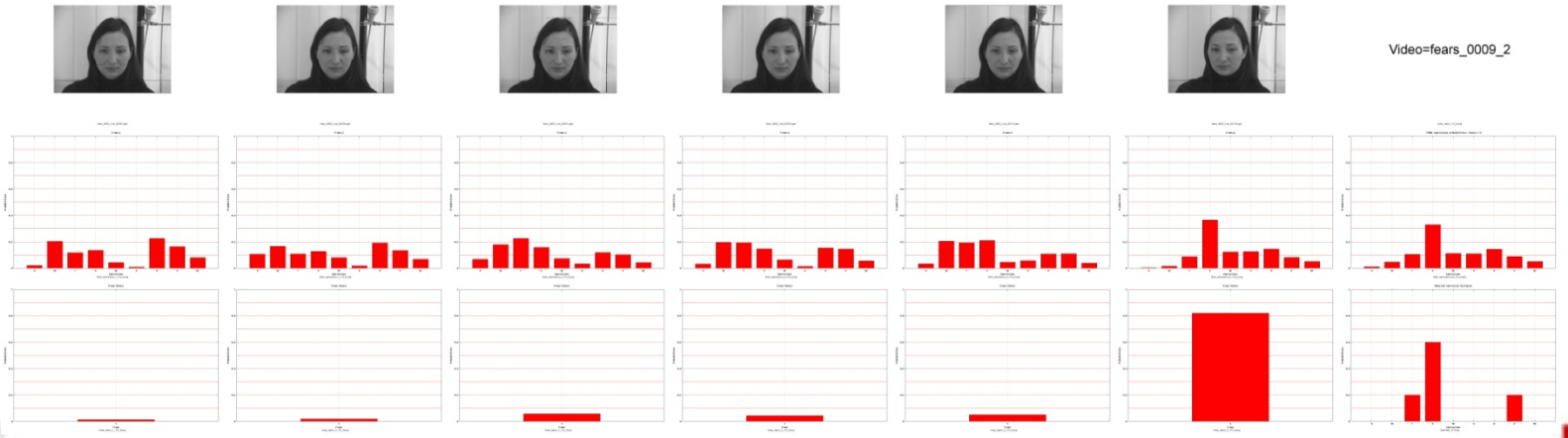
Model 3: prediction 3

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

Model 1:



Model 3:



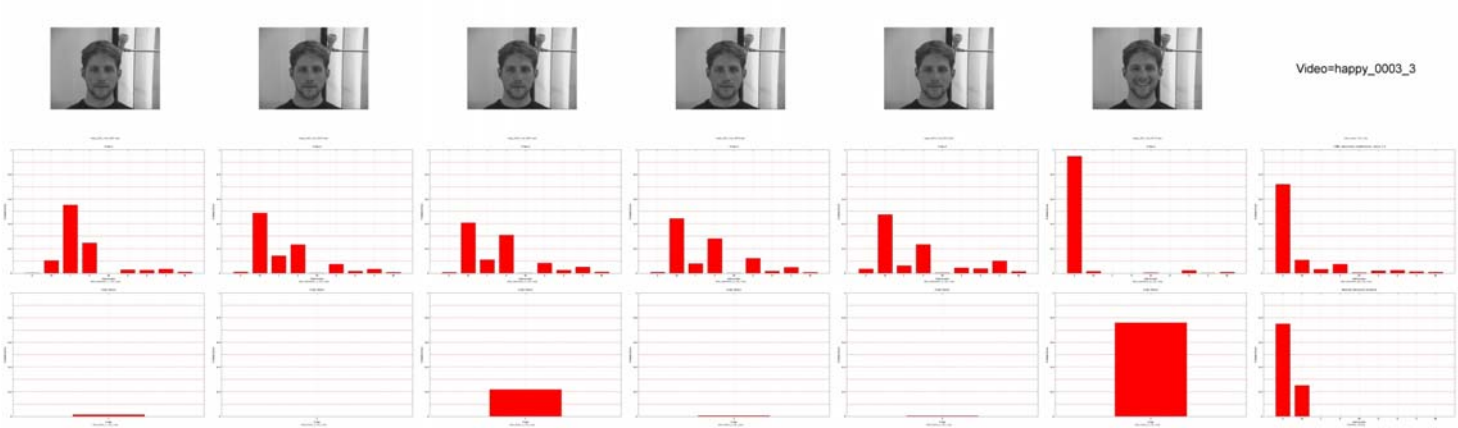
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

