National Technical University of Athens
Department of Electrical and Computer Engineering

A Genetic Algorithm for Efficient Video Content Representation

Anastasios Doulamis, Yannis Avrithis,
Nikolaos Doulamis and Stefanos Kollias

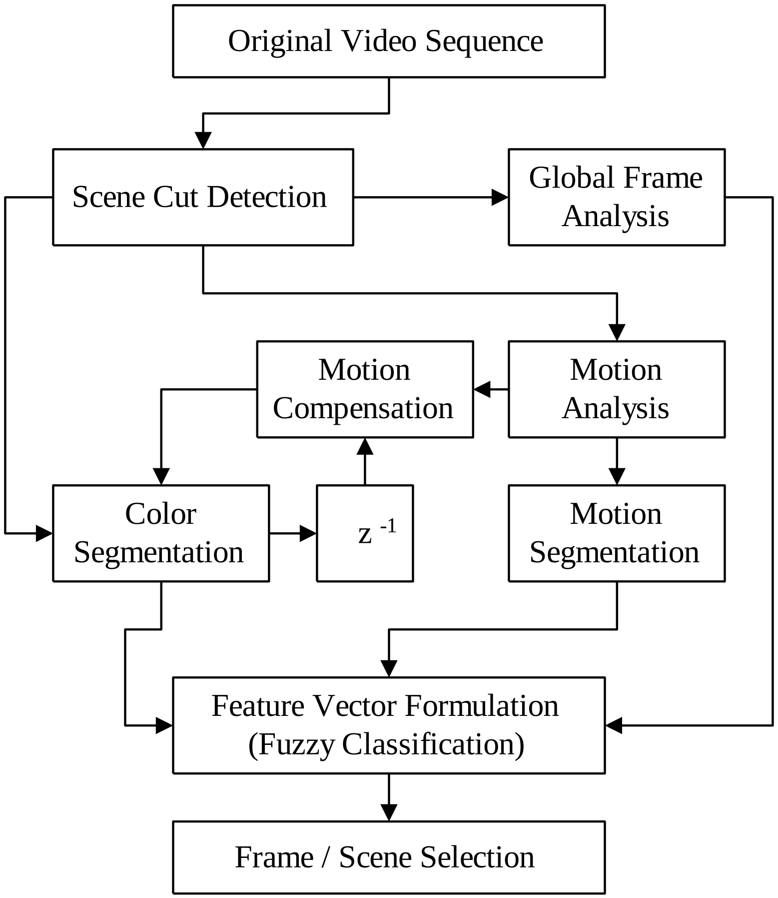
Objectives

- Automatic selection of a limited number of *key frames and scenes* from video databases
- Key frames and scenes provide sufficient information about the video content
- Representation of video sequences by using multi-dimensional *feature vectors* of key frames and scenes
  - color & motion information
- *Video queries* applied directly on the extracted key frames and scenes

Applications

- Multimedia database management
  - reduction of storage requirements for search capabilities
  - direct content-based retrieval
  - faster and more efficient video queries
  - improvement of user interface
- Multimedia interactive services
  - production of low resolution video clip previews (trailers) or still image mosaics
  - browsing of video databases on web pages

