Automatic Analysis of Facial Expression

Maja Pantic

iBUG @ Imperial College London

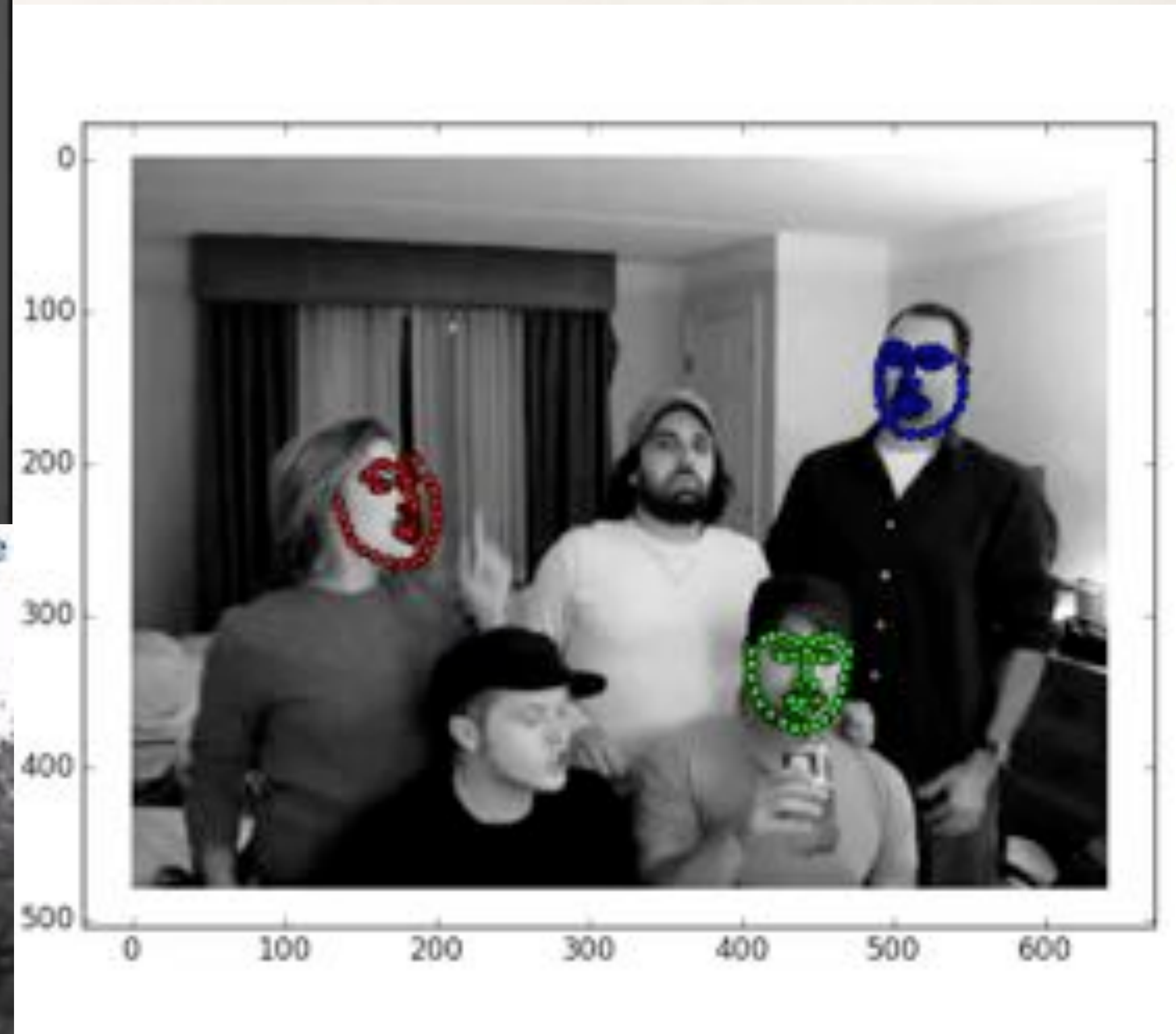
Identity

Age

Gender

Beauty

Personality



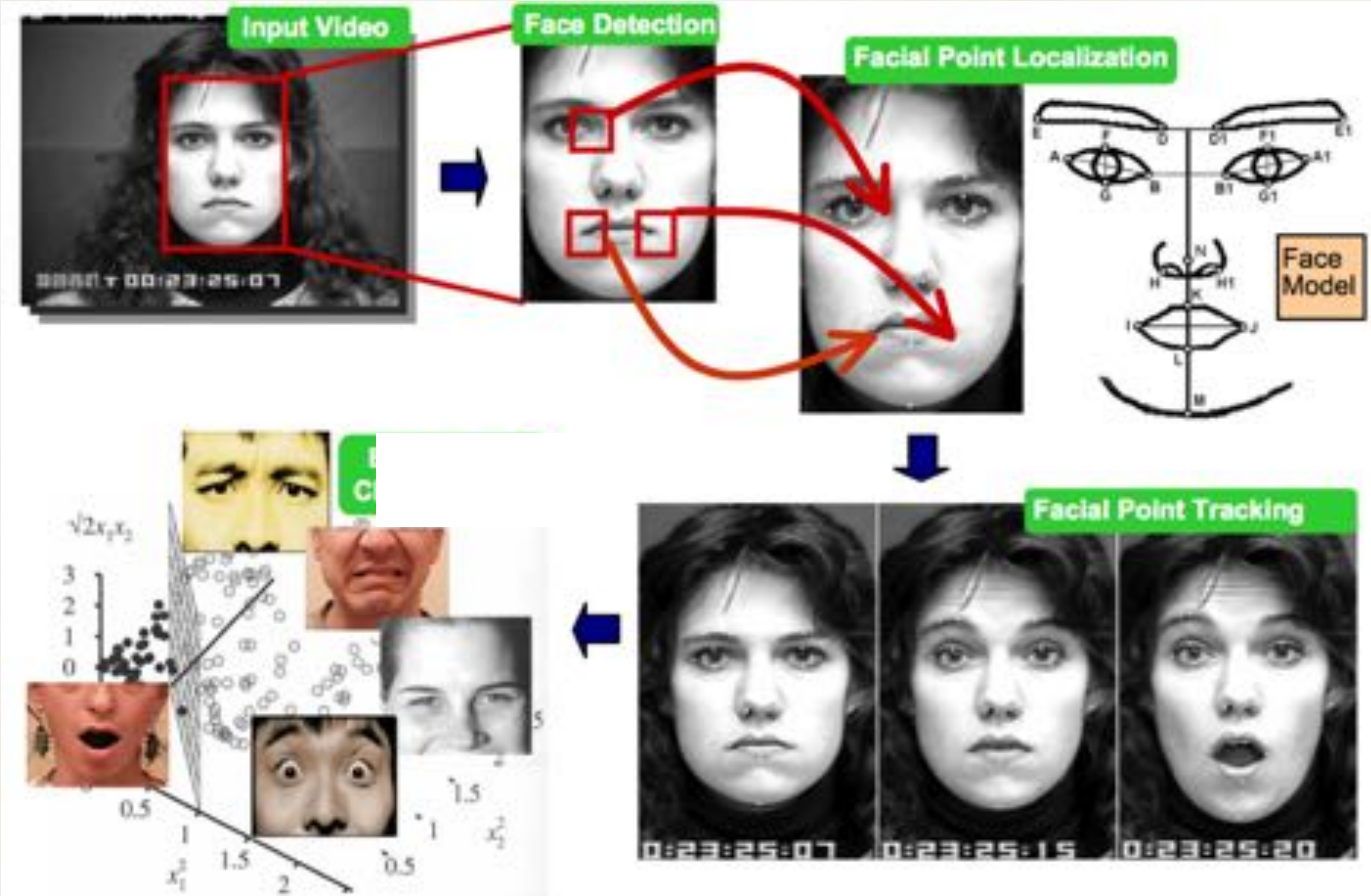
Motivation:

People emote and react on external stimuli all the time

Motivation:

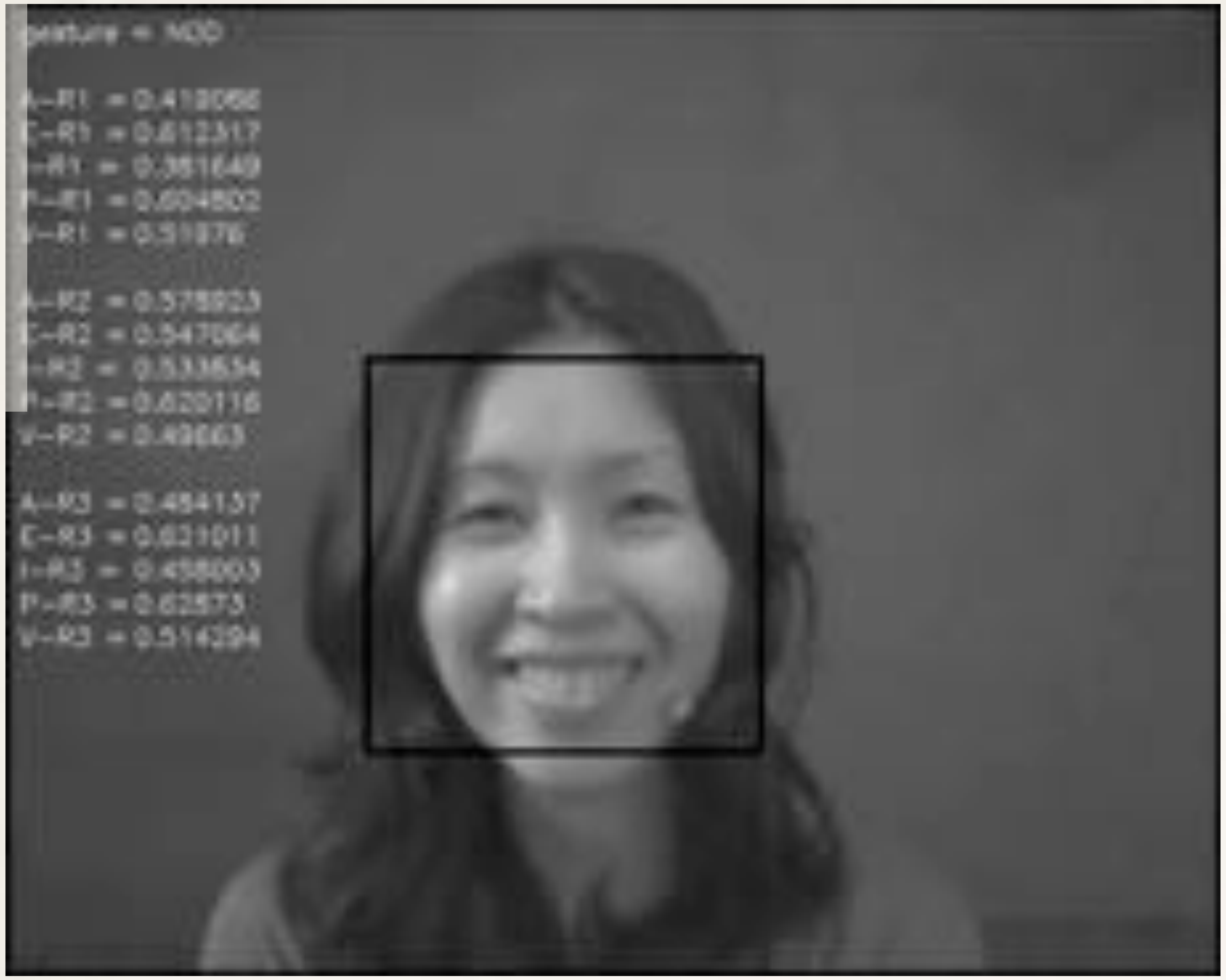
Faces = observable windows into the inner world of emotions, stances, moods

