A Study of the Region Covariance Descriptor

Impact of Feature Selection and Image Transformations

Hayden Faulkner, Ergnoor Shehu, Zygmunt L. Szpak, Wojciech Chojnacki, Jules R. Tapamo, Anthony Dick and Anton van den Hengel

THE UNIVERSITY of ADELAIDE

INTRODUCTION

A modern computer vision pipeline for generic image classification and recognition consists of three broad conceptual steps:

- selecting suitable (region covariance descriptors)
- defining a measure of similarity between feature descriptor (distance between covariance matrices)
- learning a classification rule that uses the feature descriptors and corresponding similarity measure to determine what the image represents

MOTIVATION

- Region covariance descriptor has proven to be useful in numerous computer vision applications.
- The properties of the descriptor are not well understood or documented.

REGION COVARIANCE DESCRIPTOR

Ω	image
x	spatial coordinates of a pixel in image
R	rectangular region of interest in image
$\phi : \Omega \rightarrow \mathbb{R}^n$	mapping from pixels to length-n feature vectors
Λ_R	n-by-n covariance matrix

$$\Lambda_R = \frac{1}{|R|-1}\sum_{x \in R}(\phi(x)-\mu_R)(\phi(x)-\mu_R)^{\top}$$

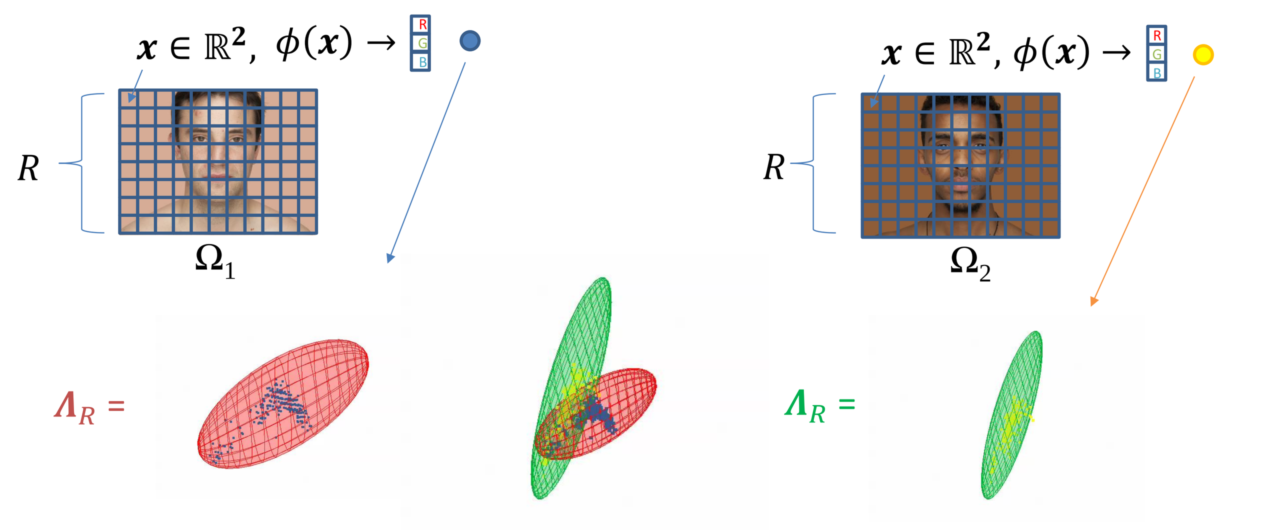
$\mu_R = \frac{1}{|R|}\sum_{x \in R}\phi(x)$ mean feature

$|R|$ number of pixels in R

FEATURE MAPPINGS

spatial x coordinate	magnitude of second-order partial derivative in horizontal direction
spatial y coordinate	magnitude of second-order partial derivative in vertical direction
red channel	magnitude of second-order mixed partial derivative
green channel	magnitude of edge response
blue channel	edge orientation
magnitude of first-order partial derivative in horizontal direction	luminance (LAB colour space)
magnitude of first-order partial derivative in vertical direction	a channel (LAB colour space)
	b channel (LAB colour space)

