



**PROFESSUR  
MEDIENINFORMATIK**

## **05 Video Retrieval**

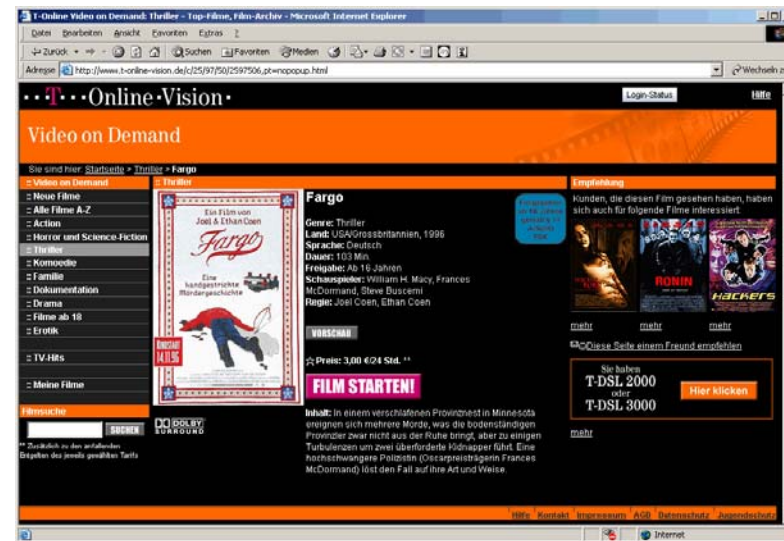
Lecture Media Retrieval

Maximilian Eibl, Medieninformatik, TU Chemnitz



# Einsatz

- Bildarchive von Fernsehanstalten
- Lokalfernsehen
- E-Learning
- Video-on Demand



Kiron Filmarchiv

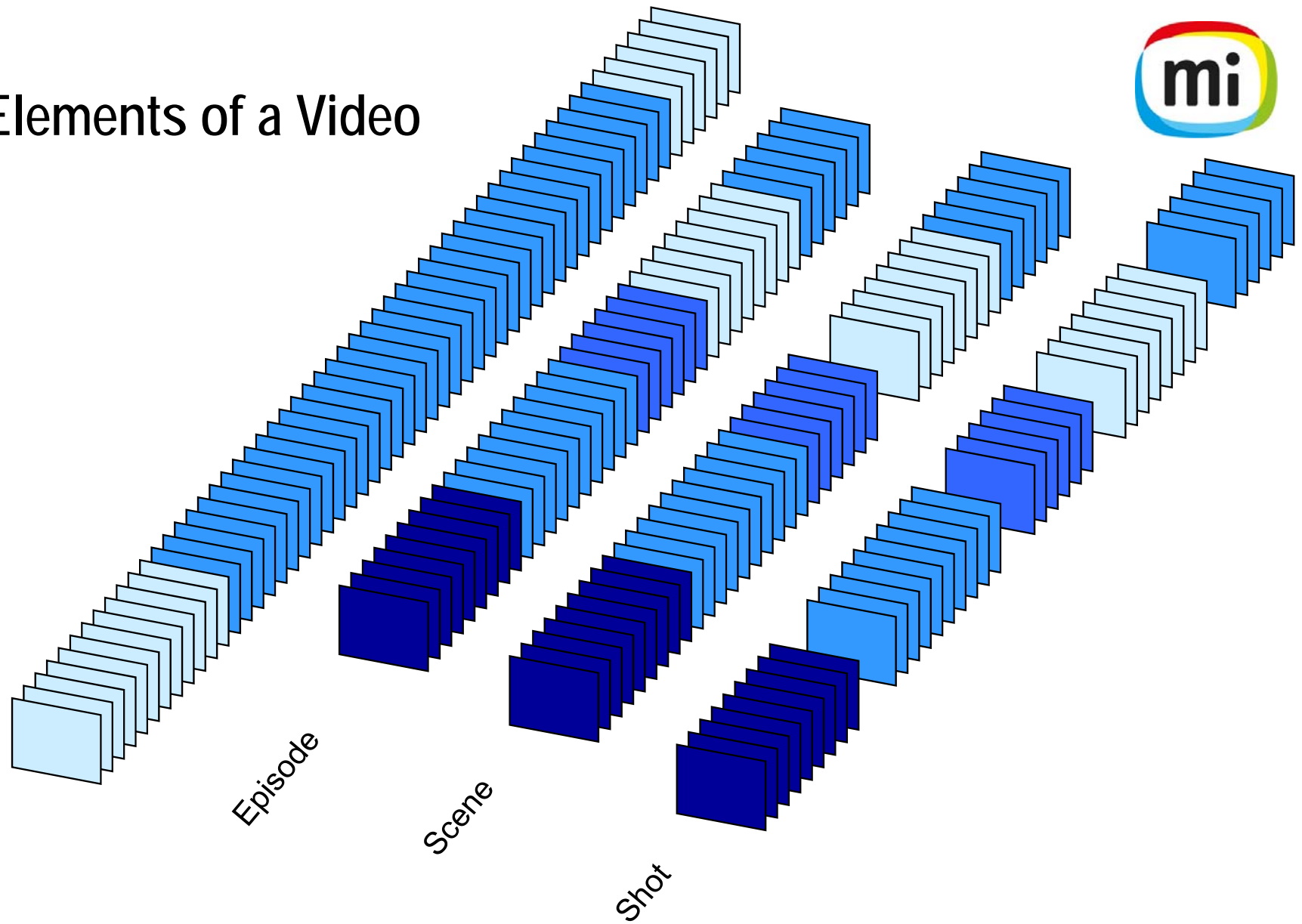
[http://www.sueddeutsche.de/wirtschaft/artikel/4837476/zoom\\_0\\_0/](http://www.sueddeutsche.de/wirtschaft/artikel/4837476/zoom_0_0/)



## Elements of Content Based Video Retrieval

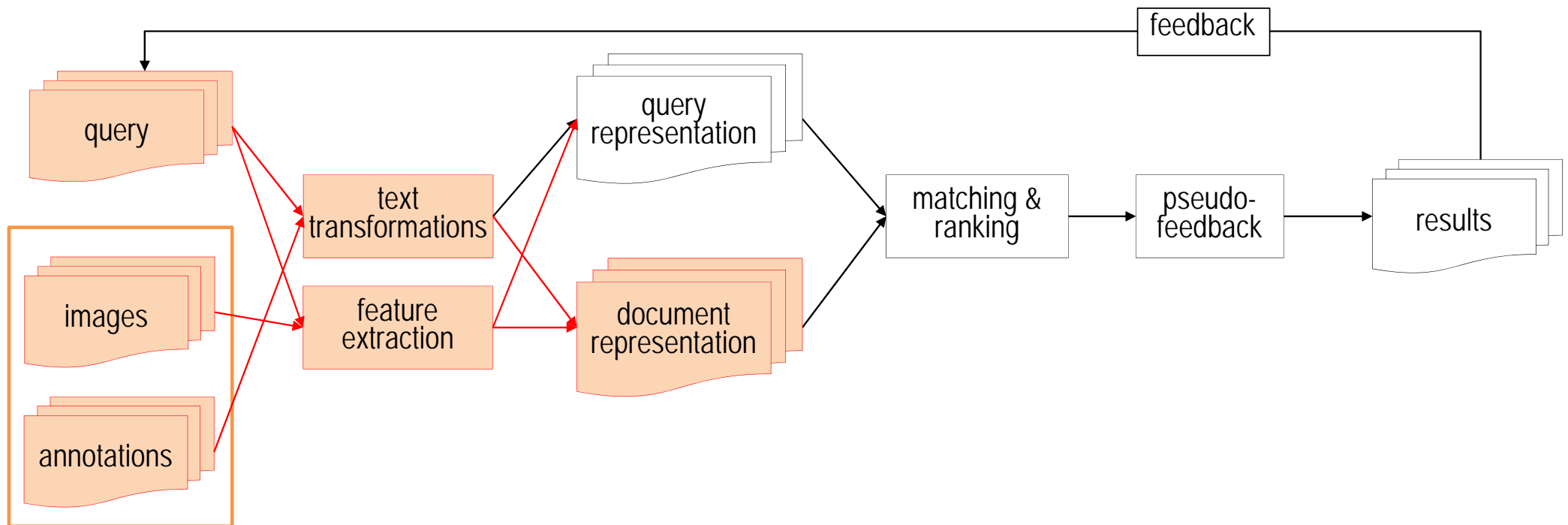
- Elements of CBIR
- Segmenting
- Key Frames
- Camera movement
- Object tracking
- Text layer

# Structural Elements of a Video

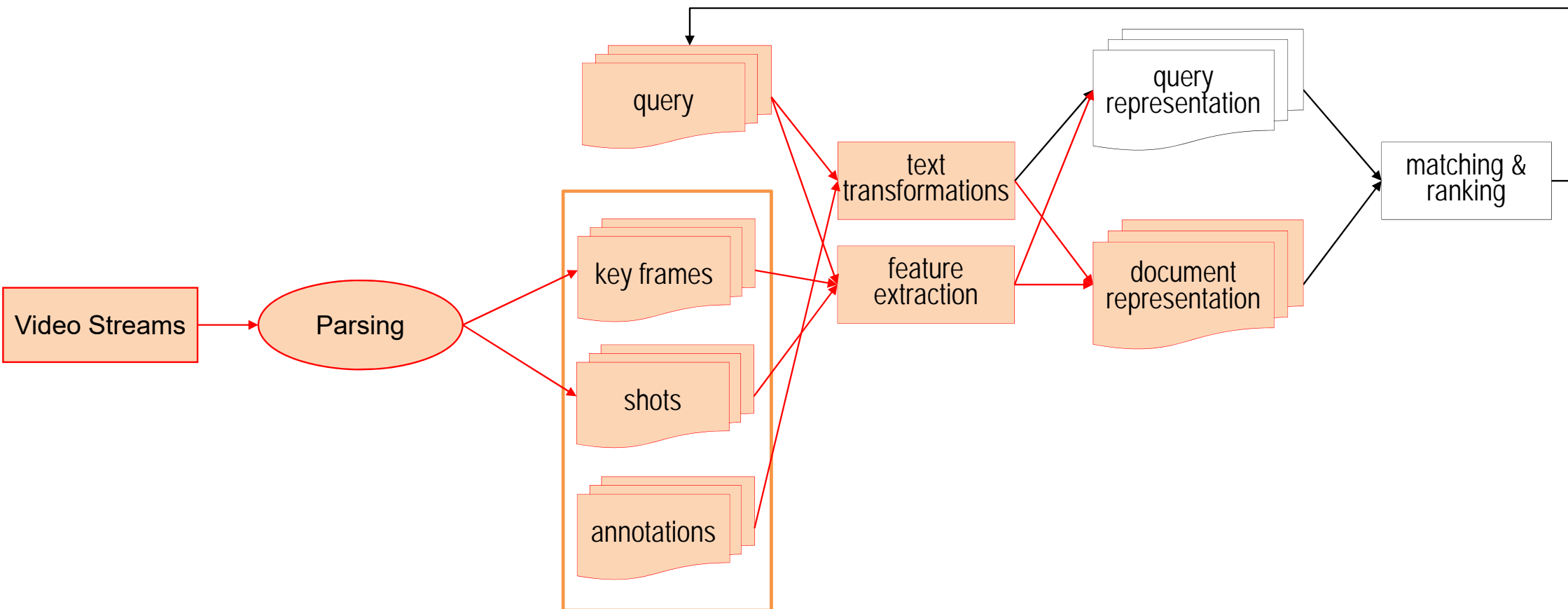




# IRS Components: Image Retrieval



# IRS Components: Image Retrieval





# Segmenting

- Hard cuts smooth transition

Hard cut



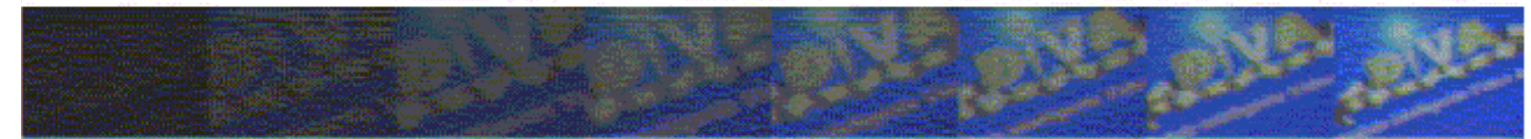
Dissolve



Wipe



Fade-In





## Problems concerning Segmenting

- Smooth transitions
- Fast objects
- Fast camera movements
- Picture noise
- Overlay in TV news like interview / stock ticker

"The ideal metric for automatic video partitioning will be able to differentiate between the following three image changes: (1) shot change, either abrupt or gradual, (2) motions, introduced by both camera operation and object motion; and (3) luminosity changes and noise." (Zhang 2005:29)





# Segmentierung: Harte Schnitte I

- Intensity change

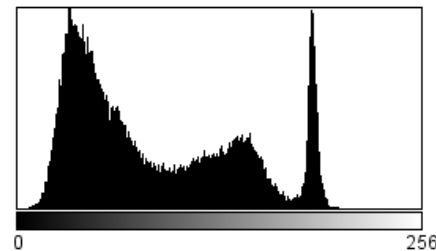
$$D_{cut} = \sum_{x,y} |I(x, y, t) - I(x, y, t+1)|$$

- Histogram change

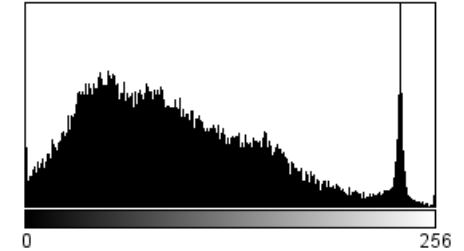
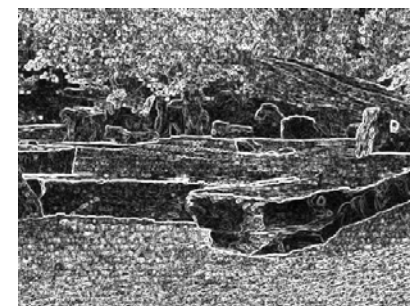
$$D_{cut} = \sum_j |H(j, t) - H(j, t+1)|$$

- Edge Change Ratio ECR

$$ECR_1 = \max \left( \frac{E_{in}}{S_t}, \frac{E_{out}}{S_{t+1}} \right)$$



Count: 49920      Min: 0  
Mean: 87.432      Max: 230  
StdDev: 53.690      Mode: 32 (761)



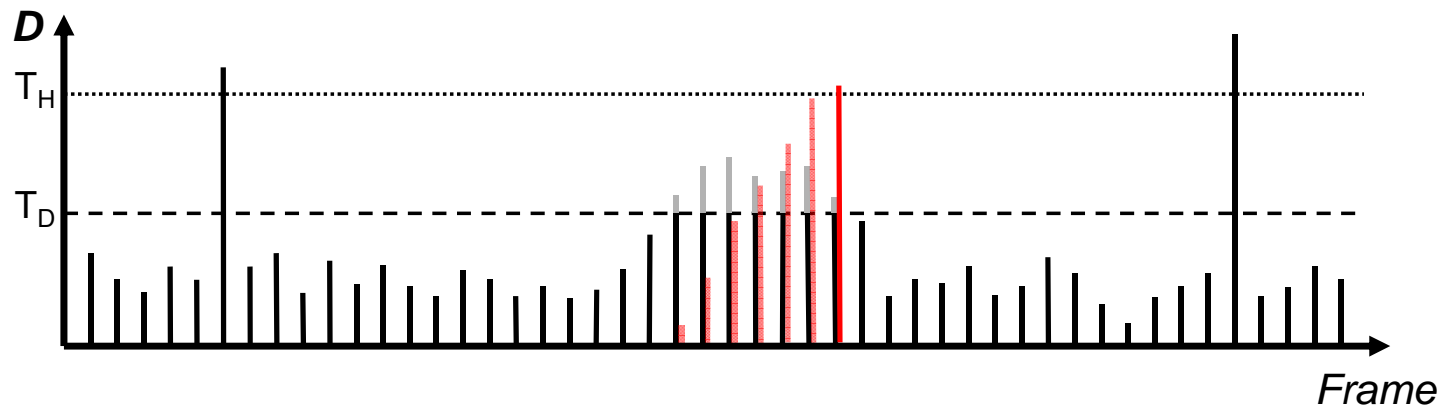
Count: 48573      Min: 0  
Mean: 96.502      Max: 255  
StdDev: 59.454      Mode: 234 (816)



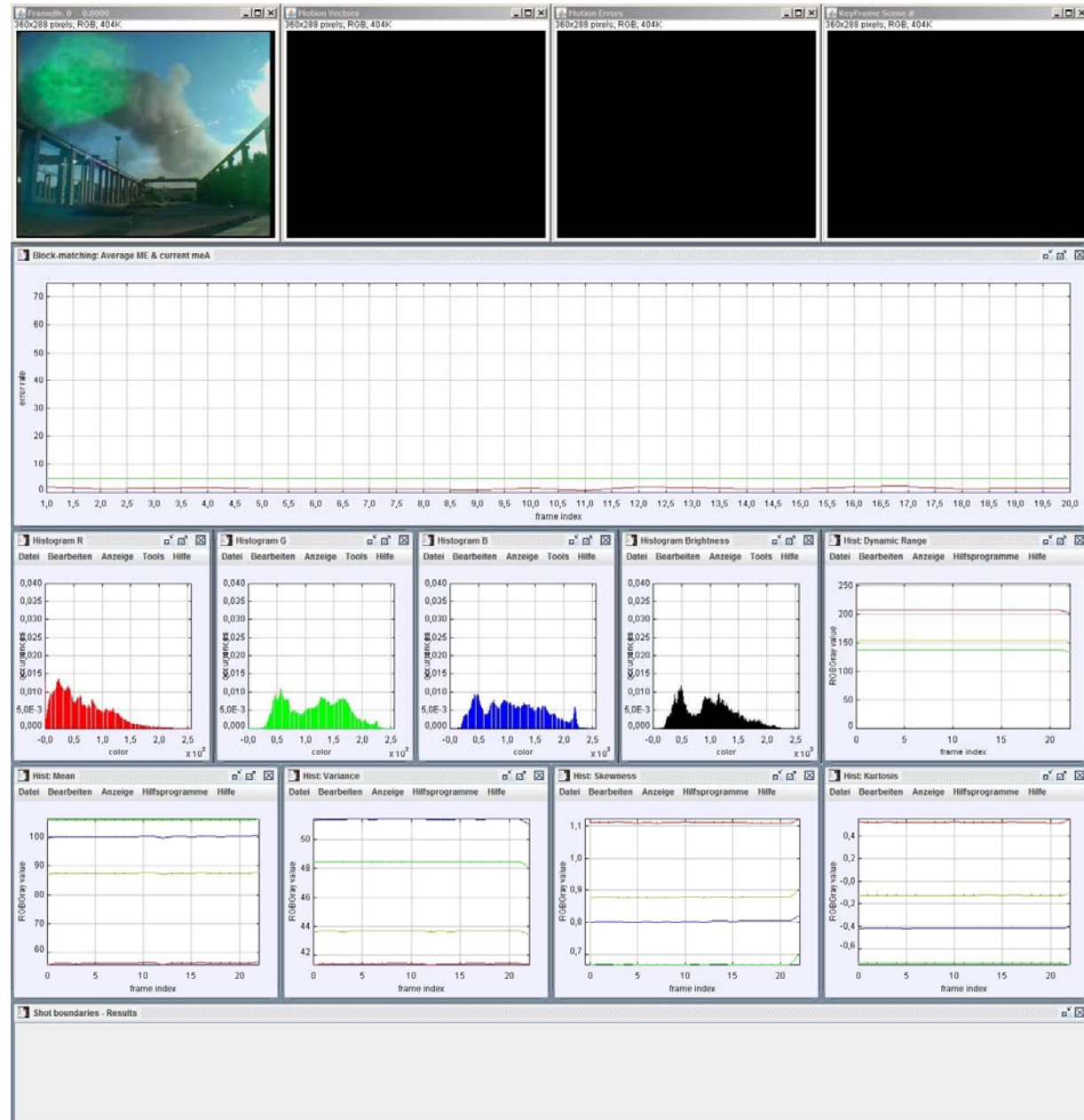


## Transitions: *twin-comparison*

- Comparison 1: Difference of two consecutive frames
- Comparison 2: Accumulated difference over a frame sequence
- Differenzwert: 
$$D_{cut} = \sum_{x,y} |I(x, y, t) - I(x, y, t + 1)|$$










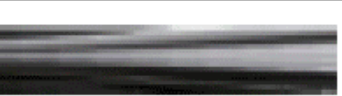


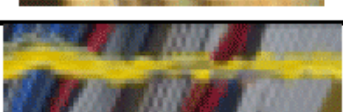

# Videoretrieval: Beispiel *shot detection*

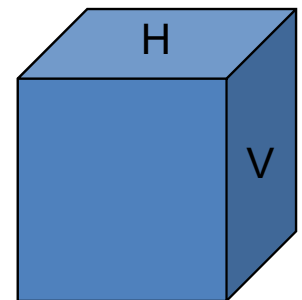




# Shot based Representation I: Global motion feature

- Representation of global / dominant movements like:
  - Zoom
  - Camera tracking
  - Camera pan
  - ...

Motion type	Horizontal Slice	Vertical Slice
<i>static</i>		
<i>pan</i>		
<i>tilt</i>		
<i>zoom</i>		
<i>object motion</i>		
<i>tracking</i>		