Aim: Automatic analysis of human behaviour



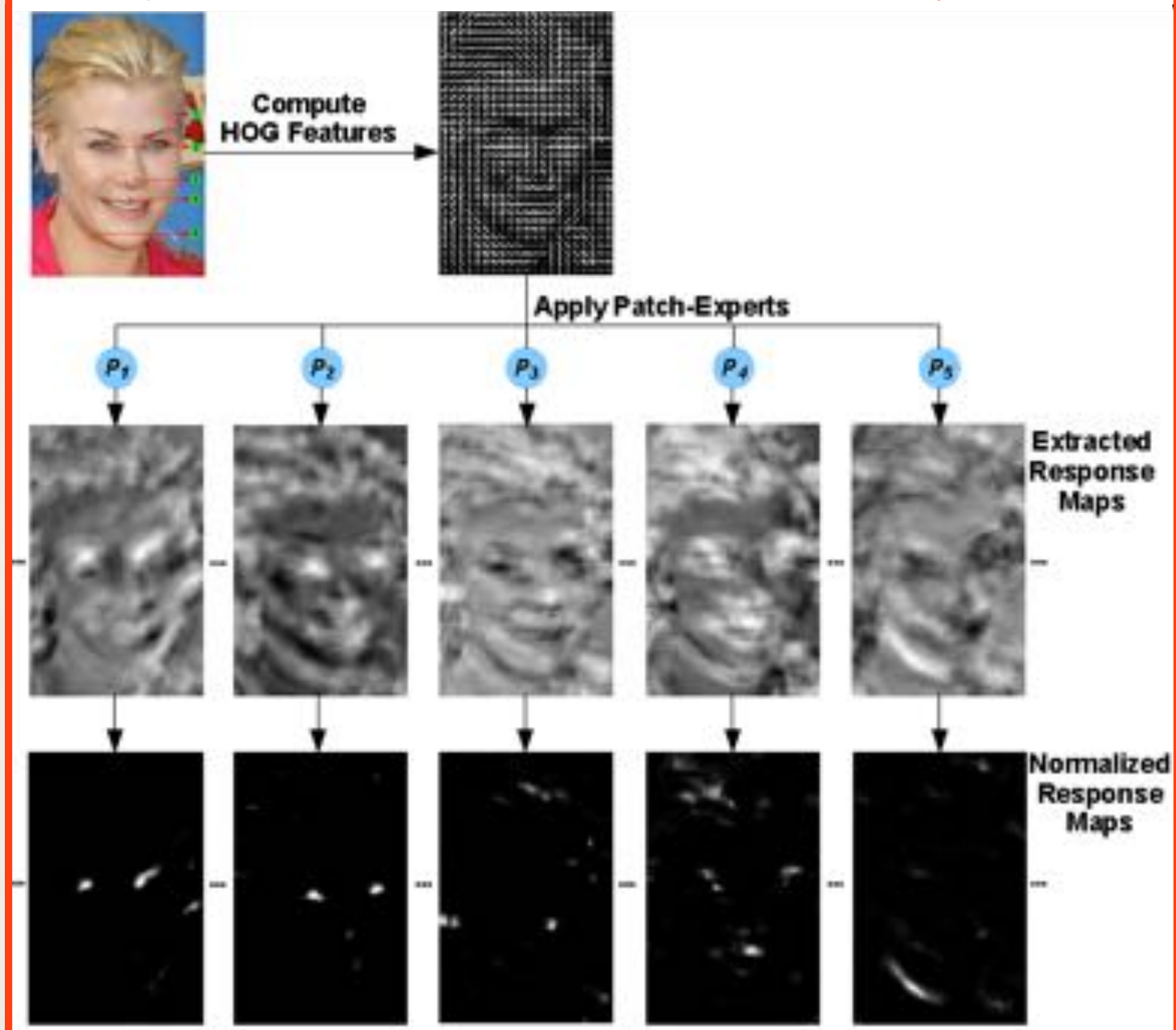
Input Video
Face Detection
Facial Point Localization
Face Model
Facial Point Tracking
00:23:25:07
0:23:25:07
0:23:25:15
0:23:25:20

Face and Facial Point Detection & Tracking

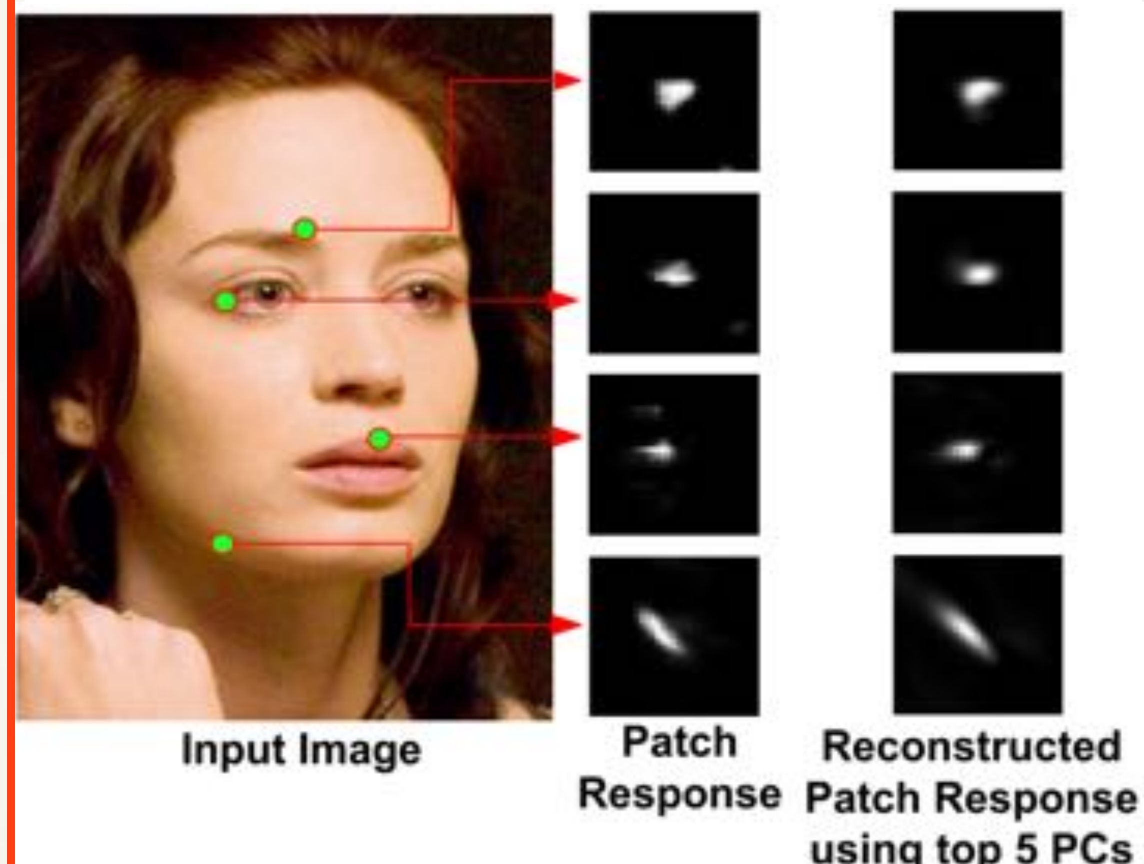


Face and Facial Point Detection & Tracking

HoG features & convolve with HoG filters



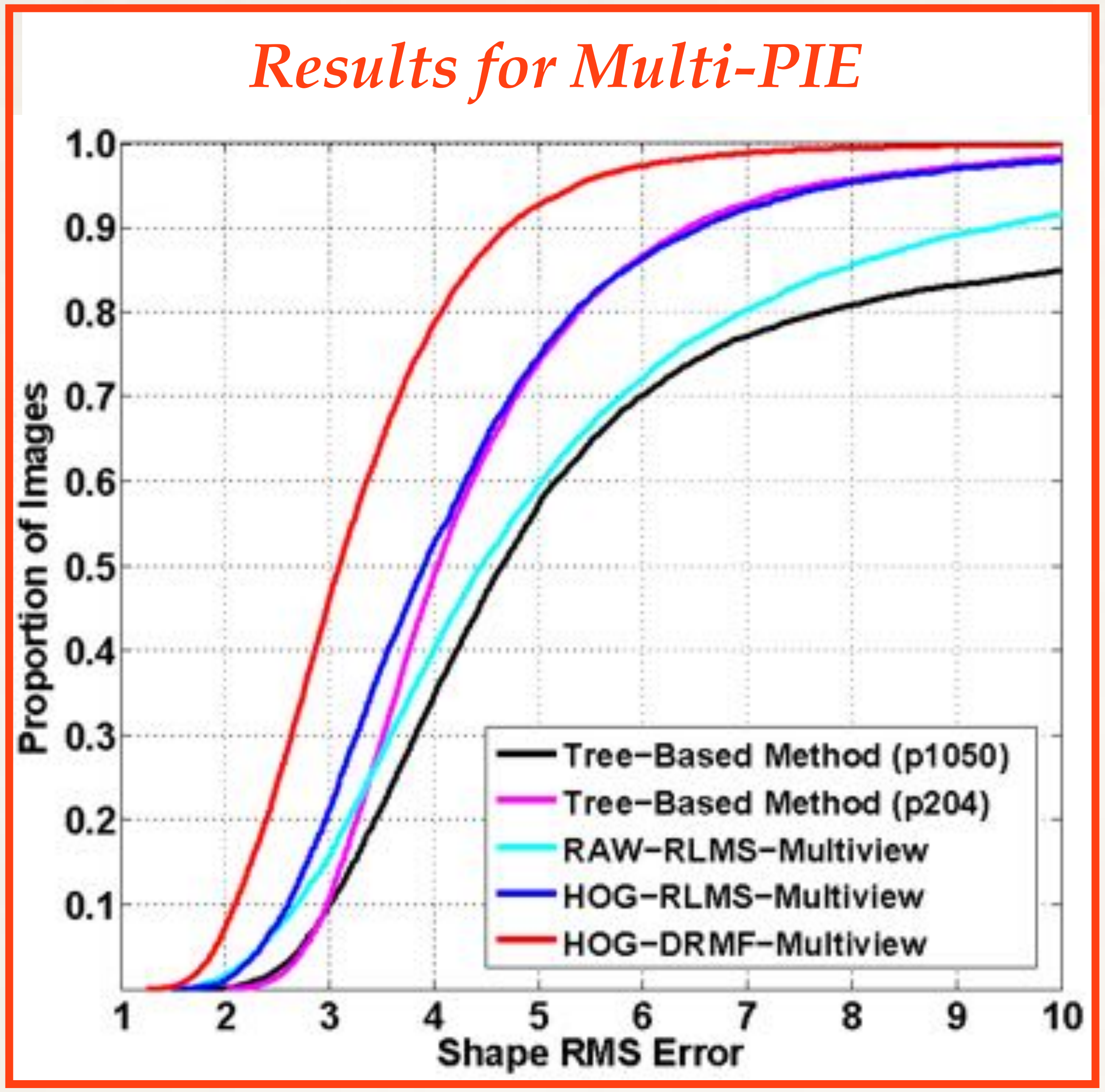
train a weak learner U for each point

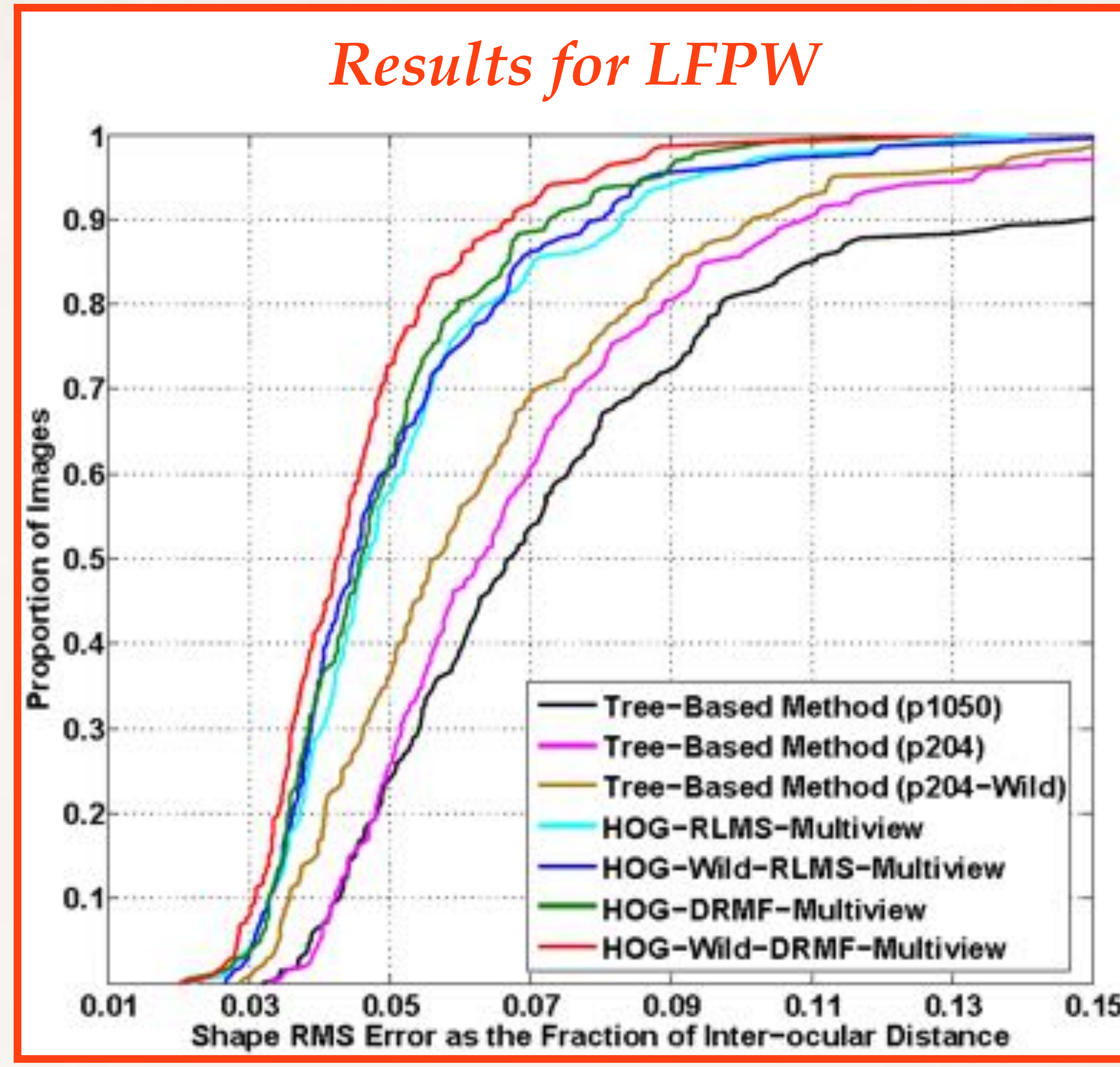


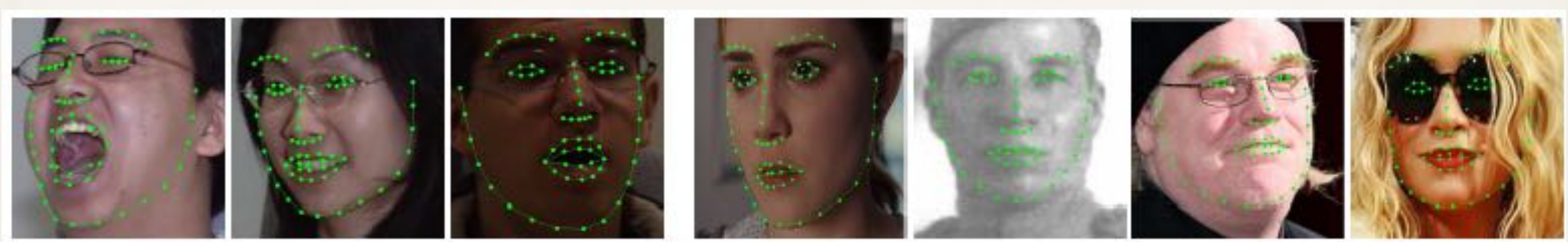
fit & update a set of weak learners U

HoG-PCA & Constrained Local Models (CVPR 2013 and 2014, PAMI 2015)

