Face Spoofing Detection through Partial Least Squares and Low-Level Descriptors

William Robson Schwartz, Anderson Rocha, Helio Pedrini
{schwartz, anderson.rocha, helio}@ic.unicamp.br
Institute of Computing, University of Campinas

Introduction

Problem: 2-D image-based facial verification or recognition system can be spoofed with no difficulty (a person displays a photo of an authorized subject either printed on a piece paper)

Idea: anti-spoofing solution based on a holistic representation of the face region through a robust set of low-level feature descriptors, exploiting spatial and temporal information

Advantages: PLS allows to use multiple features and avoids the necessity of choosing before-hand a smaller set of features that may not be suitable for the problem

Partial Least Squares

➢ PLS deals with *a large number of variables and a small number of examples*

➢ Data matrix X and response matrix Y

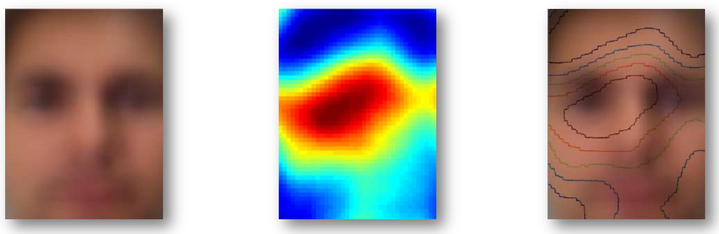
$$X_{n\times N} = T\,P^T + E$$
$$Y_{n\times M} = U\,Q^T + F$$

Scores, Loadings, Residuals

➢ Practical Solution: NIPALS algorithm

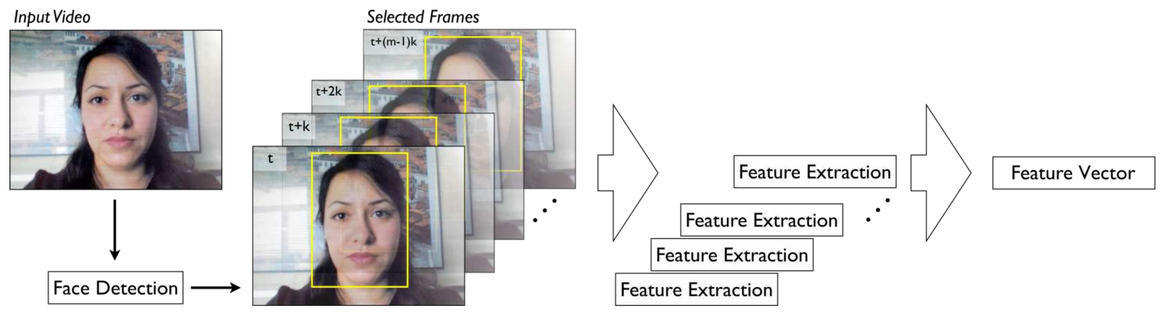
Iterative approach to calculate PLS factors

➢ PLS weights the feature descriptors and estimates the location of the most discriminative regions



