Image/Video Segmentation for Content-Based Functionalities

Dengsheng Zhang
dengsheng.zhang@infotech.monash.edu.au

Gippsland School of Computing and Information Technology
Monash University
Australia

Outline

- Motivations
- Objectives
- Image segmentation techniques
- Issues and problems relating to motion
- Video segmentation techniques
- Proposed scheme
- Conclusion

Motivations

- Content-based multimedia information retrieval
- Content-based functionalities of MPEG-4 video

Content-based multimedia information retrieval

- High demand for content-based information retrieval
- Limited capability provided by current multimedia database
- Object-based or shape-based retrieval is better than other methods
- Decomposition of image into objects for content-based purposes (storage and representation)

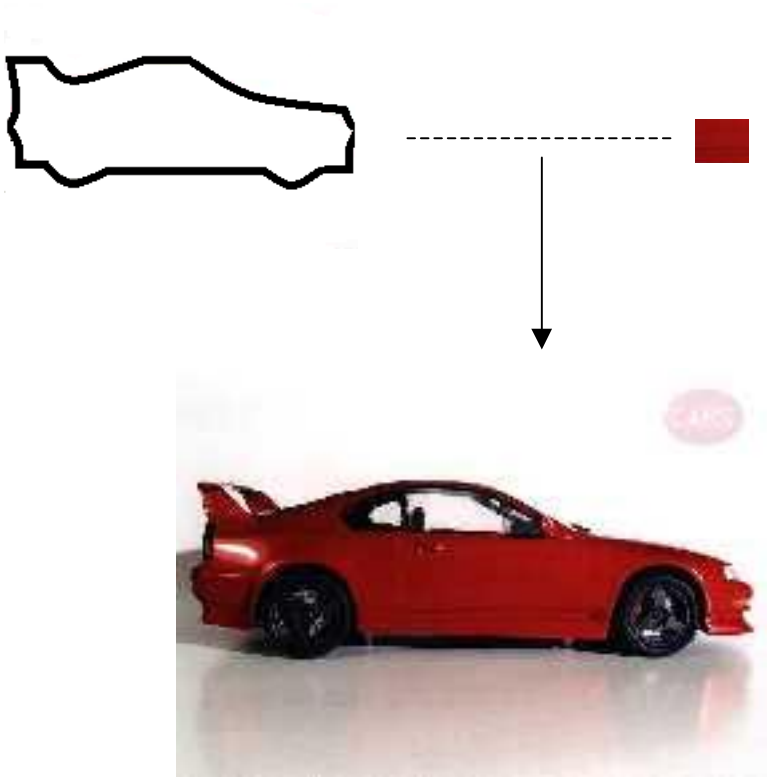
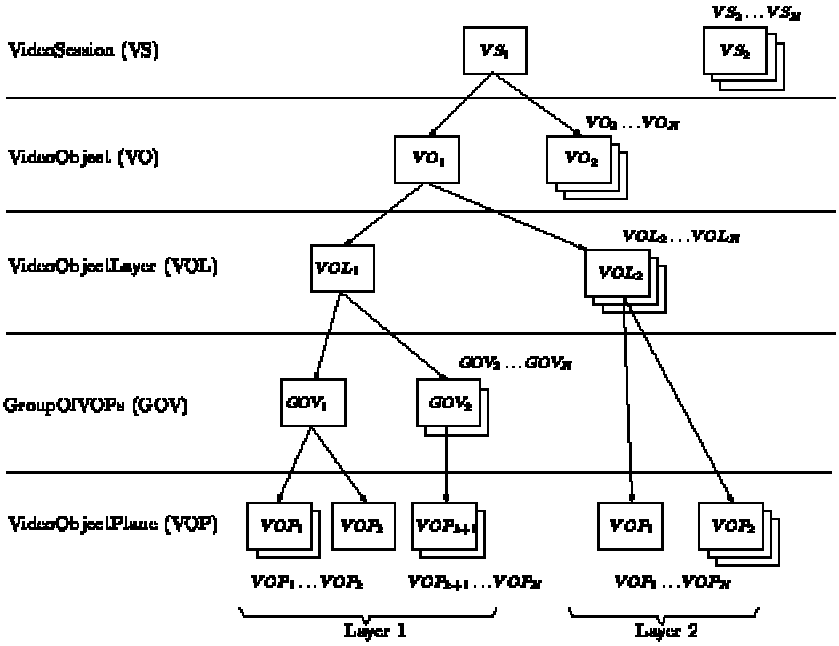