Face and Facial Point Detection & Tracking





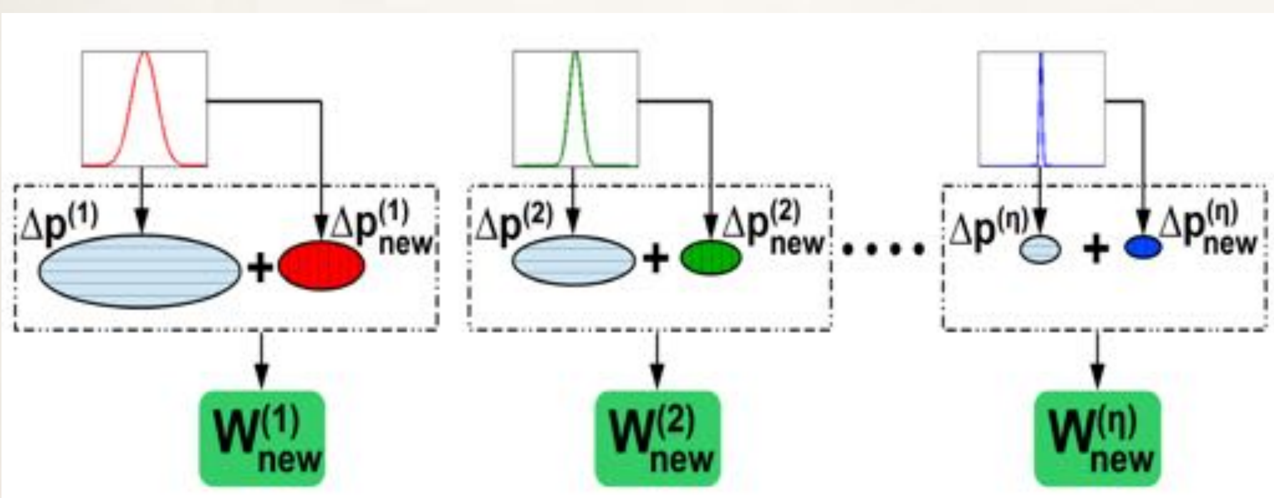


- C/CUDA implementation: 30 fps (***ibug.doc.ic.ac.uk/resources***)

HoG-PcA & constrained Local Models (cVPR 2013 and 2014, PAMI 2015)

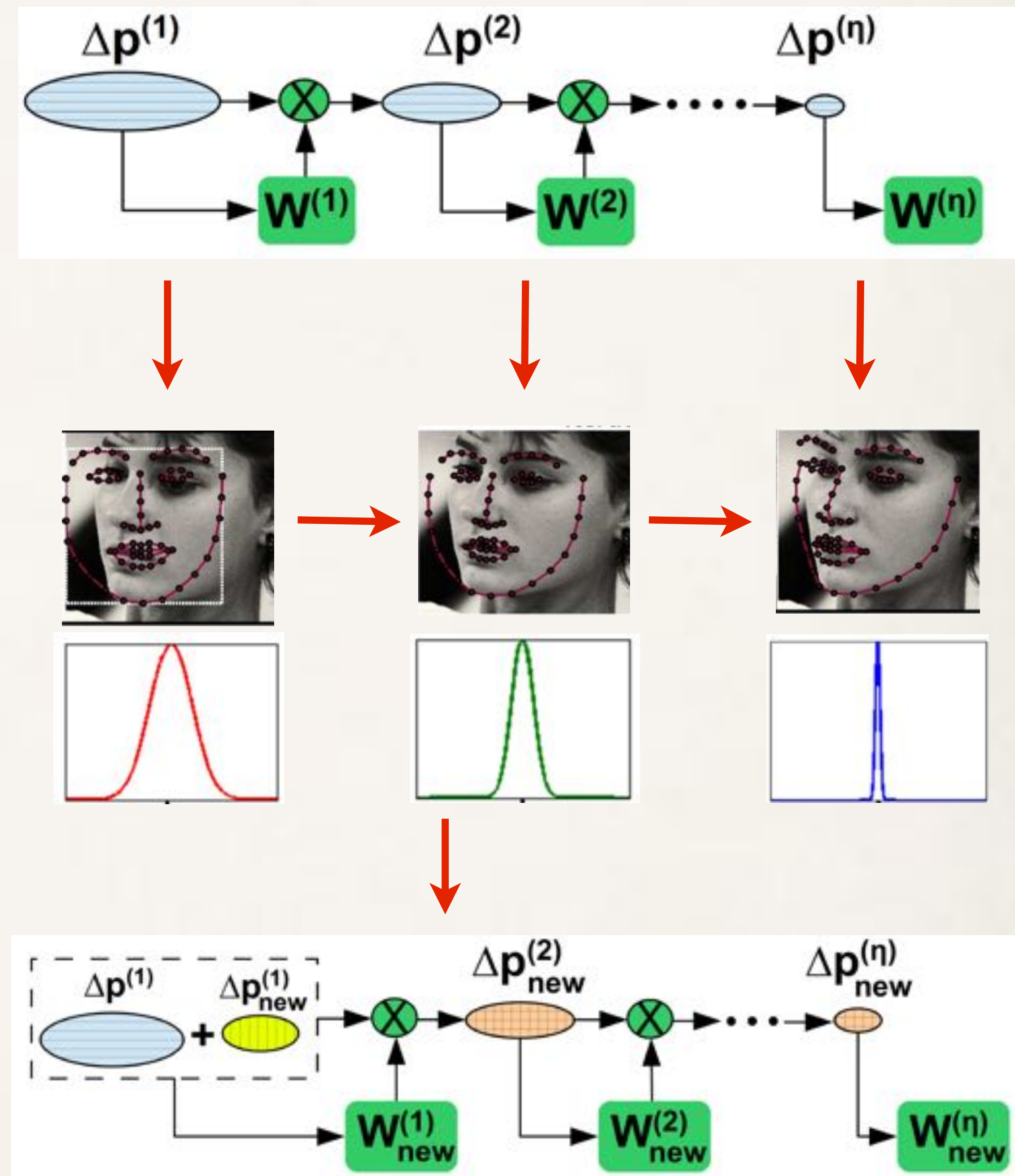
Face and Facial Point Detection & Tracking

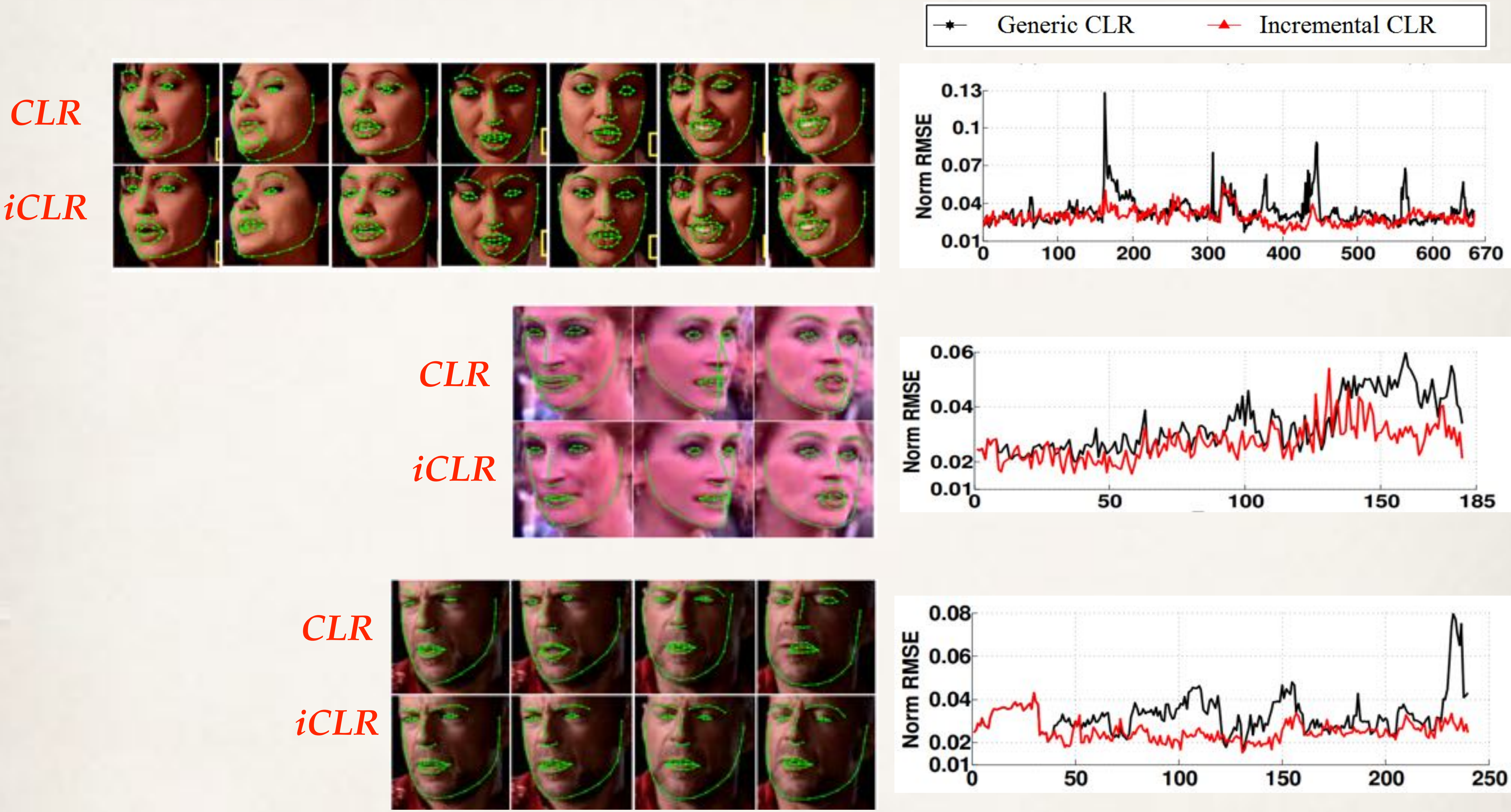
Face and Facial Point Detection & Tracking



Incremental cascade Linear Regression
(Asthana et al, CVPR 2014)

CLR by Zhou & De La Torre (CVPR'13)

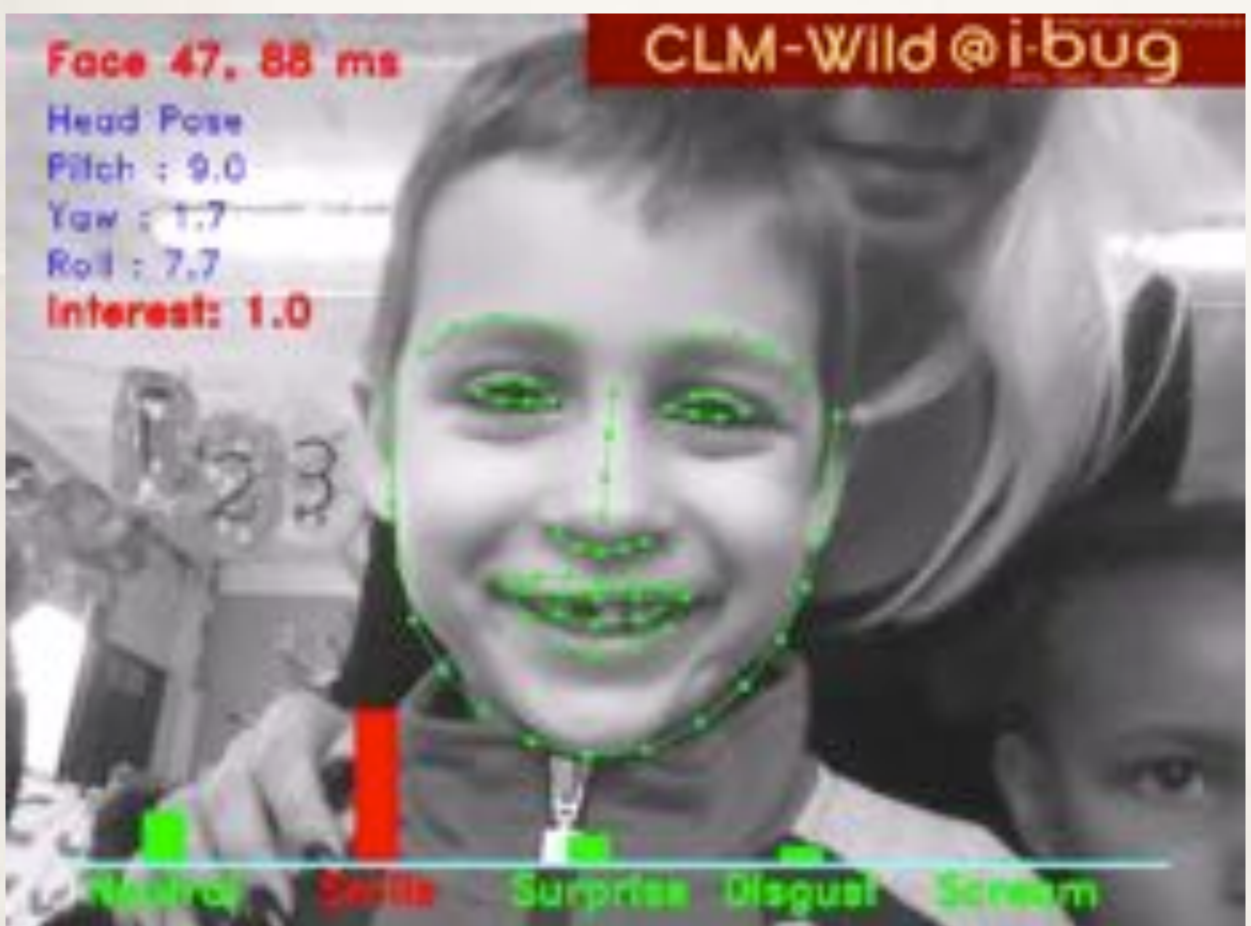




Incremental cascade Linear Regression (Asthana et al, CVPR 2014)

