Image Statistics Based on Diffeomorphic Matching

Guillaume CHARPIAT [1], Olivier FAUGERAS [2], Renaud KERIVEN [3]

[1] Projet Odyssée, École Normale Supérieure, France

[2] Projet Odyssée, INRIA Sophia Antipolis, France

[3] Projet Odyssée, École Nationale des Ponts et Chaussées, France

Introduction

- **Goal:** non-supervised object recognition, image classification
  $\hookrightarrow$ automatic description of shape and colour variations ?
- **Method:** compute a ***mean*** object and ***characteristic modes of variation***
- **Basic Tool :** non-supervised diffeomorphic image matching
  $\hookrightarrow$ natural definition of mean and modes
- **Advantages:** non-supervised, non-specific to a particular dataset
- **Example:** Yale face image dataset
- **Classification Application:** face expression recognition task

Diffeomorphic Image Matching

- **Initial Image:** A
- **Target Image:** B
- **Method:** search for a diffeomorphism h such that $A \circ h \simeq B$

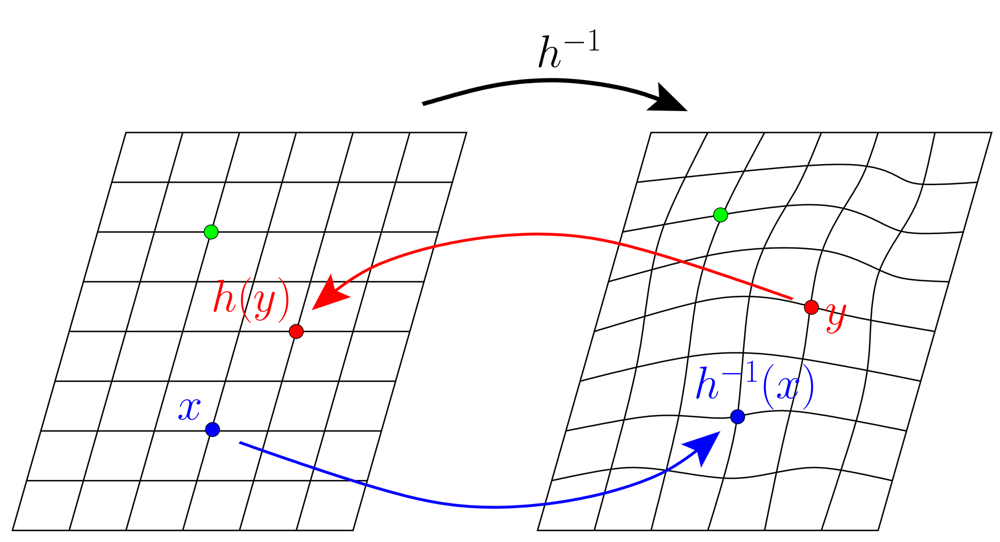


Image A Image $A \circ h$

- **Image Similarity Criterion:** Local Cross-Correlation (LCC)

$$LCC(A,B) = \int_{y\in\Omega} \frac{v_{AB}(y)^2}{v_A(y)\, v_B(y)}\, dy$$

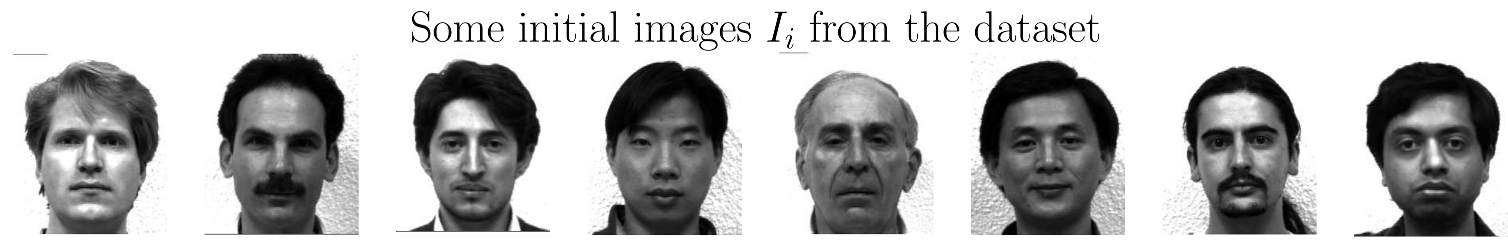
with $v_A(y)$ local variance of image A in a gaussian neighborhood of point y, and $v_{AB}(y)$ the local covariance