Figure 1. Illustration of shape-based retrieval

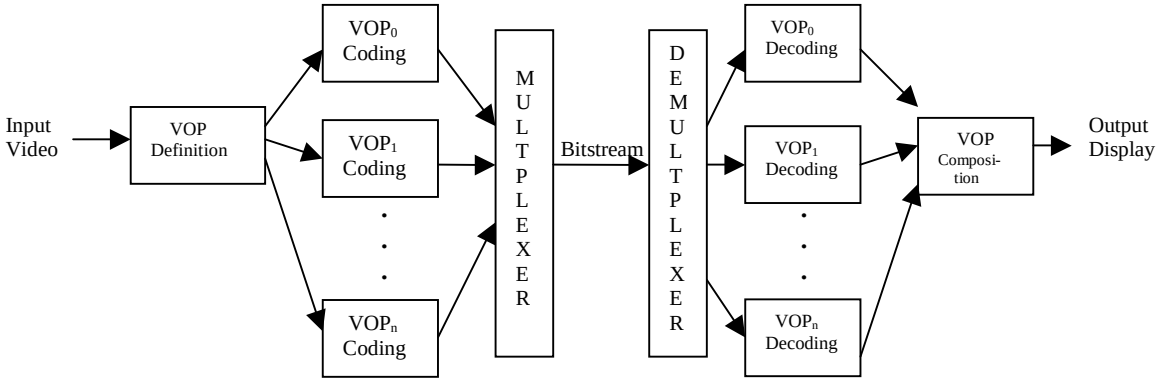
Content-based functionalities of MPEG-4 video

- MPEG-4—a standard for multimedia applications
- MPEG-4 video functionalities
  - Compression efficiency
    - Object coding providing subjectively better audio-visual quality
    - Coding of multiple concurrent data streams (as required for steroscopic and multiviews video applications)
  - Content-based interactivity
    - Content-based accessing of data using tools such as indexing, hyperlinking, querying, or browsing.
    - Separately decoding and manipulating video objects
    - Editing bitstream
    - Hybrid natural and synthetic data coding
    - Improving the temporal random access of frames and objets in video sequences
  - Universal access
    - Robustness in error-prone environments
    - Content-based scalability (temporal and spatial)

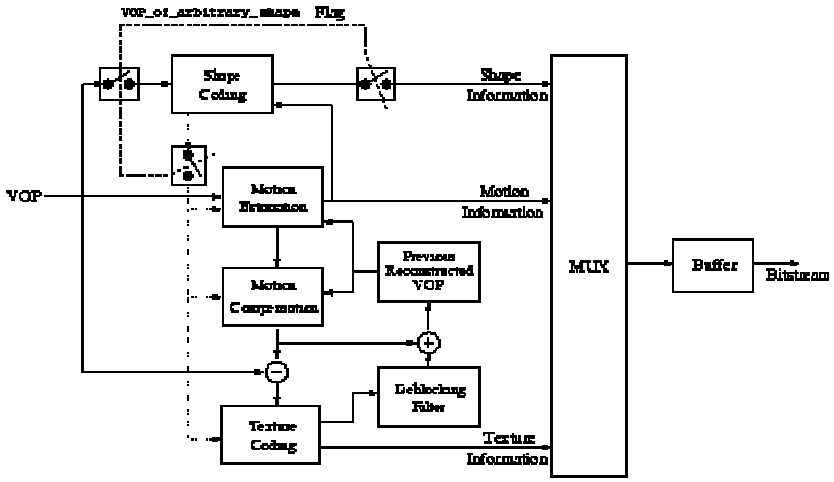
An overview of MPEG-4 video VM



a. MPEG-4 video bitstreams



b. MPEG-4 codec



c. MPEG-4 video encoder for the VOP class

Figure 2. Illustration of MPEG-4 video Verification Model

Where are these objects or VOPs from?

- VOP creation not specified in MPEG-4
- Effective VOP extraction algorithms do not exist
- VOP source for current MPEG-4 video encoders

  a. Manual extraction
  b. Semi-automatic
  c. Application-oriented—video conference
  d. Synthetic object—computer generated picture

Objectives

The proposed project aims at investigating and developing a novel and practical segmentation scheme to automatically extract objects from images and video sequences. The segmented objects

- Should be semantically meaningful objects which coherently correspond to real world objects recorded on the images and can eventually be encoded as VOPs in MPEG-4.
- Boundaries of each segment should be simple, not ragged, and must be spatially accurate.
- Should be able to attach MPEG-7 information about the objects like shape, color, texture, hyperlinks, URLs etc to facilitate automatic extraction of multimedia information.
- Should agree with what humans perceive.
- Should be able to handle birth/death of objects.