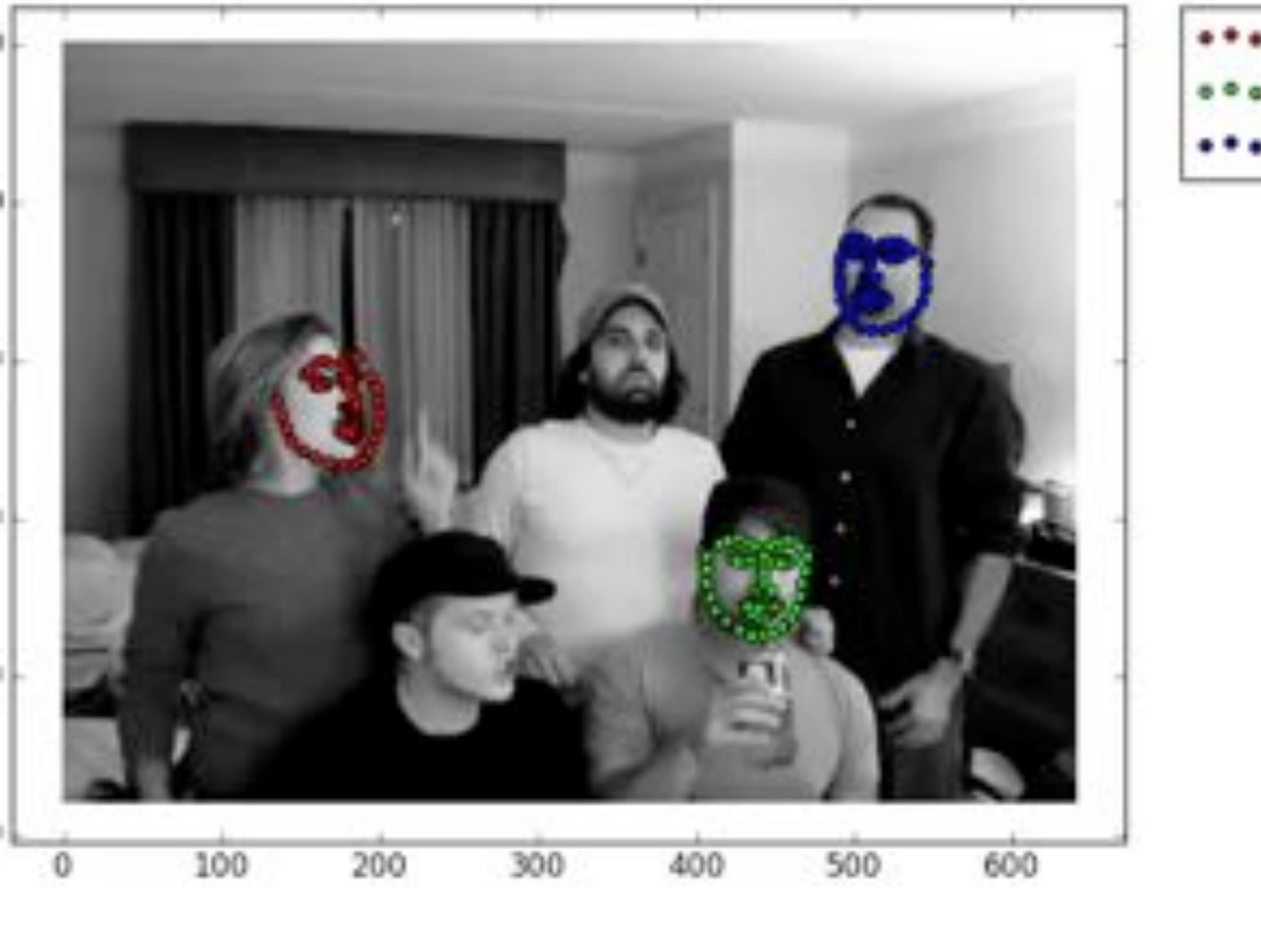
Applications

DE-ENIGMA



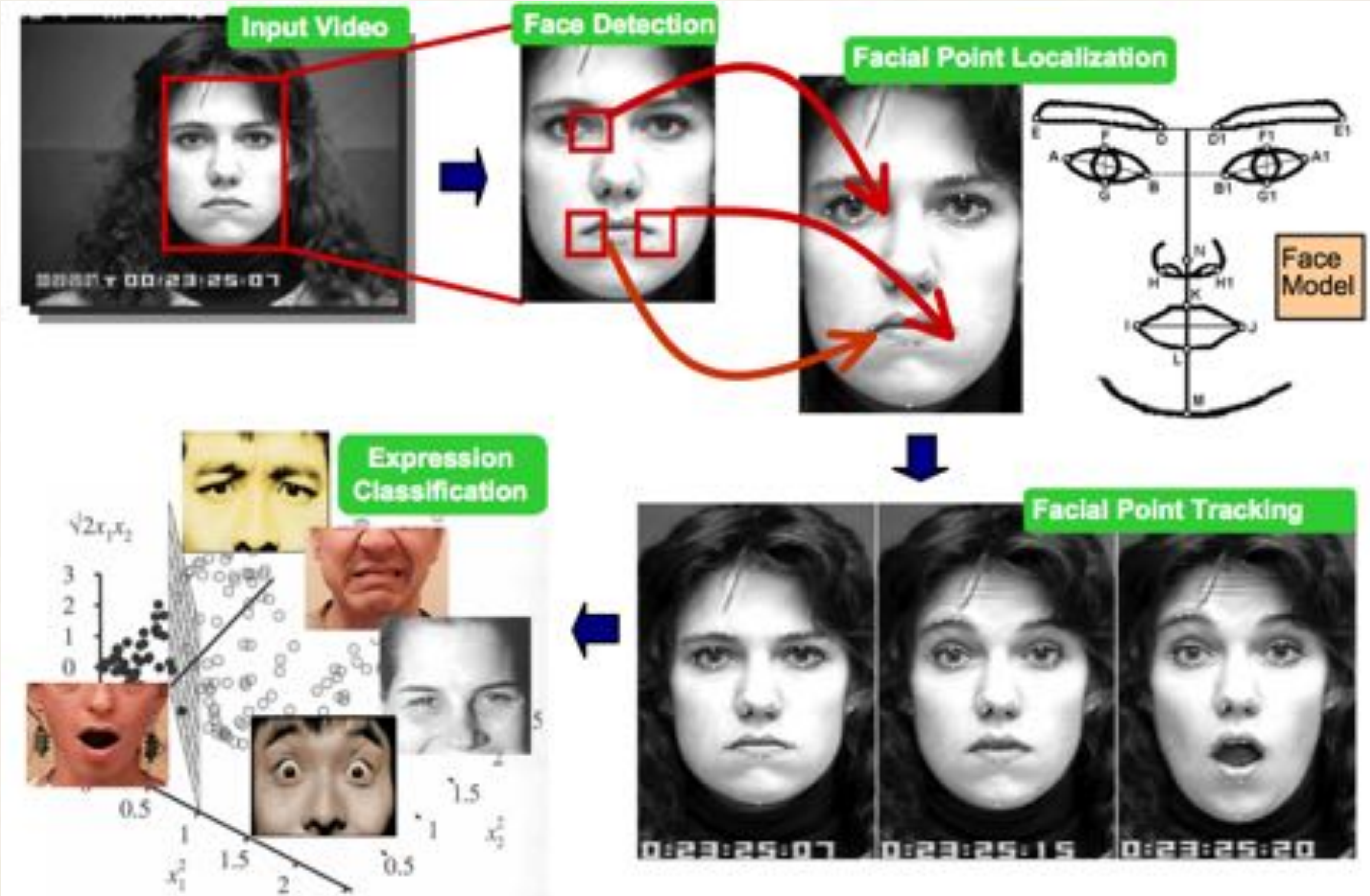
300-VW @ ICCV 2015

(ibug.doc.ic.ac.uk)

