FACE RECOGNITION IN REAL WORLD IMAGES

Xavier Fontaine, Radhakrishna Achanta and Sabine Süsstrunk

School of Computer and Communication Sciences
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

ivrl.epfl.ch
{xavier.fontaine, radhakrishna.achanta, sabine.susstrunk}@epfl.ch

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Introduction

We design a computationally-efficient method for face recognition adapted to **real-world conditions** that can be trained using **very few training examples**. This is done by performing a novel **alignment process** followed by classification using sparse representation techniques.

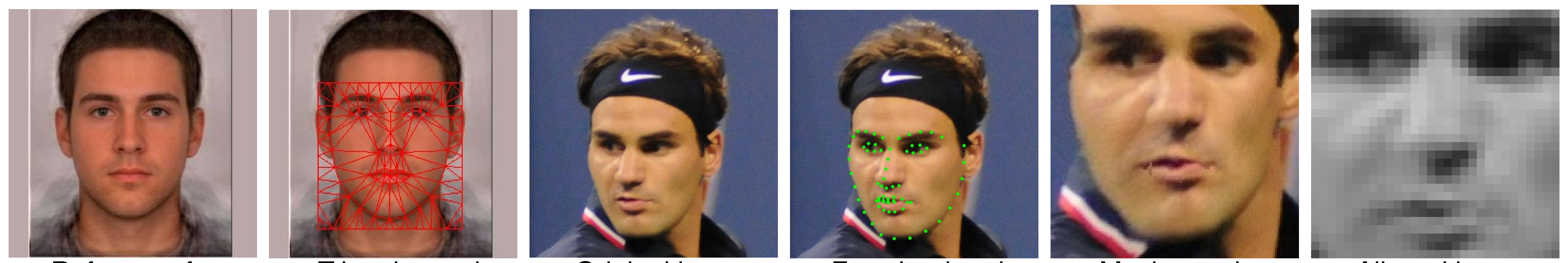
Mesh Warping and Global Alignment

► We achieve a global alignment of all images as follows:

- **Fig. 1 (a)** Choose a reference face with full frontal views, no occlusions and no strong shadows.
- **Fig. 1 (b)** Detect the face and its **landmarks** then compute the **Delaunay triangulation mesh** of the reference face image, by adding equally-spaced points on the border of the face square.
- **Fig. 1 (c)** Detect the face on a test image.
- **Fig. 1 (d)** Detect the **facial landmarks** on the test image.
- **Fig. 1 (e)** Use the reference triangulation with the landmarks of the test face to warp each triangles of the test face to **map the corresponding ones on the reference triangulation** with affine transformations.
- **Fig. 1 (f)** Rotate, scale, translate and crop the warped image to fix the eye positions and the inter-eye and eye-chin distances. Finally convert the image to gray-scale.

Main Result

The performances of a Face Recognition system can be considerably improved using **preprocessing** techniques to **align the images with each other**. Alignment can be done in **nearly real-time** by using **mesh-warping techniques**. A **triangulation mesh** is extracted based on **facial landmarks** for a reference face. Then all images are **deformed** to match their features with the corresponding features of the reference image. Finally recognition is performed via a **modified RSC algorithm**.



(a) Reference face (b) Triangle mesh (c) Original image (d) Face landmarks (e) Mesh warping (f) Aligned image

Figure 1: *The steps of alignment using our method on an image of Roger Federer*

Robust Sparse Coding Algorithm [1]

- Training images: dictionary $D = [r_1; \dots; r_n]$.
- Assume iid coding residuals $e = y - D\alpha = [e_1; \dots; e_n]$ (pdf f_θ).
- **Goal:** maximize the likelihood $L_\theta(e) = \Pi_{i=1}^{n} f_\theta(e_i)$.
- Add a **sparsity constraint** on α.
- Choose f_θ based on a logistic function.

► **Weighted-LASSO problem:**

$$\min_{\alpha} \left\| W^{1/2}(y - D\alpha) \right\|_2^2 \text{ s.t. } \|\alpha\|_1 \leqslant \epsilon$$

► **We use the ℓ^2-norm instead of ℓ^1-norm**

$$\min_{\alpha} \left\| W^{1/2}(D\alpha - y) \right\|_2^2 + \lambda \|\alpha\|_2^2$$

and obtain $\alpha = (D^T W D + \lambda I)^{-1} D^T W y$.

This leads to a **4-fold speedup** without significant loss of accuracy.

