# Video Rushes Summarization by Adaptive Acceleration and Stacking of Shots

Marcin Detyniecki Christophe Marsala

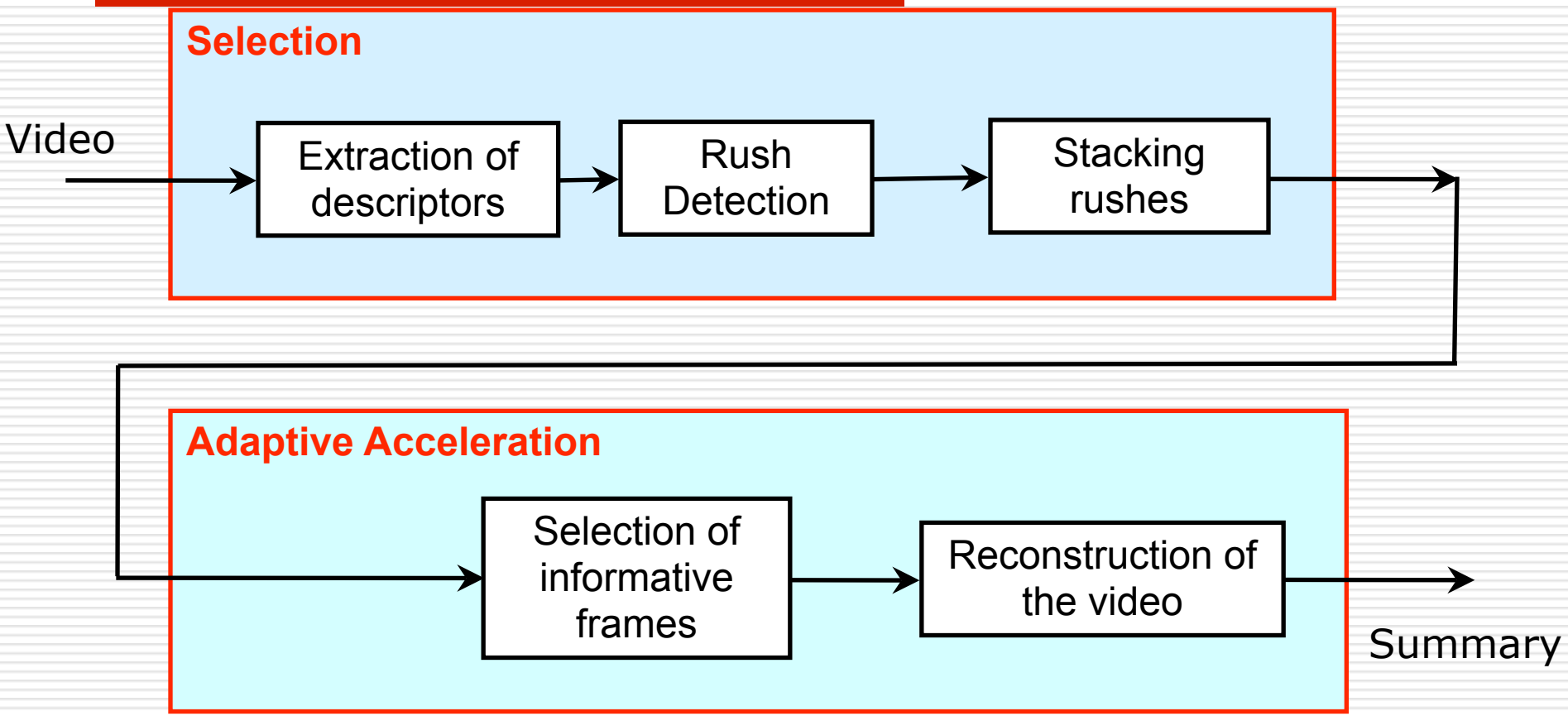
Laboratoire d'Informatique de Paris 6 (LIP6)