- **Smoothing Term:** for example $R(h) = \|h - \mathrm{Id}\|_{H^1}$
- **Image Matching:** multiscale gradient descent with respect to warping field h of

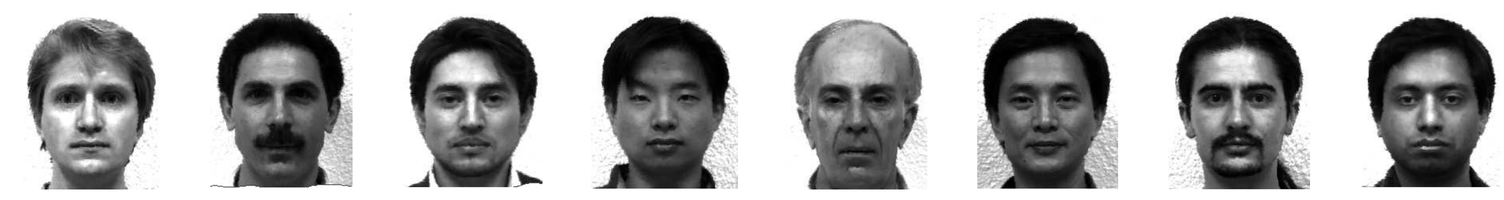
$$E(A,B,h) = R(h) - LCC(A \circ h, B)$$

Computation of the Mean Image

Some initial images I_i from the dataset



Together warped images $I_i \circ h_i$



Mean image M
(sum of the warped images)
$M = \frac{1}{n}\sum_i I_i \circ h_i$

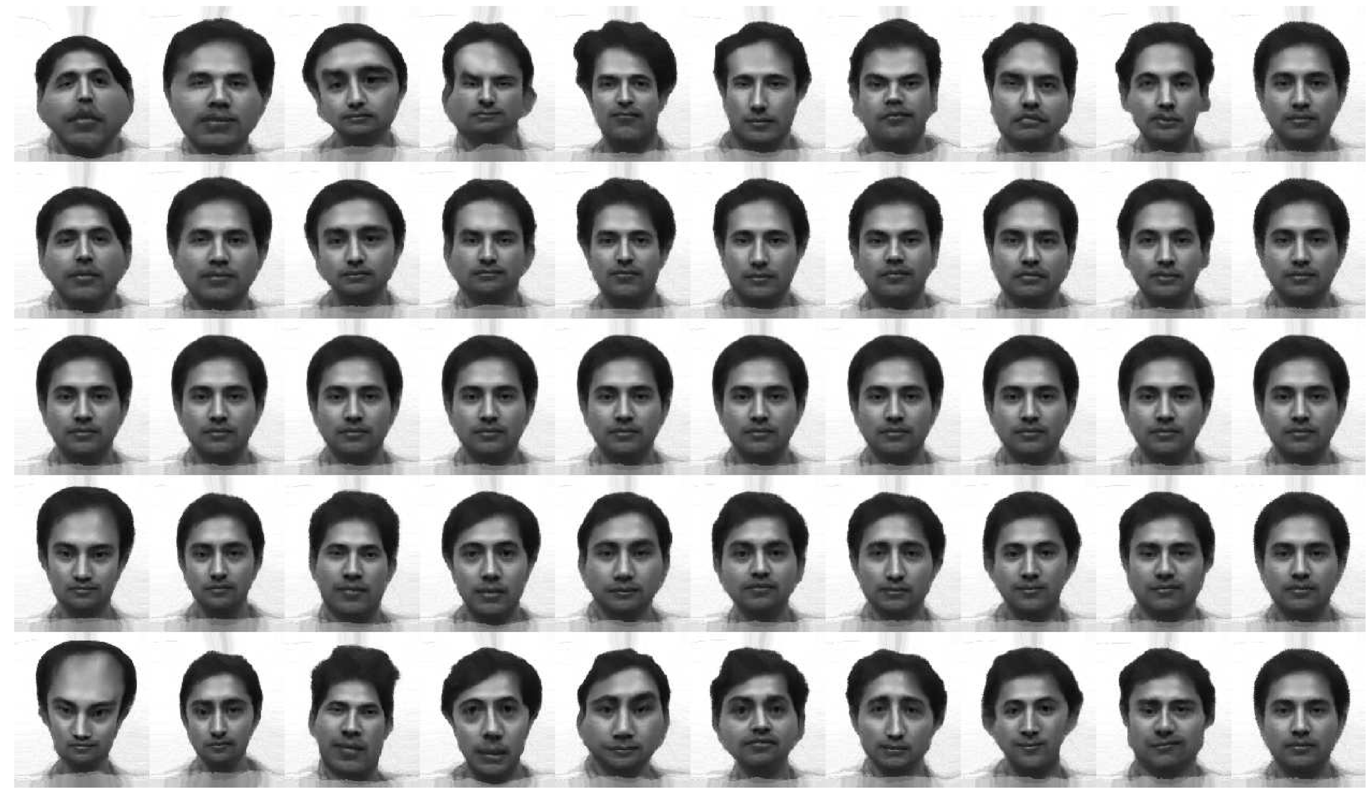
- **Data:** set of n images I_i
- **Goal:** computation of a mean image M
- **Problem:** gradient descent involving M intensity leads to local minima
  $\hookrightarrow$ creation of new contours instead of moving existing ones
- **Solution:** use diffeomorphisms
  $\hookrightarrow$ associate to each image I_i a diffeomorphism h_i
  $\hookrightarrow$ minimize with respect to the n fields h_i the similarity between all warped images