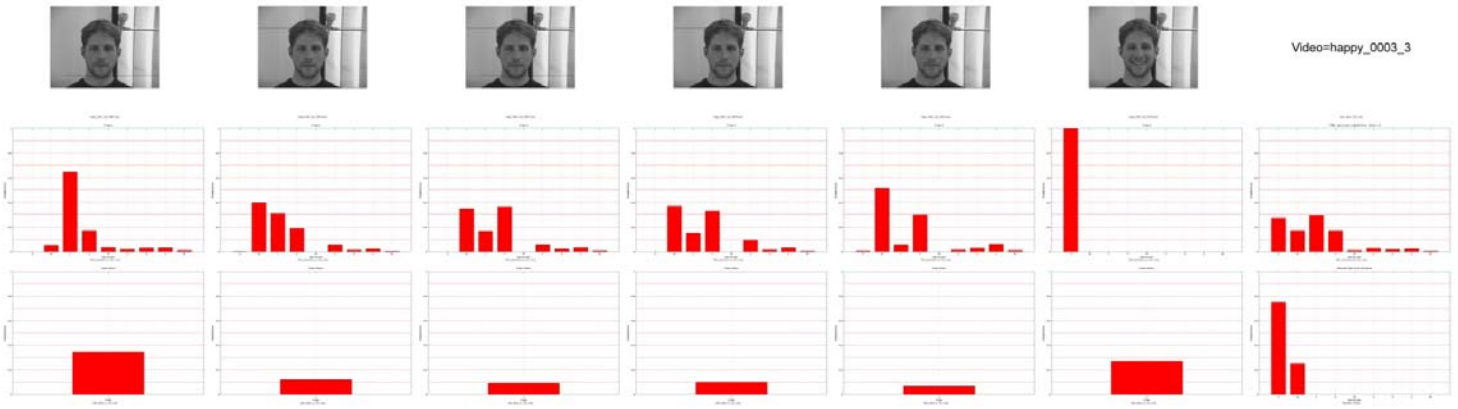
Model 4: prediction 2

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

Model 1:



Model 4:



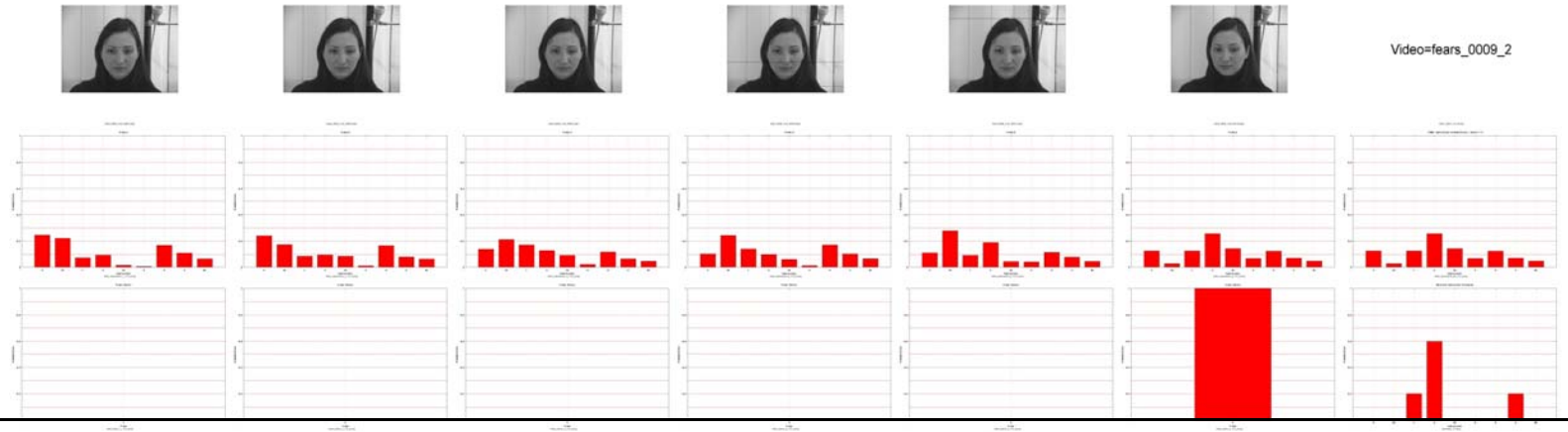
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