Université Pierre et Marie Curie – CNRS

Paris, France

TRECVID BBC Rushes Summarization Workshop - 2007

# Our approach



# Pre-treatment

- Sampling of the video
  - Only a small number of frames are kept
- All the works is done on a subset of frames
  - We have tested 5 frames/sec.
  - Study on devel video
    - All parameters of the proposed methods are linked
- A set of descriptors will be extracted from each frame

# Selection: frame descriptors

Global Histogram

4 Area Histograms

Several kinds of similarities between frames

# Selection: shot detection

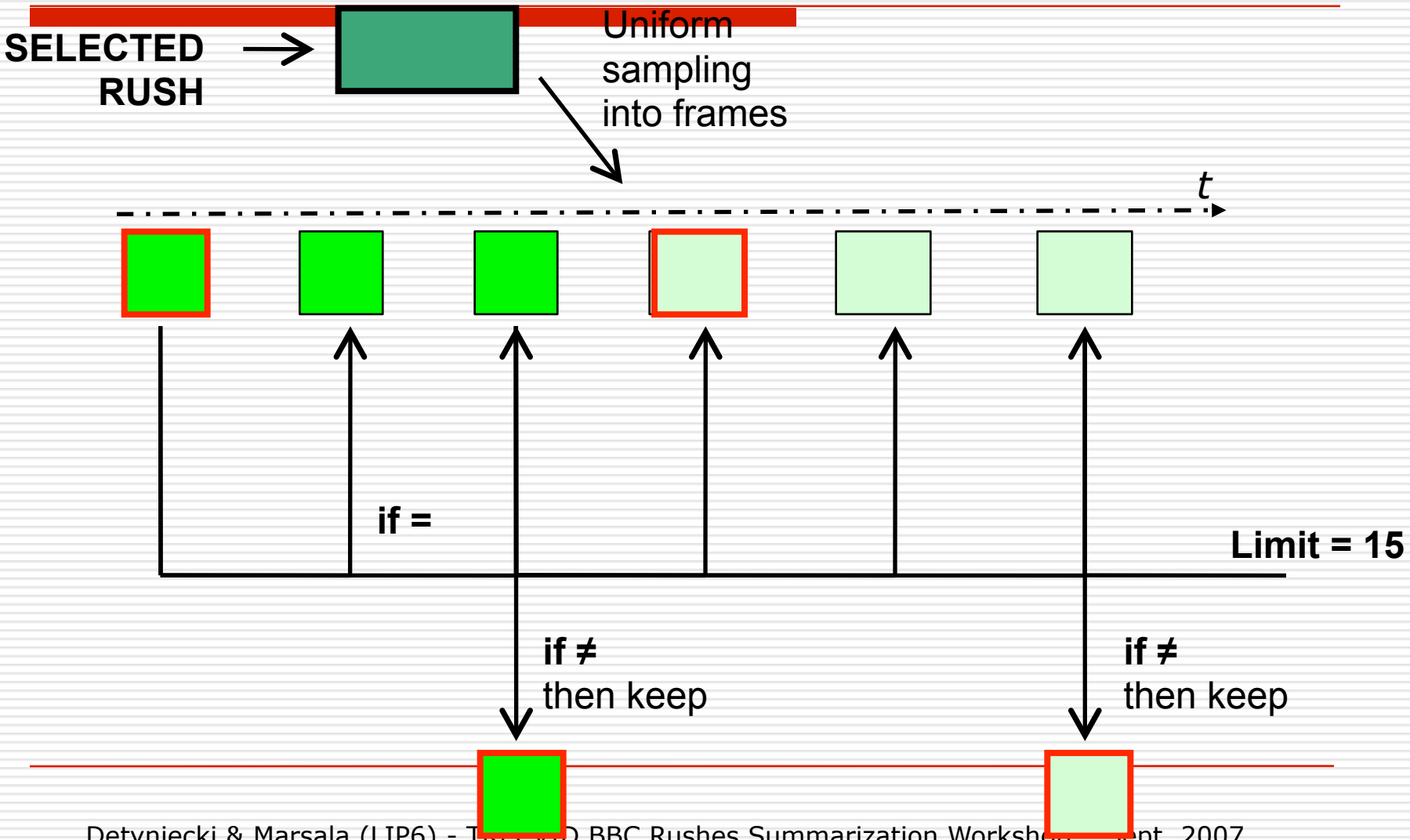

F5 beginning of a new shot:
sim1 > mean(sim2,sim3,sim4) + *Threshold*