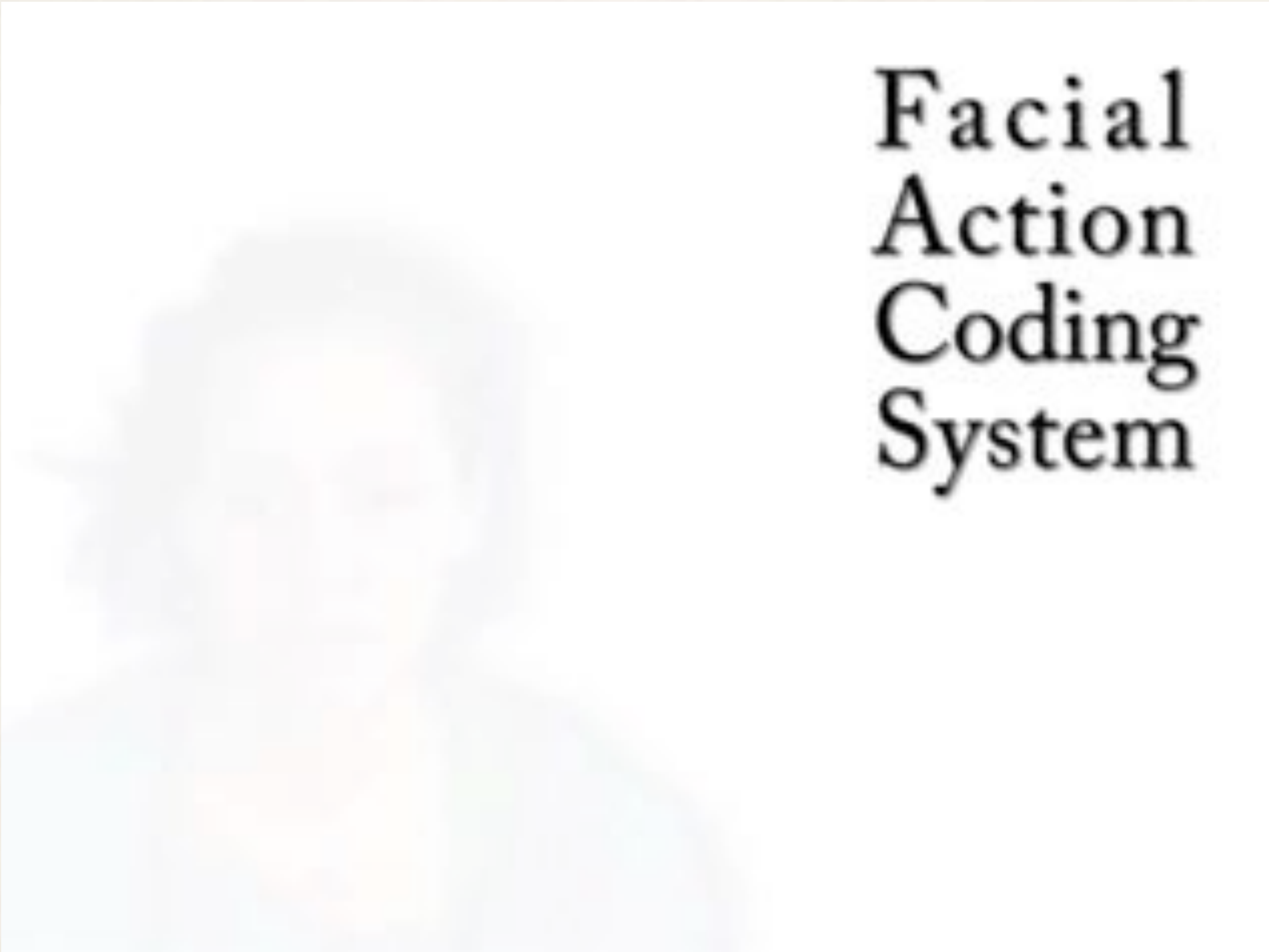


Automatic Sentiment Analysis in the Wild

Facial Expression Classification

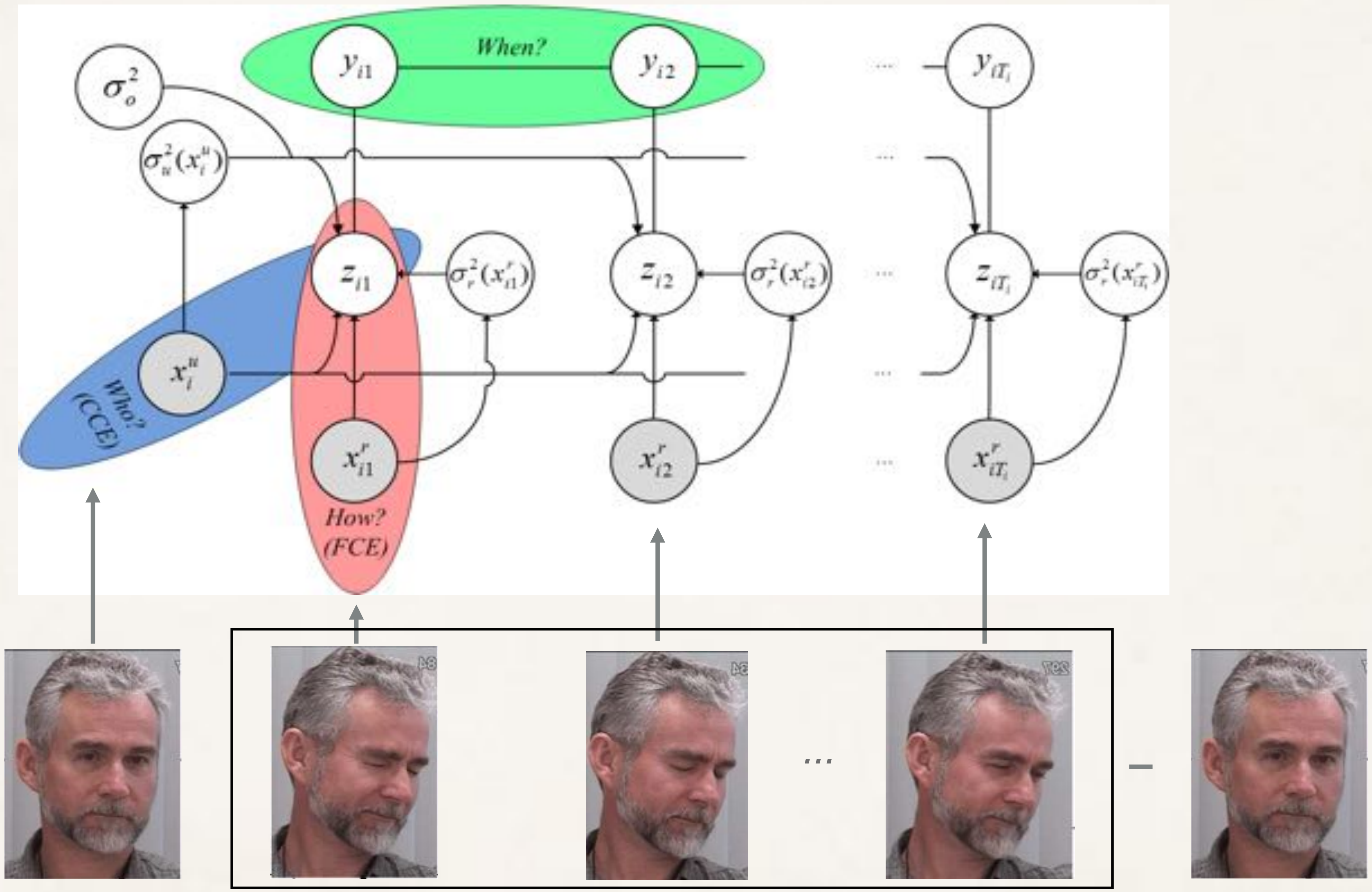


Facial Expression Classification



Facial Expression Classification

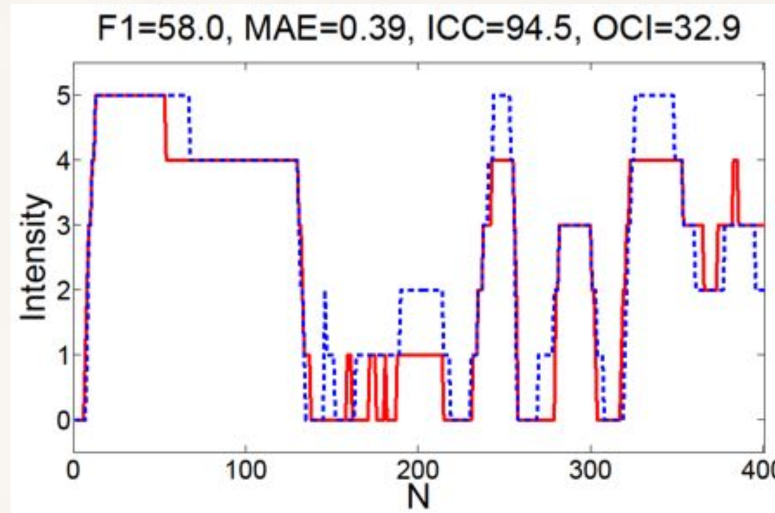
context (W5+): who? when? where? what? how? why?



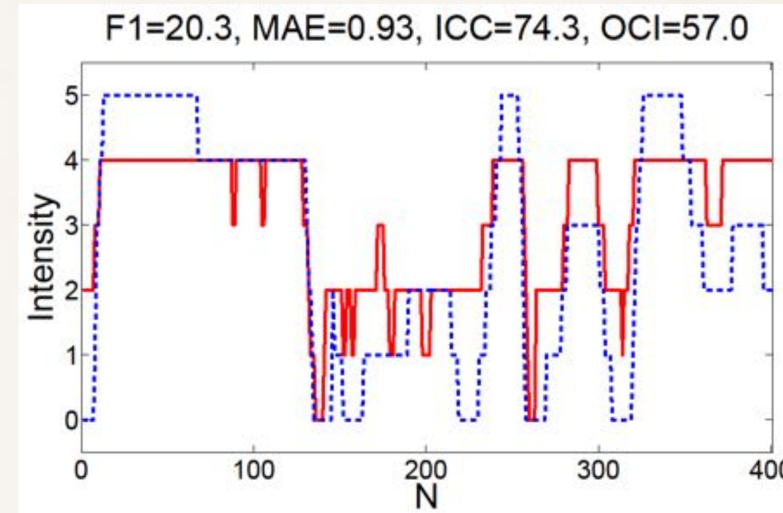
Feature-based, context-sensitive cORF approach to AU intensity estimation

(Rudovic, Pavlovic & Pantic, PAMI 2015)

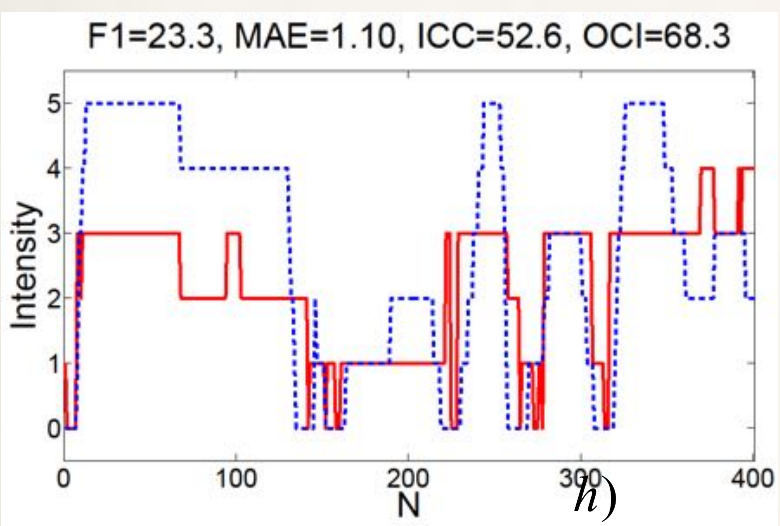
context (W5+): who? when? where? what? how? why?



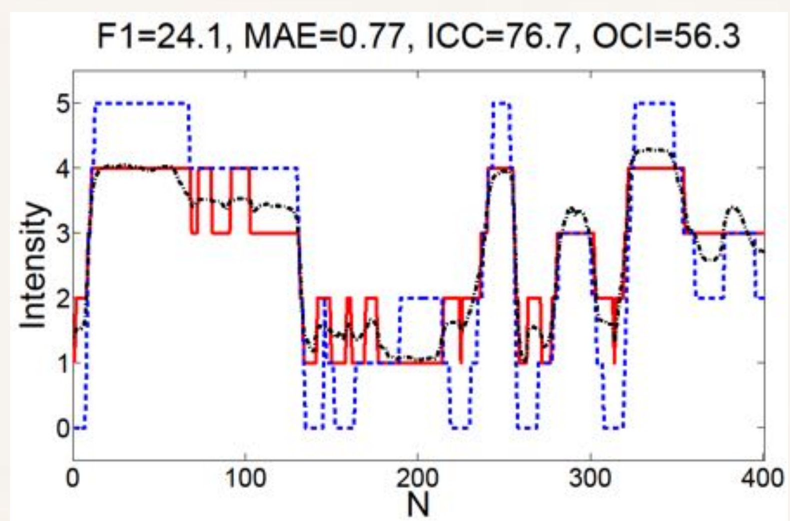
$cs - CORF(w + h)$



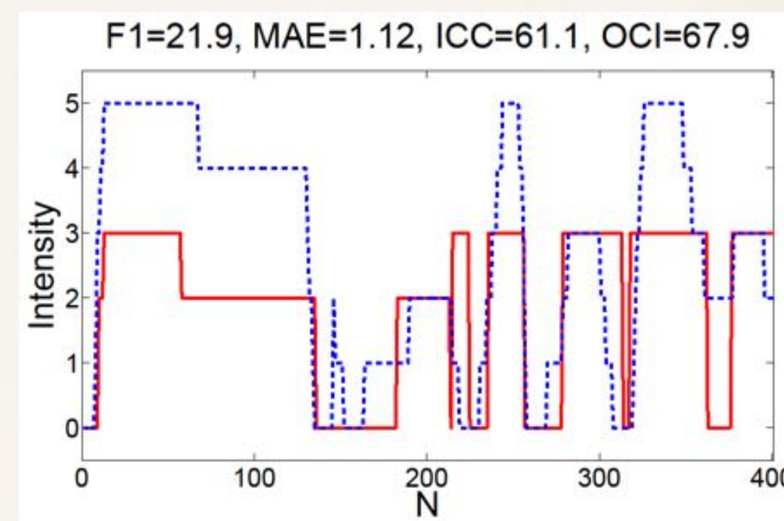
$CORF(w + h)$



$CRF(w)$



RVM



$SR + SVM$

Feature-based, context-sensitive CORF approach to AU intensity estimation

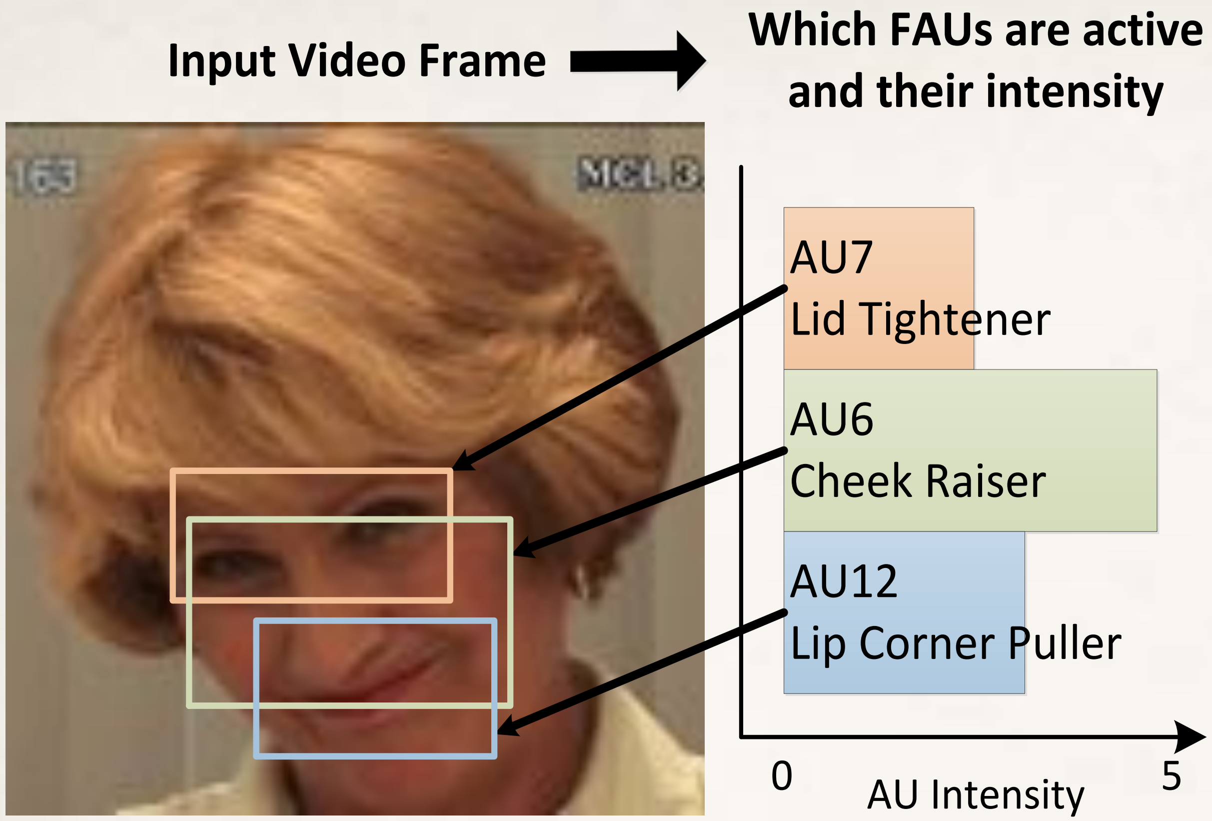
(Rudovic, Pavlovic & Pantic, PAMI 2015)

