Color-Based Retrieval of Facial Images

Yannis Avrithis, Nicolas Tsapatsoulis and Stefanos Kollias

Image, Video and Multimedia Lab.

Dept. of Electrical and Computer Engineering

National Technical University of Athens

e-mail: {iavr,ntsap}@image.ntua.gr

Presenter: Anastasios Doulamis

Overview

- Content-Based Retrieval
- A Working Scenario
- Color Segmentation
- Skin-Tone Color Distribution
- Shape Processing
- Retrieval Result Ranking
- Experimental Results

Content-Based Retrieval

- New tools for summarization, content-based query, browsing, indexing and retrieval required for the emerging multimedia applications
- Existing systems use color, motion, texture, shape information as well as spatial and temporal relation between objects
- Extraction of *semantic* information requires *a priori* knowledge and can only be achieved in the context of specific applications
- Growing interest in retrieval of images containing *human faces*: face detection and segmentation required

Face Detection for Multimedia Applications

- In many cases it is enough to detect the presence of a face in a picture / video sequence
  - i.e. detect the anchorperson
- Fast Implementations (real-time performance is desirable)
  - example: news summarization
- Color should be exploited
  - convenience with dedicated content-based indexing /retrieval algorithms

The Proposed Technique

- Combine color segmentation and color based face detection for facial image retrieving
- M-RSST segmentation algorithm employed; average color components, size, location, shape and texture extracted.
- Adaptive 2-D Gaussian density function used for modeling skin-tone color distribution; exploit shape characteristics to discriminate face from skin segments
- Query-by-example framework proposed for interactive, configurable and flexible content-based human face retrieval

A Working Scenario

- Images in database segmented and color chrominance components, size and shape information stored
- *Query-by-example* : User presents a facial image; system performs face detection and ranks existing images according to several criteria
- Retrieval based on *color similarity*, *facial scale* or *number of face segments* possible
- Retrieved images returned to user; further manual selection used to *adapt* skin-color probabilistic model

Color Segmentation: M-RSST

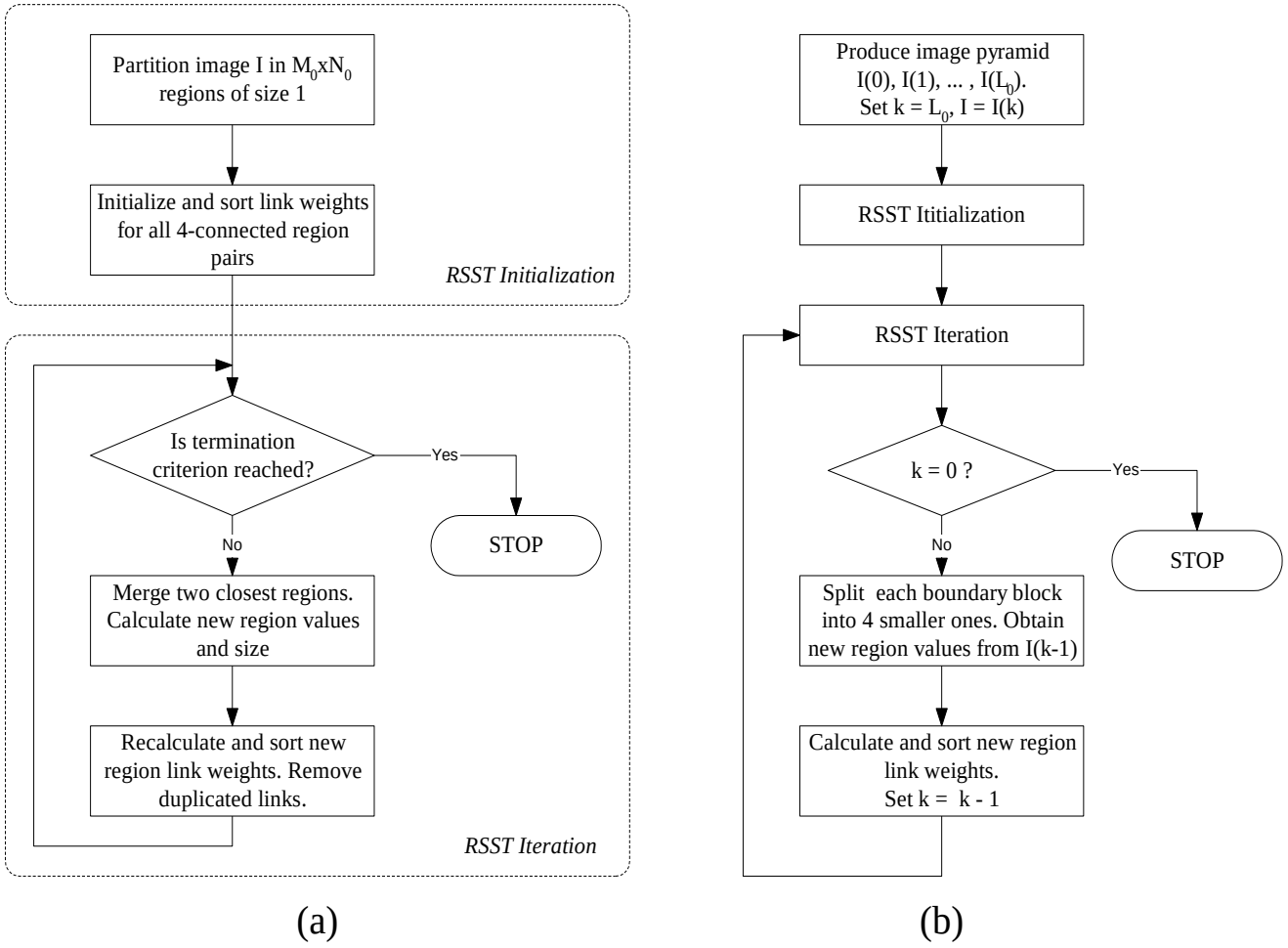
- *Multiresolution decomposition* and construction of a truncated image pyramid
- All 4-connected region pairs assigned a *link weight* equal to the distance measure