$$\sum_i R(h_i) \;-\; \frac{1}{n-1}\sum_{i\neq j} LCC(I_i \circ h_i, I_j \circ h_j)$$

- **Mean Image:** $M = \frac{1}{n}\sum_i I_i \circ h_i$

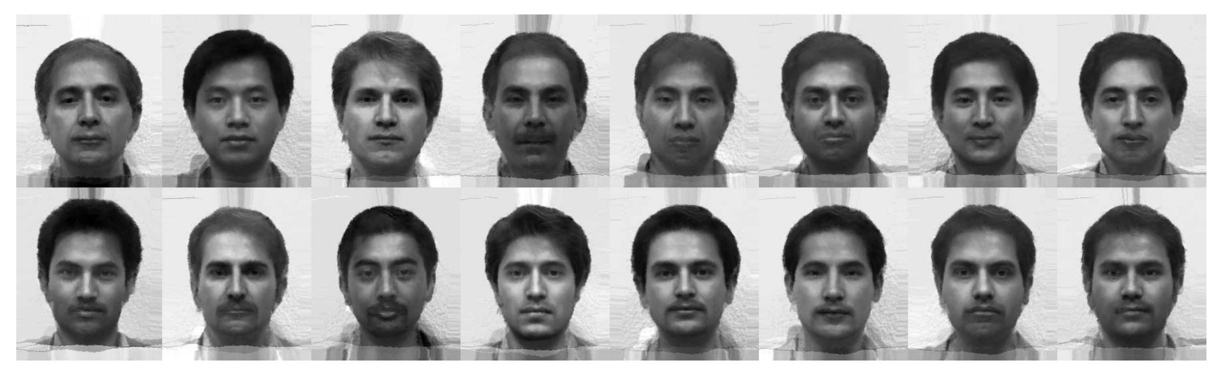
Modes of Shape Variation

- **Data:** n images I_i, n warping fields h_i
- **Method:** warping field statistics
- **Correlation between 2 fields:** $\langle h_i | h_j \rangle = \int_\Omega h_i(y) \cdot h_j(y)\, dy$
- **Modes:** diagonalization of correlation matrix, extraction of eigenmodes P_k and their eigenvalues σ_k
- **Example:** application of the modes to the mean image M with different amplitudes



Modes of Shape and Intensity Variation

- **Intensity Variations:** $v_i = I_i \circ h_i - M$
- **Intensity Variation Correlation:** $\langle v_i | v_j \rangle = \int_\Omega v_i(y) v_j(y)\, dy$
- **Combined Shape-Intensity Correlation:** $\langle h_i, v_i | h_j, v_j \rangle = \langle h_i | h_j \rangle + \langle v_i | v_j \rangle$
- **New Modes:** with 2 parts (deformation field and intensity modification field)



Classification Task: Face Expression Recognition

- **Goal:** associate, to any new face, its expression (among 5 given expressions)
- **Data:** a training set, a new face with and without expression
- **Method:** compare the warping field between the 2 new images to known labeled fields (after alignment to the mean face). Result: 53/65. Incorrect labels are marked.

- **Algorithm Comparison:** SVM on images leads to 17 errors instead of 12.
- **Without Normal Face:** use warping field from the mean face to the new face. Result: 41/65 (SVM on images: 38).