Image segmentation

- Feature-based segmentation
  - Segmenting image into pixel groups according to characteristics such as color, intensity, or hue.
  - Result: regions of coherent features, no symbolic meaning
- Physics-based segmentation
  - Identifying coherent regions of an image according to model of object appearance.
  - Goal: find regions that correspond to symbolic scene elements
  - Result: regions of coherent reflection models such as material, illumination, reflectance

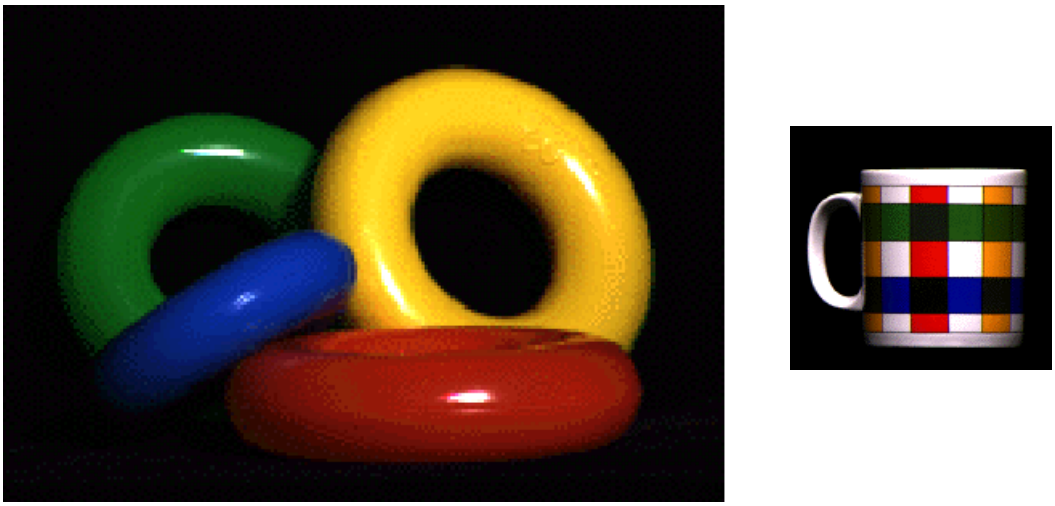


Figure 3. Illustration of difference between feature-based and physics-based segmentation

Techniques for image segmentation

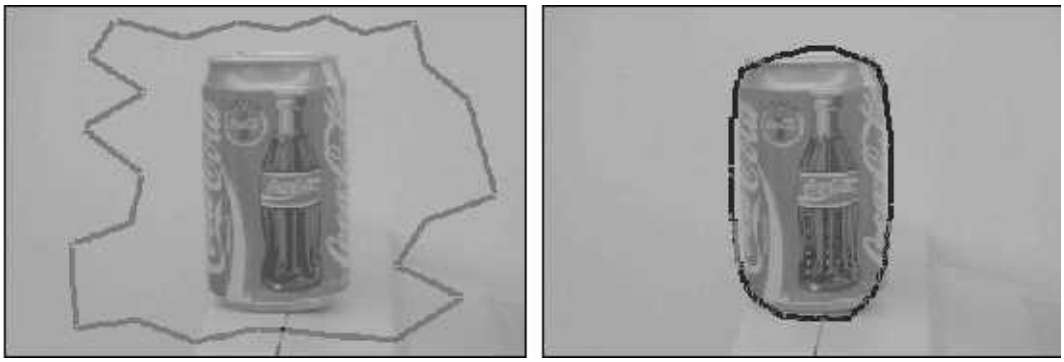
- Region-based segmentation

  Thresholding, Region growing, Texture analysis, Spectral analysis, Shadow, occlusion, highlights
- Boundary-based segmentation
  - Ridge detection
  - Edge detection
- Morphological segmentation

  Image simplification—marker extraction—watershed algorithm
- Bayesian segmentation

  $MAP(P(X|O))$

  $P(X|O) \propto P(O|X)P(X)$
- Model-based approaches
  - Active contour (snakes—physics-based)