Facial Expression Classification

TERESA
Telepresence Reinforcement-learning Social Agent

Feature-based, context-sensitive CORF approach to AU intensity estimation
(Rudovic, Pavlovic & Pantic, PAMI 2015)

Facial Expression Classification

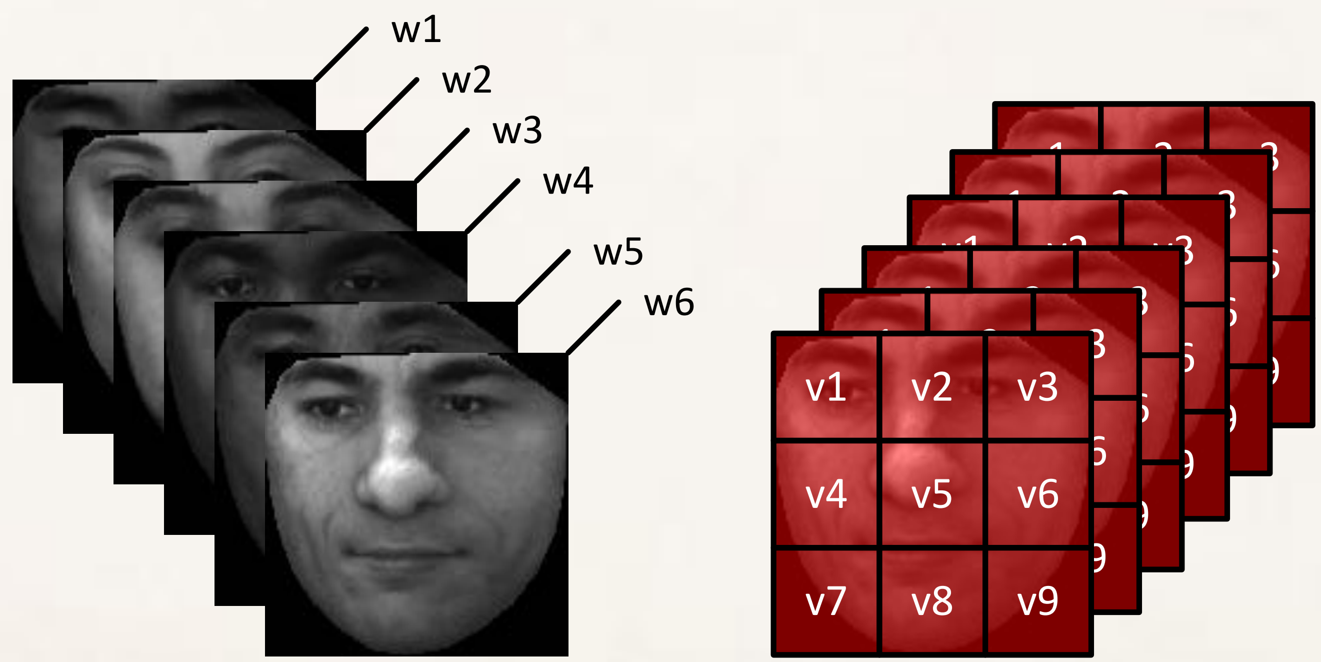


DSRVM prediction function:

$$y(\mathbf{x};\mathbf{w},\mathbf{v}) = \sum_{m=1}^{M} \sum_{k=1}^{K} w_m v_k \kappa_k(\mathbf{x}, \mathbf{x}_m)$$

$\kappa_k(\cdot,\cdot)$ Gaussian kernel defined on patch k

w: weights over training instances

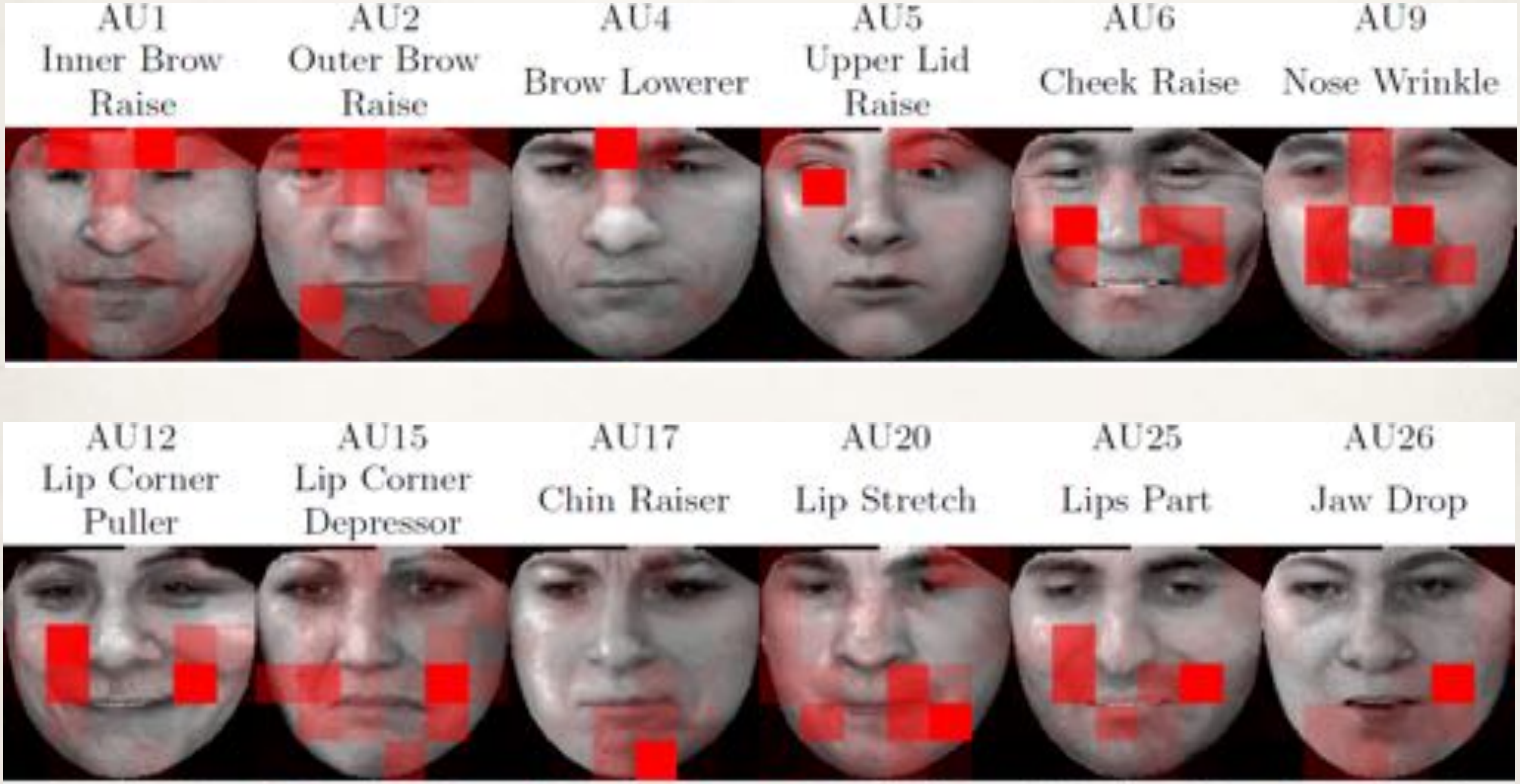


v: weights over facial patches

Doubly Sparse RVM for pain and AU intensity estimation

(Kaltwang, Todorovic & Pantic, PAMI 2016)

Facial Expression Classification



Doubly Sparse RVM for pain and AU intensity estimation
(Kaltwang, Todorovic & Pantic, PAMI 2016)

Facial Expression Classification

Method	AU1	AU2	AU4	AU5	AU6	AU9	AU12	AU15	AU17	AU20	AU25	AU26	AVG
DSRVM	0.31	**0.28**	**0.54**	0.17	**0.57**	**0.43**	**0.80**	0.32	**0.40**	**0.23**	**0.66**	**0.42**	**0.43**
RVM all	0.26	**0.29**	0.41	0.12	0.46	0.32	0.75	0.33	**0.40**	0.19	0.62	0.40	0.38
p-value	.11	1.00	.00	.05	.02	.00	.05	.49	.80	.07	.34	.59	.01
RVM best	0.18	0.15	0.46	**0.19**	0.40	0.30	0.73	0.18	0.21	0.16	0.61	0.17	0.31
p-value	.00	.01	.07	.16	.03	.03	.01	.01	.00	.00	.11	.00	.00
RVM sep	**0.35**	0.27	0.50	0.17	0.47	0.34	0.78	0.34	0.38	0.21	0.61	0.40	0.40
p-value	.42	.24	.18	.68	.03	.03	.06	.16	.61	.14	.08	.49	.04
SMKL	0.26	0.20	0.45	0.17	0.49	0.36	**0.81**	0.26	0.36	0.22	**0.66**	0.40	0.39
p-value	.06	.01	.02	.94	.06	.02	.26	.05	.14	.43	.93	.47	.00
mRVM	0.21	0.16	0.32	0.10	0.39	0.40	0.69	0.23	0.32	**0.24**	0.59	0.23	0.32
p-value	.25	.17	.02	.22	.02	.57	.00	.22	.17	.97	.00	.03	.00
SR+SVM	0.13	0.13	0.33	0.04	0.40	0.39	0.75	0.21	0.29	0.18	0.47	0.36	0.31
p-value	.01	.00	.00	.02	.01	.24	.04	.08	.08	.14	.00	.09	.00

Measure:
Pearson Correlation Coefficient

Comparison Methods:
RVM all – RVM using all patches with equal weights
RVM best – RVM using single best patch
RVM sep – RVM using all patches as separate basis functions
SMKL – Simple Multiple Kernel Learning (A. Rakotomamonjy et al., JMLR'08)
mRVM – multi-kernel RVM (T. Damoulas et al., ICMLA'08)
SR+SVM – Spectral Regression combined with SVM (S. Mavadati et. al., TAC'13)

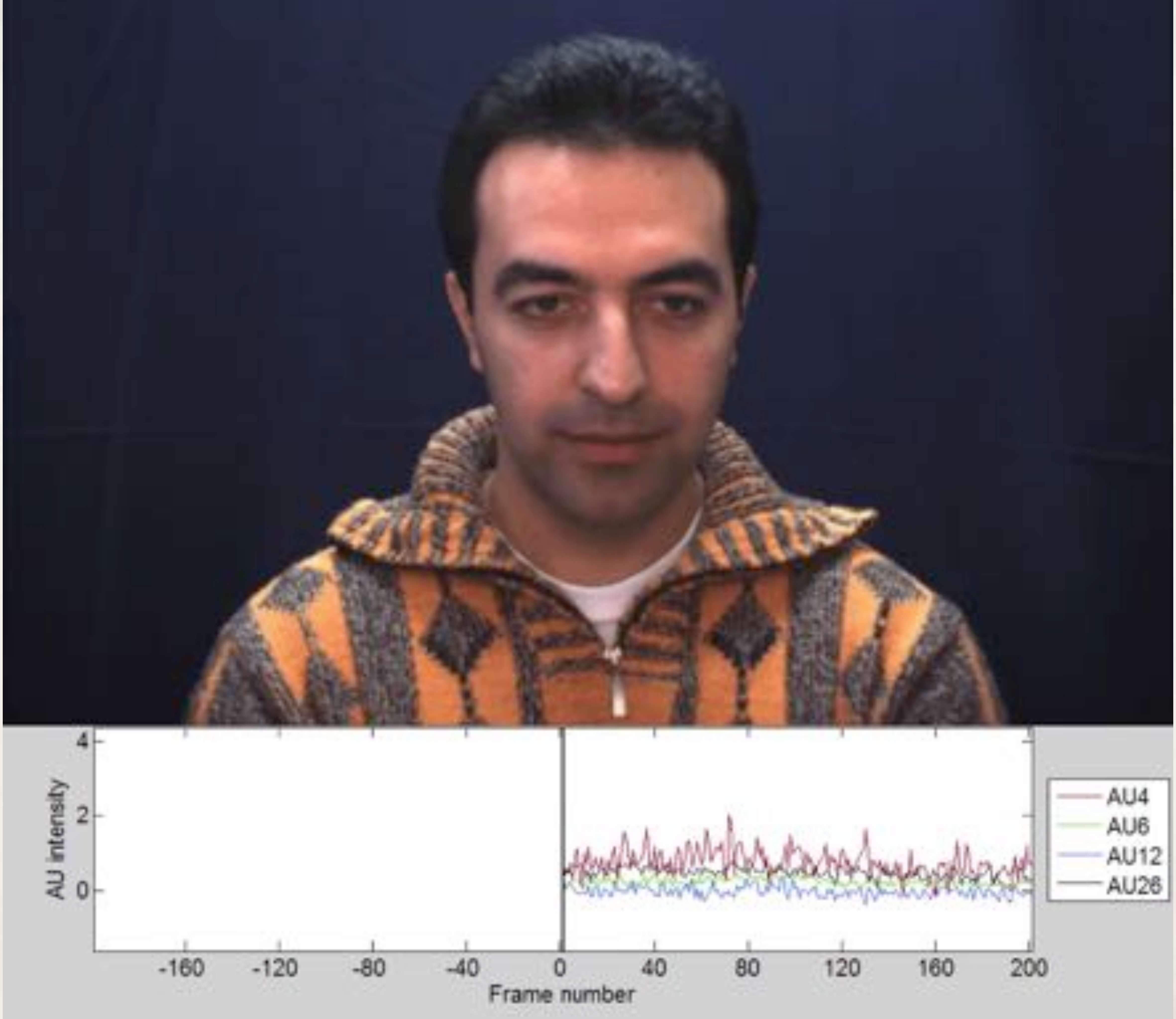
Doubly Sparse RVM for pain and AU intensity estimation
(Kaltwang, Todorovic & Pantic, PAMI 2016)