# Selection: shots and stacking

# Adaptive Acceleration

- Idea: same as browsing a video
  - normal play of interesting sequences
    - Goal in a soccer video
    - Action scenes in movies
  - fast forward of uninteresting sequences
    - Break time in soccer video
    - Transitional scenes in movies
  - Skimming
- Interestingness
  - Information changes in the frames
  - Same frames => no additional information

# Selecting informative frames



# Example

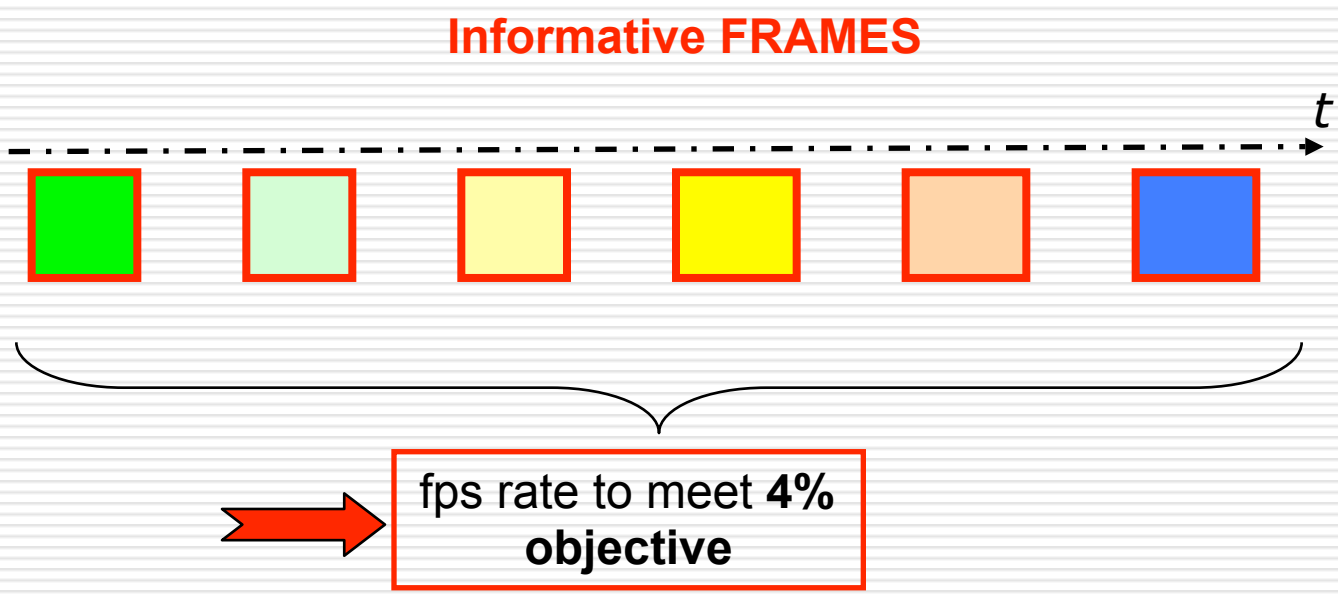
$t$

Action scene with moving persons

Static scene with rare changes

.....

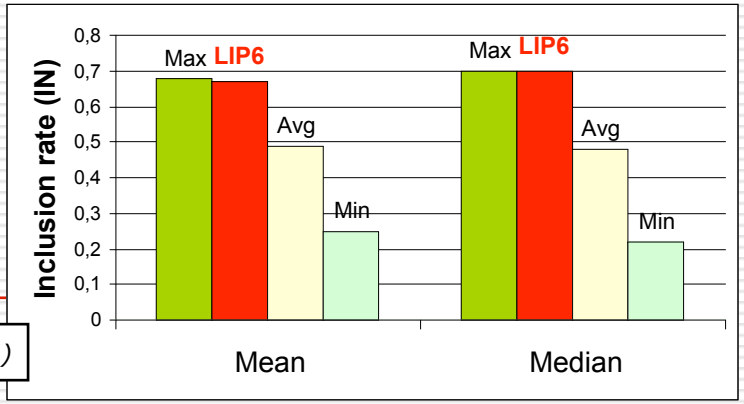
# Reconstructing the video




Given the number of informative frames and the required duration of the video, it is easy to reconstruct the summary

# Discussion (IN)

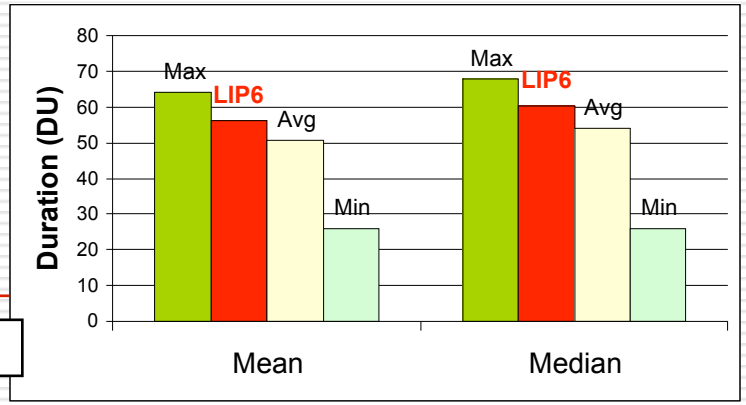
(0-1)




- Inclusion rate
- Main advantages
  - Stacking preserves greatly the information
    - Only duplicate shots are forgotten
  - Acceleration reduces redundancy
    - Does not greatly modify the informative content
    - Only a visual alteration is done not a loss of information

# Discussion (DU)

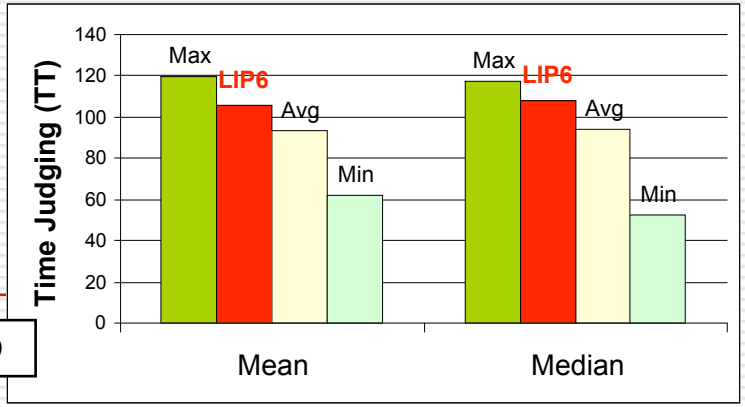
*(in seconds)*

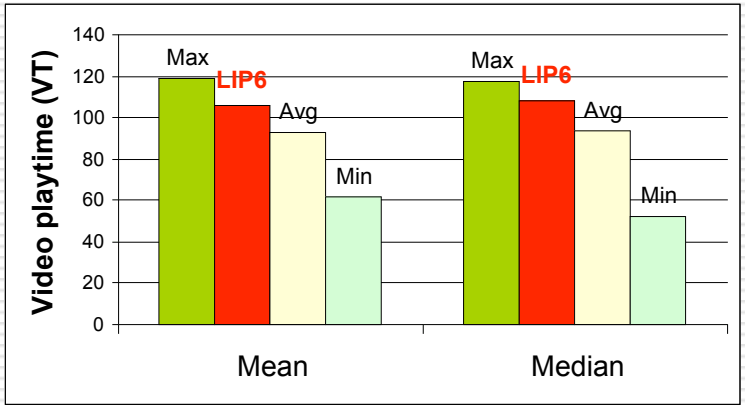



- Duration
- 4% duration constraint
- (XD) very low *(should be 0?)*
  - Rounding value
  - Video tool used: ffmpeg

# Discussion (TT/VT)

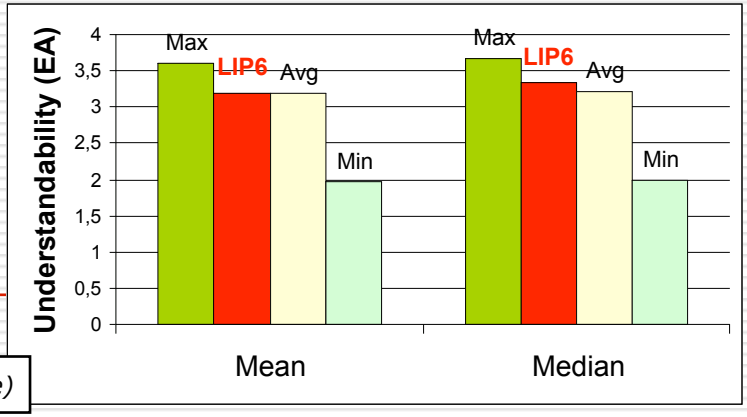
(in seconds)



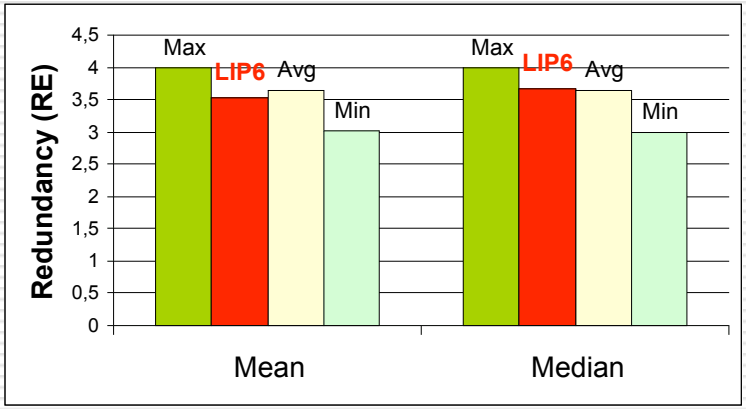




- Time judging
  - Total time spent (TT)
  - Video play time (VT)
- Depends on
  - Duration of summaries
  - Understandability
  - Number of inclusions [to be] detected (?)

