Efficient Face detection for Multimedia Applications

N. Tsapatsoulis, Y. Avrithis and S. Kollias

Image, Video and Multimedia Lab.

Dept. of Electrical and Computer Engineering

National Technical University of Athens

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

Face Detection: Is it only for Face Recognition?

- A few years earlier
  - Face Detection <=> Face Recognition

- Present Applications of Face Detection
  - Face Recognition
  - Content based Video Indexing and Retrieval
  - Video Scene Classification / Annotation
  - News Summarization

Face Detection as a part of Face Recognition Schemes

- High Accuracy is required
  - Remarkable results are obtained only if we pose rigorous constraints
  - Algorithms are concentated in gray-scale images
  - Template matching or low level feature detection
- Time consuming procedures

Face Detection and Multimedia Applications

- In many cases its 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 Scheme

- Combine color segmentation and skin color characteristics
- Use M-RSST as a general purpose segmentation algorithm.
- Associate each segment with a skin color probability obtained by an adaptive 2-D Gaussian density function used for modeling skin-tone color distribution;
- Exploit shape characteristics to discriminate face from skin segments => face probability
- Query-by-example framework proposed for interactive human face retrieval

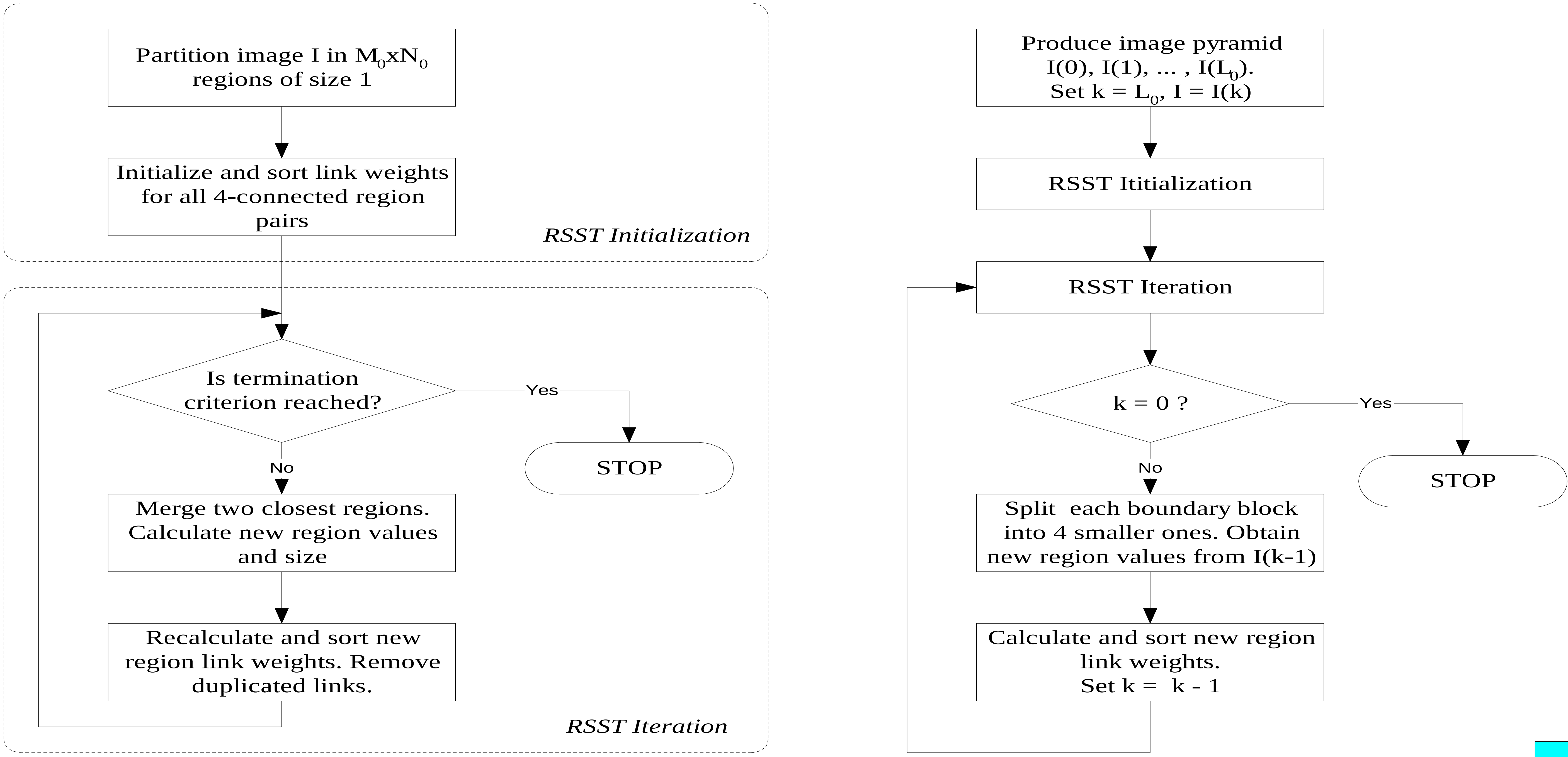
Color Segmentation: M-RSST

- *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

M-RSST Flowchart



(a) (b)

The *YCrCb* color space and the 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
  - The influence of *Y* channel is small
- However, post processing steps are required:
  - Other objects have skin like color
  - *Y* channel influence not totally negligible
  - Compact objects desirable => Filtering

Skin Color Modeling Issues

- Skin color subspace covers a small area of the *Cr-Cb* plane but:
  - it cannot be modeled in such a general way to be efficient for all images that include faces
  - ‘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

- Skin color characteristics are modeled via a 2D-Gaussian distribution

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

x: input pattern (mean chrominance components of an image block)

μ_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 : the estimated from the current image / frame mean vector

  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

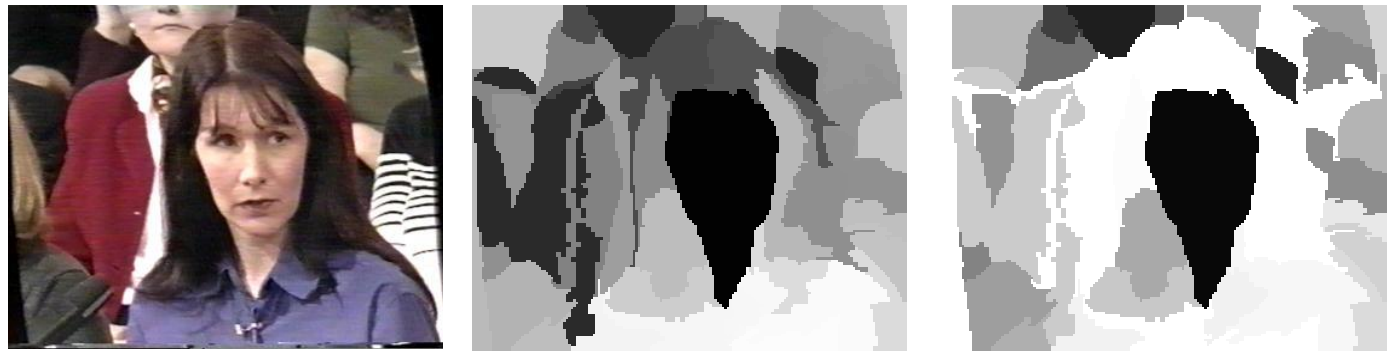
Shape Processing

- *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

Segmentation and probability assignment

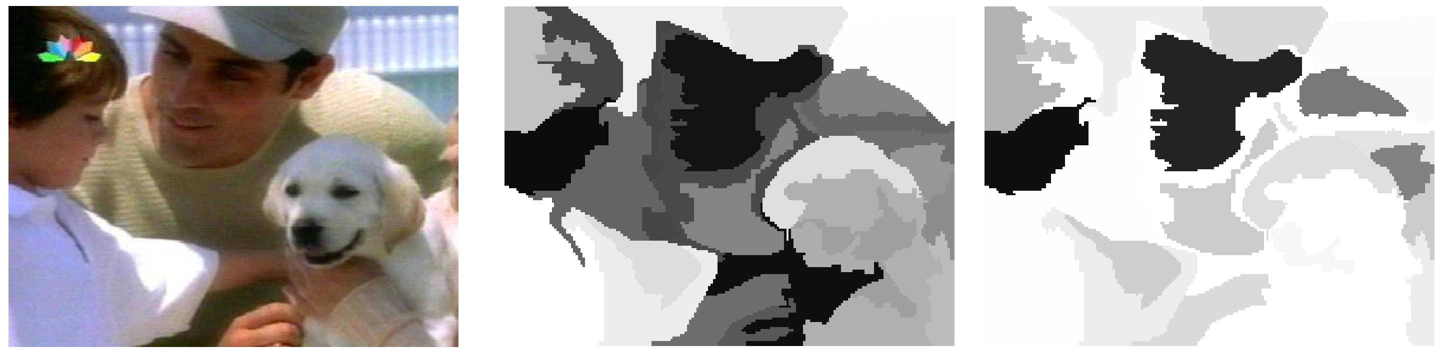
Face detection in a variety of situations

