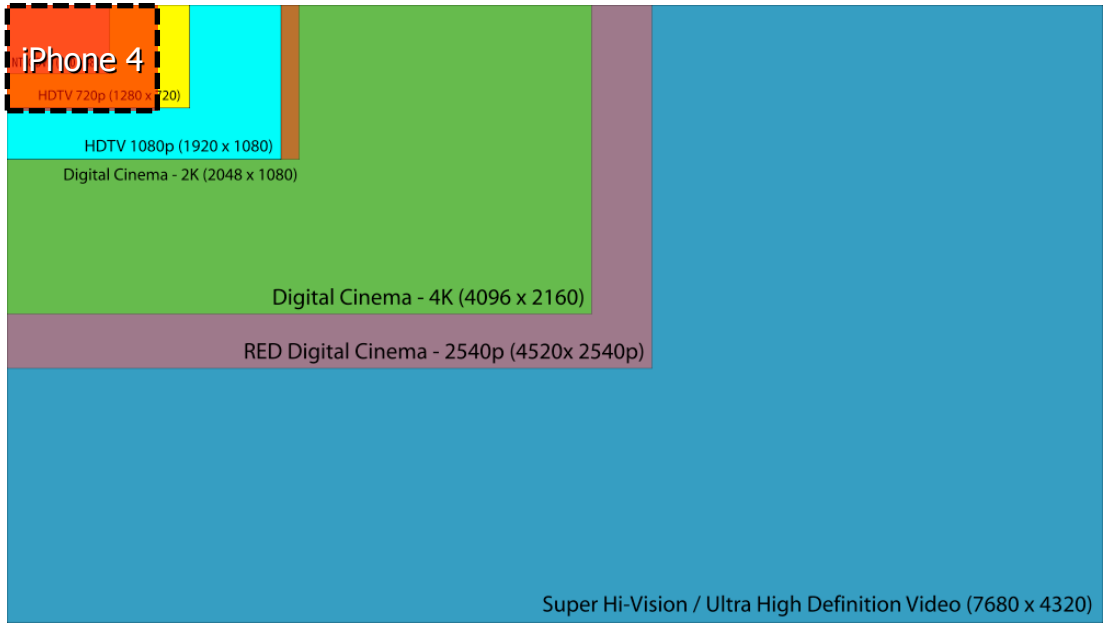
Crowdsourced Automatic Zoom and Scroll for Video Retargeting

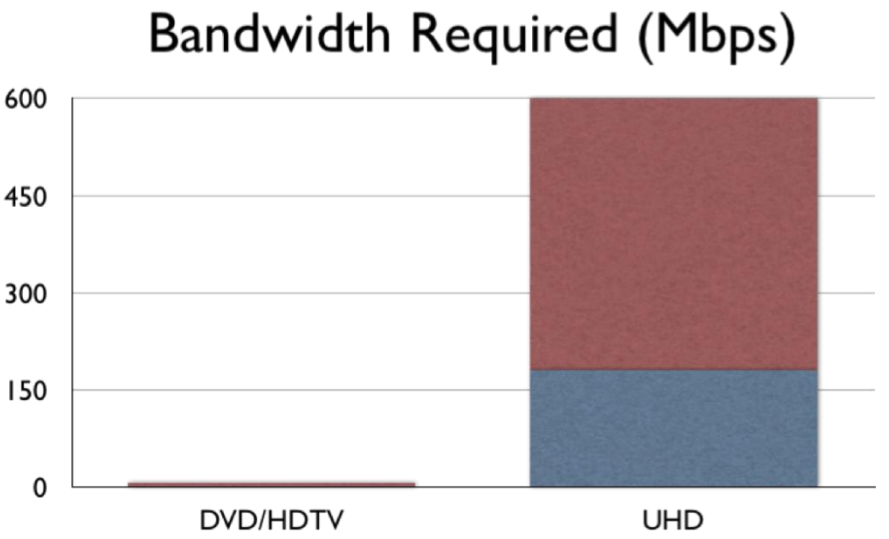
Axel Carlier, Wei Tsang Ooi

NUS (Singapore)

Vincent Charvillat, Romulus Grigoras, Geraldine Morin

IRIT (Toulouse, France)





Video retargeting:

Making a large video fit into a smaller screen and available with network capacities.

One simple way:

Scale down the video to the resolution of the screen

Example

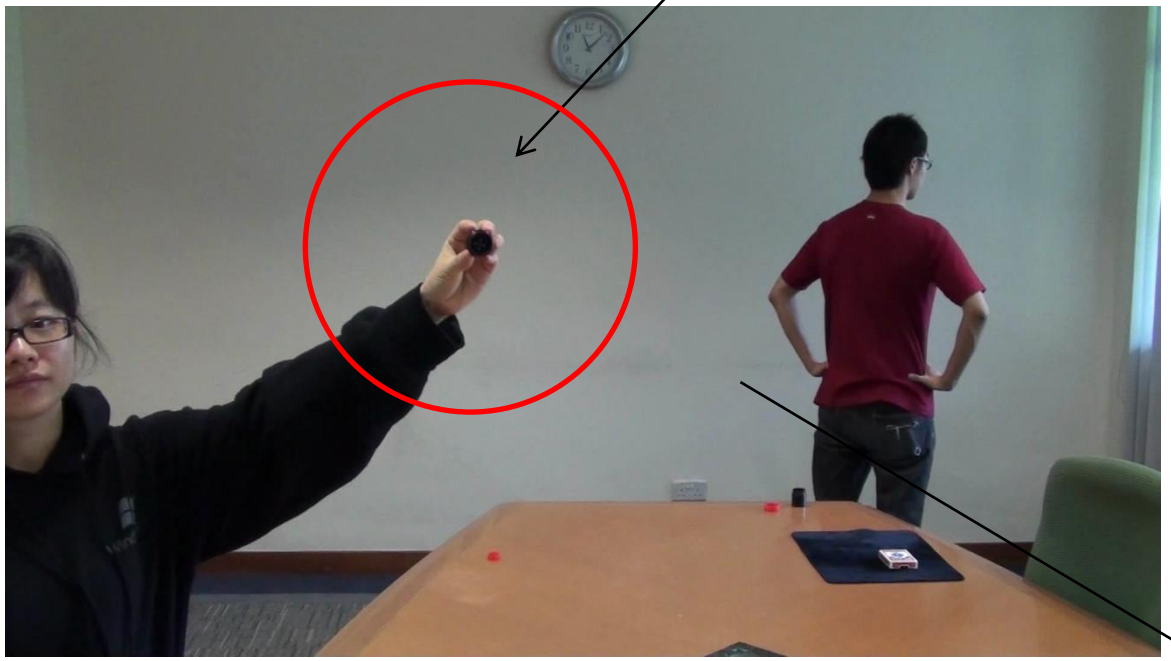
Drawback:

Important details may not be visible

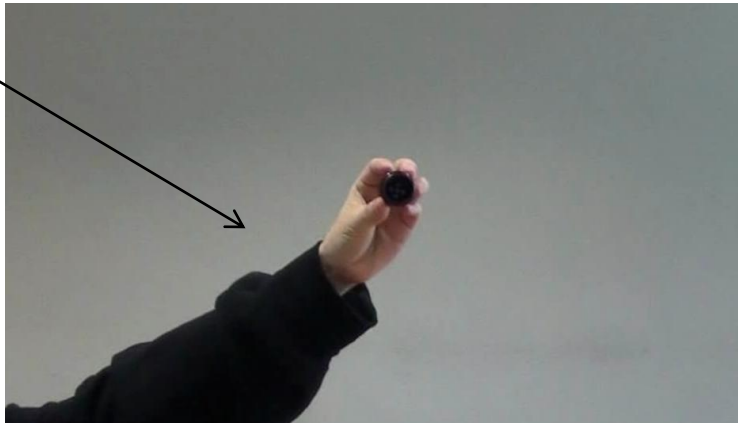
What makes a good video retargeting ?

- Good comprehension of the video content
- The video is aesthetically satisfying

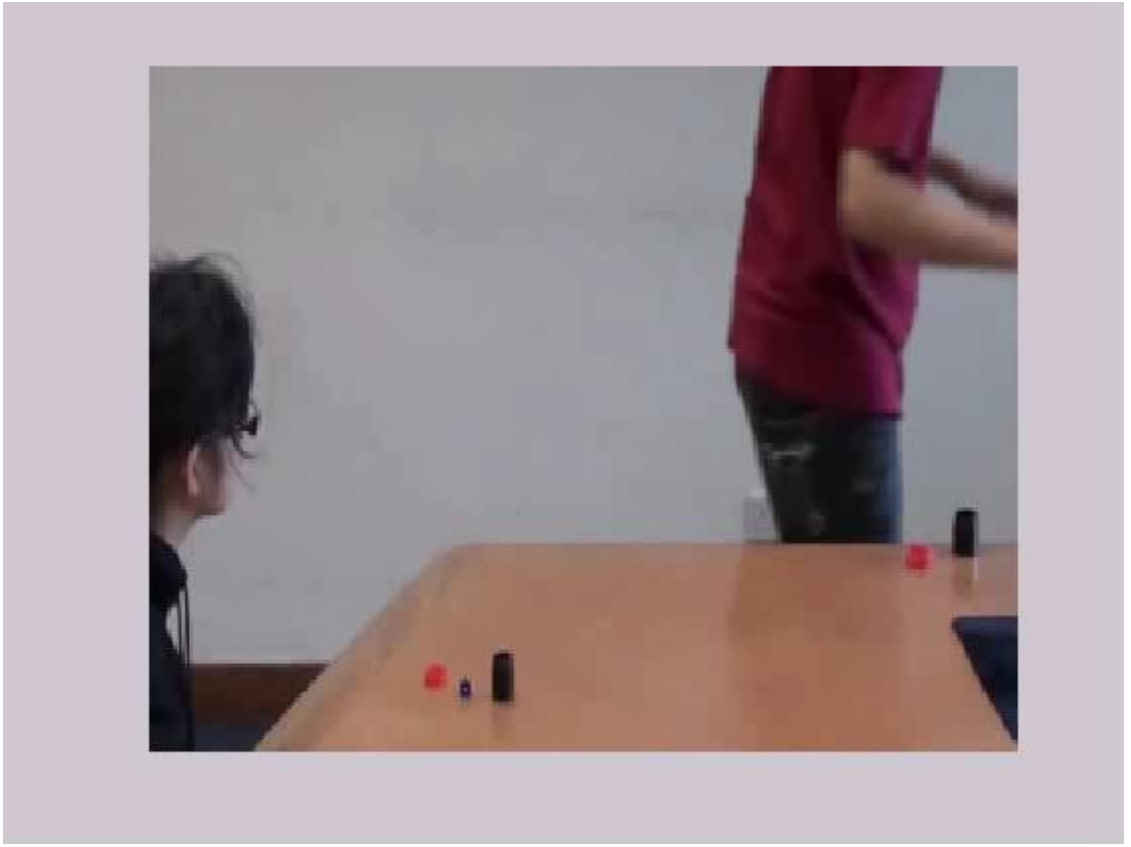
Important region



Zoom

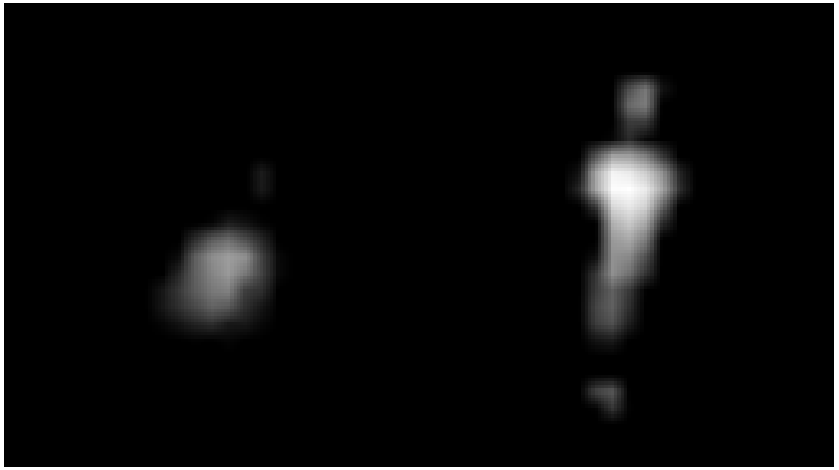


Idea: Zoom and scroll to show only interesting regions of the video



Problem: how to find the region **automatically** ?