Method		#RV	#RK	TRN	TST
		# Relevance vectors	# Relevant kernels	training time	testing time
DSRVM		55.5	21.4	21.3	0.4
RVM all	[47]	120.6	-	1.2	1.3
RVM best	[47]	77.7	-	1.3	0.1
RVM sep	[3], [49]	217.6	-	36.2	3.1
SMKL	[35]	1914.2	32.5	149.8	7.8
mRVM	[6]	52.6	36.0	77.4	1.2
SR+SVM	[25]	464.2	-	0.1	0.1

Measure:
Pearson Correlation Coefficient

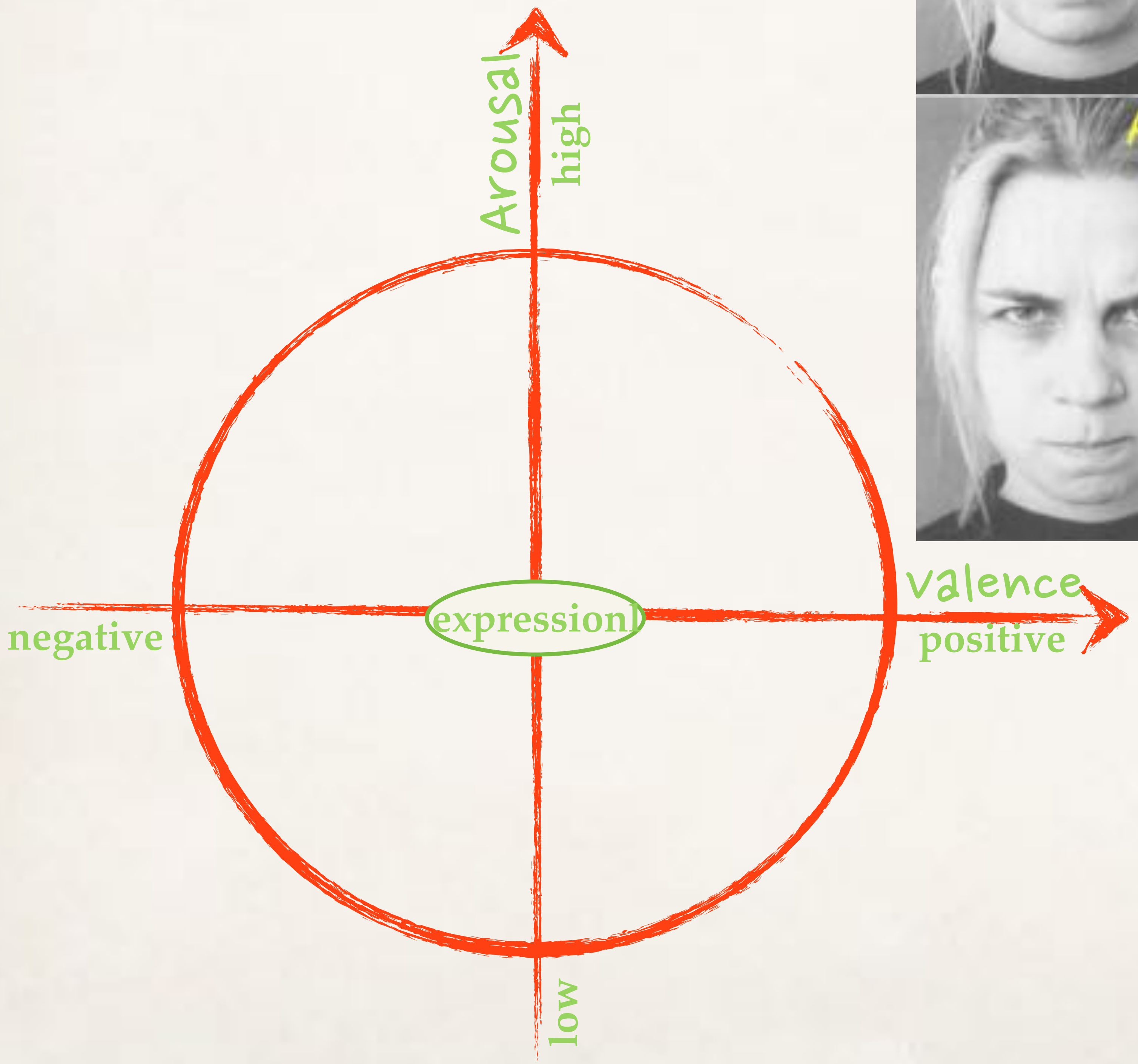
Comparison Methods:
RVM all – RVM using all patches with equal weights
RVM best – RVM using single best patch
RVM sep – RVM using all patches as separate basis functions
SMKL – Simple Multiple Kernel Learning (A. Rakotomamonjy et al., JMLR'08)
mRVM – multi-kernel RVM (T. Damoulas et al., ICMLA'08)
SR+SVM – Spectral Regression combined with SVM (S. Mavadati et. al., TAC'13)

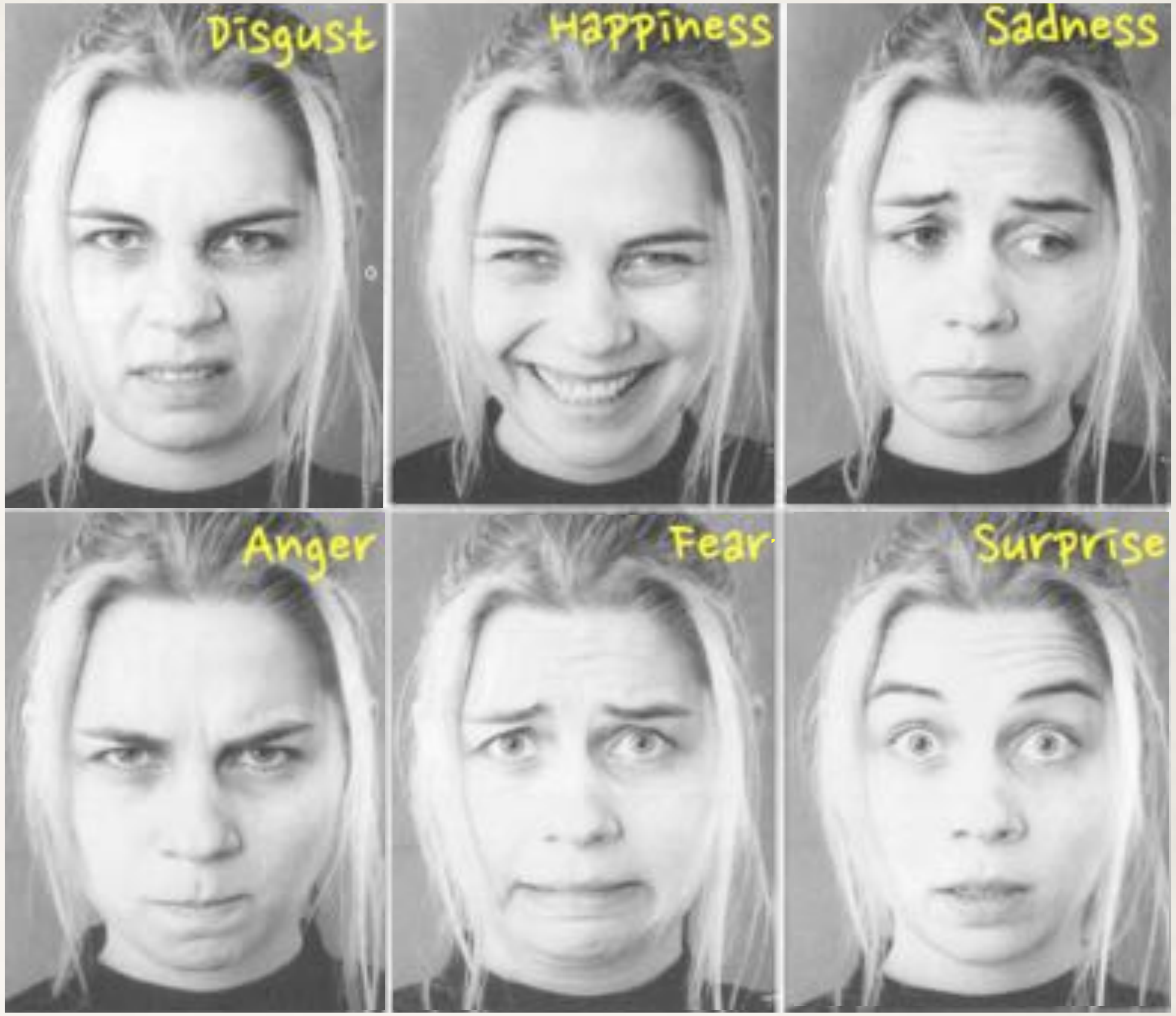
Doubly Sparse RVM for pain and AU intensity estimation
(Kaltwang, Todorovic & Pantic, PAMI 2016)



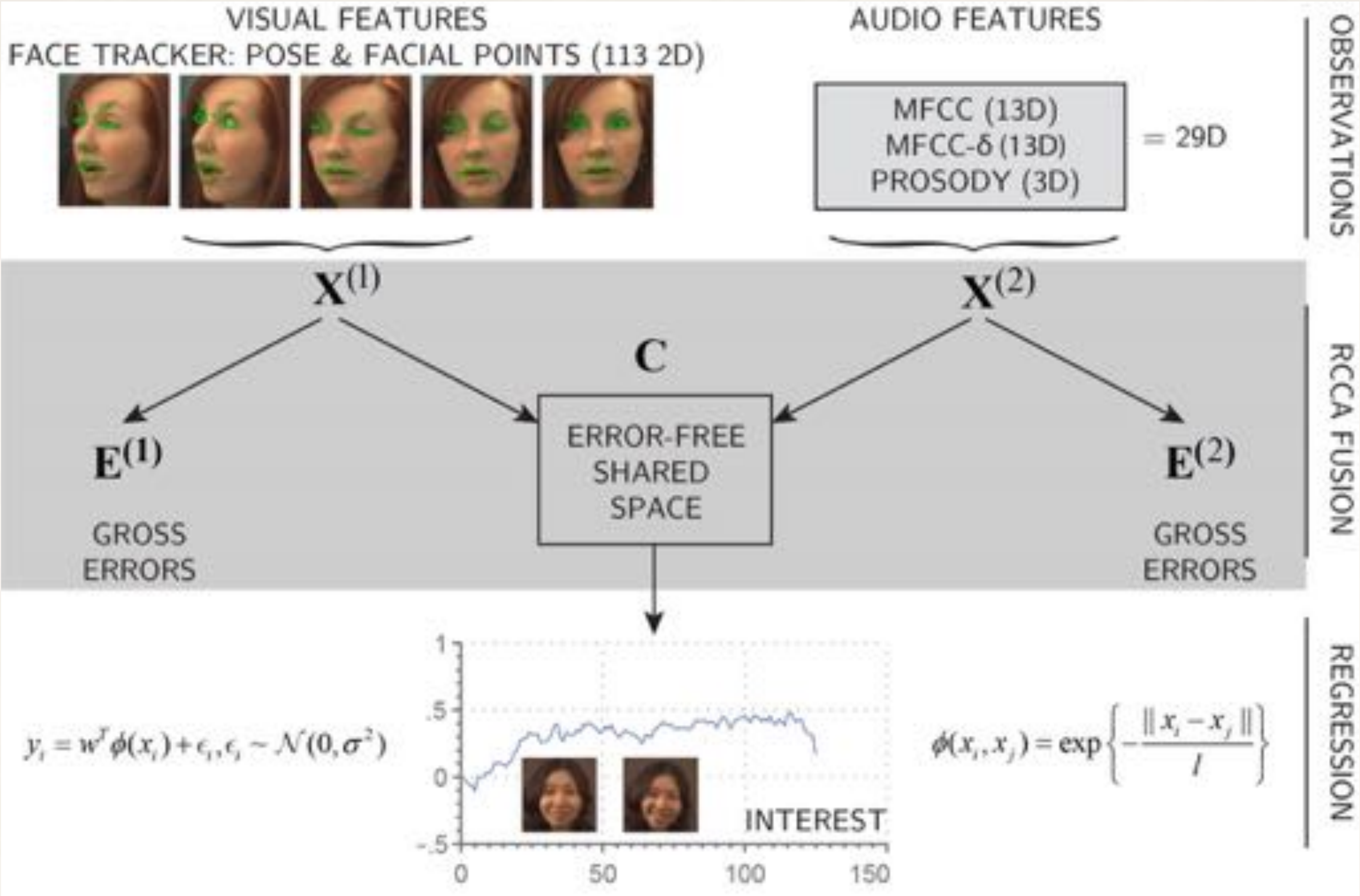
AU intensity
Frame number
AU4
AU6
AU12
AU26