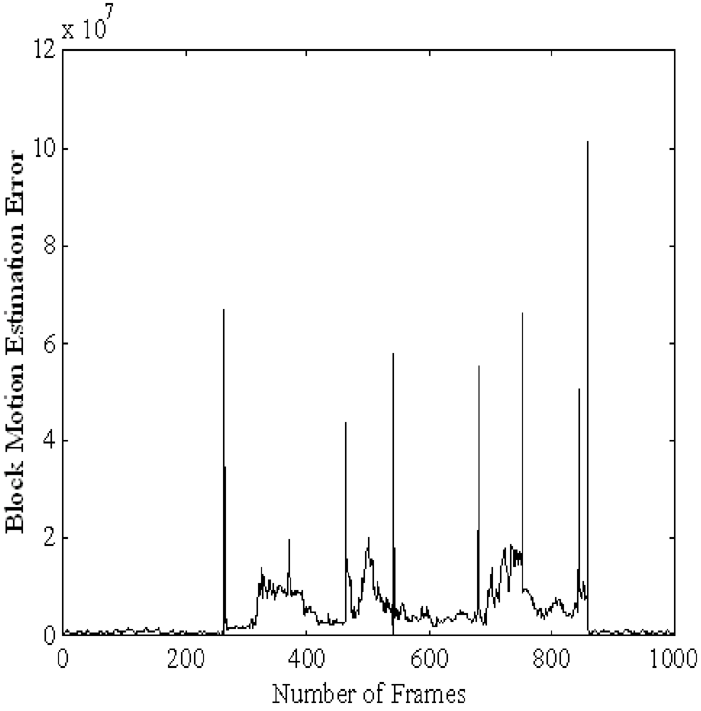
Proposed System Architecture



- Scene cut detection
- Feature extraction for each frame
- Formulation of scene feature vectors
- Selection of the most representative scenes
- Extraction of key frames for each scene

Scene Cut Detection

- Computation of the sum of the block motion estimation error
- Selection of frames for which the sum exceeds a certain threshold
- Computations applied directly to MPEG coded sequences

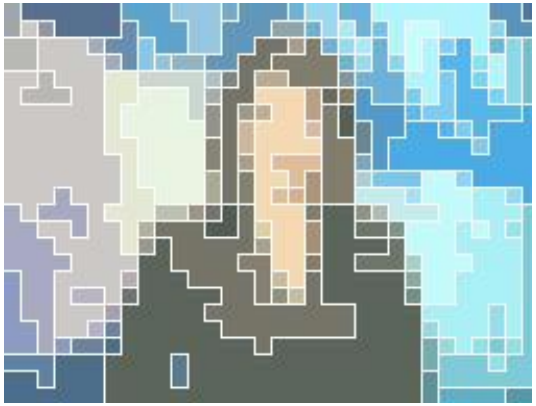


Color Segmentation

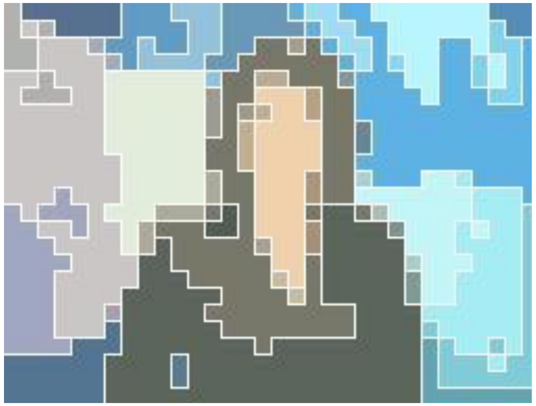
- Segmentation according to *spatial homogeneity*
- *Block resolution* (reduction of computational time, exploitation of MPEG information)
- *Hierarchical merging* of similar segments (depending on color homogeneity & segment size)
- *Object tracking:* comparison with motion compensated segmentation results of previous frames (connected regions are encouraged to remain connected in successive frames)
- *Color features*: number of segments, location, size & mean color of each segment

