A Unified Approach for the Detection of Gradual Transitions in Historic Material

Markus Seidl

Problem description and motivation

Task: Find all shot boundaries in a video.

Shot boundary types:
- Abrupt (cut)
- Gradual (GT)
- Fade in / Out
- Dissolve
- Wipe

State of the art
- SBV is (was?) an active research field
- TRECVid Conference: Yearly evaluation of approaches
- TRECVid 2007 GT detection mean:
  - frame precision (fp) = 0.82; frame recall (fr)=0.63

Approaches
- Unified: One detector for all GT types
- Specialized: One detector for each GT type

Main challenges
- Camera / object motion
- Illumination changes
- Camera / object motion

Method

Input
- Single frames
- Ground truth for cuts

Feature Extraction
- Single frames, GT cuts

Future Features
- Global luminance histogram
- Local luminance histograms
- Global edge histogram
- Local edge histograms
- Local DCT Coefficients

Construction of continuity signal

Intermediate feature creation

Classification
- SVM Classification
  - Class label 1 for transitional frames
  - Class label 0 for non-transitional frames

Post processing

Results

Experimental Study
- Three different training data sets
- Systematic evaluation of different configurations
  - Features
  - Fusion strategies
  - Post processing

Best results

Historic material
- Single Features
- DCTCoeffs, Global EdgeHist, Local LumHist combined with late fusion; Median filter with a window size of 31
- Contemporary reference material:
  - TRECVid SBD task 2007
  - fp=0.44; fr=0.88

Best features different
- Local EdgeHist vs. Global LumHist

Contemporary material

Best features

Lessons learned
- Best fusion strategy for combination of features: Late fusion
- Larger intermediate feature creation kernel sizes
- Assumption: Due to longer gradual transitions
- Median filter: Better fr and fp
- Begin-end matching and KLT verification: No improvement
- Contemporary material:
  - Luminance features better than colour features (!)