# Discussion: (EA / RE)

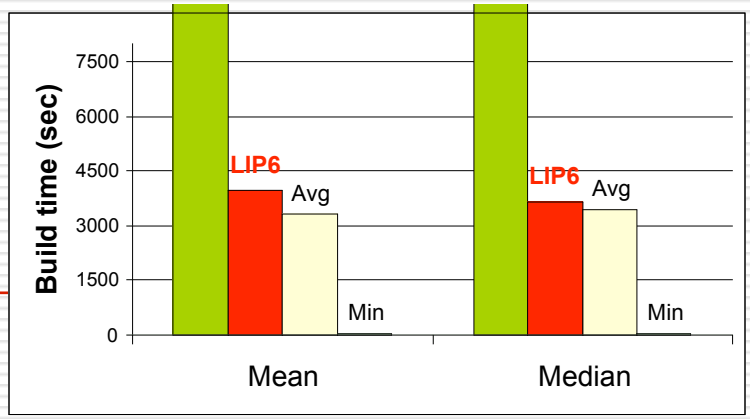



*(1 strongly disagree - 5 strongly agree)*



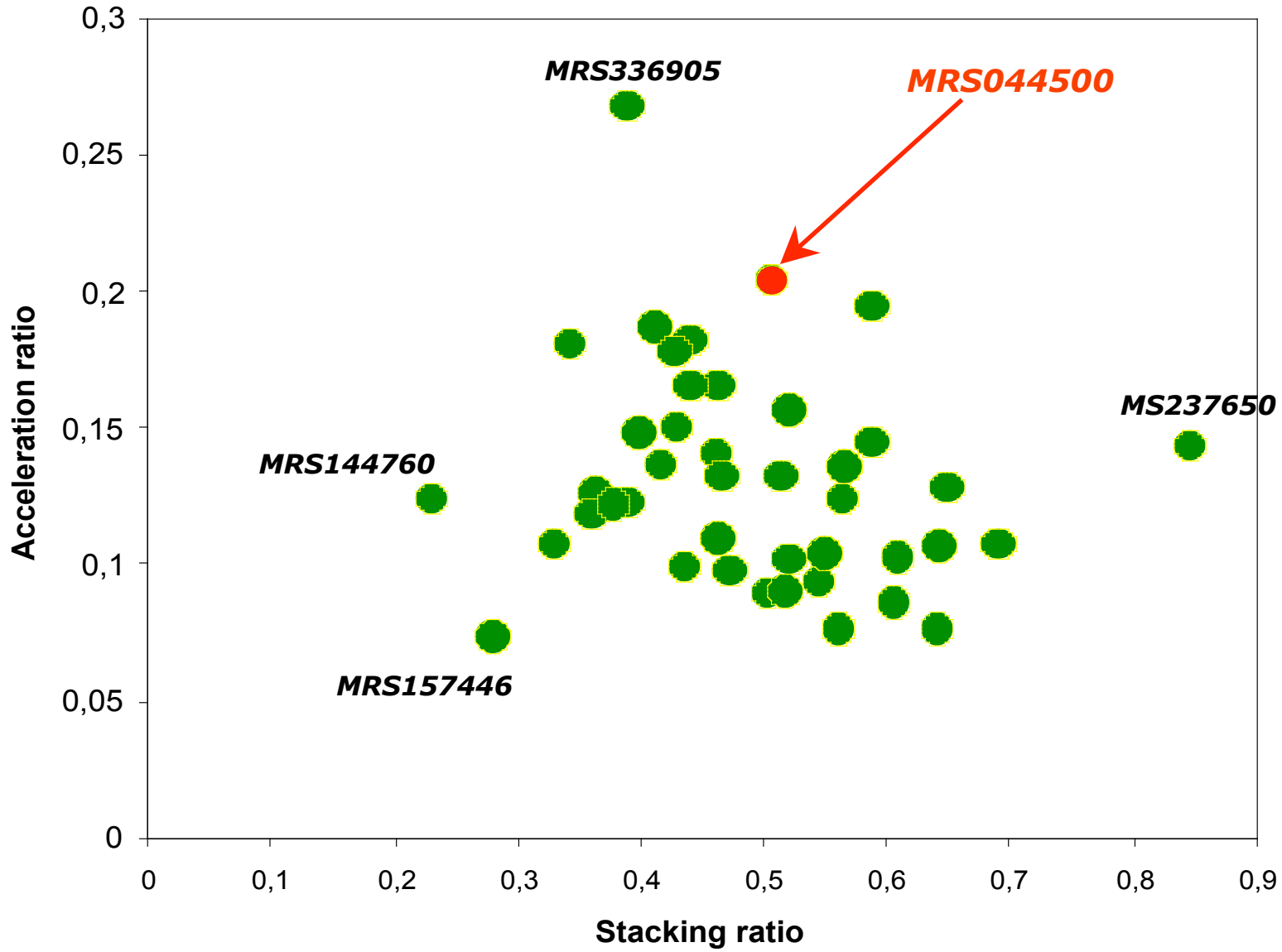

- Understandability
  - Easy to understand (EA)
  - Redundancy (RE)
- Drawbacks
  - Adaptive acceleration may spoil the understandability (too fastly displayed!)
  - Redundancy relies on the stacking step
    - Shot comparison should be enhanced

# Discussion: time




- Building time
- Rem
  - Now: extraction of descriptors
    - 99% of the processing time
    - non optimized tools (perl-based)
  - Stacking / acceleration steps
    - Fast steps: frames handled in a row

# Results: Stacking vs Acceleration



# Conclusion and future work

- Proposed approach
  - Two complementary steps
    - Stacking shots and Adaptive acceleration
  - Specifically developed for the challenge
  - Few loss of information but some redundancy
- Future work
  - Parameters to tune
  - Speed up the descriptor extraction step
  - Application to any video