DARTMOUTH

VOX CLAMANTIS IN DESERTO

1769

Facial Recognition

[*Computer Science 183 – Computer Vision*]

Project Members: Haider Syed & Travis Peters
Project Advisor: Lorenzo Torresani

Project Review...

- Face Recognition Problem
- Testing with the *Faces in the Wild*[1] dataset
- Primary objectives - address general problems for face recognition:
  - Complex background
  - Varying subject pose
  - Varying image scale
  - Varying illumination conditions

[1] Labeled Faces in the Wild database: http://vis-www.cs.umass.edu/lfw/

Proposed Approach…

- Gaussian Mixture Model
- GrabCut
- SIFT (descriptors only)
- kNN Classification

Current Implementation…

- Gaussian Mixture Model (segmentation)

$$p(x \mid \lambda) = \sum_{i=1}^{M} w_i \cdot g(x \mid \mu_i, \Sigma_i) \qquad g(x \mid \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{d}{2}} (\det(\Sigma_i))^{\frac{1}{2}}} \times \exp\left(-\frac{1}{2} (x-\mu_i)^T \Sigma_i^{-1} (x-\mu_i) \right)$$

where

$$\lambda = \{ w_i, \mu_i, \Sigma_i \} \qquad i = 1, \ldots, M$$

and

$$w_i = \frac{1}{T} \sum_{t=1}^{T} \Pr(i \mid x_t, \lambda) \qquad \mu_i = \sum_{t=1}^{T} \Pr(i \mid x_t, \lambda) \cdot x_t \Big/ \sum_{t=1}^{T} \Pr(i \mid x_t, \lambda) \qquad \Sigma_i = \left(\sum_{t=1}^{T} \Pr(i \mid x_t, \lambda) \cdot x_t^2 \Big/ \sum_{t=1}^{T} \Pr(i \mid x_t, \lambda) - \mu_i^2 \right)$$

  - Train skin model (skin/non-skin identification)
  - Expectation-Maximization used to estimate model parameters
  - Thresholding based on ratio of skin/non-skin probabilities
  - Erode noise, fill "holes" in skin, face = largest contiguous region in image.

- SIFT [1] (descriptors)
  - Using SIFT to extract features and create feature matrices for each image set

- kNN Classification
  - Each set of N patient images are used to create a single patient feature set.
  - Use kNN to match probe image to closest patient image set for classification.

---

1 Currently using Andrea Vedaldi's implementation of SIFT from: http://www.robots.ox.ac.uk/~vedaldi/code/sift.html

Testing

Faces in the Wild Dataset

- 13,000+ images
- 1680 have 2 or more images
- We chose 100 image sets
  - 4 images per subject
  - minimal/no occlusions
  - pose variations roughly within 60 degrees
  - no faces within close proximity of primary subject

1 : Labeled Faces in the Wild http://vis-www.cs.umass.edu/lfw/

SGM vs. GMM

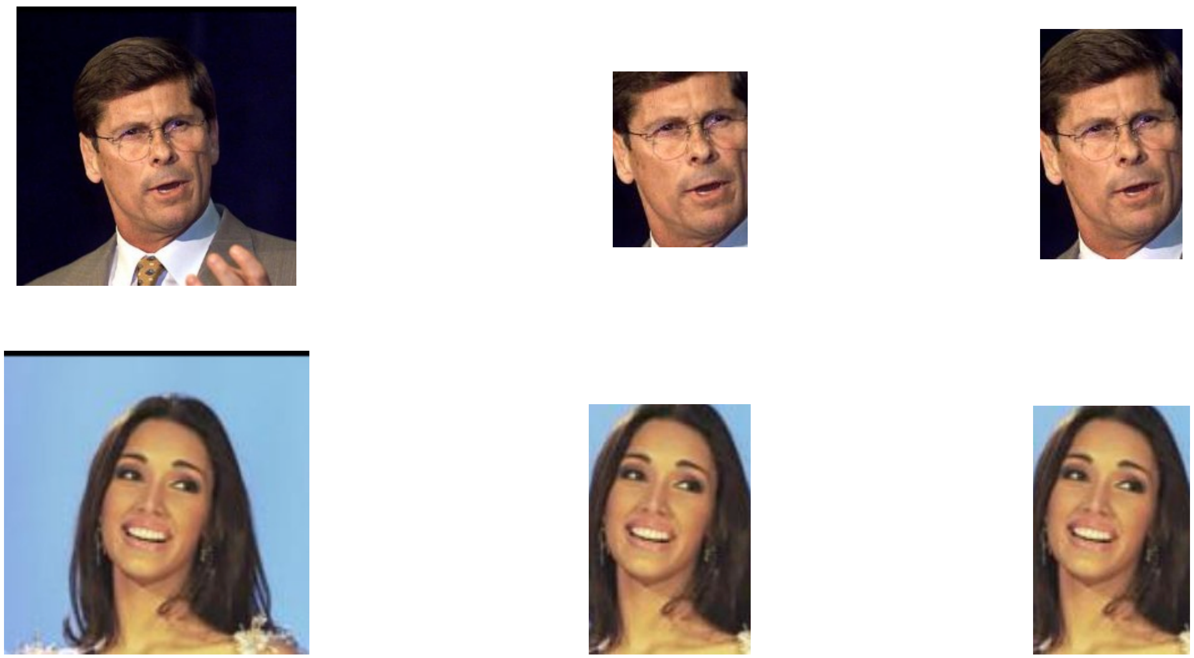
- Performance results from a qualitative comparison of Single Gaussian Model vs. Gaussian Mixture Model in skin region detection…

