

Facial Action Recognition using **sparse appearance descriptors** and their **pyramid** representations

By Bihan Jiang , Michel Valstar, Maja Pantic
iBUG group, Imperial College London
29th Oct, 2011

What the Face Reveals?

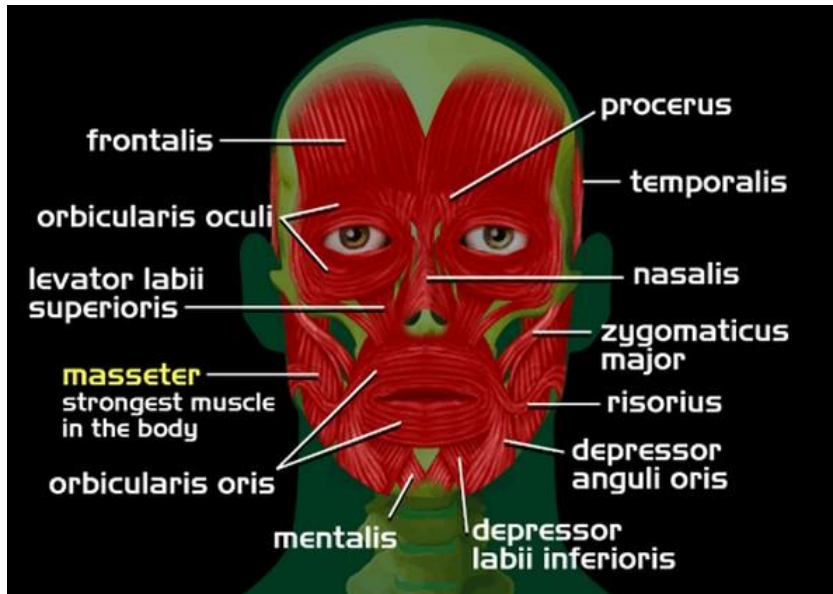


What is she thinking?

Is he lying?



Facial Action Coding System (FACS)