Experiments

► ***General Pipeline:***

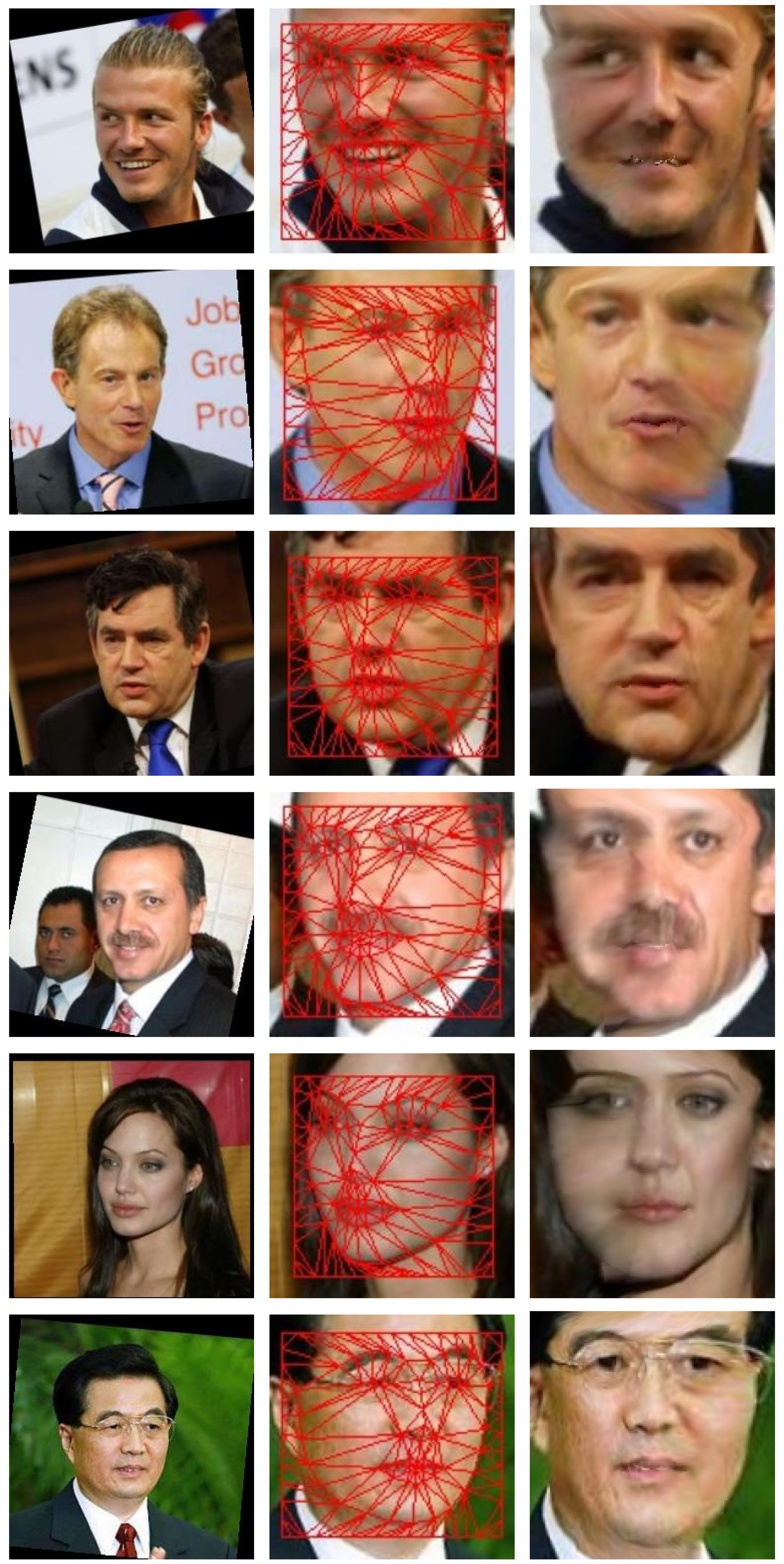
- Subject every image to the **same alignment** process: **mesh warping** ("frontalization") and global alignment.
- Apply the **modified** RSC algorithm.

Table 1 shows the results of applying the modified RSC algorithm on 3 kinds of images corresponding to the experiments:

- **Exp 1**: Use raw LFWa images.
- **Exp 2**: Use the faces detected by Viola-Jones face detector.
- **Exp 3**: Perform recognition on the faces **preprocessed by our alignment algorithm**.

Note the **considerable performance gain** due to **our preprocessing algorithm**.

Examples



(a) Original image (b) Triangulation (c) Warped

Figure 2: *Mesh warping examples on LFWa dataset [2] pictures of David Beckham, Tony Blair, Gordon Brown, Recep Tayyip Erdogan, Angelina Jolie and Hu Jintao*

Results

	Exp. 1	Exp. 2	Exp. 3
Recognition rate (%)	19.6	28.8	**76.4**
Time for one image (s)	3.2	3.0	**1.6**

Table 1: *Recognition rates on the LFWa database with 7 training images and 3 test images*

Method	2	5
SRC [3]	24.4 ± 2.4%	44.1 ± 2.6%
CRC [4]	27.4 ± 2.1 %	42.0 ± 3.2 %
MSPCRC [5]	35.0 ± 1.6%	41.1 ± 2.8%
Ours	**51.1** ± 2.9%	**74.2** ± 2.5%
Our time	*0.15 s*	*0.85 s*

Table 2: *Recognition rates on the LFWa dataset with 2 and 5 training samples per person*

Method	Rate
NN	10.6%
SRC [3]	22.3%
SVDL [6]	30.2%
Ours	**33.3** ± 3.4%
Our time	*0.02 s*

Table 3: *Recognition rates on the LFWa dataset with a single training sample per person*

References

[1] Meng Yang, Jian Yang, and David Zhang. Robust Sparse Coding for Face Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 625–632, June 2011.

[2] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.

[3] John Wright, Allen Y. Yang, Arvind Ganesh, S. Shankar Sastry, and Yi Ma. Robust Face Recognition via Sparse Representation. *IEEE Transactions on pattern analysis and machine intelligence*, 31(2):210–227, February 2009.

[4] Lei Zhang, Meng Yang, and Xiangchu Feng. Sparse representation or collaborative representation: Which helps face recognition? In *2011 International Conference on Computer Vision*, pages 471–478. IEEE, 2011.

[5] Pengfei Zhu, Lei Zhang, Qinghua Hu, and Simon C. K. Shiu. Multi-scale patch based collaborative representation for face recognition with margin distribution optimization. *European Conference on Computer Vision*, 2012.

[6] Meng Yang, Luc Van Gool, and Lei Zhang. Sparse variation dictionary learning for face recognition with a single training sample per person. In *The IEEE International Conference on Computer Vision*, December 2013.