Color Segmentation Results

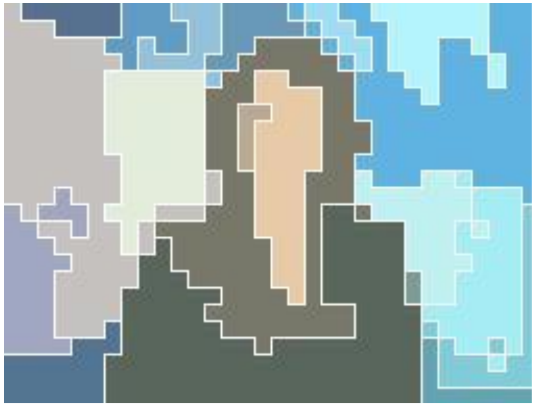
Original frame



1st iteration



3rd iteration

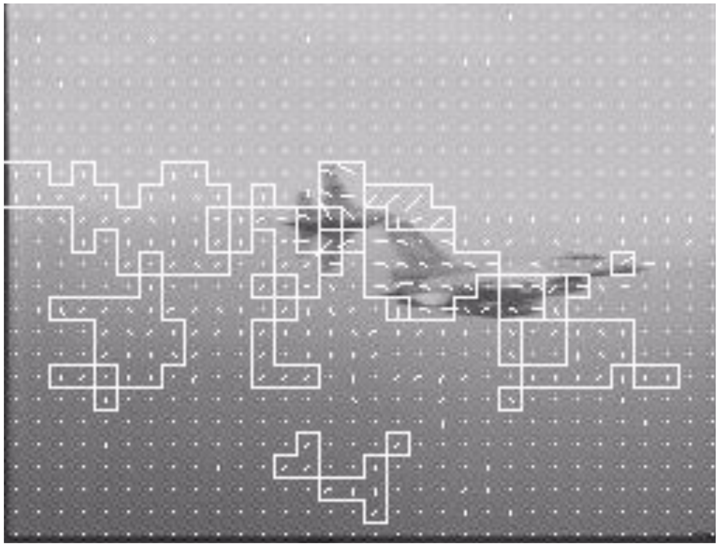


Final Result

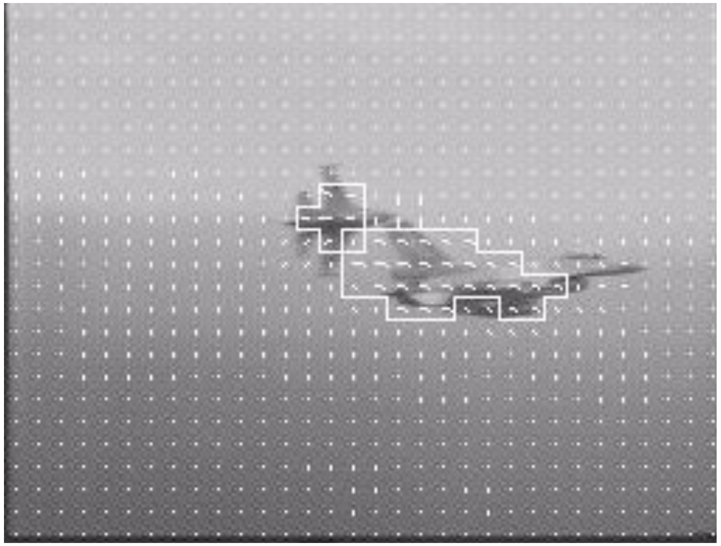
Motion Segmentation

- Segmentation according to *motion homogeneity*
- *Block resolution* (reduction of computational time)
- Motion vectors derived from motion analysis
- *Median filtering* of derived motion vectors: elimination of “noise”, preservation of “edges”
- *Motion features*: number of segments, location, size & mean motion vector of each segment

Motion Segmentation Results



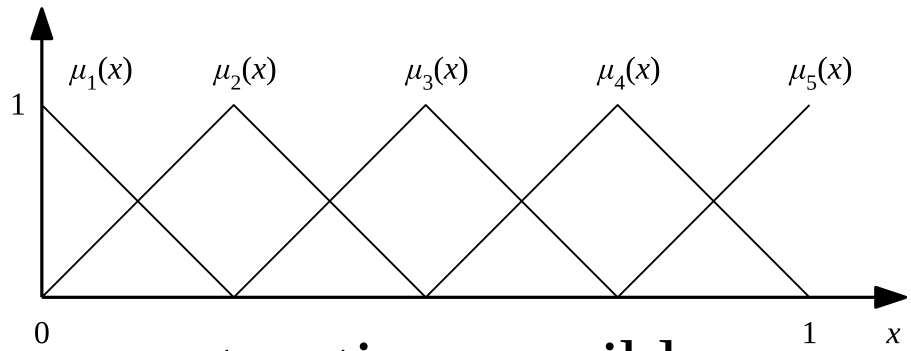
Motion segmentation without filtering



Motion segmentation with filtering

Feature Vector Formulation

- *Multidimensional "histogram"*: classification of color and motion segments into pre-determined classes
- *Fuzzy classification:* normalization of each feature x to [0,1], and partitioning into Q classes defined by membership functions $\mu_n(x) \in [0,1],\ n = 1,\ldots,Q$



- Histogram construction possible even with small number of samples x.