Facial Expression Classification





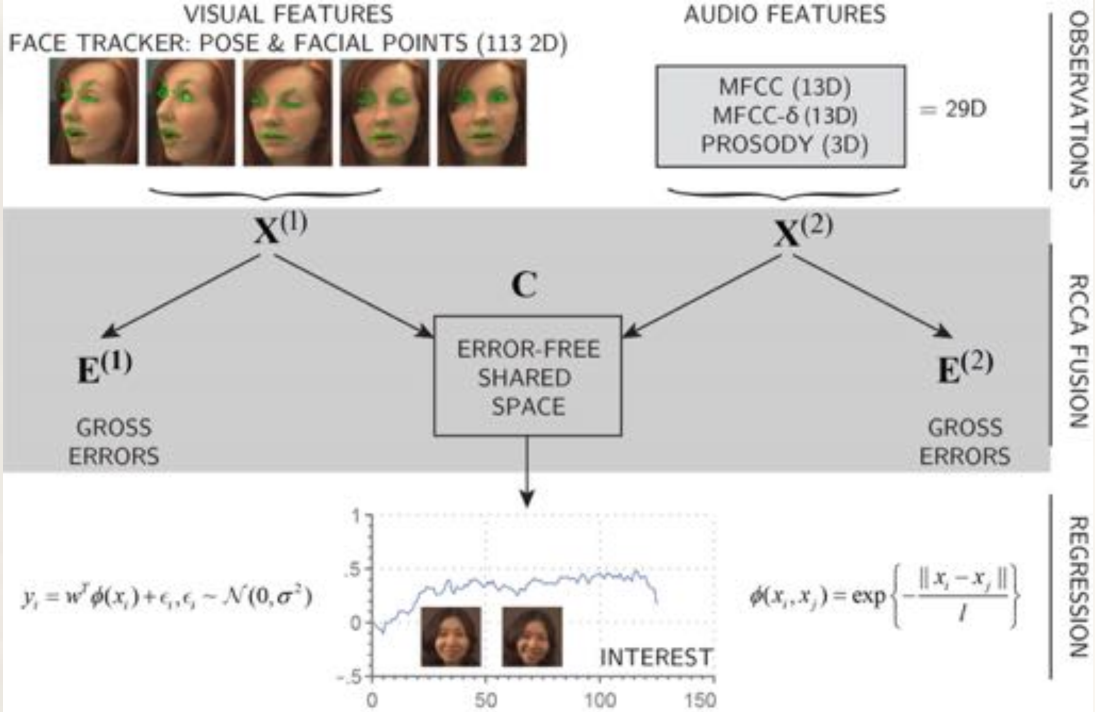
Facial Expression Classification

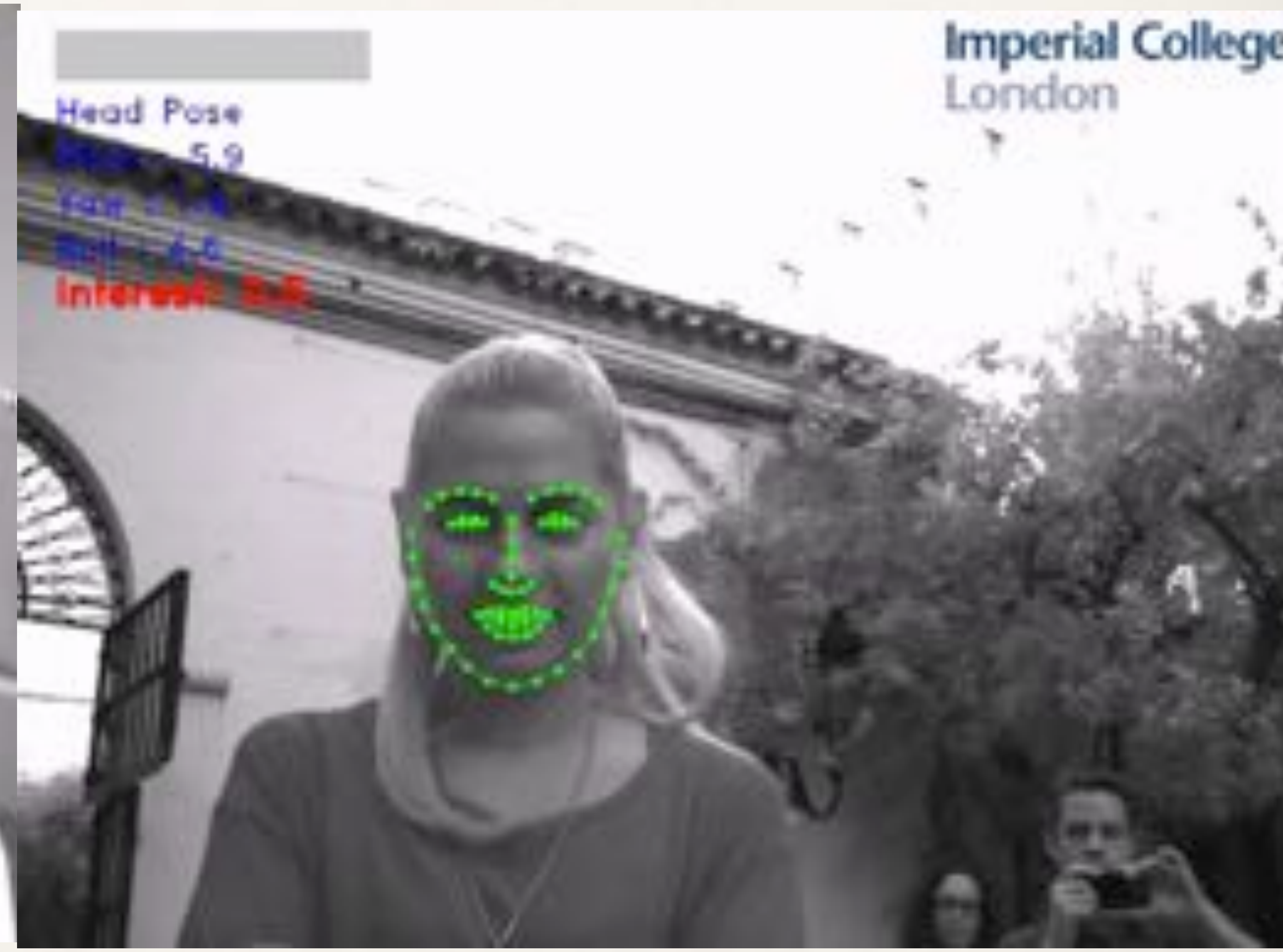


	Face	Audio	F_l	$RCICA_f$	$JIVE_f$	$COBE_f$	CCA_f	$LS\text{-}CCA_{\ell 1,f}$	$LS\text{-}CCA_{\ell 2,f}$
MSE	0.033	0.031	0.031	0.029	0.030	0.030	0.031	0.031	0.032
COR	0.432	0.460	0.443	**0.490**	0.460	0.463	0.458	0.480	0.464

Robust correlated & Individual component Analysis for V/A/I Prediction
(Panagakis / Nicolaou et al, CVPR'13, ICASSP'14, PAMI 2016)

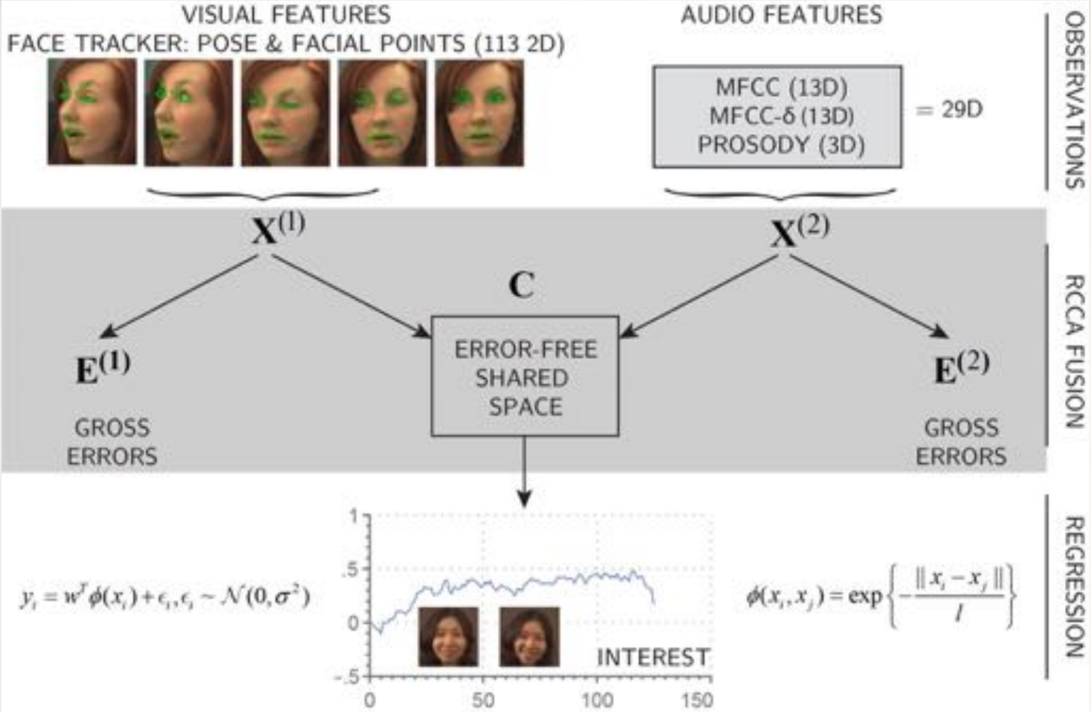
Facial Expression Classification



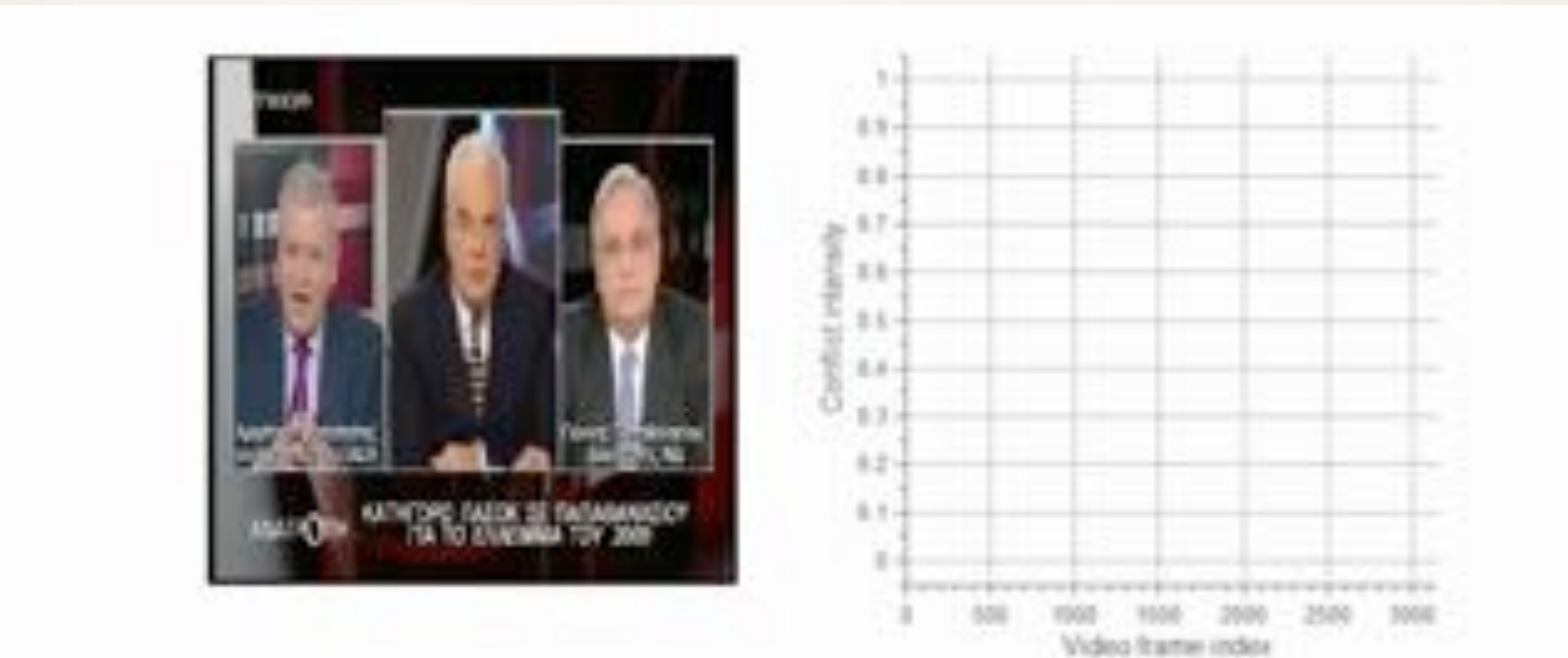


(RCICA - Panagakis / Nicolaou et al, CVPR'13, ICASSP'14, PAMI 2016)

Facial Expression Classification



CONFER



(RCICA - Panagakis / Nicolaou et al, CVPR'13, ICASSP'14, PAMI 2016)

Lessons Learned

Lessons Learned

1) work with data obtained in naturalistic (unconstrained) conditions!
2) context is the key to behaviour understanding!
3) work on expression intensity estimation!
4) stop working on single-person scenarios only!

MAHNOB Mimicry DB - (Bilakhia, Petridis & Pantic, PRL 2015) ibug.doc.ic.ac.uk