Approaches using Content Analysis

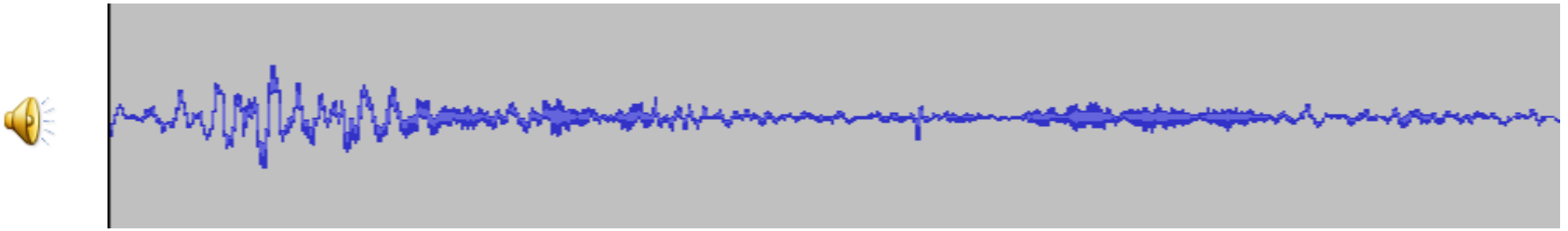


Saliency map

Motion detection

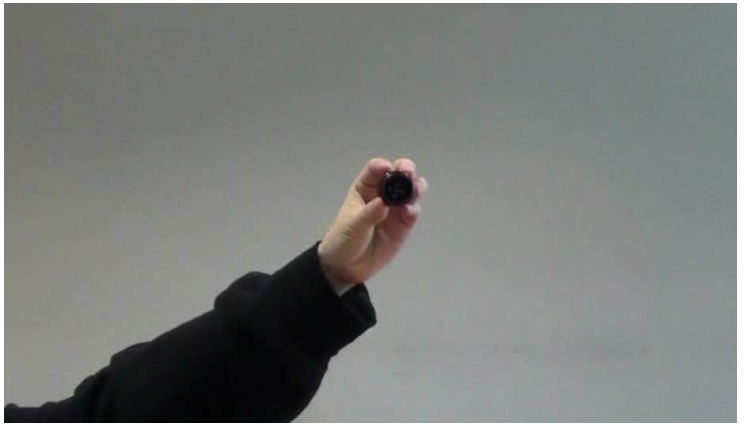
Too many regions

Liu, Gleicher MM 06
Avidan, Shamir Commun. ACM 09



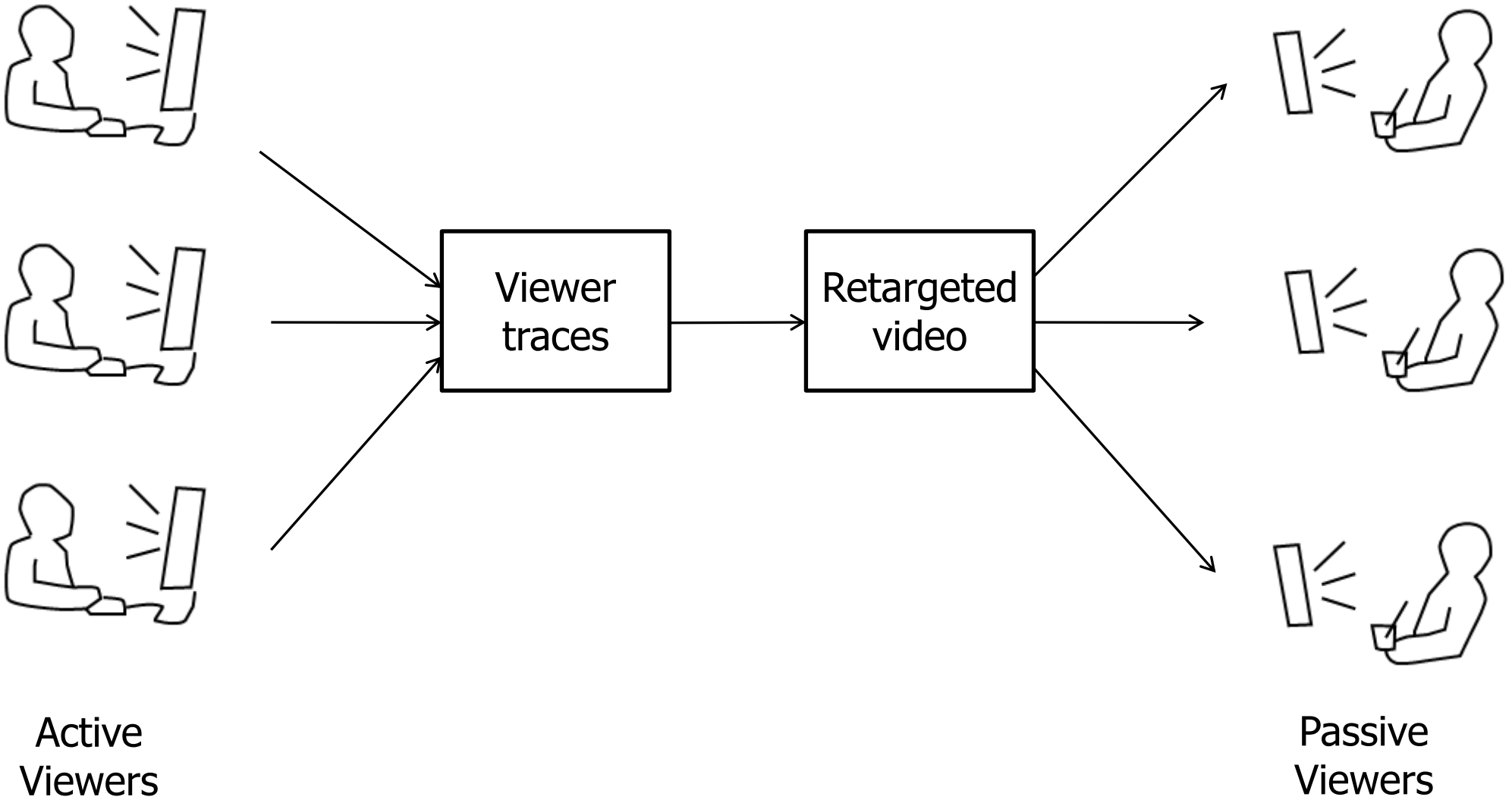
Speech Recognition + Natural Language processing

Object recognition



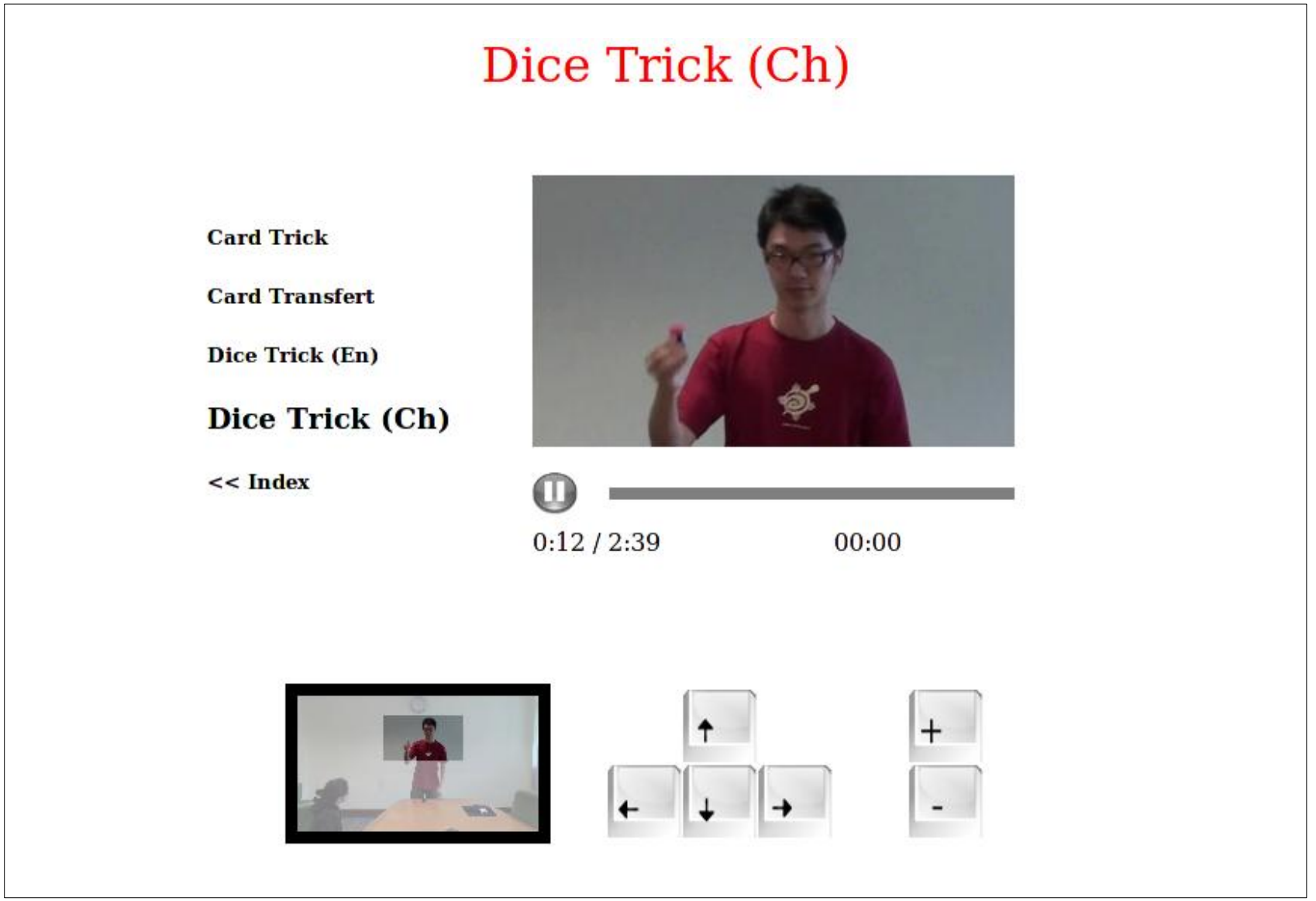
Our idea: Crowdsourcing

Identifying regions of interest by gathering implicit input from users.