Multidimensional Fuzzy Classification

- *Degree of membership* allocated to each class

  $$F(n_1,\ldots,n_L) = \sum_{i=1}^{K}\left\{\prod_{j=1}^{L}\mu_{n_j}(f_j^{(i)})\right\}$$

  where $n_j \in \{1,2,\ldots,Q\}$: classification index for jth feature, Q: no. of partitions, L: no. of features, K: no. of segments, $f_j^{(i)}$: jth feature of ith segment, $\mu_n(f)$: degree of membership of feature f in partition n
- Feature vector formed by degrees of membership for all $M=Q^L$ combinations of $n_1,\ldots,n_L \in \{1,2,\ldots,Q\}$
- *Global frame characteristics* included in feature vector (color histogram, etc.)

Representative Scene Selection

- *Scene feature vector* constructed based on frame feature vectors over duration of a scene
- *Clustering* of similar scene feature vectors $\mathbf{s}_i \in \Re^M$, $i=1,\ldots,N_S$ and selection of cluster *representatives* $\mathbf{c}_i$, $i=1,\ldots,K_S$
- *Average distortion* $D(\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{K_S}) = \sum_{i=1}^{K_s} \sum_{\mathbf{s} \in Z_i} d(\mathbf{s}, \mathbf{c}_i)$ should be minimized, where

$$Z_i = \{\mathbf{s} \in S : d(\mathbf{s}, \mathbf{c}_i) < d(\mathbf{s}, \mathbf{c}_j)\ \forall\ j \neq i\}$$

  is the *influence zone* of $\mathbf{c}_i$
- Minimization performed with *generalized Lloyd* or *K*-means algorithm

Optimization Methods for Frame Selection

- Minimization of a *correlation criterion* (key frames should not be similar to each other)
- *Correlation measure* of feature vectors $\mathbf{f}_i,\ i = x_1, \ldots, x_{K_F}$

$$R(\mathbf{x}) = R(x_1, \ldots, x_{K_F}) = \left(\sum_{i=1}^{K_F - 1} \sum_{j=i+1}^{K_F} (\rho_{x_i, x_j})^2 \right)^{1/2}$$

where ρ_{ij}: correlation coefficient of vectors $\mathbf{f}_i$, $\mathbf{f}_j$ and $\mathbf{x} = (x_1, \ldots, x_{K_F})$: index vector corresponding to a set of selected frame numbers

- Minimization of $R(\mathbf{x})$ w.r.t. $\mathbf{x}$ implemented by *logarithmic search* or *genetic algorithm* (exhaustive search is unfeasible)

Genetic Approach for Frame Extraction

- Chromosomes are represented by index vectors

  $\mathbf{x} = (x_1, \ldots, x_{K_F}) \in V^{K_F}$

- An *initial population* of chromosomes is then generated by selecting sets of frames whose feature vectors reside in extreme locations of the feature vector trajectory
  - local maximization of the magnitude of the second order derivative of feature vector trajectory
- The correlation measure $R(\mathbf{x})$ is used as an objective function to estimate the performance of all chromosomes for a given population

Genetic Approach for Frame Extraction

- The fitness function follows a linear normalization scheme
  - chromosomes are ranked in ascending order of $R(\mathbf{x}_i)$
- If F_B is an arbitrary fitness function, then the best chromosome is chosen by

