Color video segmentation by dissimilarity based on edges

Lucía Ramos[1] Jorge Novo[1] José Rouco[1]
Antonio Mosquera [2] Manuel G. Penedo [1]

VARPA Group, Department of Computer Science, University of A Coruña, Spain

Artificial Vision Group, Department of Electronics and Computer Science, University of Santiago de Compostela, Spain

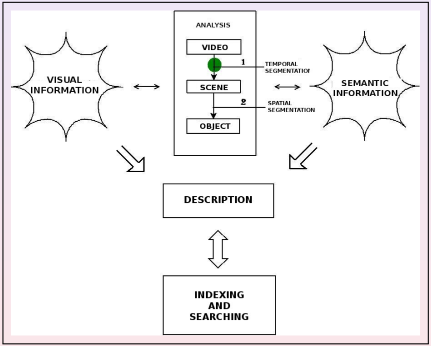
24 de junio 2010

Index

Contents

Introduction > Motivation

Content-based multimedia retrieval



- Manual processes
    - Time-consuming
    - There are no uniform criteria

Solution

Automate

Introduction > Objectives

- Study of dissimilarity measures based on edges
  - Space domain
  - Frequency domain
  - Improved extensions to color-space
- Experimentation
  - Experiments design
  - Results

Contents

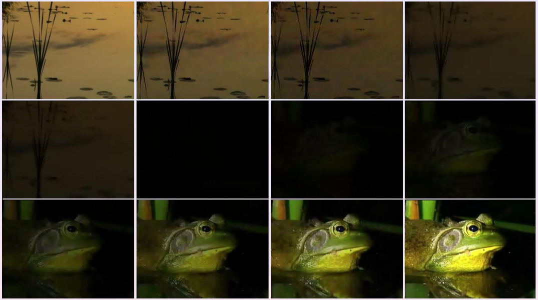
Shot change types

- Sudden, ocurring an abrupt change between two consecutive frames
  - Cut

- Gradual, ocurring slowly over successive frames

Dissolve

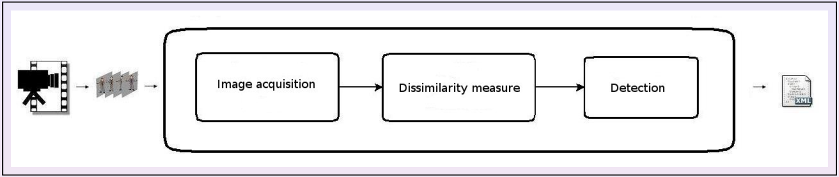
Frames of different scenes are superimposed, fading of the previous scene to defining the extent of the new.



Fade

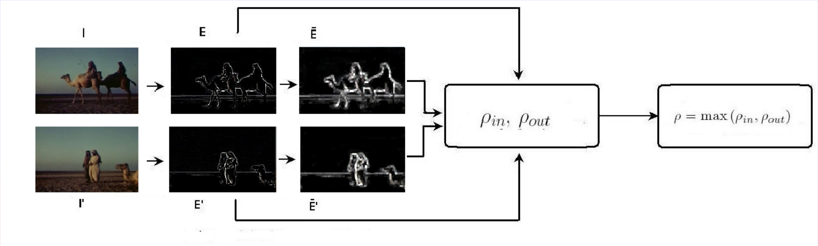
Special case of dissolve where a monochrome frame is intercalated between consecutive scenes.

General scheme of temporal video segmentation



- Acquire successive frames composing the video sequence
  - Grayscale representation
  - Color representation
- Obtain edge-based distance
  - Dissimilarity on space domain
  - Dissimilarity on frequency domain
- Decide the existence of shot changes
  - Fixed threshold
  - Adaptive threshold

Dissimilarity measures based on edges > Space domain



$$\rho_{out} = 1 - \frac{\sum_{x,y} E[x,y] \overline{E'}[x,y]}{\sum_{x,y} E[x,y]}$$

$$\rho_{in} = 1 - \frac{\sum_{x,y} \overline{E}[x,y] E'[x,y]}{\sum_{x,y} E[x,y]}$$