  - Kalman-filter

Criteria of solving over-segmentation problem

- Domain-based or application-oriented approach (e.g. video conference, surveillance, traffic tracking)
- Combining high level information (e.g. lines)
- Combining motion information

Issues and problems relating to motion

- Rigid motion and non-rigid motion
- Apparent motion and real motion

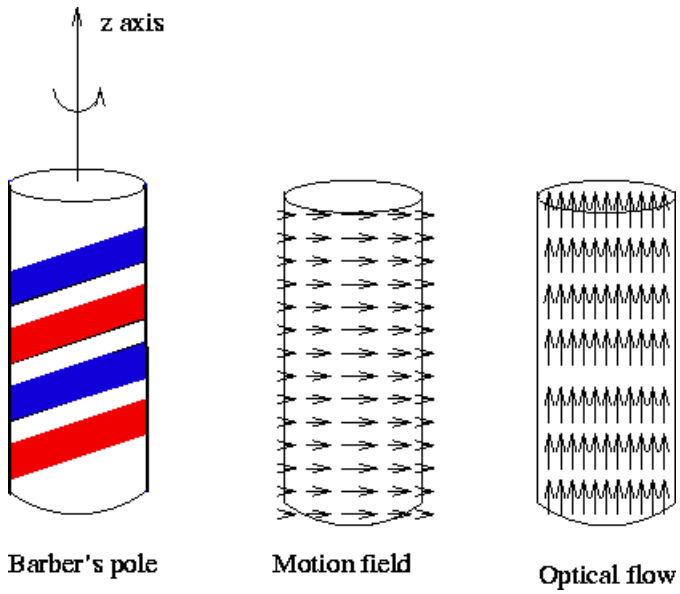


Figure 3. Illustration of apparent motion and real motion

- Local motion and global motion
- Correspondence problem
  - Optical flow—an ill-posed problem

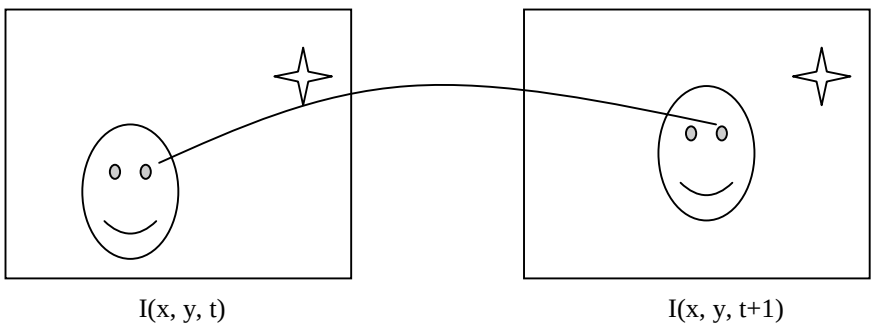


Figure 4. Correspondence problem

Video segmentation techniques

- Segmentation for robotic navigation and tracking of moving object
- Segmentation for second generation video coding or object-based video coding
- Segmentation for MPEG-4 VOP

Video segmentation techniques

- Segmentation for robotic navigation and tracking of moving object
  - Choosing data primitives to represent the scene
  - Correspondence establishment
  - Geometry estimation
  - Classification
  - Constrains or assumptions: rigid motion, MCSO or SCMO
  - Result: a sparse feature map or a fixed image window of the moving object

Video segmentation techniques

- Segmentation for second generation video coding or object-based video coding
  - Change detection
  - Affine motion estimation and intra-frame segmentation
  - Segmentation on coherent motion
  - Result: regions with different motion
  - Limitations: over-segmented, no accurate boundaries

Video segmentation techniques

- Segmentation for MPEG-4 VOP
  - Mosaicing—represent images with layers
  - Segmentation using relative depth
  - A semiautomatic approach
  - AMOS

VOP segmentation

- Mosaicing—represent images with layers

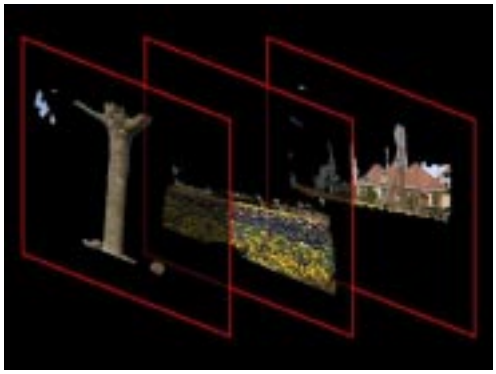
a. Original frames

b. Warped frames

c. Accumulated layer

Algorithm:

- Local motion estimation (optical flow)
- Segmentation by affine model fitting
- Layer synthesis



d. Layered representation

Figure 3. Representing of moving images into layers

VOP segmentation

- Segmentation using relative depth
  - Bottom level segmentation
    - Intra-frame – segmentation of image into region with homogenous grey level
    - Inter-frame— time recursive projection
  - Top level: relative depth estimation
    - Parametric motion estimation
    - Motion parameters comparison
    - Overlapping computation (depth estimation)
    - Depth level assignation

VOP segmentation

- A semiautomatic approach
  - Global motion estimation using optical flow and affine model –Background identification
  - Change detection using morphological filtering—IMO identification
  - IMO model initialization – Canny operator
  - Model update—object tracking using Hausdorff distance
  - VOP extraction—contour link and boundary approximation
  - Limitations: Single IMO, two different motions, semi-automatic, no motion no VOP

VOP segmentation

- AMOS--An active system for MPEG-4 Video object segmentation
  - Initial object identification by user
  - Creating of feature maps – edge, color and motion field
  - Region segmentation by fusion of the features
  - Affine motion estimation for the segmented regions
  - Object tracking by motion projection and inter-frame segmentation

A scheme for segmenting image sequence into semantically meaningful objects

- Pre-processing-- smoothing, morphological filtering, gamma correction
- Intra-frame segmentation-- edge detection, morphology, and region growing
- Inter-frame segmentation-- motion estimation, occlusions, relative depth
- Shape and VOP extraction-- snakes

The proposed scheme intends to improve current techniques in two aspects:

(1) Automatic segmentation

(2) Applicable to more general purposes

Conclusion

- Solving the General Vision Problem is the Holy Grail of vision research.
- We are going to take a small step towards it.

References

[Maxwell96] Bruce A Maxwell *Segmentation and Interpretation Using Multiple Physical Hypotheses of Image Formation* Ph.D thesis 1996 Carnegie Mellon University (CMU-RI-TR-96-28)

[Koenen99] Rob Koenen *MPEG-4 Overview – (Seoul Version)* ISO/IEC JTC1/SC29/WG11 N2725

[MPEG98] Requirements Group *MPEG-7: Context and Objectives (version - 10 Atlantic City)* ISO/IEC JTC1/SC29/WG11 N2460

[BNC99] Paul Bocheck, Yasuyuki Nakajima and Shih Fu Chang *Real-time Estimation of Subjective Utility Functions for MPEG-4 Video Objects* Proceedings of the Packet Video'99 (PV'99), New York, USA, April 26-27, 1999

[Smith97] M. Smith *Reviews of Optic Flow, Motion Segmentation, Edge finding and Corner Finding* Technical Report TR97SMS1 Oxford University, UK

[Torr95] Philip Hilaire Sean Troo *Motion Segmentation and Outlier Detection* Ph.D thesis, University of Oxford, 1995

[Meier98] Thomas Meier *Segmentation for Video Object Plane Extraction and Reduction of Coding Artifacts* Ph.D thesis, Dept. Of Electrical and Electronic Engineering, The University of Western Australia, 1998

[ZC98] Di Zhong and Shih-Fu Chang AMOS: *An Active System For MPEG-4 Video Object Segmentation* International Conference on Image Processing, Oct. 1998, Chicago, Illinois, USA.

[Pardás97] M. Pardás *Relative depth estimation and segmentation in monocular schemes* Picture Coding Symposium, PCS 97, Berlin, Germany, Sept. 1997, pp.367-372

[WA94] John Y. A. Wang and Edward H. Adelson *Representing Moving Images with Layers* IEEE Trans. on Image Processing vol.3, no. 5, pp.625-638, Sept. 1994

[Diehl91] Norbert Diehl *Object-oriented motion estimation and segmentation in image sequences* Signal Processing: Image Communication 3, Elsevier, 1991, pp.23-56

[Pichumani97] Ramani Pichumani *Construction of A Three-Dimensional Geometric Model for Segmentation and Visualization of Cervical Spine Images* Ph.D thesis, 1997, Stanford University

[RHB99] Andre Redert, Emile Hendriks and Jan Biemond *Correspondence Estimation in Image Pairs* IEEE Signal Processing Magazing 16(3):29-46 May, 1999