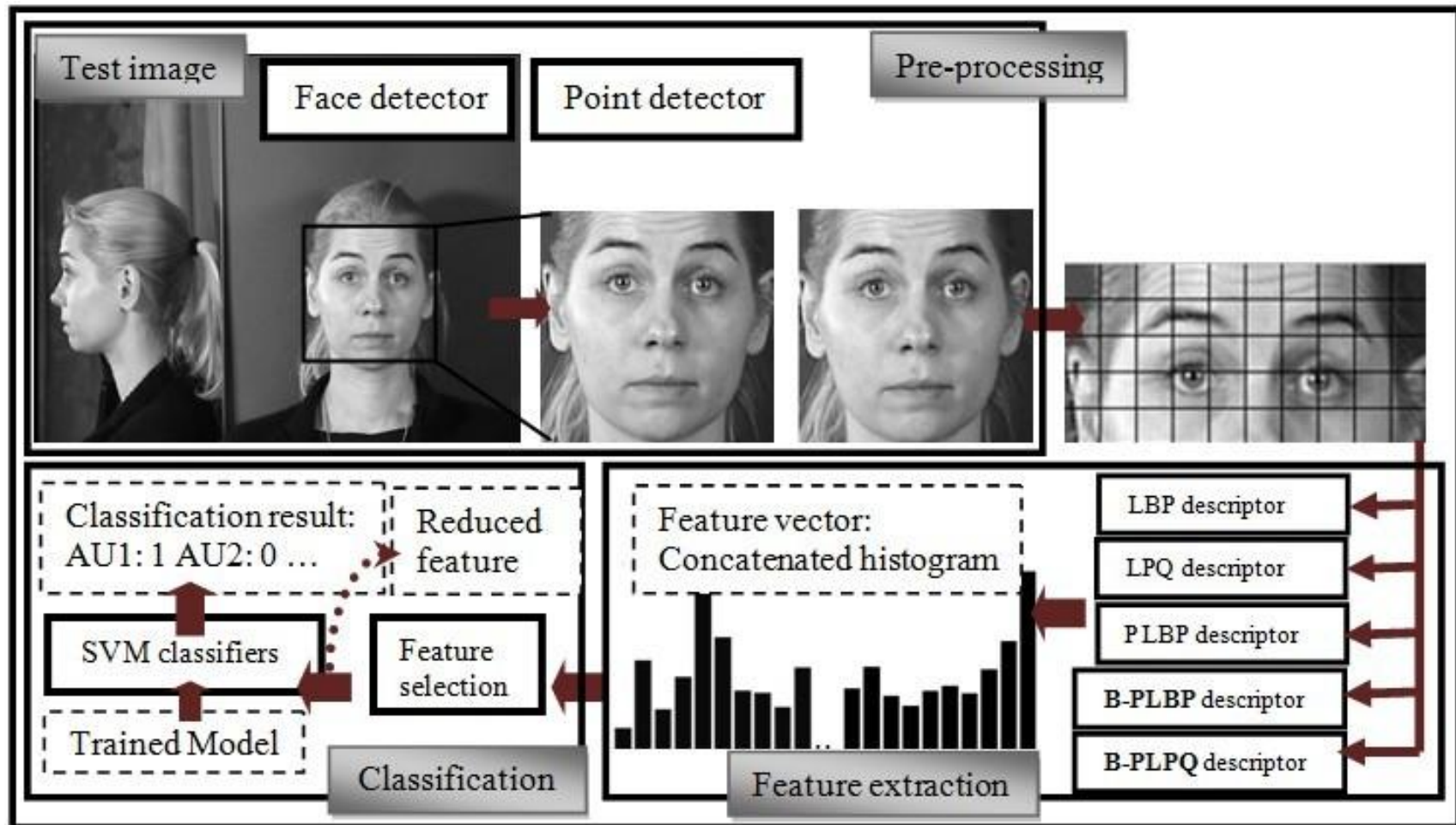


9 upper face AUs
18 lower face AUs
5 others

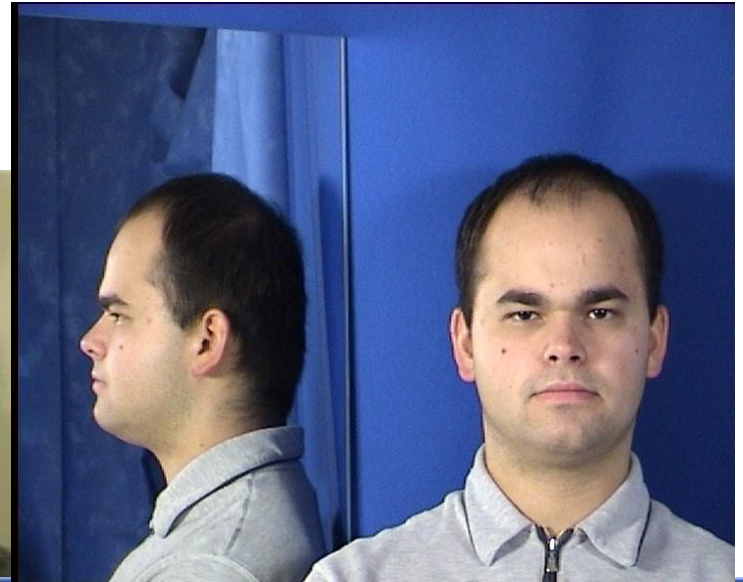
e.g. Happiness:
AU6 (cheek raiser) + AU12 (lip corner puller) + AU25 (lip part)





System outline



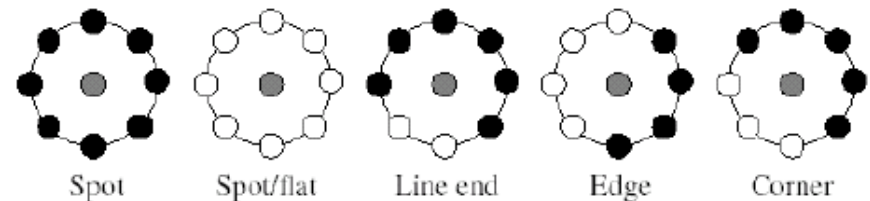
The MMI Database



Local Binary Pattern (LBP)

example	thresholded	weights
6 5 2	1 0 0	1 2 4
7 6 1	1  0	128  8
9 8 7	1 1 1	64 32 16

Pattern = **11110001**
LBP = 1 + 16 + 32 + 64 + 128 = **241**



For the pixel (x_c, y_c) , then the coordinates of his P neighbours with radius R are respectively:

$$\begin{cases} x_p = R \times \cos(2\pi p/P) + x_c \\ y_p = R \times \sin(2\pi p/P) + y_c \end{cases}$$