Proportion of exiting pixeles

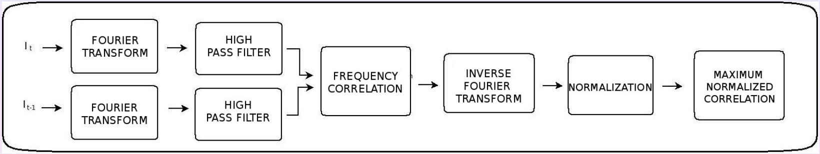
Elevated for a cut, a fade-out or in the first half of a dissolve

Proportion of entering pixeles

Elevated for a cut, a fade-in or in the second half of a dissolve

Scene breaks are detected by looking for peaks in ρ

Dissimilarity measures based on edges > Frequency domain



Normalized correlation on frequency domain

Measuring the similarity between two consecutive frames. The peak represents the best match

$$\rho(\xi) = \frac{TF^{-1}\left\{\widehat{x}_1(\omega)\,\widehat{x}_2^*(w)\right\}}{\sqrt{\int|\widehat{x}_1(\omega)|^2 d\omega \cdot \int|\widehat{x}_2(\omega)|^2 d\omega}}$$

Ideal high-pass filter

Enhance the contribution of the regions that exhibit the characteristics of changes

Improved extensions > Color dissimilarity on space domain

Objective

Get edges due to color variations

$\Downarrow$

Adapt edges extractor

Canny

- Smoothing
- Derivation
- Non-maximum suppression
- Hysteresis

Improved extensions > Color dissimilarity on space domain

RGB space

Not capture all possible color differences

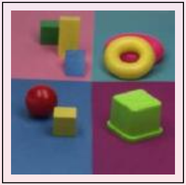
Ratios of Gevers

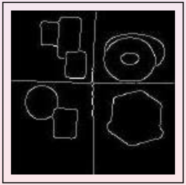
For calculating perceptual color differences independently of the luminosity

$$m_i\left(C_1^{\overrightarrow{k_1}}, C_1^{\overrightarrow{k_2}}, C_2^{\overrightarrow{k_1}}, C_2^{\overrightarrow{k_2}}\right) = \frac{C_1^{\overrightarrow{k_1}} C_2^{\overrightarrow{k_2}}}{C_1^{\overrightarrow{k_2}} C_2^{\overrightarrow{k_1}}}$$

Derivation

- ⋆ Finite difference between the red-green, red-blue and green-blue channels





Improvements > Color dissimilarity on frequency domain

Objective

Correlation based on color variations

⇓

Adapting high-pass filter

Fourier Transform works with a complex representation

Color space ⇒ Several dimensions

- RGB
  - ⋆ Get 3 complex transformations and combine processed
  - ⋆ Luminosity can not be separated

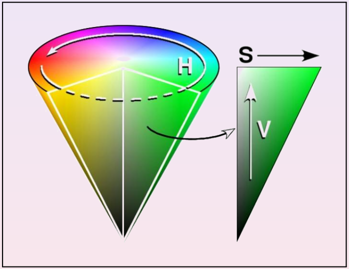
Improvements > Color dissimilarity on frequency domain

HSV space

- Complex representation of the hue-saturation space

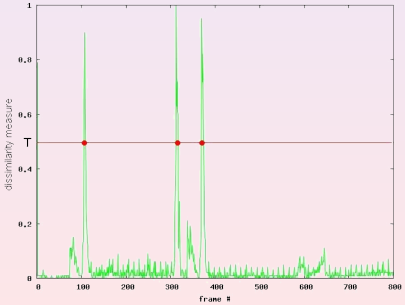
$$b(x, y) = S(x, y) \cdot e^{iH(x,y)}$$

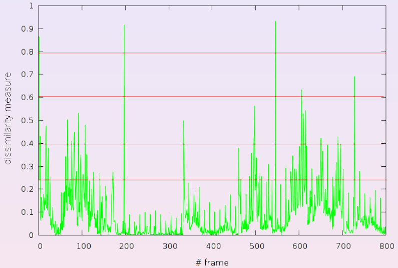
- The lightness component is discarded



Improvements > Adaptative threshold

- Fixed threshold
  - It is based on experimental results
  - Remains fixed throughout the sequence