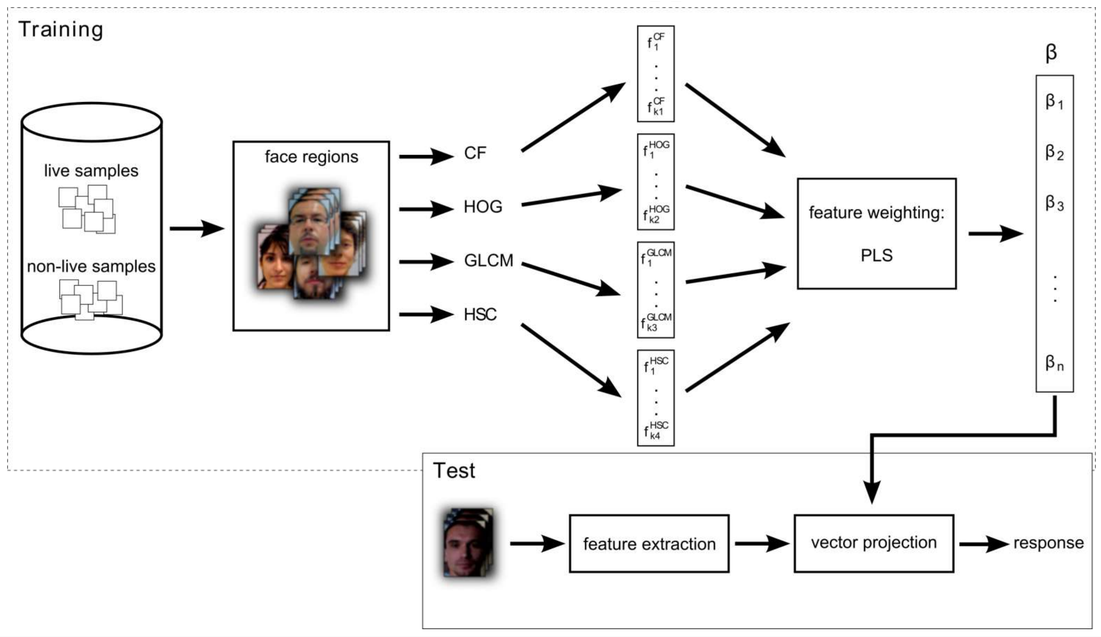
Anti-Spoofing Proposed Solution

➢ A video sample is divided into *m* parts, feature extraction is applied for every *k*-th frame. The resulting descriptors are concatenated to compose the feature vector



➢ PLS is employed to obtain the latent feature space, in which higher weights are attributed to feature descriptors extracted from regions containing discriminatory characteristics between the two classes

➢ The test procedure evaluates if a novel sample belongs either to the live or non-live class. When a sample video is presented to the system, the face is detected and the frames are cropped and rescaled



Experimental Results

Print-Attack Dataset

➢ *Dataset*: 200 real-access and 200 printed-photo attack videos [1]

➢ *Setup*: face detection, rescale to 110 x 40 pixels, 10 frames are sampled for feature extraction (HOG, intensity, color frequency (CF) [2], histogram of shearlet coefficients (HSC) [3], GLCM)

➢ *Classifier evaluation*: SVM type C with linear kernel achieved EER of 10%. PLS method achieved EER of 1.67%

Name	# descriptors	EER (%)
HOG	326,880	11.67
Intensity	154,000	8.33
CF	27,240	6.67
GLCM	159,360	6.67
HSC	581,120	4.33
Combination	1,094,600	1.67

Feature combination

Team	FAR (%)	FRR (%)
IDIAP	0.00	0.00
UOULU	0.00	0.00
AMILAB	0.00	1.25
CASIA	0.00	0.00
SIANI	0.00	21.25
Our results	1.25	0.00

Comparisons

NUAA Dataset

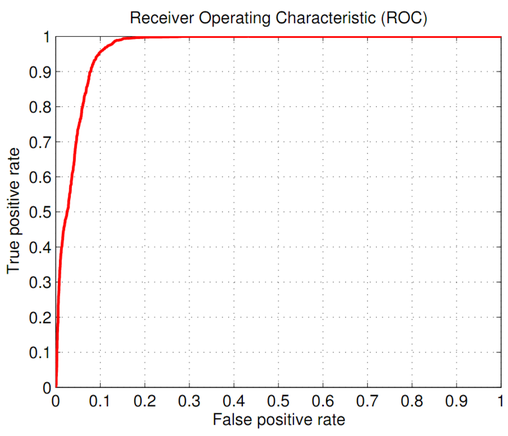
➢ *Dataset*: 1743 live images and 1748 non-live images for training. 3362 live and 5761 non-live images for testing [4]

➢ *Setup*: faces are detected and images are scaled to 64 x 64 pixels

➢ *Comparison*: Tan et al. [4] achieved AUC of 0.95

Name	# descriptors	EER (%)	AUC
Intensity	4,096	52.20	0.425
HOG	6,984	16.80	0.908
HSC	12,416	12.40	0.944
GLCM	3,552	9.60	0.960
Combination	22,952	8.20	0.966

Feature combination



[1] https://www.idiap.ch/dataset/printattack
[2] W. R. Schwartz, A. Kembhavi, D. Harwood, and L. S. Davis. *Human Detection Using Partial Least Squares Analysis*. In IEEE ICCV, pages 24–31, 2009.
[3] W. R. Schwartz, R. D. da Silva, and H. Pedrini. *A Novel Feature Descriptor Based on the Shearlet Transform*. In IEEE ICIP, 2011.
[4] X. Tan, Y. Li, J. Liu, and L. Jiang. *Face liveness detection from a single image with sparse low rank bilinear discriminative model*. In ECCV, pages 504–517, 2010.