Where $p = 1 \dots P$

$$\text{LBP}_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) \times 2^p$$

$$H_i = \sum_{x,y} I(f(x,y) = i) \quad , i = 0, \dots, n - 1$$

where n is the possible labels produced by LBP operator and

$$I(A) = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{otherwise} \end{cases}$$

Local Phase Quantisation (LPQ)

The local phase information is extracted using a short-term Fourier transform (STFT) computed over a rectangular M -by- M neighbourhood N_x at each pixel position x of the image $f(x)$ defined by

$$F(\mathbf{u}, \mathbf{x}) = \sum_{\mathbf{y} \in N_x} f(\mathbf{x}) \mathbf{w}_R(\mathbf{y} - \mathbf{x}) e^{-j2\pi \mathbf{u}^T \mathbf{y}} = \mathbf{w}_u^T \mathbf{f}_x$$

Frequency points: $u_1 = [a, 0]^T$, $u_2 = [0, a]^T$, $u_3 = [a, a]^T$, and $u_4 = [a, -a]^T$,

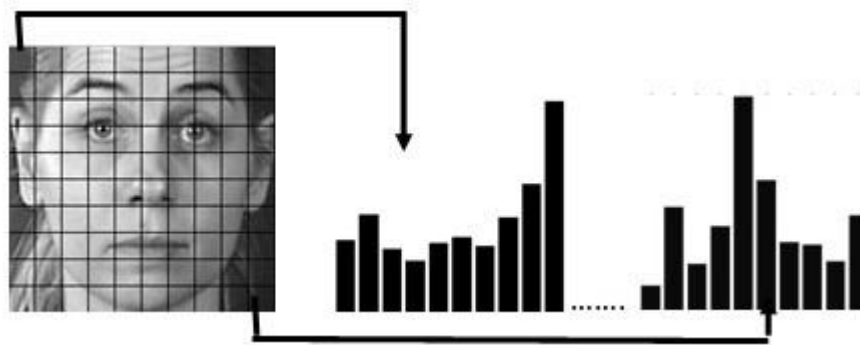
$$F_x = [F(u_1, x), F(u_2, x), F(u_3, x), F(u_4, x)]$$

$$q_j = \begin{cases} 1, & \text{if } g_j \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

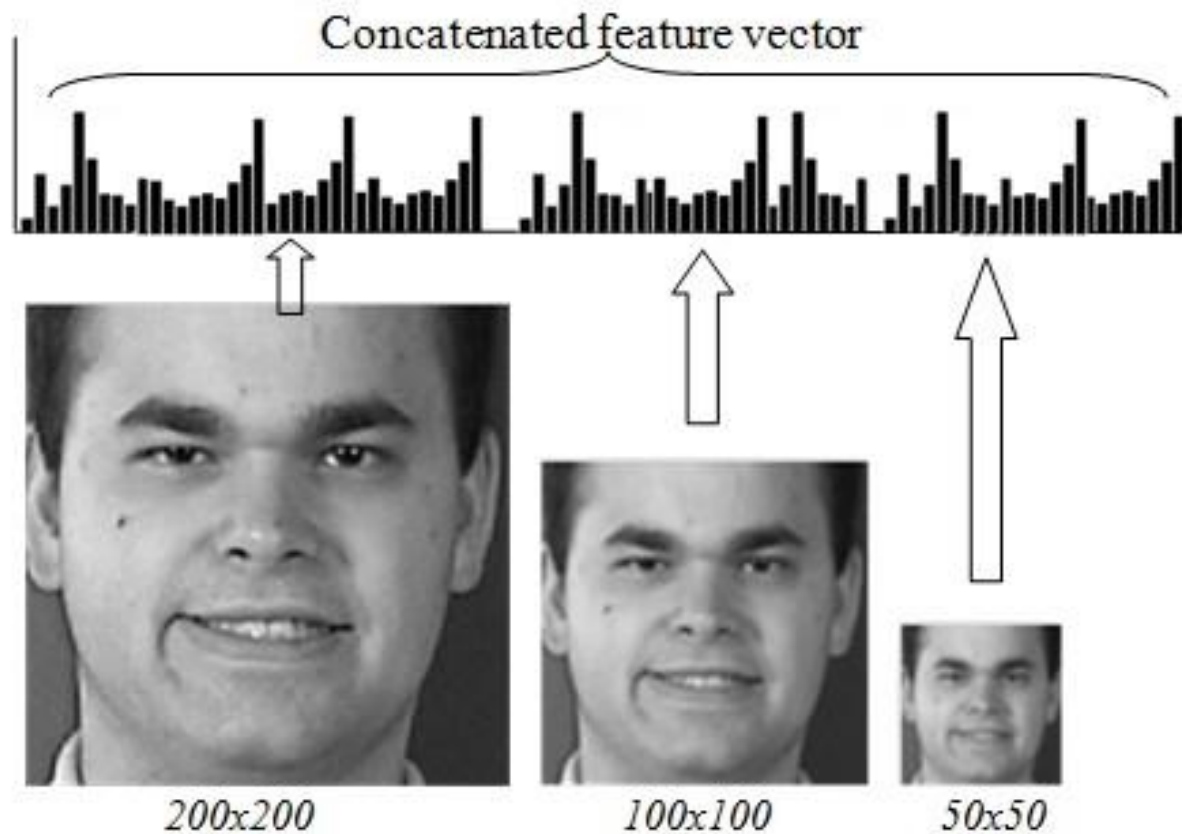
where $g_j(x)$ is the j th component of the vector $Gx = [\text{Re}\{F_x\}, \text{Im}\{F_x\}]$.

$$f_{LPQ}(x) = \sum_{j=1}^8 q_j 2^{j-1}$$

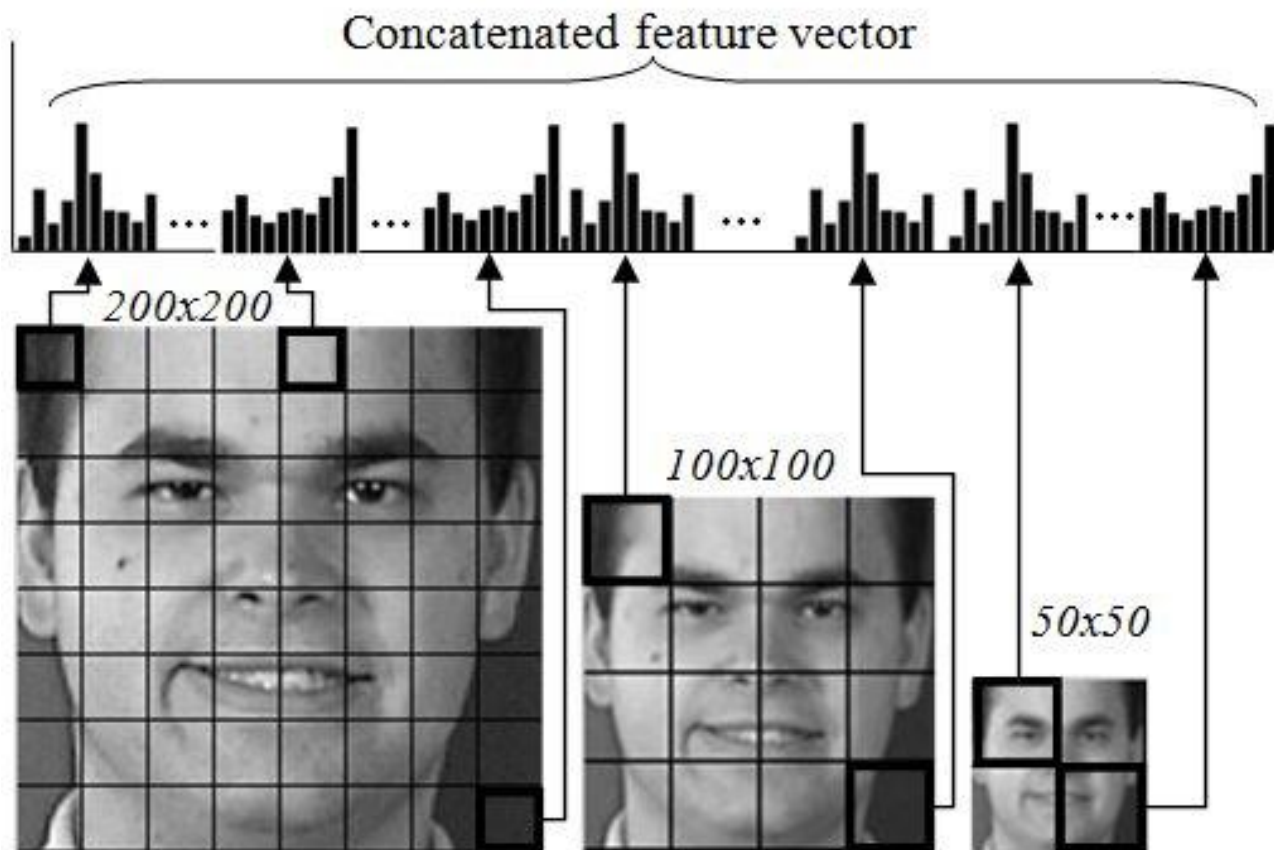
Pyramid representations



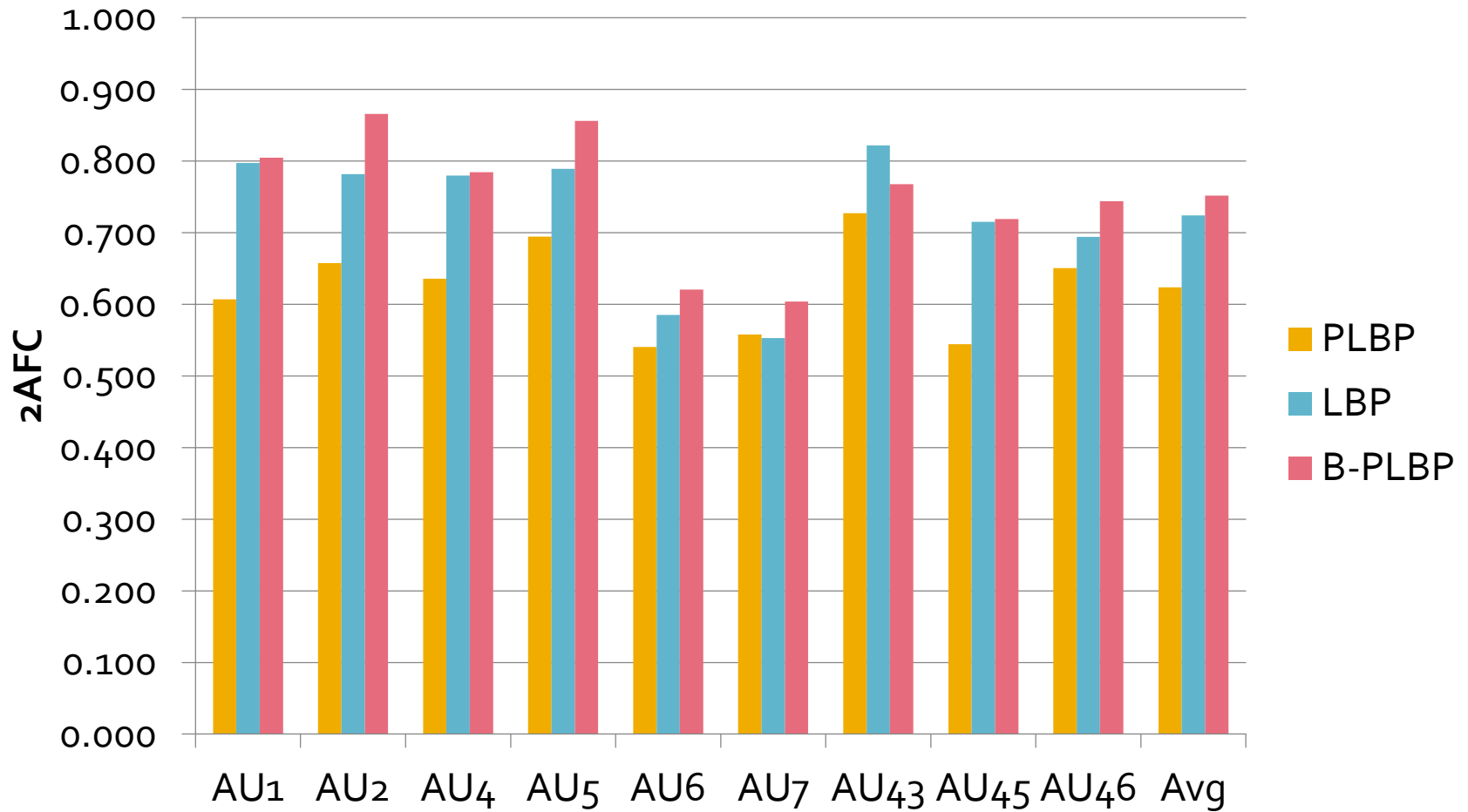
Pyramid representations



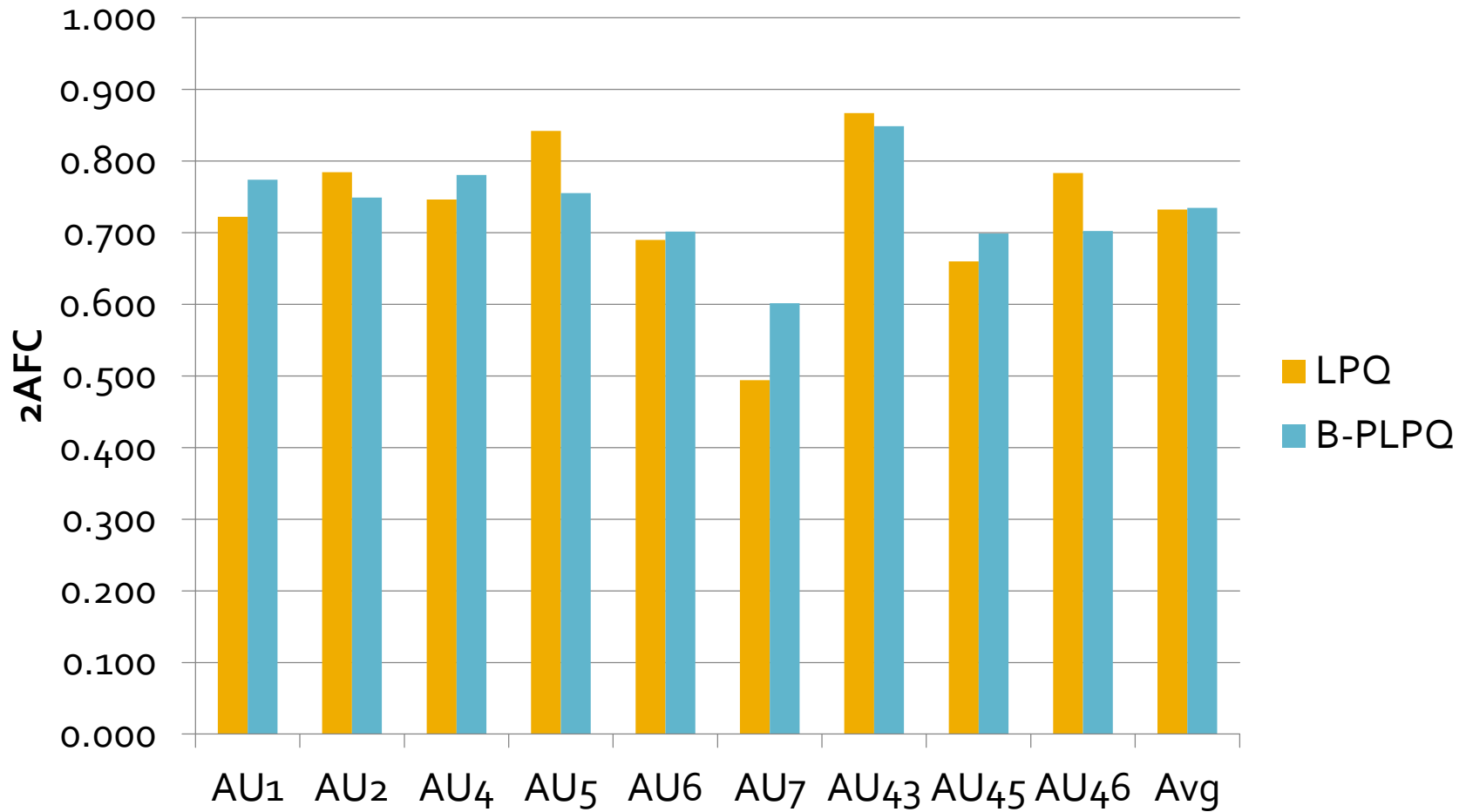
Pyramid representations



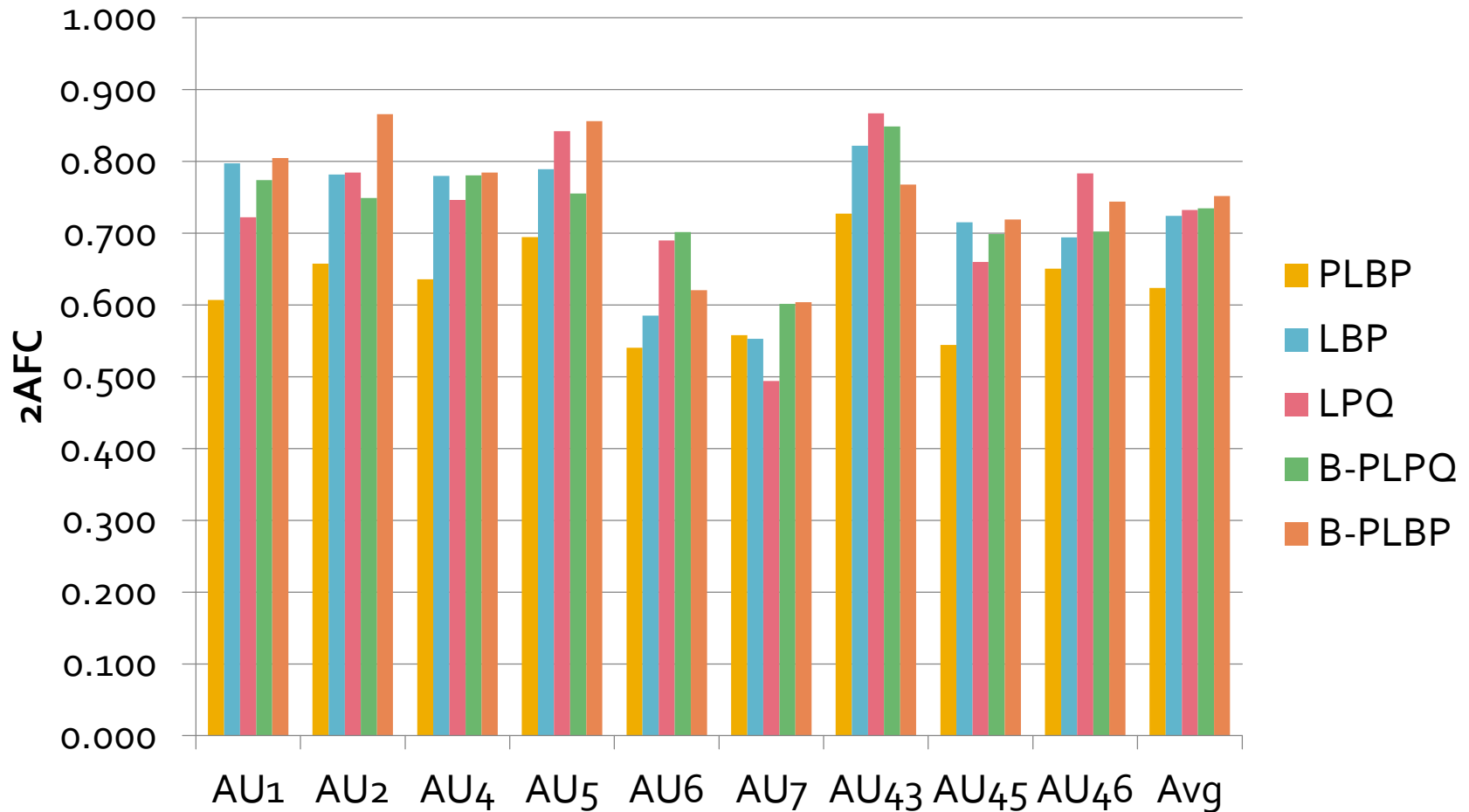
Results analysis



Results analysis



Results analysis



References

- [1]. T. Ahonen, A. Hadid, and M. Pietikainen. Face recognition with local binary patterns. In European Conference on Computer Vision, pages 469{481, 2004.
- [2]. V. Ojansivu and J. Heikkila, "*Blur insensitive texture classification using local phase quantization*", In Proc. Int. Conf. on Image and Signal Processing (ICISP'08), pages 236–243, 2008.
- [3]. B. Jiang, M. Valstar, and M. Pantic. Action unit detection using sparse appearance descriptors in space-time video volumes. In Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition (FG'11), Santa Barbara, CA, USA, March 2011.
- [4]. X. Qian, X. Hua, P. Chen, and L. Ke. Plbp: An effective local binary patterns texture descriptor with pyramid representation. *Semi-Supervised Learning for Visual Content Analysis and Understanding*, 44(10):2502{2515, 2011.

Questions?