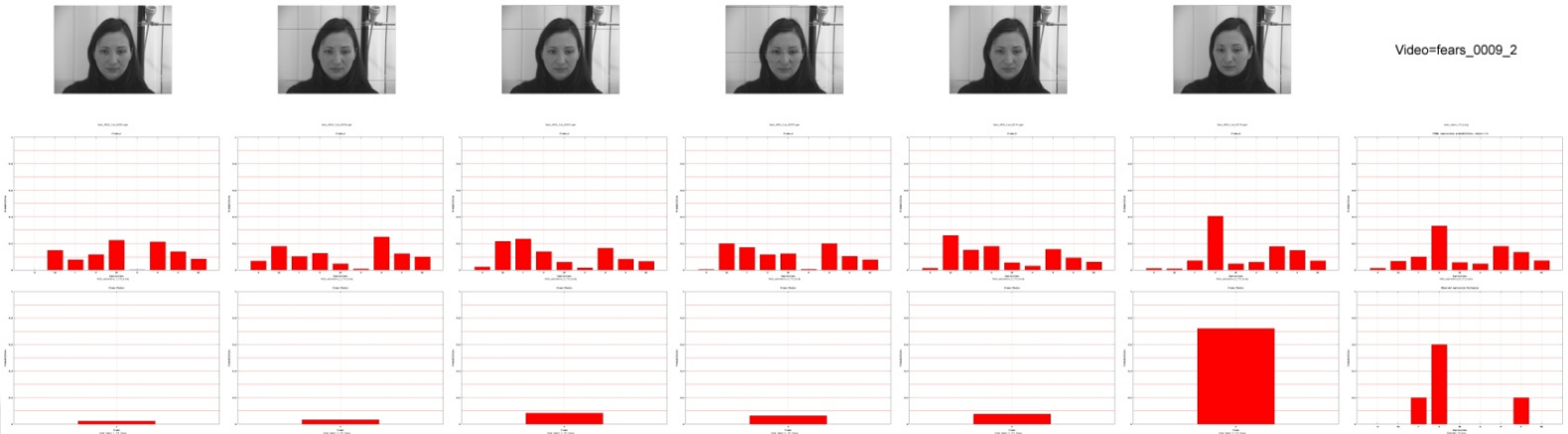
Model 4: prediction 3

- Expressions order: **H, SU, F, D, SA, A, N, O, DK**

Model 1:



Model 4:



TRANSP-OR

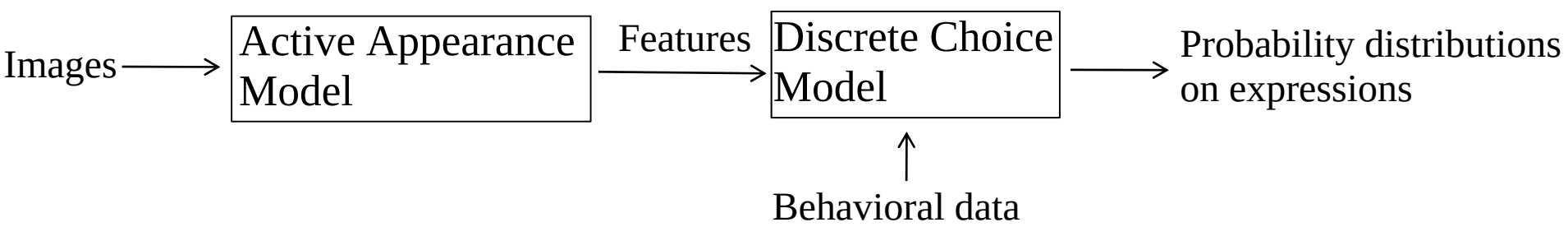
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Introduction

- Static version of the work:

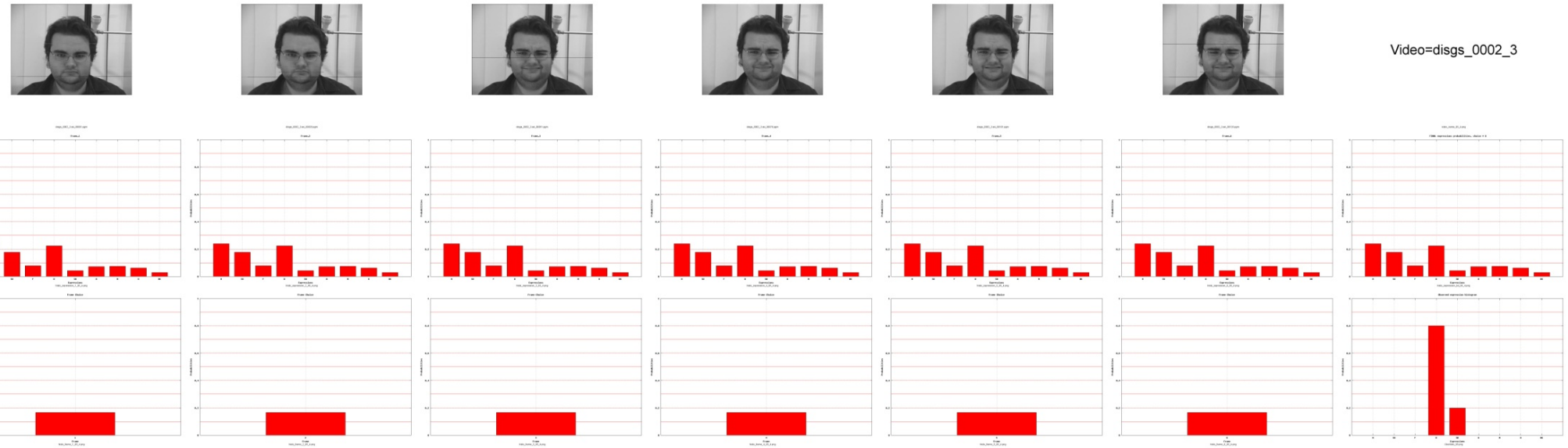
Sorci, M., Antonini, G., Cruz, J., Robin, T., Bierlaire, M. and Thiran, J.-P. (n.d.). Modelling human perception of static facial expressions, *Image and Vision Computing* . Accepted for publication.

- Images: Cohn-Kanade database
- Behavioral data: internet survey



- **Static** → **Dynamic**

ASC model



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE