8.5 Application Examples

8.5.1 Genre Recognition

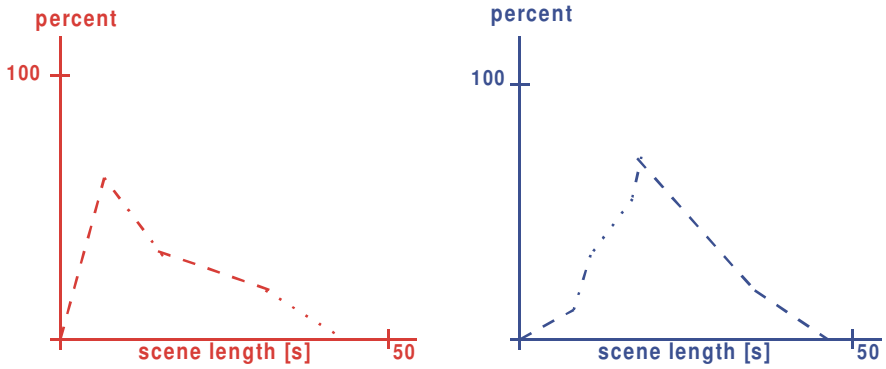
Goal

Assign a genre to a given video, e.g., movie, newscast, commercial, music clip, etc.)

Technology

Combine many parameters of the physical level to compute a characteristic signature or „fingerprint“ of the video

Example: Distribution of Scene Lengths



Observation

From the scene length distribution alone, we can distinguish a music clip from a newscast on TV.

Experimental Results

Experiment done at the University of Mannheim in 1995:

- 140 video clips from seven genres: newscast, soccer, tennis, talk show, music clip, cartoon, commercial
- The software prototype classified between 87 % (commercials) and 99 % (newscast) of the clips correctly.

Problem

The computation time for a video of 3 min was about 28 h on a Sun SparcStation 20! Reason: many different parameters had to be computed for each frame.

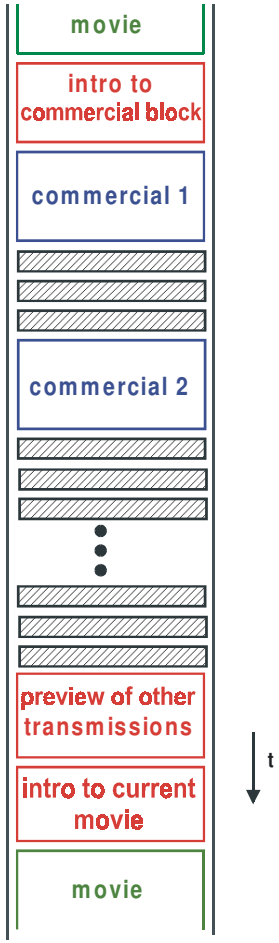
8.5.2 Finding Commercials in a Video

Motivation

- A company wants to find out if a TV station has sent all the commercials as contracted.
- A company wants to watch what the competition is doing.
- Correlate measurable parameters with the success of a commercial (colors, motion intensity, audio, etc.)
- Remove unwanted commercials from a video stream

Properties of a Block of Commercials on TV

Structure of a commerical block



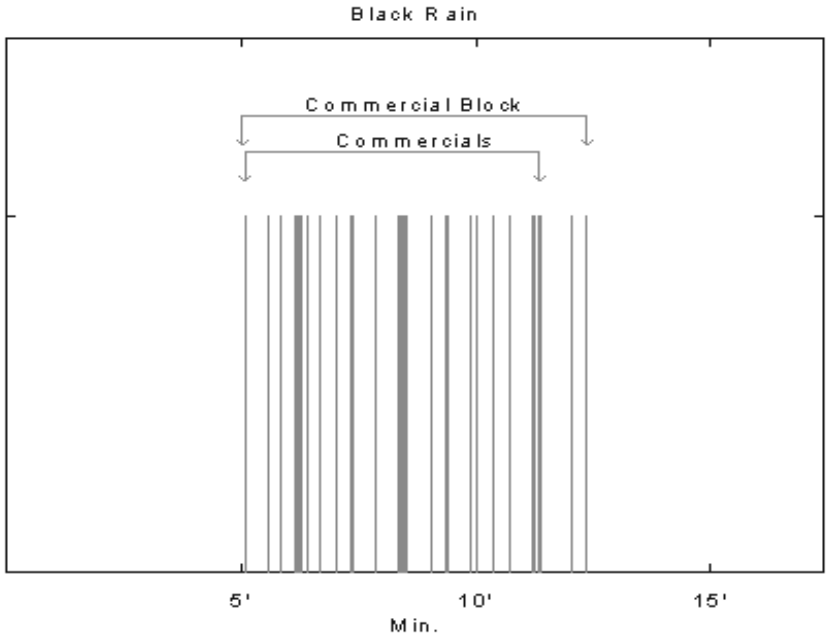
Locating Commercials in a Stream

- Locate dark monochrome frames
- Find high levels of action (a high ECR, many long motion vectors)
- Find high frequencies of hard cuts.

Dark Monochrome Frames

Example

A movie on TV interrupted by a block of commercials



Locating Commercials in Two Steps

Step 1: Locate the range of the commercial roughly

Use monochrome frames and frequent hard cuts as a pre-selector to find candidate ranges

Step 2: Determine the borders of the commercial block

Compute motion vectors and ECR to precisely determine the first and last frame of the block.

Experimental Results

Material tested

Four hours of video recorded from German television in 1996 (only movies with embedded commercial blocks)

Result

All blocks of commercials were detected correctly.

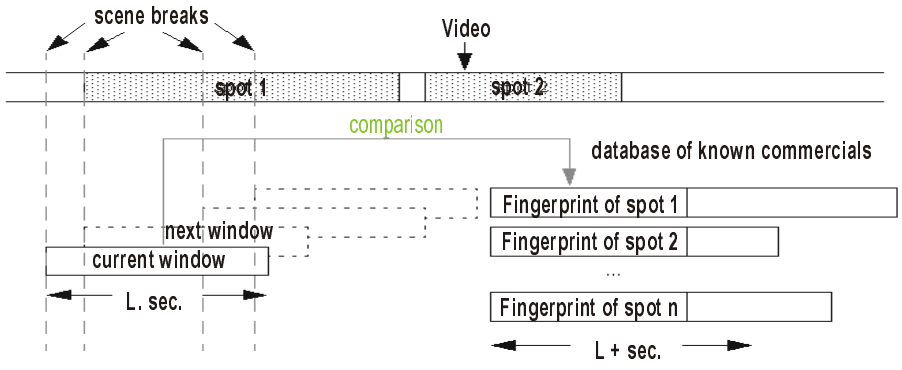
Recognizing Known Spots

Algorithm

- Compute a „fingerprint“ for each spot (a feature vector defined by a selection of basic parameters)
- Create a database of fingerprints of known commercials
- Compare the fingerprint of a commercial found in the stream under investigation with the fingerprints contained in the database

Quick Compare

- Jump to the next hard cut in the video
- Compute the fingerprint from the next *L* seconds
- Compare with the fingerprints in the database



Experimental Results

Tested with 200 clips in addition to those already stored in the database..

Result

- All known commercials were recognized
- There was no false hit
- On the average the difference between the exact begin and end of a block of commercials and the begin and end computed automatically was five frames.

8.5.3 An Intelligent Alarm System

Experiment

- Installation of a camera to observe a room
- Computation of motion vectors for the frame sequences delivered by the camera
- When motion is recognized the recording function is activated, the digital video written to disk
- **The system is programmable!** For example, when surveying a bedroom in a hospital, people dressed in white do not cause an alarm.

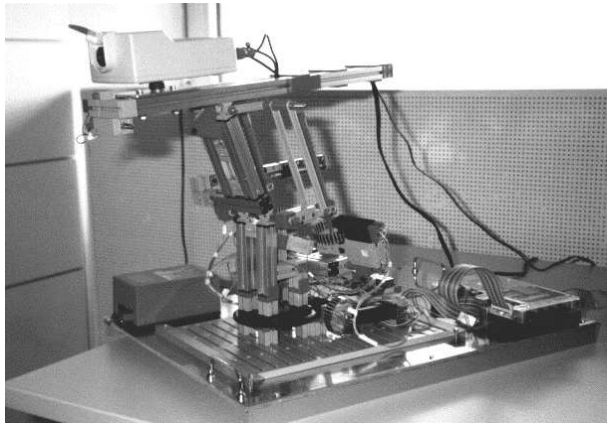
Application examples

- Protection of bedrooms in hospitals from thieves
- PC pools at universities
- Banks
- Construction sites

8.5.4 Object Tracking

Experiment

Camera robot **CaRo** at the University of Mannheim:



Algorithm

Segmentation of objects seen by the camera by edge detection and color regions. In the experiment, a red radio-controlled car moved around on a gray floor.

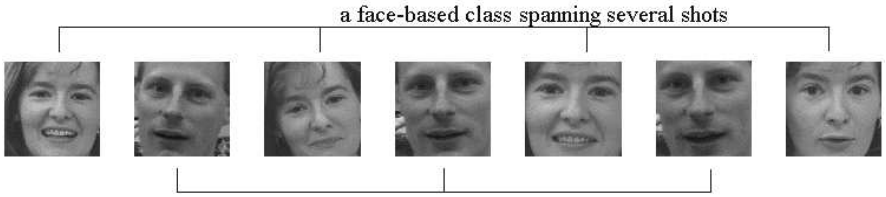
Application examples

- Follow a speaker moving around on a podium
- Trace objects in sports events, such as car races
- Traffic surveillance

8.5.5 Dialog Recognition

Algorithm

- Locate faces in the video with a neural network
- Compute a fingerprint for each face
- When faces occur in an A, B, A, B, A... Pattern a dialog is recognized



8.5.6 Video Abstracting

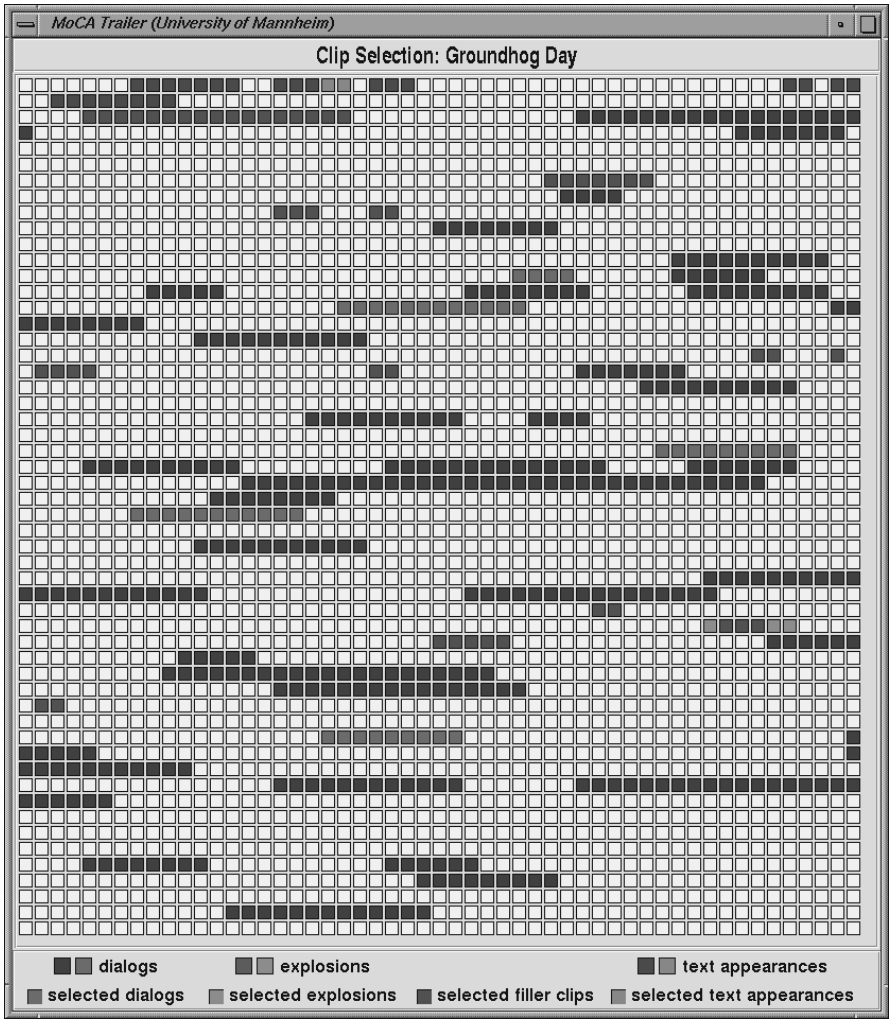
Goal

Creation of a short version of a video ("moving abstract") from the full video. The abstract should capture as much as possible of the original message of the video.

Algorithm

- Decompose the full video using cut detection
- Characterize the shots with feature vectors (a set of parameters extracted from the video), e.g.:
  - Motion intensity
  - Color content
  - Faces present?
  - Explosion present?
  - Dialog?
- Define heuristics for the priorities of shots for the abstract
- Shorten all shots to a maximum length (e.g., 6 s)
- Compose the video abstract from the shots selected, try to create smooth transitions, in particular for audio

Example for a Video Abstract



Automatic Detection of Violence or Pornography?

With the algorithms we have today the automatic detection of objectionable content, such as violence or pornography, is not possible. The semantics of “objectionable“ are so difficult that an automatic system does not work.