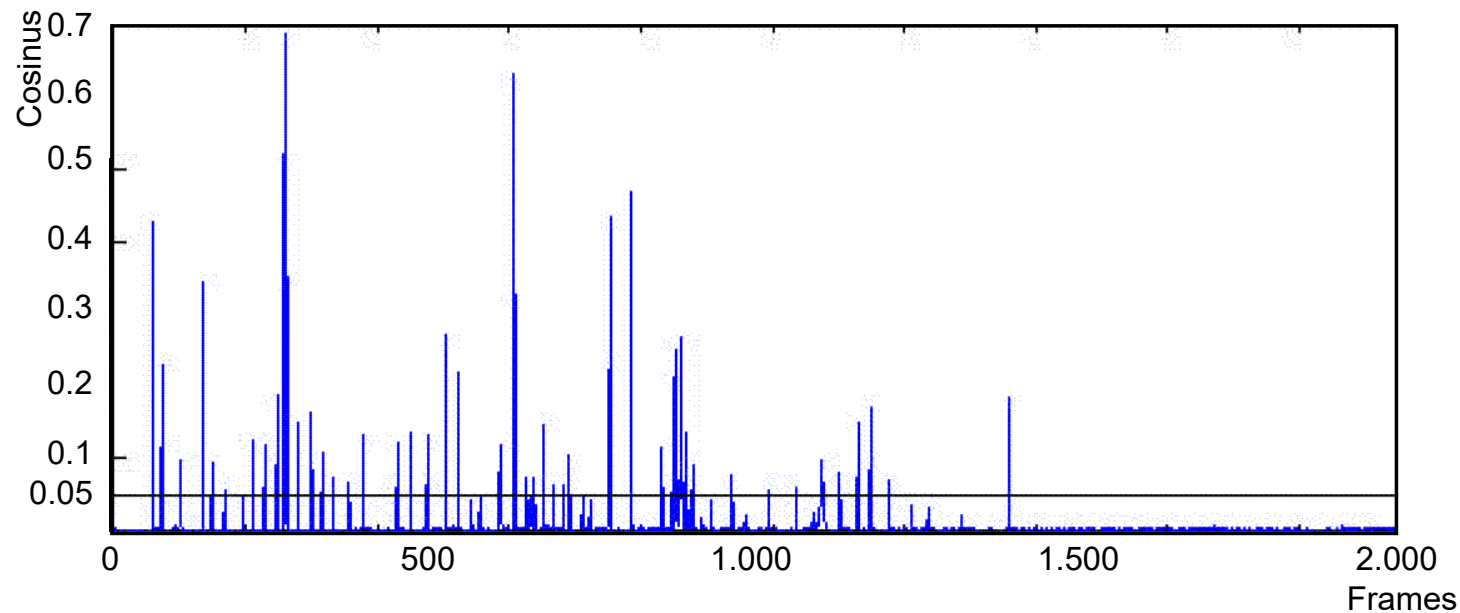
## Key Frames

- Extraction of central frames
- Extraction of representative frames
- Composition of representative frames
- Construcción of representative overview images (mosaicing)

## Study: O'Toole et al. 1999

- Basis: Histograms
- Comparison: Cosinus
- Material: 8 h TV-Program

$$D_{\cos}(A, B) = 1 - \frac{\sum_i^N (a_i \cdot b_i)}{\sum_{i=1}^N a_i^2 \cdot \sum_{i=1}^N b_i^2}$$





## Study: O'Toole et al. 1999

Video Type	# of Frames	# of Cuts	# of Gradual Transitions	Ratio of Cuts to Gradual Transitions
News and Weather	134.540	598	69	9:1
Soaps	144.958	909	94	10:1
Cookery Programs	37.370	188	42	4:1
Magazine / Chat Shows	134.985	759	64	12:1
Quiz Shows	29.093	269	4	67:1
Documentary	7.494	47	23	2:1
Comedy / Drama	110.618	839	72	12:1
Commercials	106.976	1.771	415	4:1
Total	706.034	5.380	779	15:1



## Study: O'Toole et al. 1999

	Total # of shot boundaries	# correctly identified	# falsely identified	# missed	Recall	Precision
Threshold 1 (0.010)	6.159	5.689	3.775	470	.92	.60
Threshold 2 (0.020)	6.159	5.472	1.504	687	.89	.78
Threshold 3 (0.035)	6.159	5.163	731	996	.85	.88
Threshold 4 (0.060)	6.159	4.508	431	1.651	.74	.92
Threshold 5 (0.15)	6.159	2.789	195	3.370	.45	.94





## Study: O'Toole et al. 1999

Type	Threshold	Recall	Precision
Commercials	3	.79	.74
Soaps	3	.92	.76
News	2-3	.86-.87	
Cookery	1&2	.85-.90	.35-.50
	3	.83	.71
	4&5	<.50	
Magazine	1-3	.78-.98	
Quiz	1	.98	.78
	3	.97	.97
	4-5	<.55	
Comedy / Drama	1-5		>.85
	4	.88	
	5	<.50	
Documentary	3	.09	.60
	1	.64	.52



## Study: Boreczky&Rowe 1995

- Untersuchte Methoden:
  1. Histogram
  2. Region Histogram
  3. Running Histogram
  4. Motion Compensated Pixel Differences
  5. DCT Coefficient Differences
- Vergleichsmaterial

Video Type	# of Frames	# of Cuts	# of Gradual Transitions
Television	133.204	831	42
News	81.595	293	99
Movies	142.507	564	95
Commercials	51.733	755	254
Miscellaneous	10.706	64	16
Total	419.745	2.507	506



# Study: Boreczky&Rowe 1995