Viewer traces
Retargeted video
Active Viewers
Passive Viewers

Use zoomable video

Interface

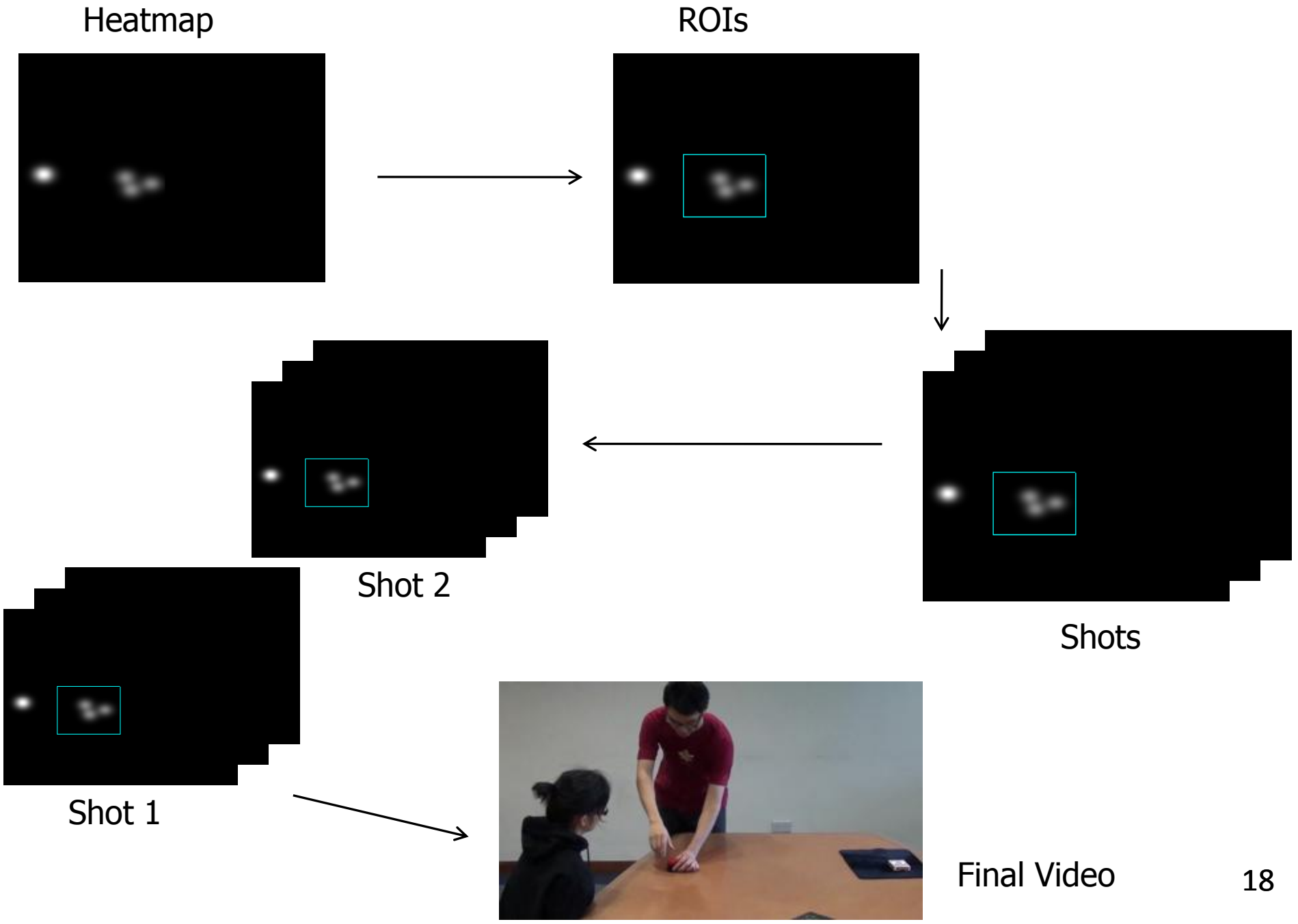

Dice Trick (Ch)
Card Trick
Card Transfert
Dice Trick (En)
Dice Trick (Ch)
<< Index
0:12 / 2:39
00:00
+
-

Example of user interaction

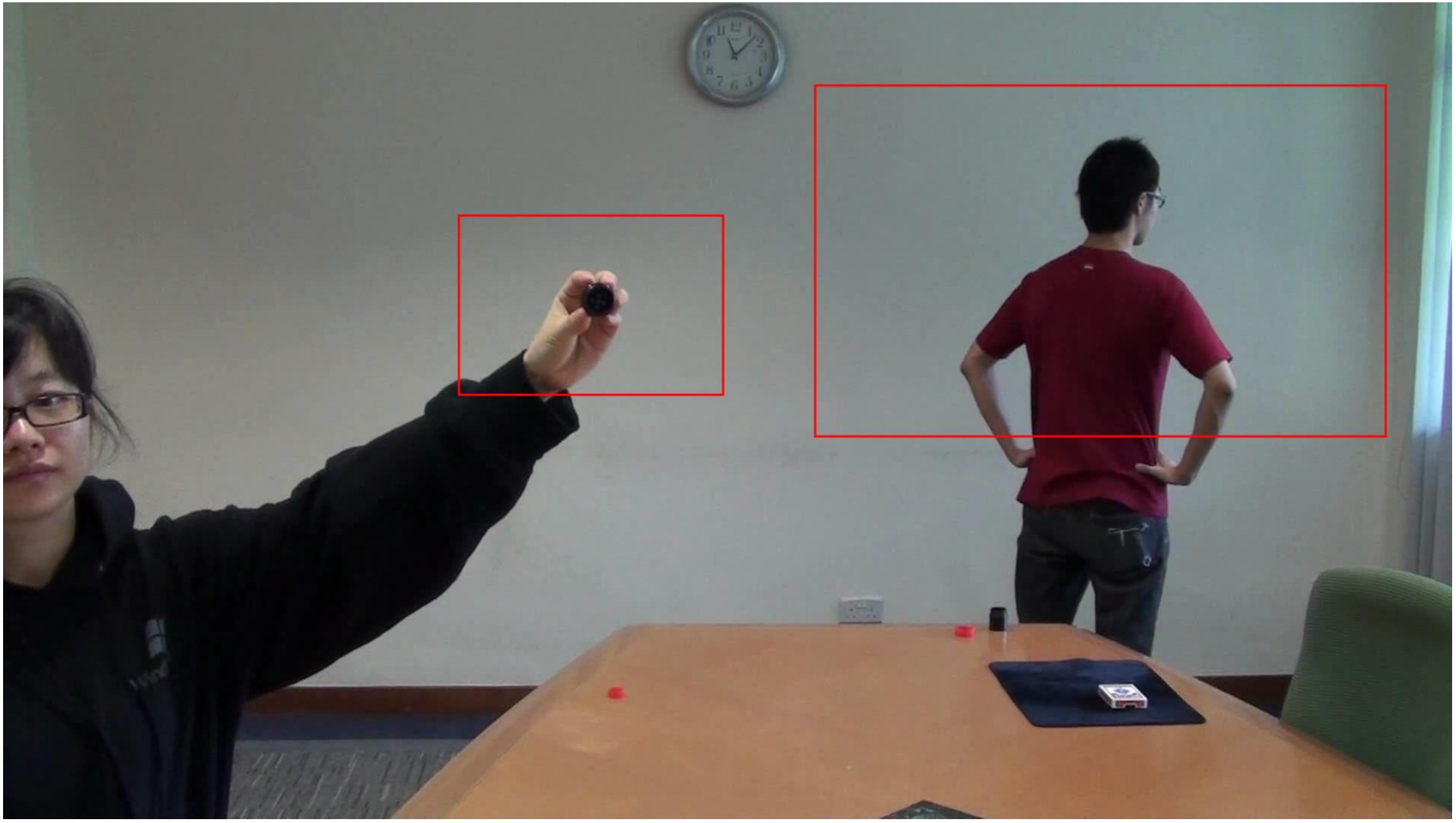
Crowdsourcing

- Tutorial: how to use the interface ?
- Magic videos :
  - HD Videos : 1920 × 1080 pixels
  - Fixed camera
  - Obvious ROIs : magician's hands, cards, dice...
- Between 7 and 12 viewers for each video
- 11,183 interaction events logged

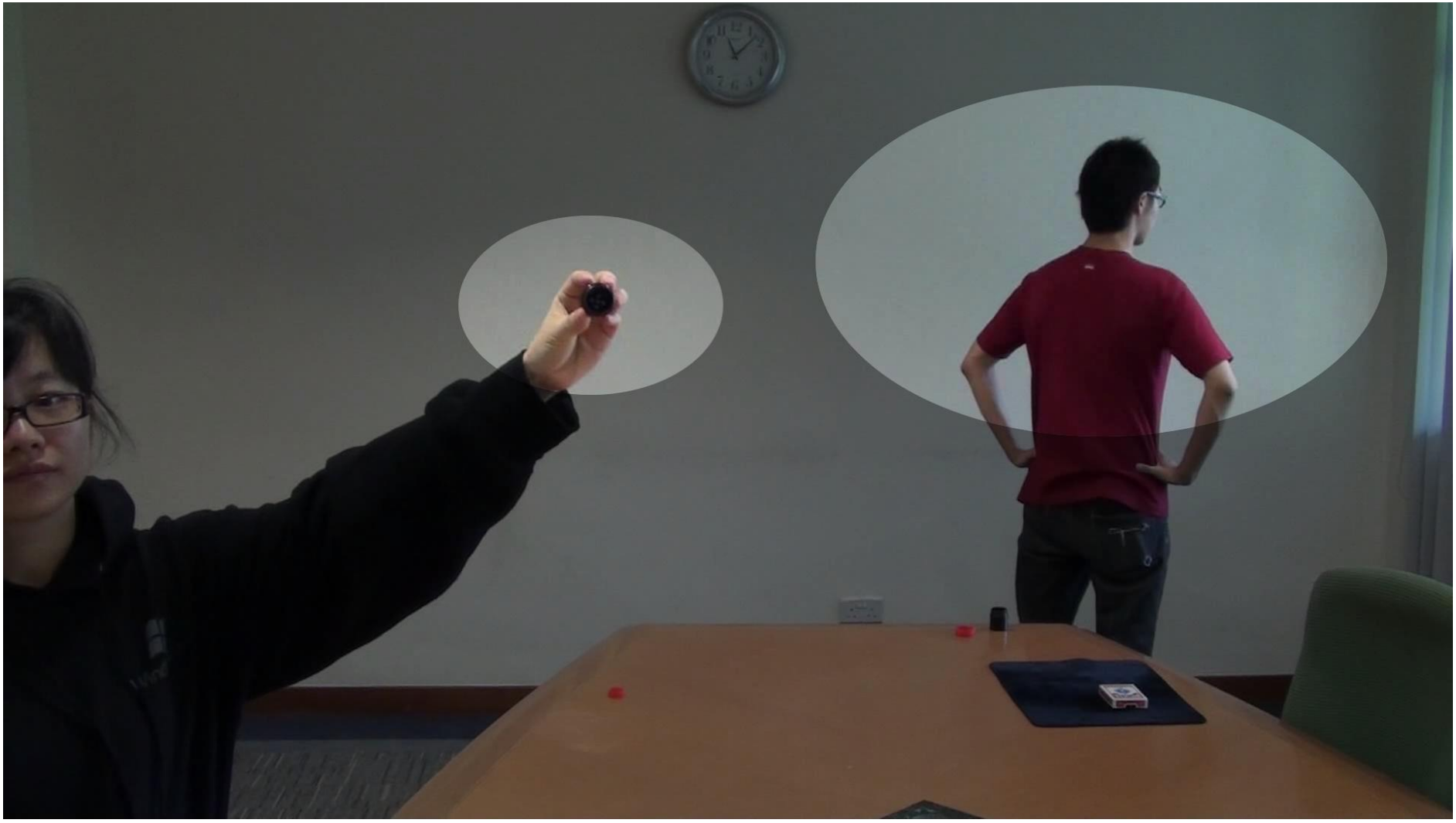
Automatic Generation of Retargeted Video