(a) (b) (c)



(a) (b) (c)

(a) (b) (c)

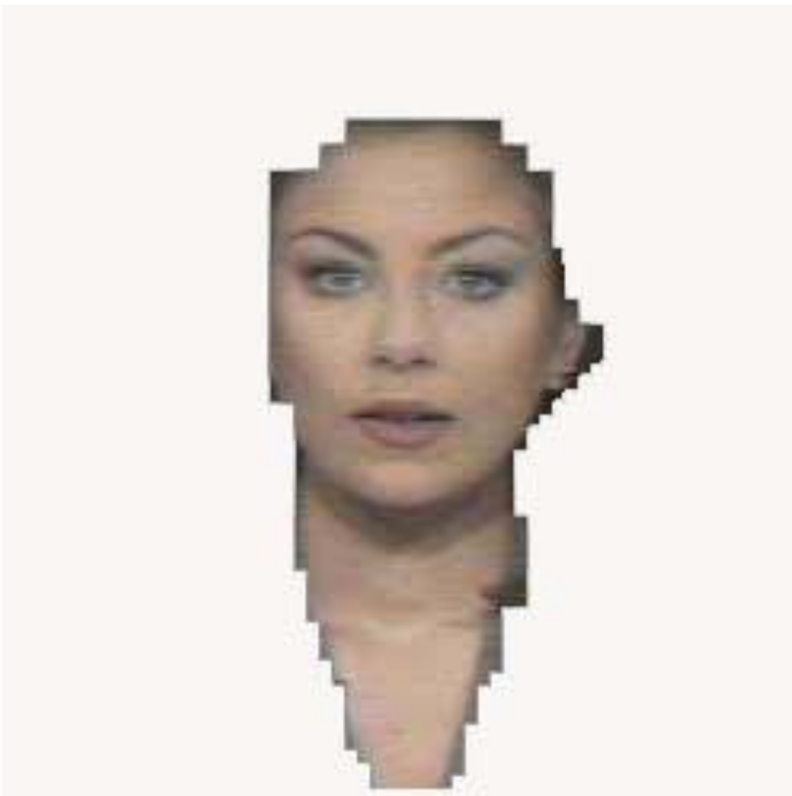


(a) (b) (c)

(a) Original images, (b) skin-color probability map, (c) final face probability map (including shape features).