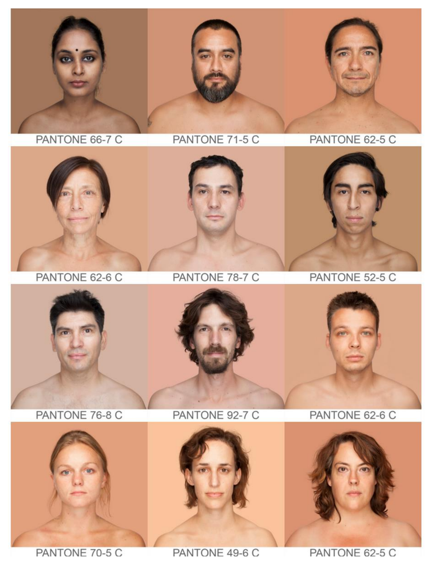
REGION COVARIANCE DESCRIPTOR EXAMPLE



$dist(\Omega_1, \Omega_2) \triangleq dist(\Lambda_R, \Lambda_R)$

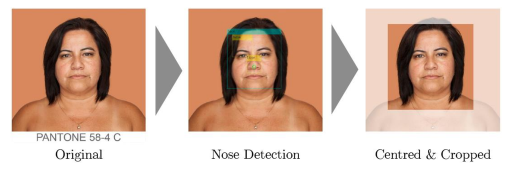
How do features and distance measures influence the similarity between two images?

Humanae © Angelica Dass



DATASET

- Diverse images of human faces 500 × 500 pixels
- Processing by centering all images on the nose and cropping to 319 × 319 pixels