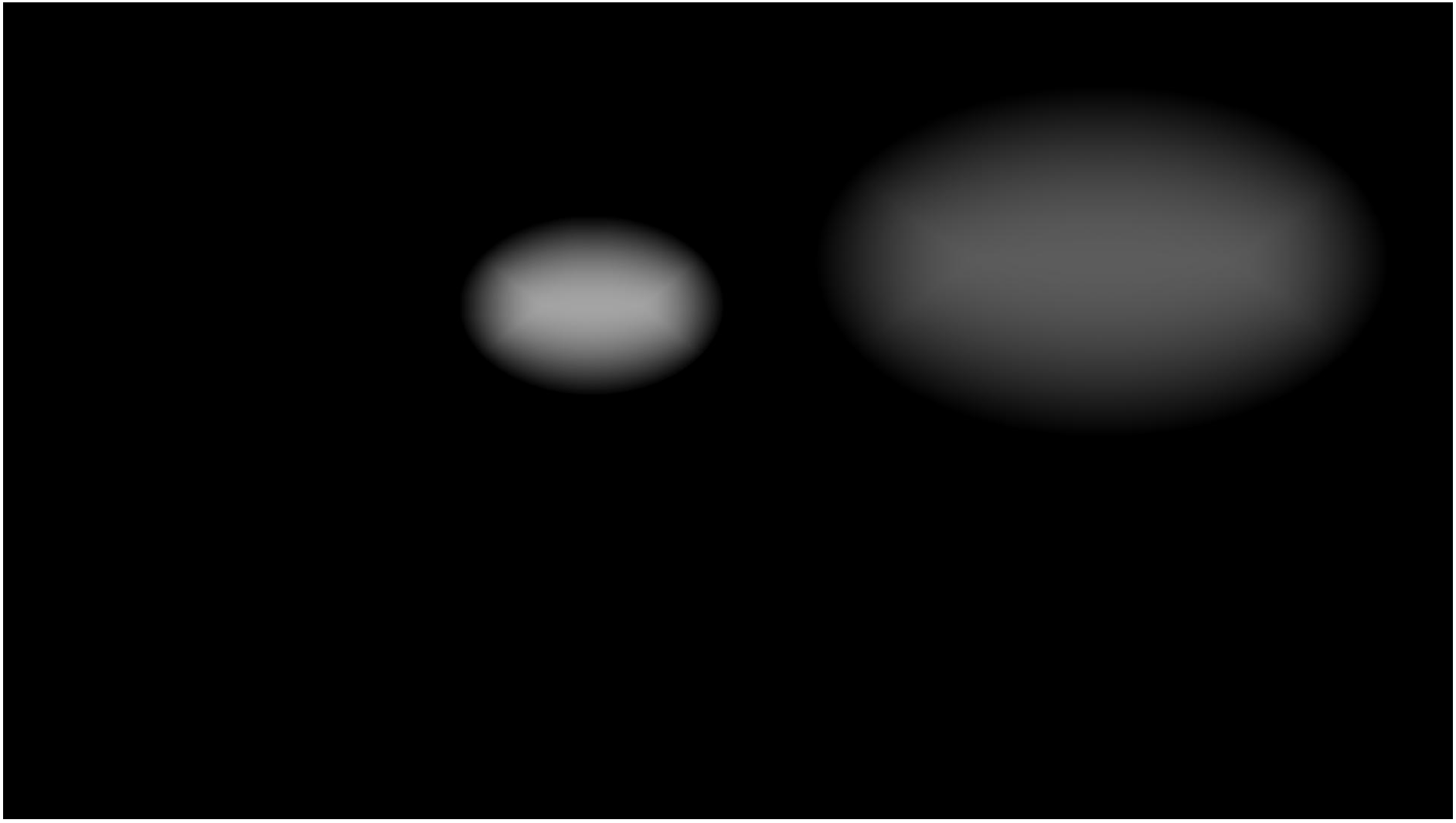
Building Heatmaps



Building Heatmaps



Building Heatmaps



Analyzing Heatmaps

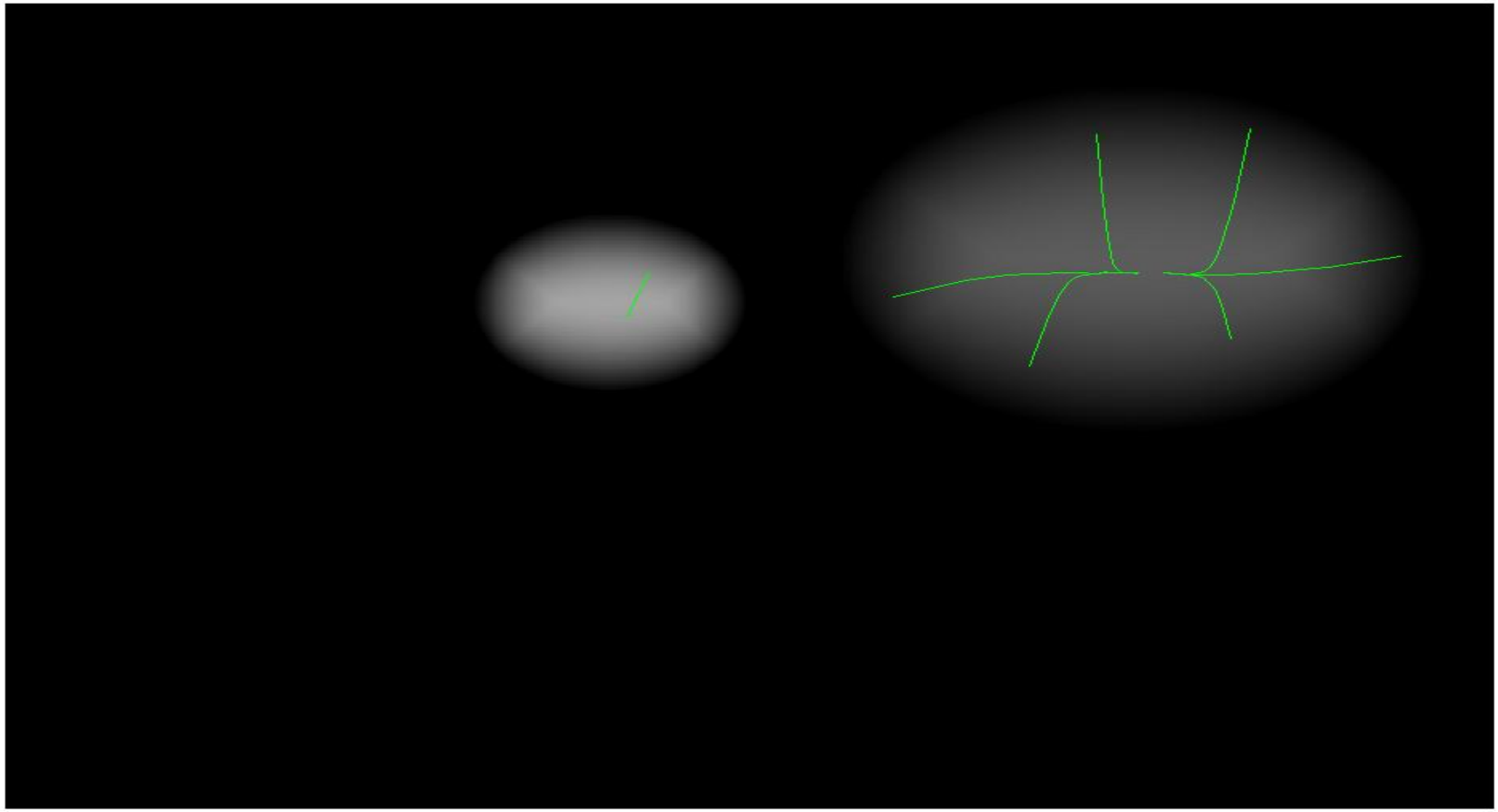
Here draw gaussians in 3d with matlab

GMM (Gaussian Mixture model)

Mean Shift
- -K
- -w_i
- -m_i
- -Σ_i

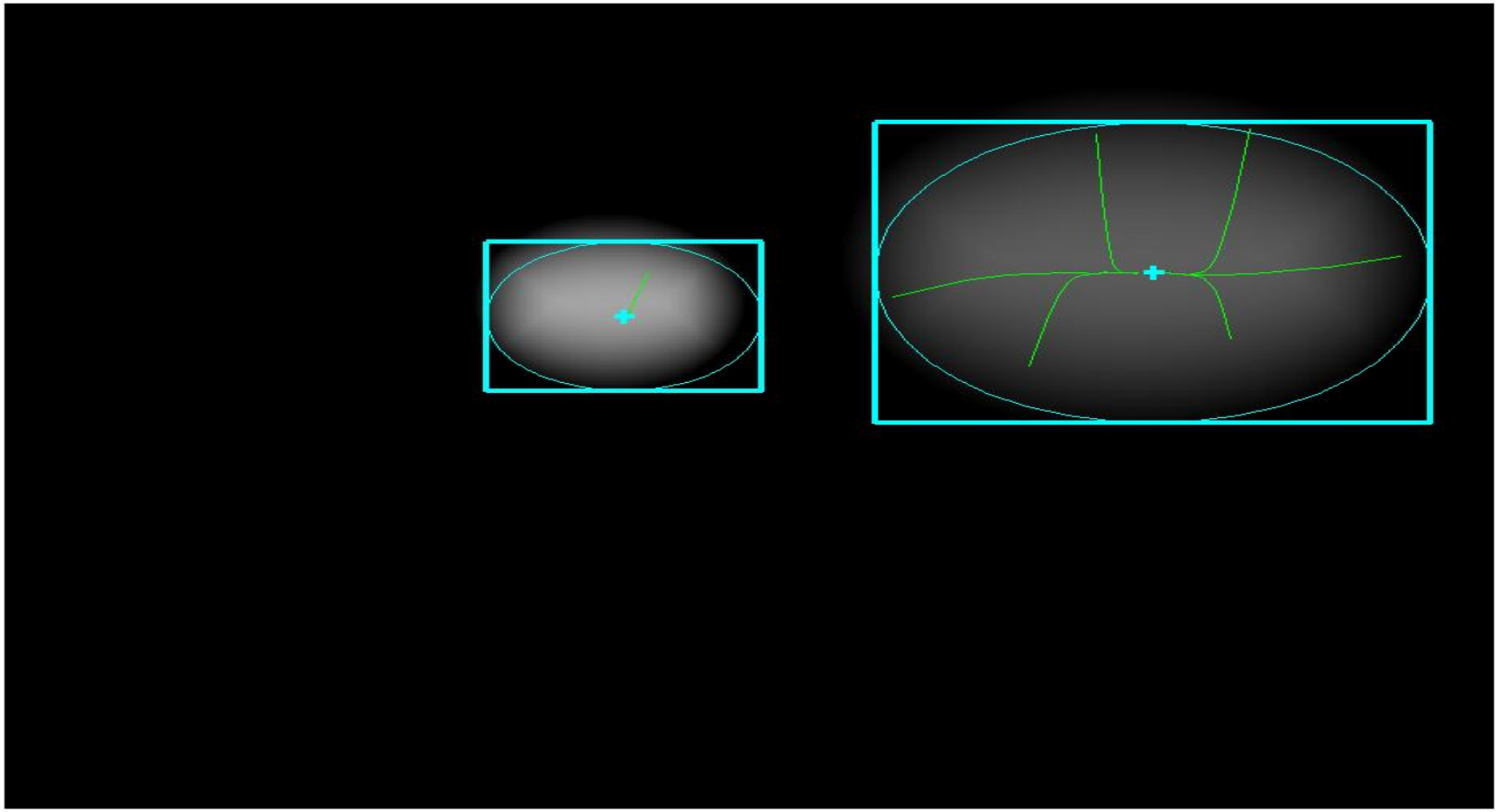
Finding Modes

Mean-Shift: Clustering algorithm
(Comaniciu, ICCV 02)

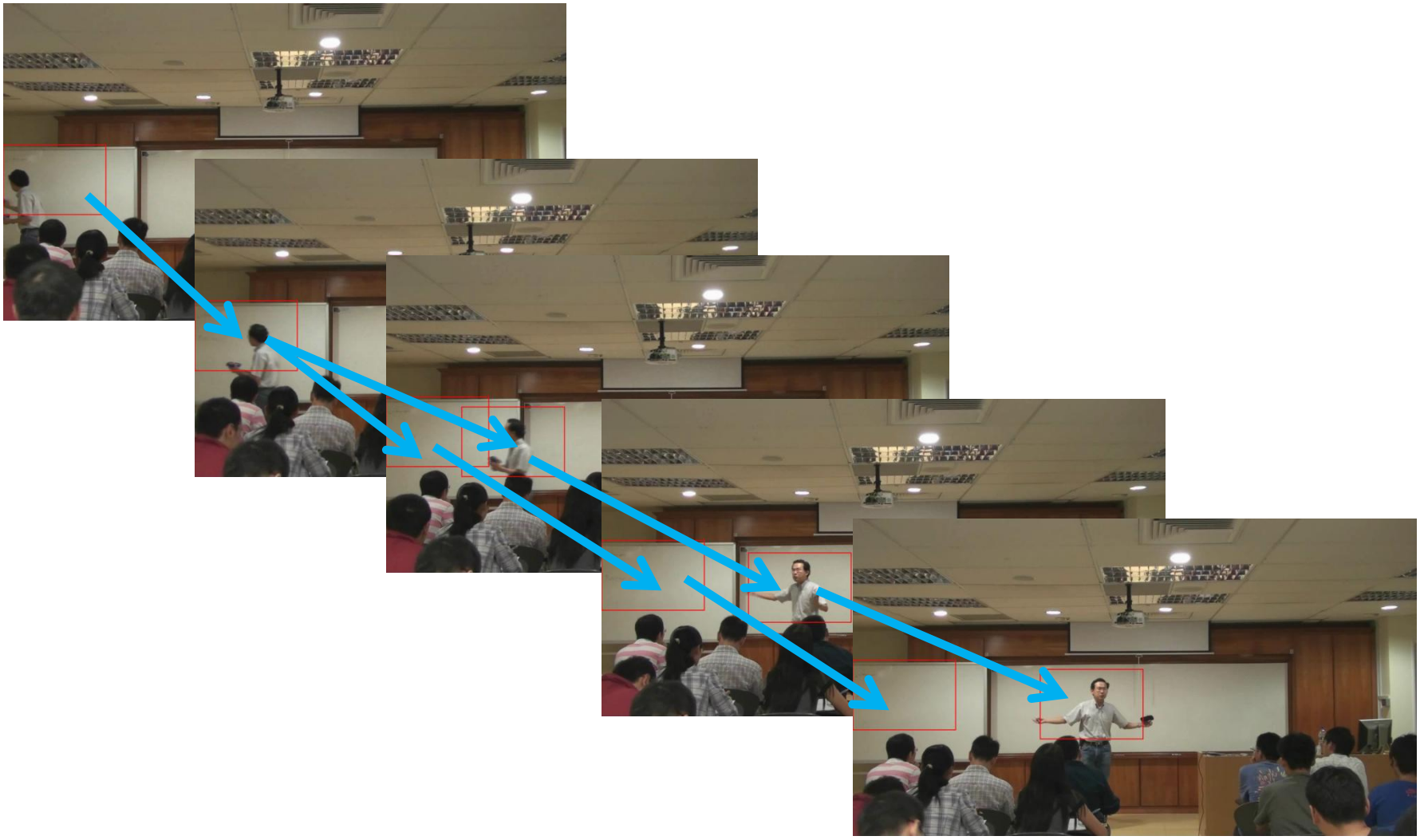


Determining ROI size

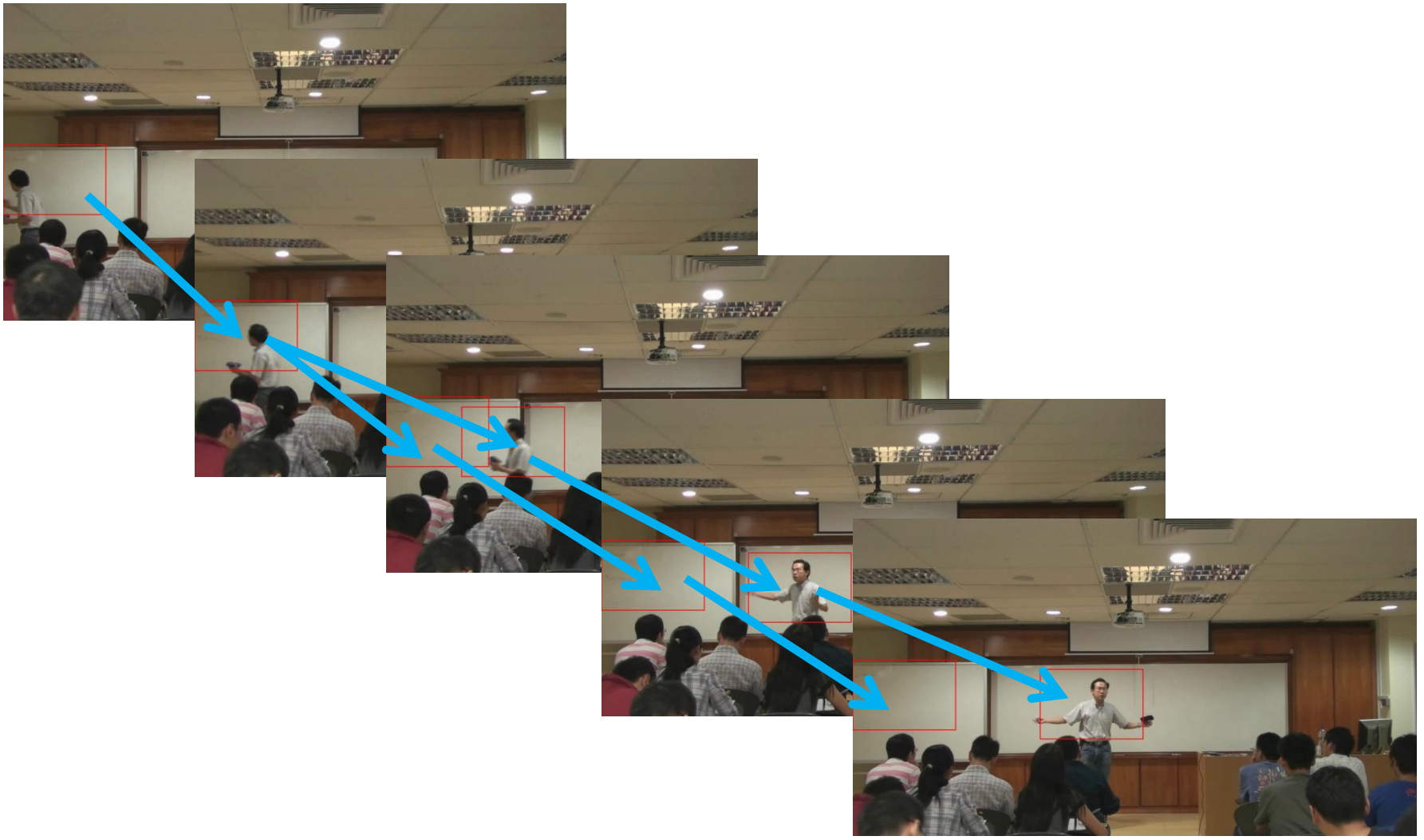
Minimum Covariance Determinant (MCD)



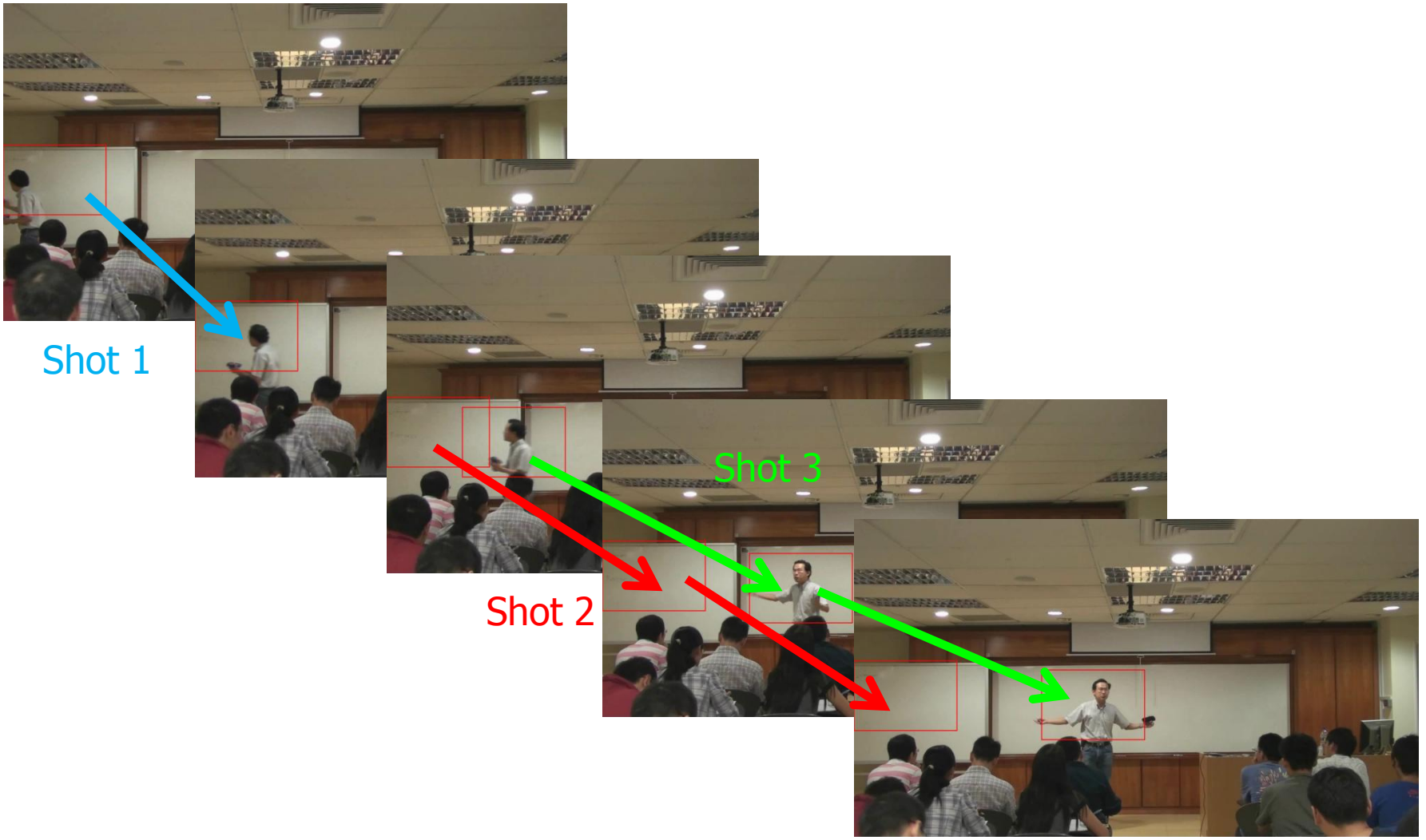
Building a ROI Dynamics Graph



Cutting the graph into shots



Shots selection



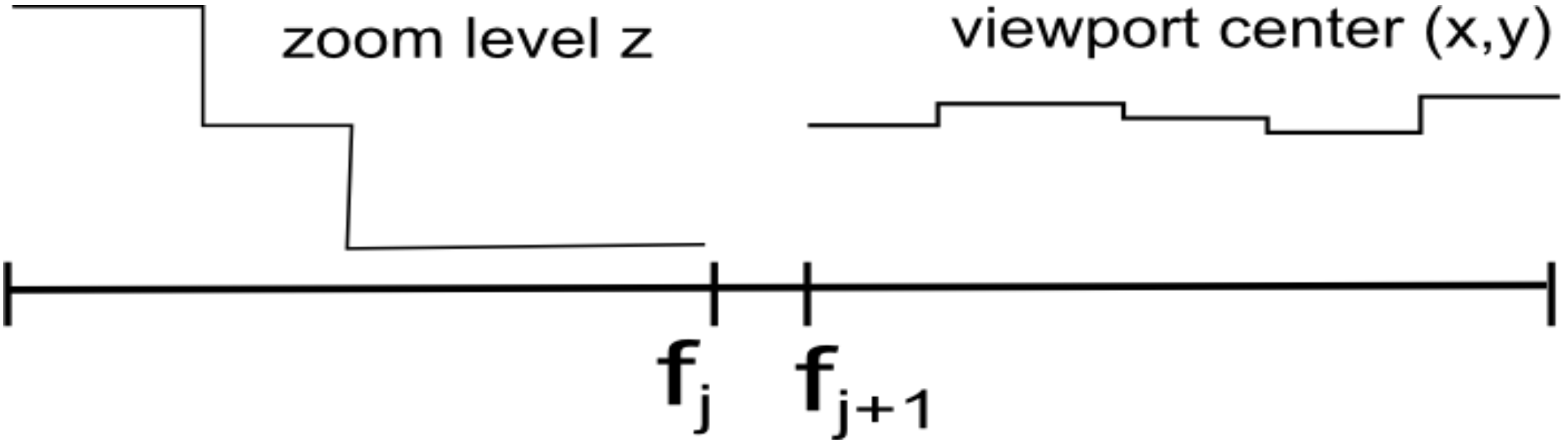
Result video

Integrating Reframing techniques

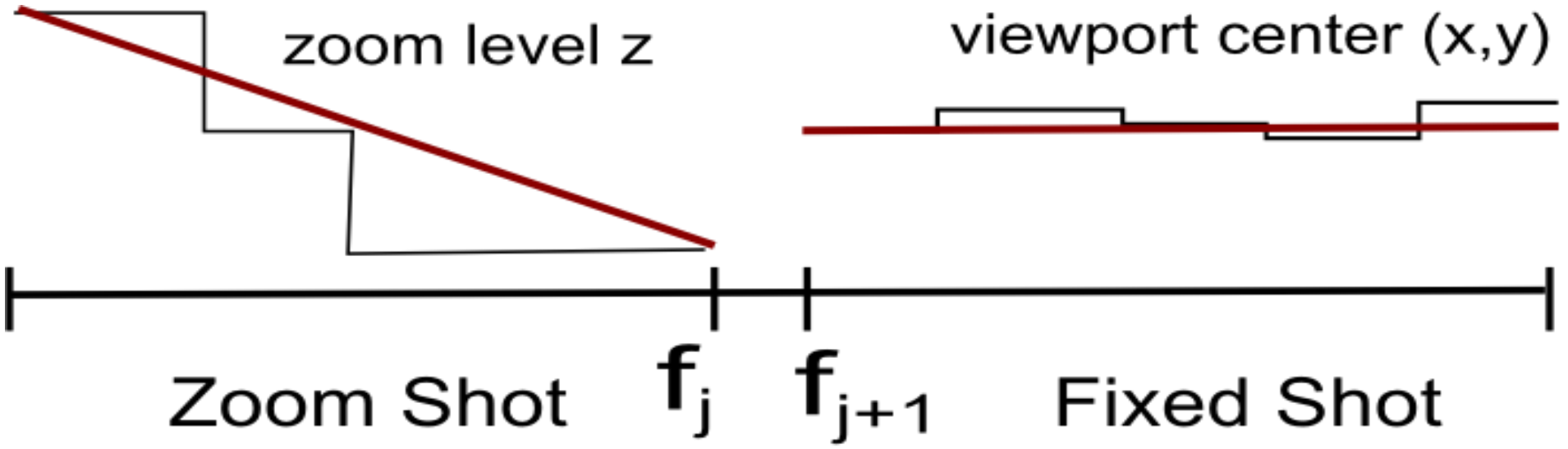
- Bottom-up reframing
  - Type of shot: fixed, zooming or dolly
  - Shot level: stabilization according to its type
  - Inter-shot level: transitions and reestablishing shots