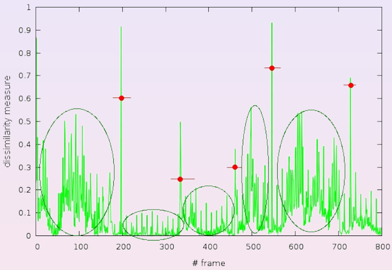


Problem

Difficult to establish the optimal threshold

Improvements > Adaptative threshold

- Adaptative threshold
  - It is based on the information contained in the video.
  - Fits over the sequence as the variability of the scene.



$$T_l = \mu + \alpha\sigma$$

It increases the rate of true positives without increasing false detections

Contents

Dataset

- TRECVID 2002

Total frames	Cuts	Fades	Dissolves	Others	Total
345306	1078	40	211	9	1338

- Heterogeneous in terms of cuts types
- Varied content
- Common cases of error in segmentation
  - Fast-moving objects
  - Camera operations
  - Changes in lighting
  - Gradual changes of different durations

Evaluation critery

- Quality detecion

$$Recall = \frac{T.P.}{T.P. + F.N.} \qquad Precision = \frac{T.P.}{T.P. + F.P.}$$

Recall

A measure of the ability to detect all relevant items

Precision

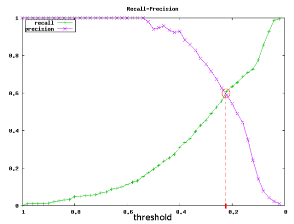
A measure of the ability to detect only relevant items

Experiments > Tuning parameter

- Representative subset of the data
- Select the parameter to optimize recall and precision

Space domain	Canny					
	T_l	T_h	σ	T	α	r
Grayscale	20	60	5	0.15	4.6	1
Color ratios	15	50	5	0.20	5	1

Frequency domain	T	α
Grayscale	0.43	4
HSV	0.28	4



Experiments > Results

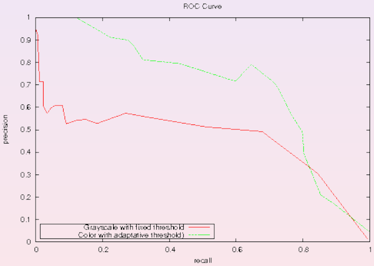
- Complete dataset
- Evaluation of each measure

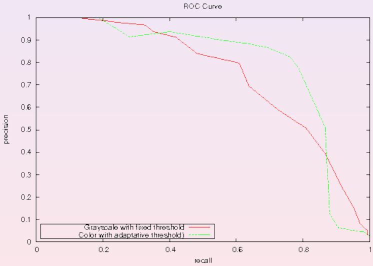
  Each cut type is considered separately

		Threshold	Recall Cuts	Recall Dissolves	Recall Fades	Recall	Precision
Edges	Grayscale	Fixed	52.75 %	14.73 %	71.50 %	51.49 %	41.91 %
		Adaptative	58.21 %	28.04 %	77.50 %	59.01 %	52.11 %
	Color ratios	Fixed	71.07 %	21.64 %	93.54 %	66.99 %	57.29 %
		Adaptative	78.43 %	41.18 %	100.00 %	70.38 %	68.50 %
Correlation	Grayscale	Fixed	74.22 %	15.33 %	86.91 %	63.24 %	65.75 %
		Adaptative	78.90 %	30.02 %	50.00 %	64.17 %	66.10 %
	HSV	Fixed	82.80 %	42.81 %	94.29 %	76.22 %	78.70 %
		Adaptative	92.41 %	50.00 %	66.98 %	82.60 %	81.90 %

Experiments > ROC Curves

- Comparison of different variants





Contents

Conclusions

- Study of segmentation techniques
  - Influence of the frame representation
  - Influence of the detection threshold
  - Sensitivity to each cut type
  - Common error situations
- Contributions
  - Color extensions
    - ⋆ Gevers ratios on space domain
    - ⋆ HSV space on frequency domain
  - Adaptive threshold

Color video segmentation by dissimilarity based on edges

Lucía Ramos[1] Jorge Novo[1] José Rouco[1]
Antonio Mosquera [2] Manuel G. Penedo [1]

VARPA Group, Department of Computer Science, University of A Coruña, Spain

Artificial Vision Group, Department of Electronics and Computer Science, University of Santiago de Compostela, Spain

24 de junio 2010