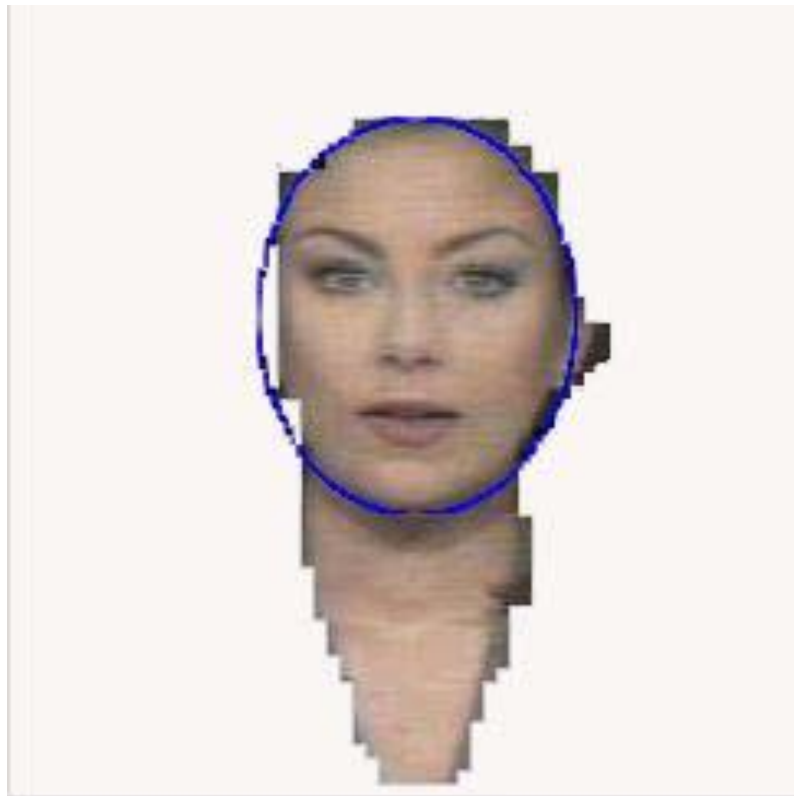
Further verification required?

Isolated skin segment



Edges

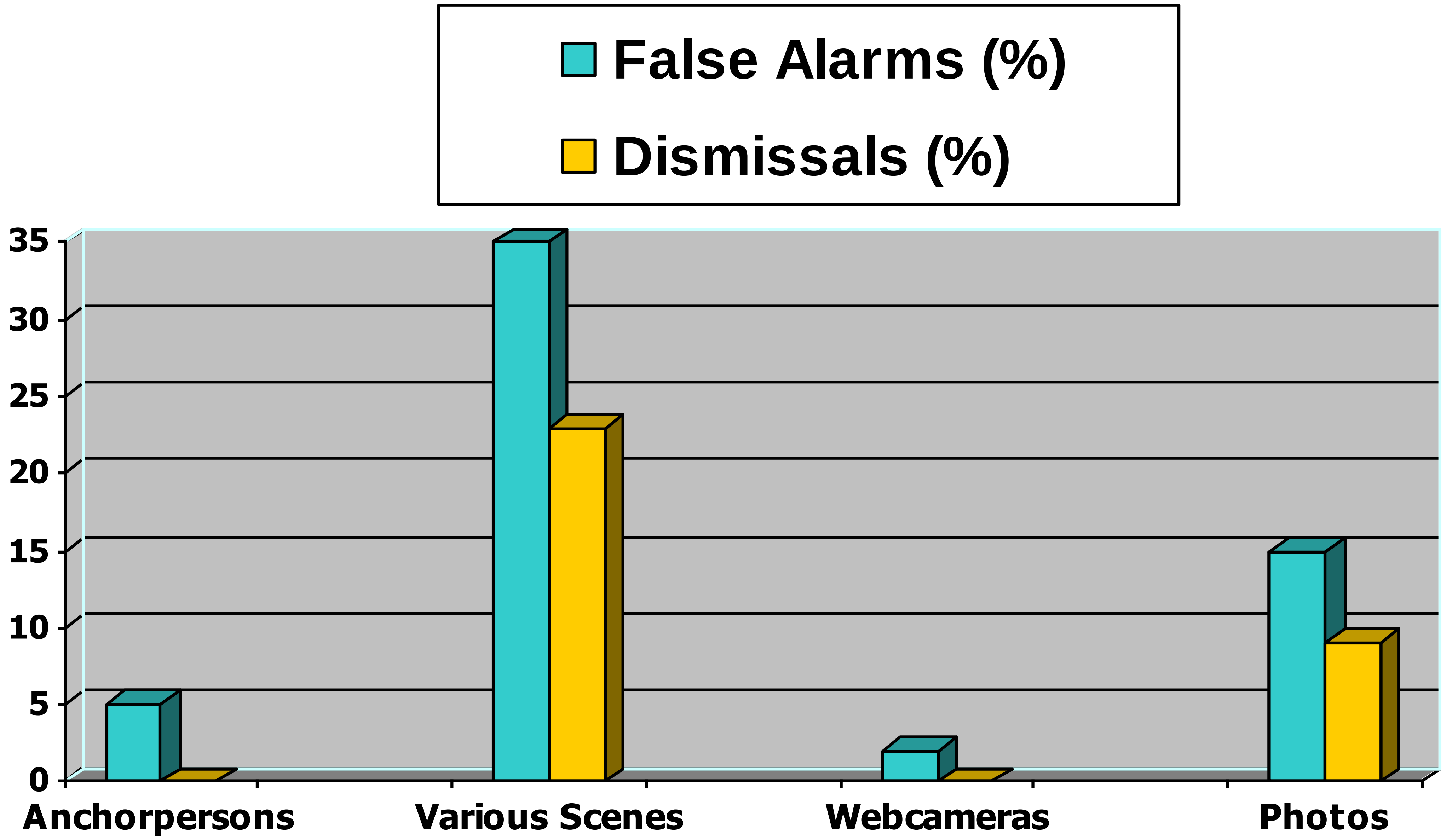
Fitted Ellipses



- Calculate the edges within the probable face segment
- Check whether an ellipses can be fitted to the edges

Experimental Results

- Anchorpersons scenes: recorded from TV news; Various scenes: recorded from TV programs; Webcameras: Shots captured using Webcameras; Photos: Regular colored photos



A Retrieval Scenario

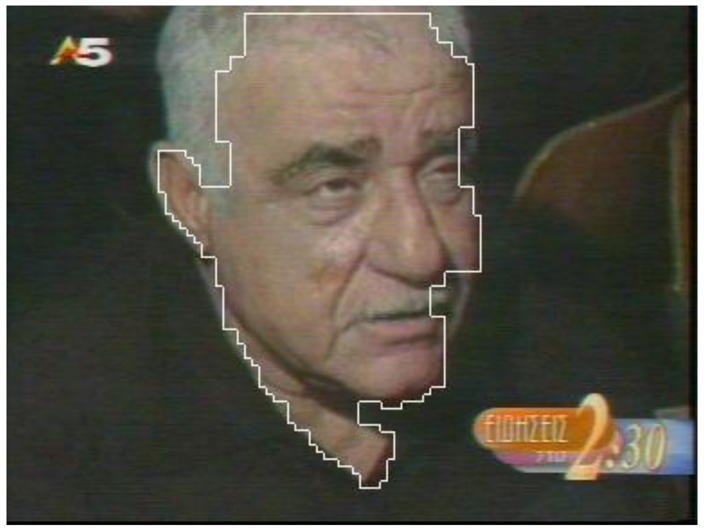
- Images in database segmented and color chrominance components, size and shape information stored
- *Query-by-example* : User presents a facial image; system performs face detection and ranks existing images according to several criteria
- Retrieval based on *color similarity*, *facial scale* or *number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

Skin Color based Retrieval

Image Presented to the system | Selected by the user segment

mem: 0.3

0.9992



0.9872

0.9735

0.9591

Retrieval based on Facial Scale

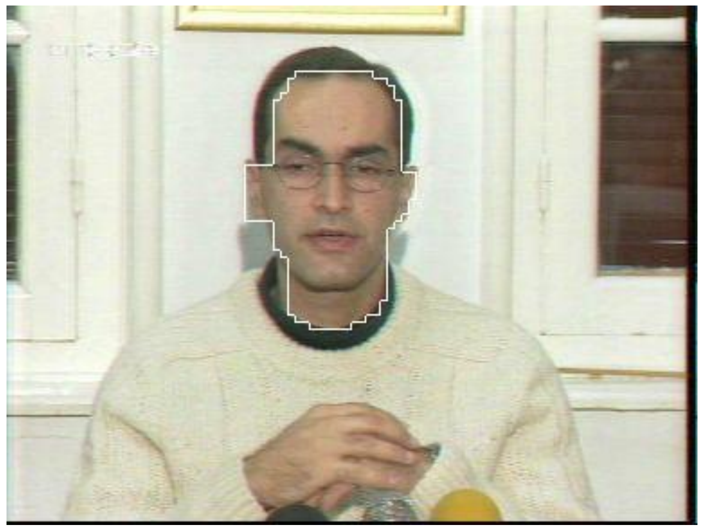
Image Presented to the system

mem: 0.8

Segmented Face

Facial area: 0.0867

0.0873



0.0883

0.0969

0.0985