$$d(X,Y) = \| \mathbf{c}_X - \mathbf{c}_Y \| \frac{a_X a_Y}{a_X + a_Y}$$

- *Recursive merging* of adjacent regions and *boundary block splitting* in each resolution level
- Fast algorithm, employed directly on MPEG streams with minimal decoding

Image, Video and Multimedia Lab.
National Technical University of Athens

M-RSST Flowchart



Image, Video and Multimedia Lab.
National Technical University of Athens

YCrCb Color Space and Human Skin

- Skin color can be modeled via the chrominance components of the *YCrCb* color model
  - Skin color covers a small part of the *Cr-Cb* plane
  - Influence of *Y* channel small
- Skin color subspace restrictions:
  - cannot be modeled in a general way for all face images
  - 'relaxing' the model => increased number of False Alarms
  - a 'rigorous' model => increased number of Dismissals
- *False Alarm*: Detection of a face in a wrong position or in frames / pictures where no faces are contained
- *Dismissal*: A failure to detect an existing face

The Proposed Skin Color Model

- Approximation of skin-tone color distribution with a *2-D Gaussian density function* on the *Cr-Cb* chrominance plane:

$$P(\mathbf{x}\,|\,\boldsymbol{\mu}_0,\mathbf{C}) = \frac{\exp\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_0)^T\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu}_0)\}}{(2\pi)^{\frac{k}{2}}\cdot|\mathbf{C}|^{\frac{1}{2}}}$$

- **x**: input pattern (mean chrominance components of an image segment)
- $\boldsymbol{\mu}_0$: mean vector, **C**: covariance matrix

Skin-Color Region Extraction

- *Re-estimation* of the mean vector based on current image / frame:

$$\boldsymbol{\mu}_0 = (1-m)\cdot\boldsymbol{\mu}_0 + m\cdot\boldsymbol{\mu}$$

μ: mean vector estimated from current image / frame

m : a memory tuning constant

- *Skin-color region merging* based on estimated skin-color probability:

$$d_C(X,Y) = [\max(1-p_X, 1-p_Y)]^2$$

- Adjacent face segments merged – remaining partition map not affected

Shape Processing

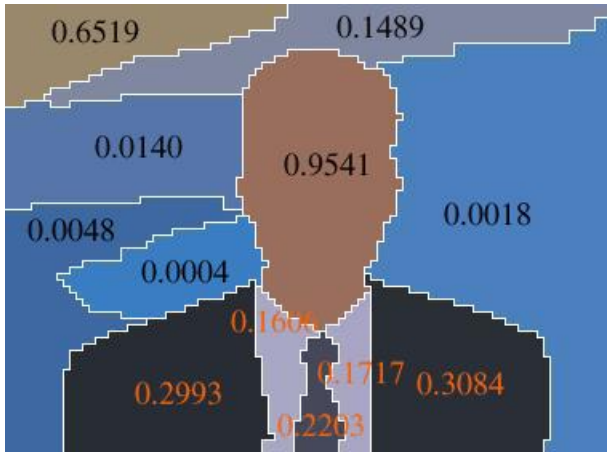
- *Global shape features* of segment contours
  - Shape *compactness* : $g_X = 4\pi\, a_X / r_X^2$
  - Shape *elongation* : $\ell_X = \sqrt{\lambda_2 / \lambda_1}$
- Both normalized in [0,1] and *invariant* to translation, scaling and rotation
- Combination with skin-color probability using non-linear functions – construction of an overall *face probability map*
- Segments with extremely irregular shape discarded