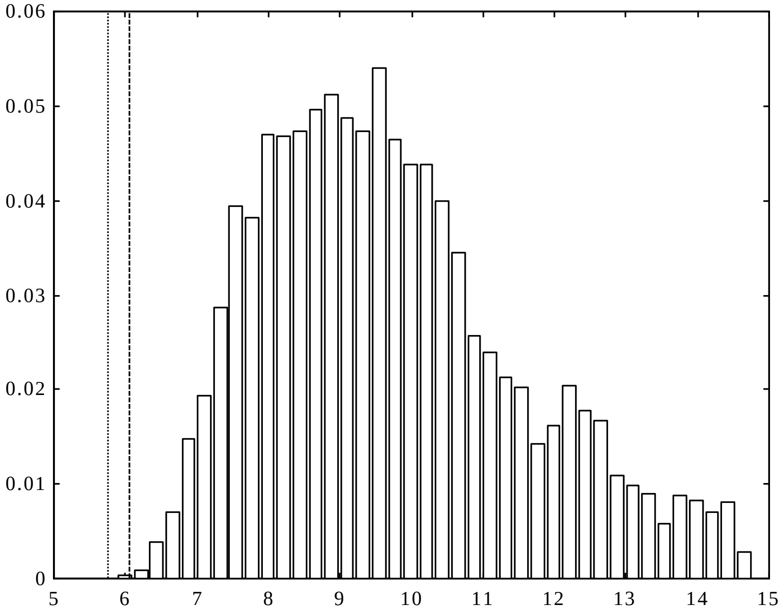
$$F(\mathbf{x}_i) = F_B - [r(\mathbf{x}_i) - 1]D, \quad i = 1, \ldots, P$$

where $r(\mathbf{x}_i) \in \{1, \ldots, P\}$ corresponds to the rank of chromosomes and D is a decrement rate

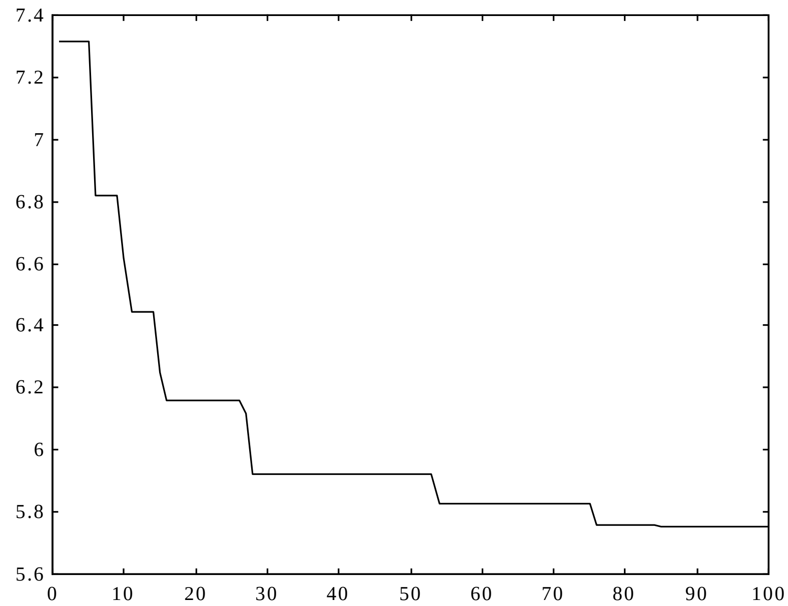
Generation of the Next Population

- *Parent selection is* applied so that a fitter chromosome gives a higher number of offspring
  - higher chance of survival in the next generation
- A *proportionate scheme*, implemented by the *roulette wheel selection* procedure, is used for parent selection
- A set of new chromosomes is produced by mating the selected parent chromosomes and applying a *crossover operator*
  - the genetic material of the parents is combined in a random way to produce the genetic material of the offspring

Experimental Results



Probability density function of the correlation measure



Genetic algorithm convergence

Experimental Results

Six selected key frames of a scene after applying the genetic algorithm

Conclusions

- *Automatic extraction* of key frames and scenes of video sequences taken from large video databases
- *Object tracking* provides smoother feature vector trajectories
- Fuzzy representation of the feature formulation makes more robust frame/scene selection
- Genetic approach for frame extraction based on correlation measure