Rates	Single Gaussian Model	Gaussian Mixture Model
Failures*	29/400 = 7.25%	57/400 = 14.25%
True-Positive	297/371 = 80.05%	274/343 = 79.88%
False-Positive	74/371 = 19.95 %	69/343 = 20.12%

- * Failure defined as an identified patch smaller than 65x65
- Results suggest that the SGM and GMM are more comparable than we previously thought.
- Potential factors to "tune" may include:
  - Number of mixture components (clusters)
  - Size of training data for skin/non-skin models

SGM vs. GMM Sample Results

Original | GMM Result | SGM Result

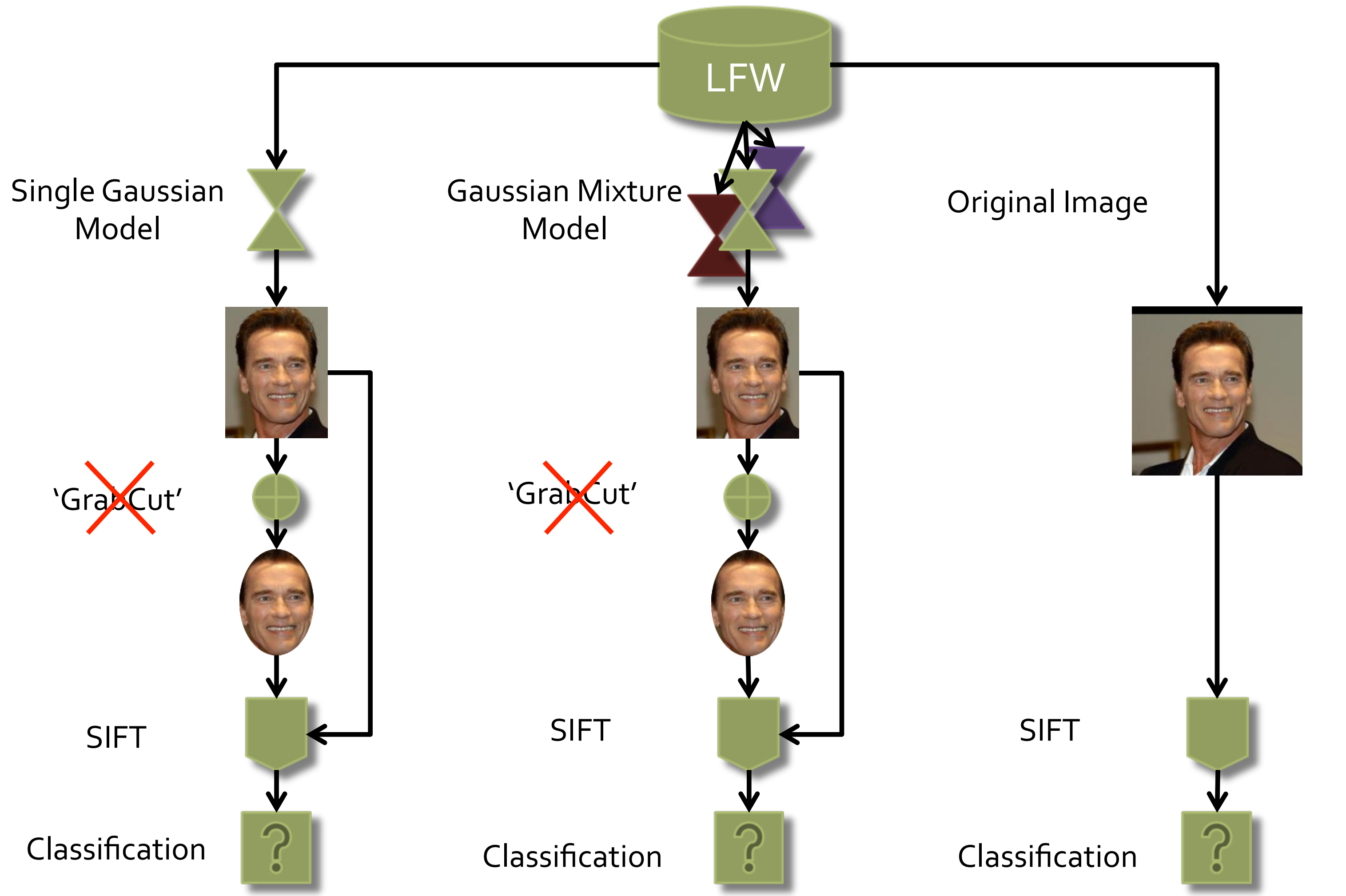


Note: *we did not "expand" bounding box – in most cases SGM/GMM was very good at capturing hair as part of the region of interest*

SGM/GMM → GrabCut

- GrabCut *does* segment the face well given a bounding box, however…
  - If we use the image generated from GrabCut, features are created from "jagged" edges
  - If we use the segmented skin to produce a "refined" bounding box, this typically degraded the quality of the bounding box created by SGM/GMM.
- Therefore, we chose to not include images that were processed by GrabCut in our final end-to-end testing.

[Comparison] End-to-End Testing



End-to-End Testing Results

- 1 test image per person
- 3 training images per person

Rates	SGM/SIFT	GMM/SIFT	SIFT
% Correct Identification	31%	*	*

** Test scripts were not quite done processing all of the information – final results will be reported during final presentation on November 19th, 2013*

Future Work…

- Explore the effectiveness of other feature descriptors
- Reduce feature correspondences based on geometric relationships – use spatial information about descriptors
- Implement “confidence factor” as a metric to avoid/reduce improper identification

Thank You!

Questions? Comments? Concerns?