Saliency-based video segmentation with graph cuts and sequentially-updated priors

Ken Fukuchi (2), Kouji Miyazato (2), Akisato Kimura (1), Shigeru Takagi (2), Junji Yamato (1)

(1) NTT Communication Science Laboratories, NTT Corporation (2) Okinawa National College of Technology

Notice!!! None of the authors cannot attend ICME2009 due to the policy of the affiliations related to the swine flu. If you have any questions and/or comments for this paper, please feel free to contact the corresponding author, Akisato Kimura <akisato@ieee.org>. Some demonstration movies can be seen in a web site http://www.brl.ntt.co.jp/people/akisato/saliency3.html. Sorry for the inconvenience.

Background:
• Extracting “important” regions from videos is a challenging and crucial task.
  ➢ Especially for video compression, object recognition, video annotation and retrieval, etc.
• It can be formulated as a problem of binary segmentation.
  ➢ Important regions = “objects”, the remaining regions = “backgrounds”
• A promising way for precise segmentation: graph-cuts based methods
  ➢ Interactive Graph Cuts [Boykov 2006], extension to videos [Kohli 2007], etc.

Problem 1: Need to provide cues for segmentation manually and carefully.

Problem 2: Segmented regions may be randomly switched as a result of the shifts of attention.

Contributions 1: Segmentation priors are provided based on visual saliency.

Contributions 2: Sequential update of priors with previous results.

Segmentation with Graph Cuts
Minimizing the energy function is equivalent to deriving the minimum cut of the following (directed) graph:

The MAP solution of the MRF for segmentation can be exactly solved in polynomial time.

Attention-based priors (Contribution 1)

Prior term
\[ \xi(A_k) = -\log p(A_k) \]

Prior density for “obj”
\[ p(A_k = \text{"obj"}) \]

Prior density for “bkg”
\[ p(A_k = \text{"bkg"}) = 1 - p(A_k = \text{"obj"}) \]

Likelihood term
\[ \phi(D|A_k) = -\log p(D|A_k) \]

Likelihood for “obj”
\[ p(D|A_k = \text{"obj"}) \]

Likelihood for “bkg”
\[ p(D|A_k = \text{"bkg"}) \]

Sequential update of priors (Contribution 2)
Assume the following two relationships:

\[ p(A_{k-1} = \text{"obj"}; t) = N(\sigma_1, \sigma_2) \]

\[ p(A_{k-1} = \text{"bkg"}; t) = N(\sigma_1, \sigma_2) \]

The prior density \( p(A_k = \text{"obj"}; t) \) at time \( t \) can be derived through Kalman filter, where the observation is \( g(A_k) = \text{"obj"}; t) \)

\[ p(A_k = \text{"obj"}; t) = \frac{\sigma_2^2(t) \phi(A_{k-1} = \text{"obj"}; t)}{\sigma_1^2(t) + \sigma_2^2(t) \phi(A_{k-1} = \text{"obj"}; t)} \]

Special thanks to Mr. Derek Pang of Simon Fraser Univ., Canada for supporting this presentation.