Dolly shot

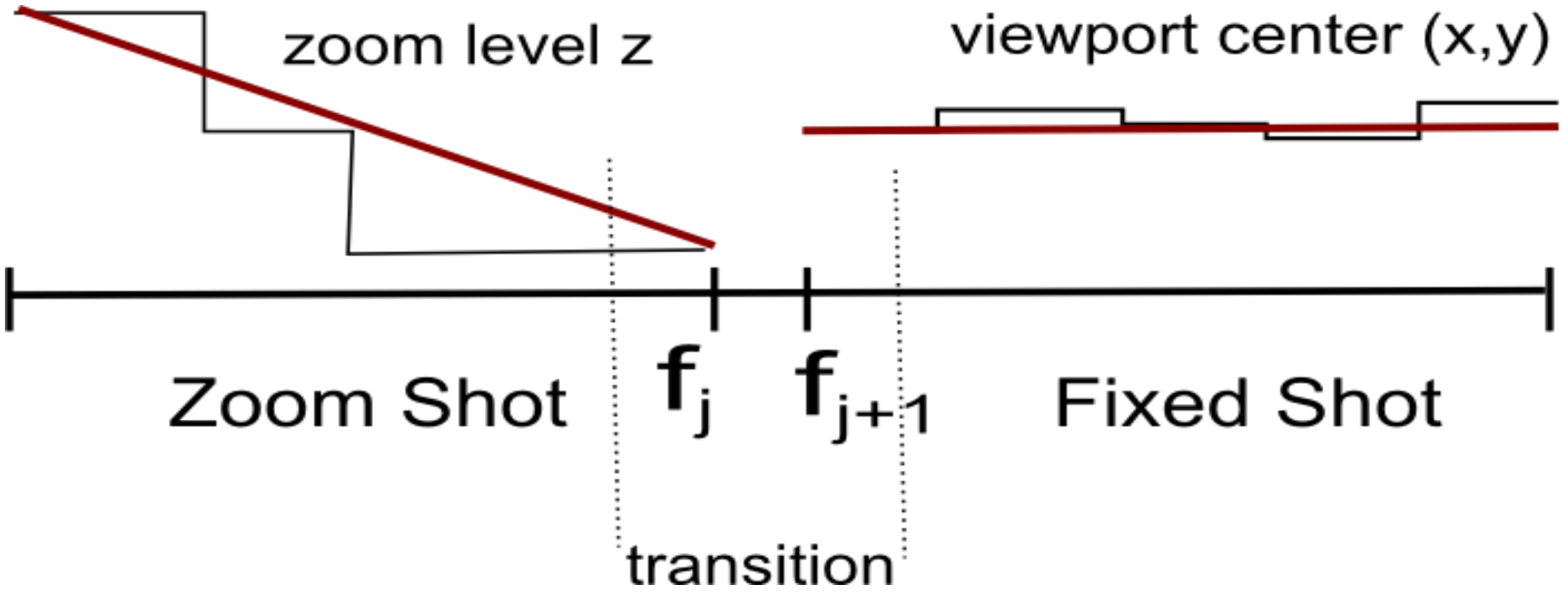
Shots



Shot stabilization



Transitions



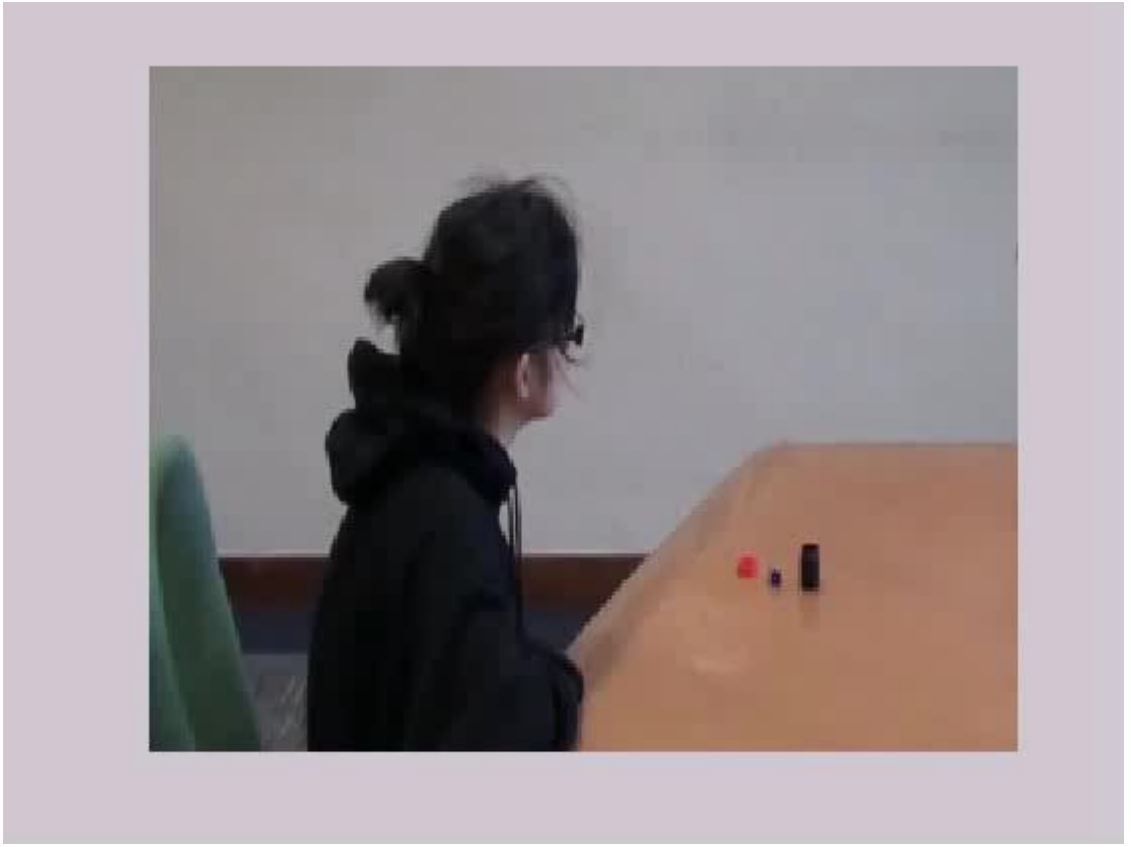
Reestablishing shots

Final Result

Results validation

- 3 poor videos:
  - User interaction (*user*)
  - Retargeted version without reframing techniques (*noRT*)
  - Original version scaled down (*nozoom*)
- Retargeted version with reframing techniques (*crowdsourced*)
- Ground truth (*expert*)

Ground Truth

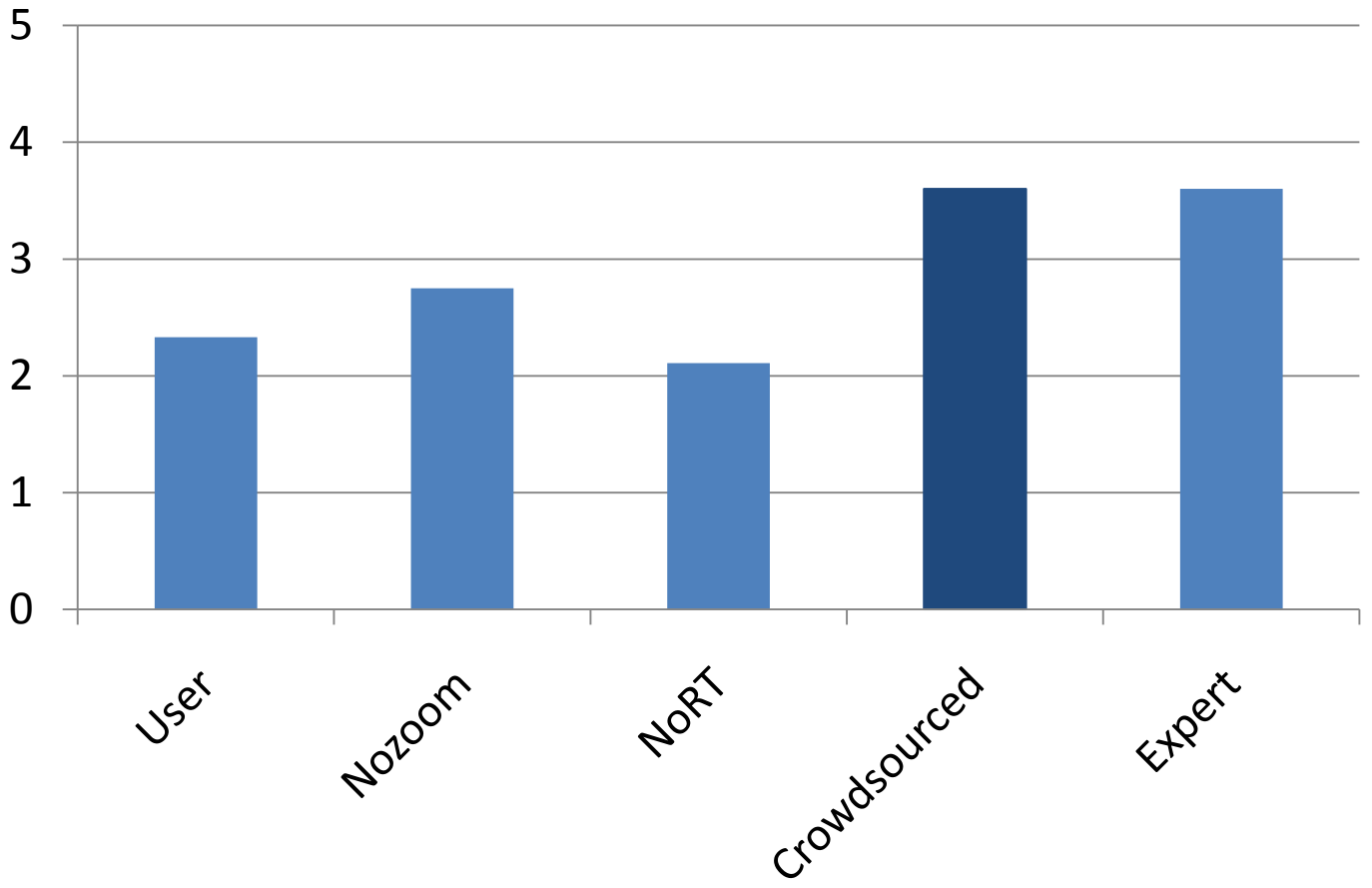


Protocol

- 48 participants divided into 3 categories
  - User – crowdsourced – expert (18)
  - NoRT – crowdsourced – expert (18)
  - Nozoom – crowdsourced – expert (12)

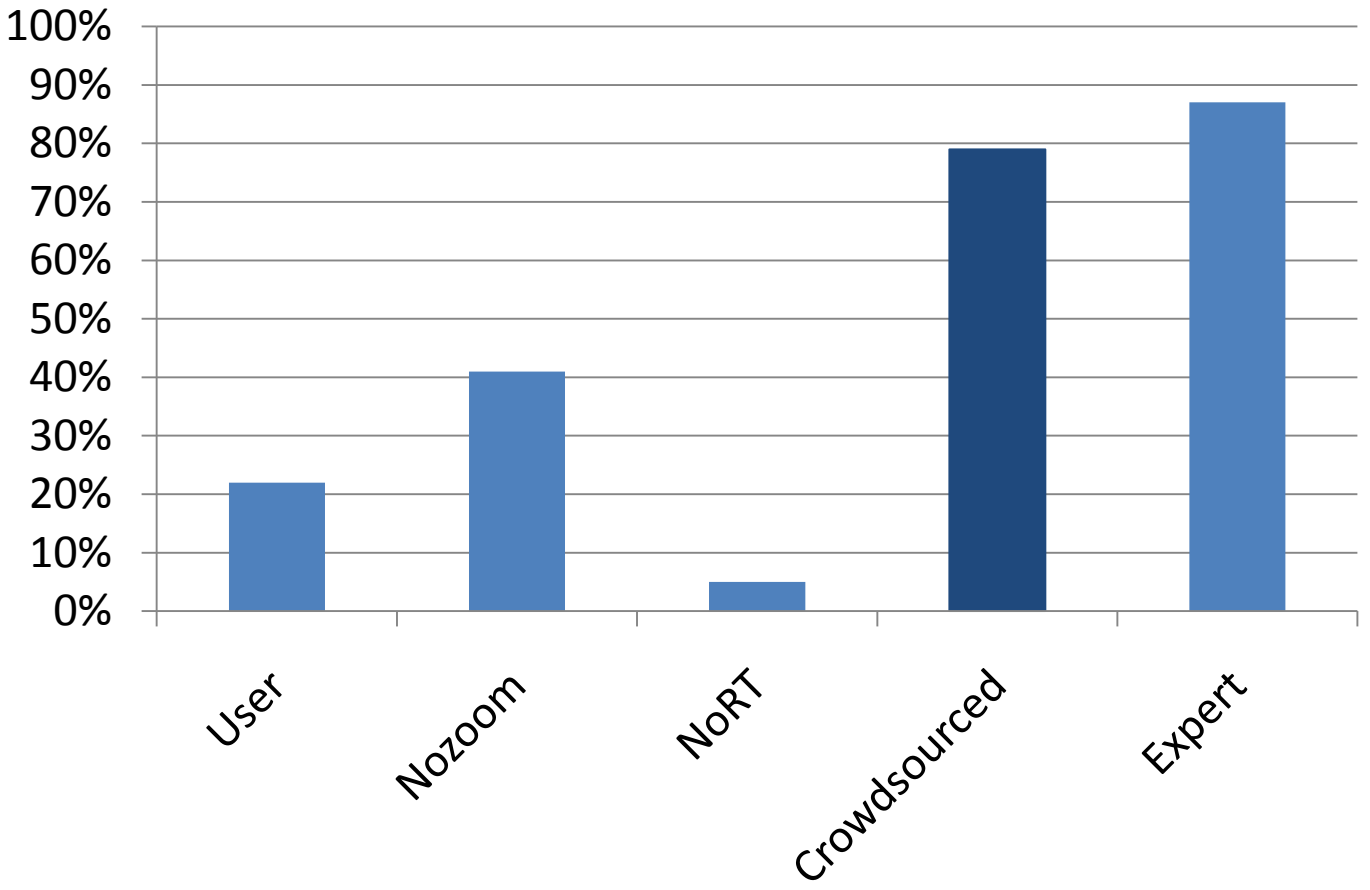
- 3 questions were asked to the participants