TRANSFORMATIONS

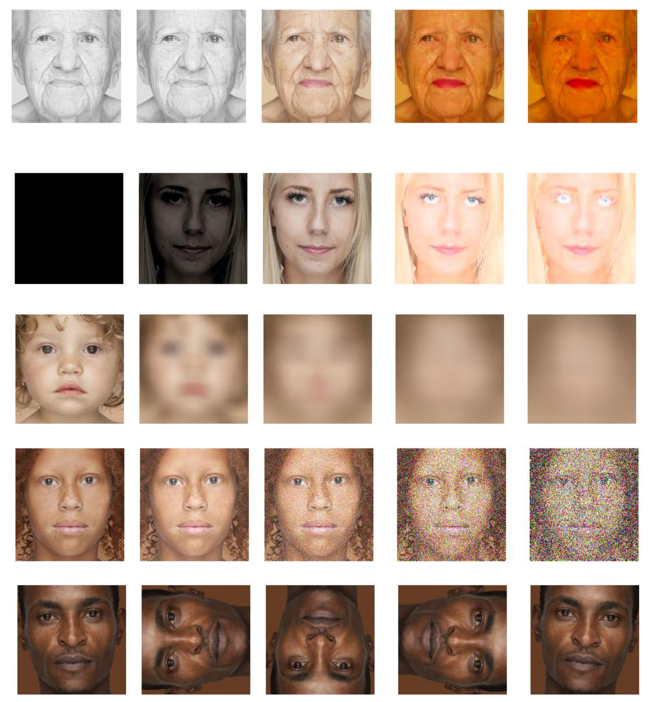
saturation

brightness

blur

noise

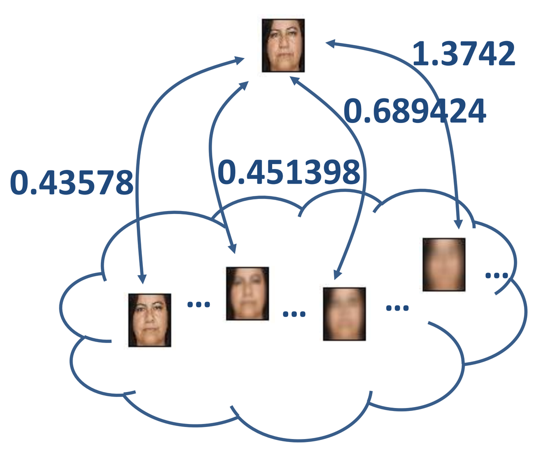
rotation

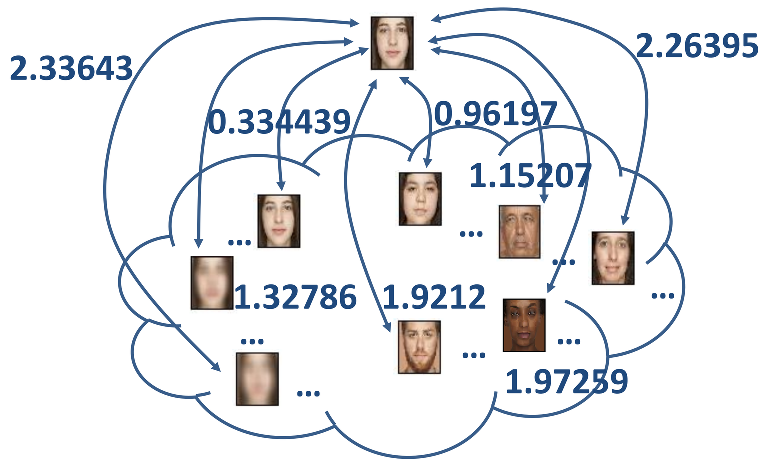


EXPERIMENTS

within: comparable set ≜ transformed base images

among: comparable set ≜ transformed base images + entire dataset





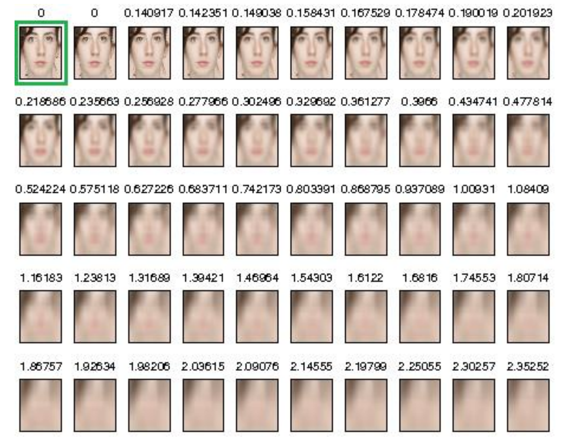
RESULTS

Features: x, y, r, g, b
Distance: Euclidean
Transform: Blur

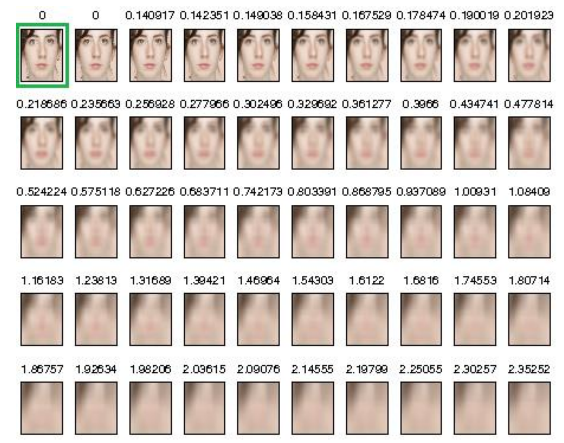
Features: Same
Distance: Same
Transform: Same

Features: x, y, r, g, b, l, a, b
Distance: Euclidean
Transform: Blur

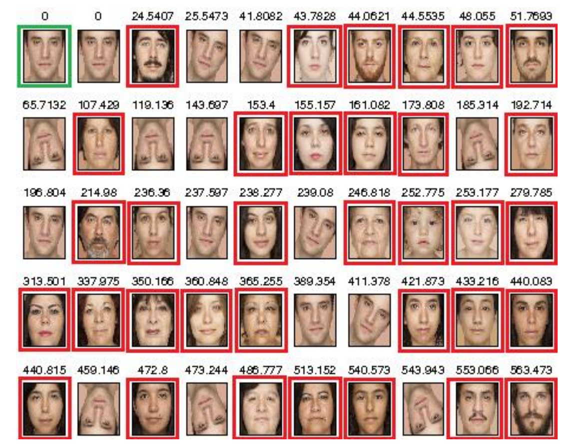
Features: x, y, r, g, b
Distance: Euclidean
Transform: Blur



Features: x, y, r, g, b
Distance: Euclidean
Transform: Blur

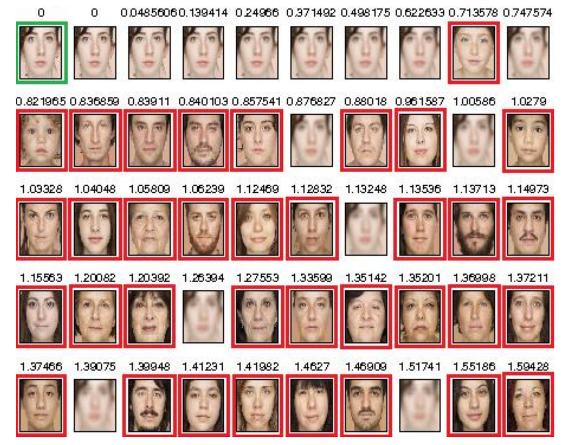


Features: $|I_{xx}|, |I_{yy}|, |I_{xy}|, \sqrt{I_x^2 + I_y^2}, tan^{-1}\left(\frac{|I_y|}{|I_x|}\right)$
Distance: Euclidean
Transform: Rotation



Features: x, y, r, g, b, $\sqrt{I_x^2 + I_y^2}, tan^{-1}\left(\frac{|I_y|}{|I_x|}\right)$, l, a, b
Distance: Same
Transform: Same

Features: Same
Distance: Log-Euclidean
Transform: Same



Features: Same
Distance: Same
Transform: Rotation

Features: Same
Distance: Affine-invariant
Transform: Same

DISCUSSION

- No distance measure works best in all situations.
- Inclusion or exclusion of a single feature can have a dramatic impact.
- Selection of features must be guided by extensive empirical analysis.
- Excellent retrieval performance observed for the $dist_E$ measure for Gaussian noise and blur transformations when the position feature (xy) was combined with a colour feature (rgb or lab).