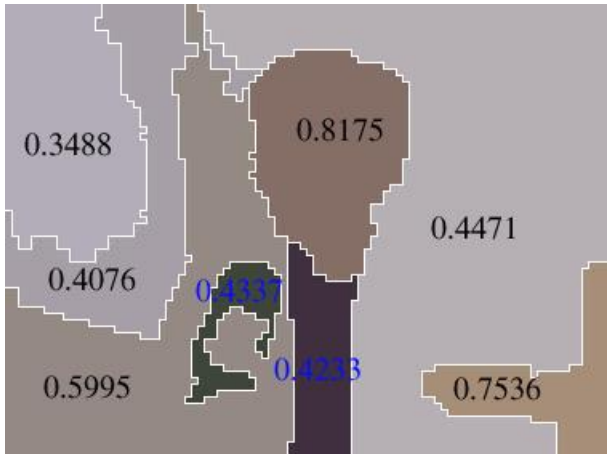
Retrieval Result Ranking

- *Query-by-example* : User presents a facial image; system performs face detection and ranks existing images according to several criteria
- *Similarity with the presented face segment* : *m* small, ranking w.r.t. segment probability
- *Facial scale* : *m* high, ranking w.r.t. percentage of image area
- *Number of face segments*: *m* high, ranking w.r.t. facial segments present in the image

Experimental Results

- Segmentation and probability assignment





Image, Video and Multimedia Lab.
National Technical University of Athens

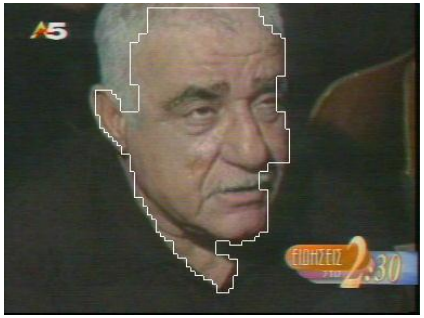
Skin Color based Retrieval

Image Presented to the system

mem: 0.3

Selected from the user segment

0.9992



0.9872

0.9735

0.9591

Image, Video and Multimedia Lab.
National Technical University of Athens

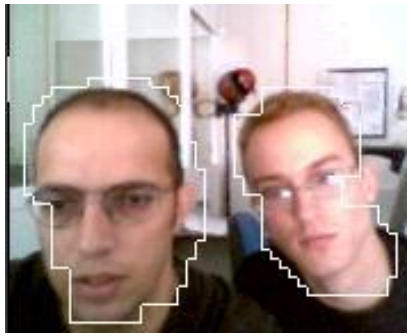
Retrieval based on number of Faces

Image Presented to the system

Segmented Faces

mem: 0.7

prob=0.6369



0.5525

0.1581

0.1224

Image, Video and Multimedia Lab.
National Technical University of Athens

Retrieval based on Facial Scale

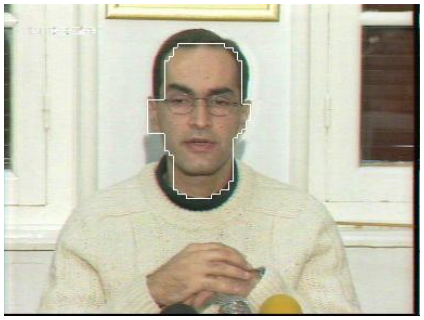
Image Presented to the system

mem: 0.8

Segmented Face

Facial area: 0.0867

0.0873



0.0883

0.0969

0.0985

Image, Video and Multimedia Lab.
National Technical University of Athens

Conclusions

- *Color segmentation* : powerful tool for object extraction, especially for human faces
- *M-RSST algorithm* : eliminates facial details and provides a single object for each face
- *Chrominance components* with a *probabilistic model* used in an efficient way for retrieving facial images from image databases
- *Interactive retrieval framework* adapts the model to user needs and leads to meaningful retrieval results