Rate the video editing of the video

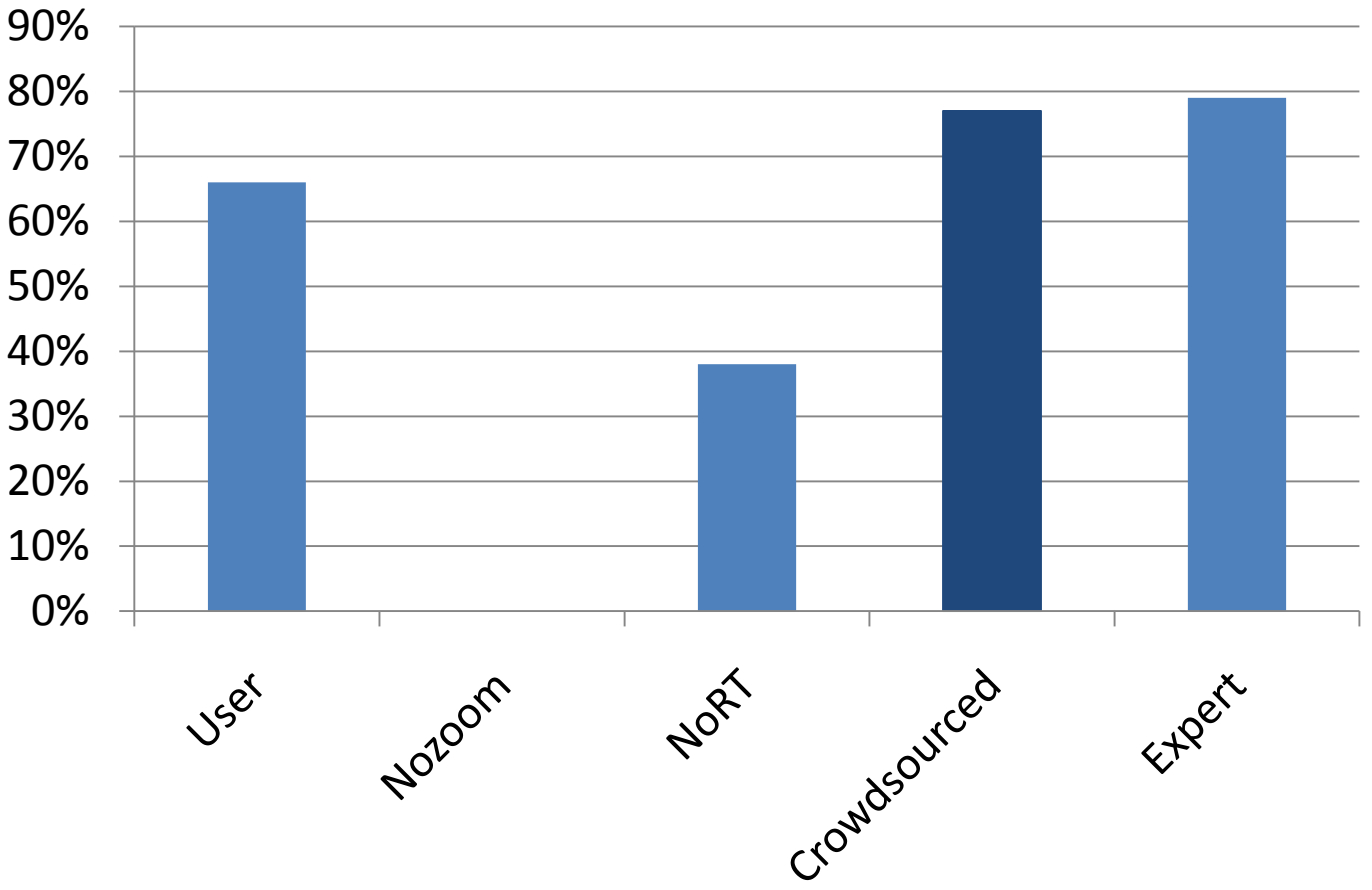


NoRT = retargeted version without reframing techniques

Is the video editing reasonable ?



Does the video manage to convey important information ?



Summary

-Gather **implicit** input from users

-No content analysis

-**In our examples** : less than 12 viewers are enough to detect ROIs

Future work

- Explore alternative methods for intermediary steps:
  - Modelling heatmaps not as a GMM
  - Adding cinematographic rules
  - Classify users into different profiles and generate a retargeted video for each profile

Questions

?

Results

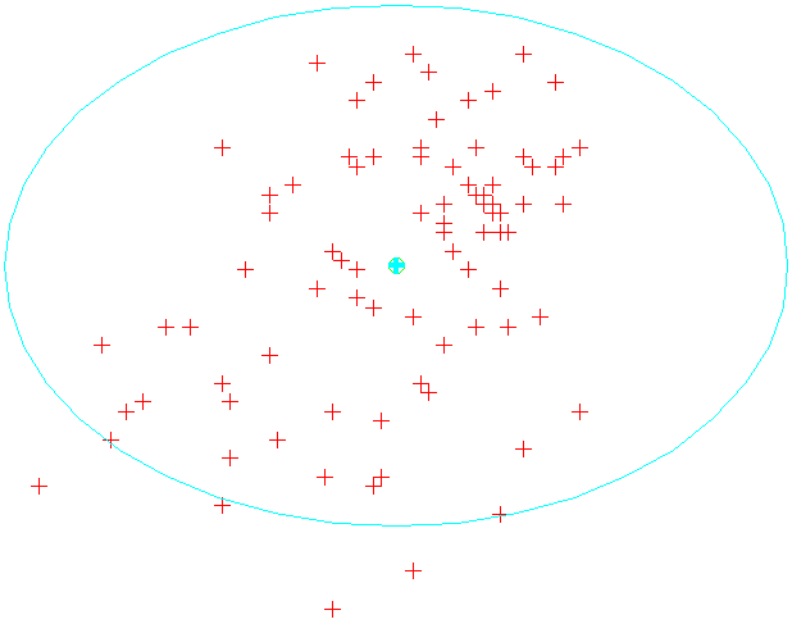
	User	Nozoom	noRT	Crowdsourced	Expert
Ratings	2.33	2.75	2.11	3.6	3.6
Reasonability	22%	41%	5%	79%	87%
Efficiency	66%	0%	38%	77%	79%

Using aesthetics

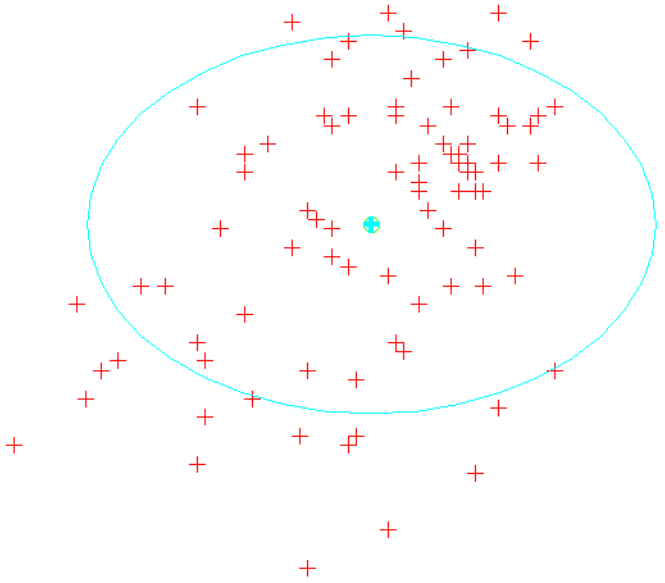
Liu, Chen, Wolf and Cohen-Or. Optimizing Photo Composition, *Computer Graphic Forum*

Luo, Yi wen and Tang, Xiaoou. Photo and Video Quality Evaluation : Focusing on the Subject, *ECCV 08*

MCD Covariance



MCD Covariance



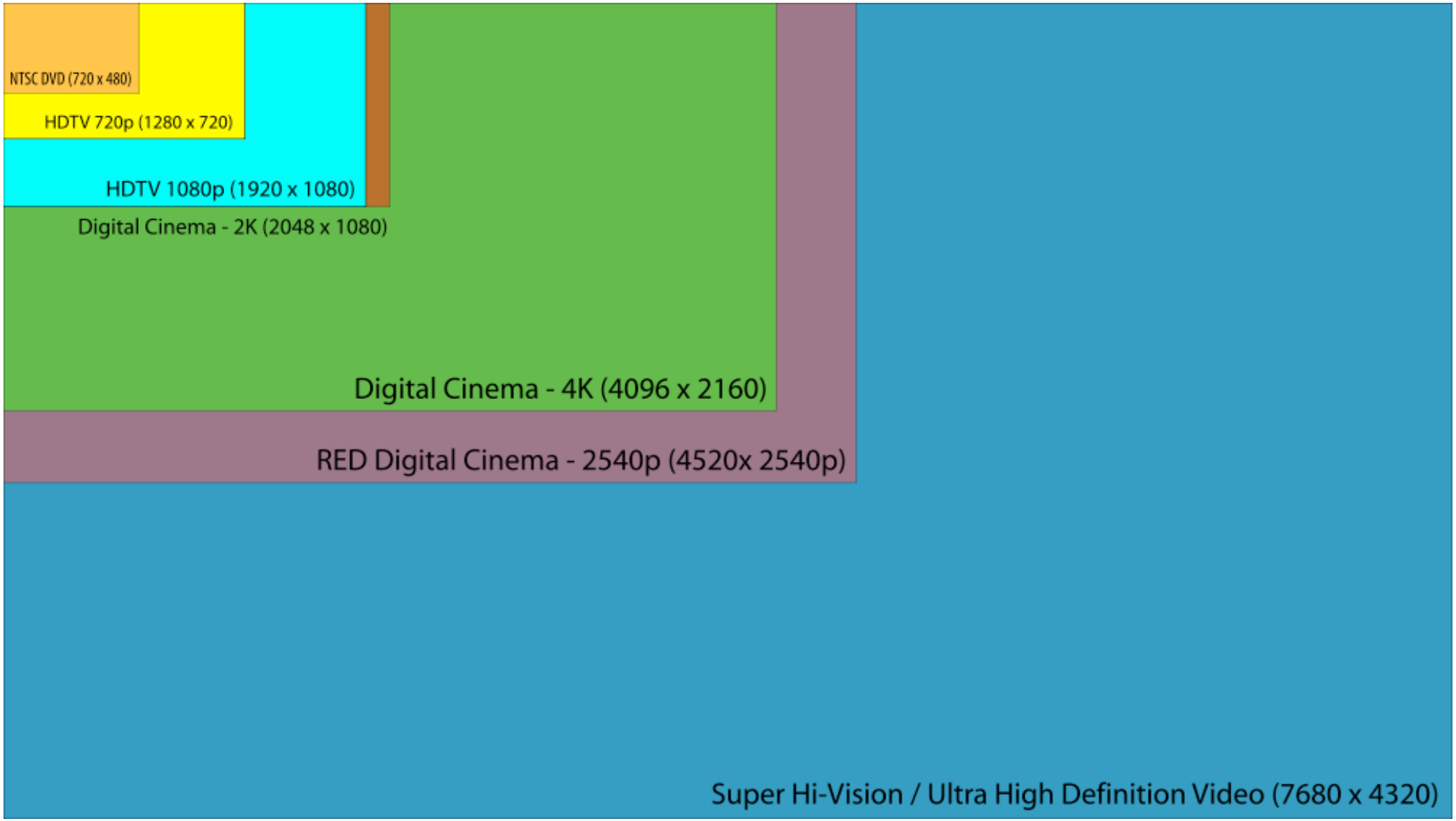
Gym Video Retargeted

Crowdsourcing

Shamma, Shaw, Shafton, Liu. Watch what I watch, *MIR 07*

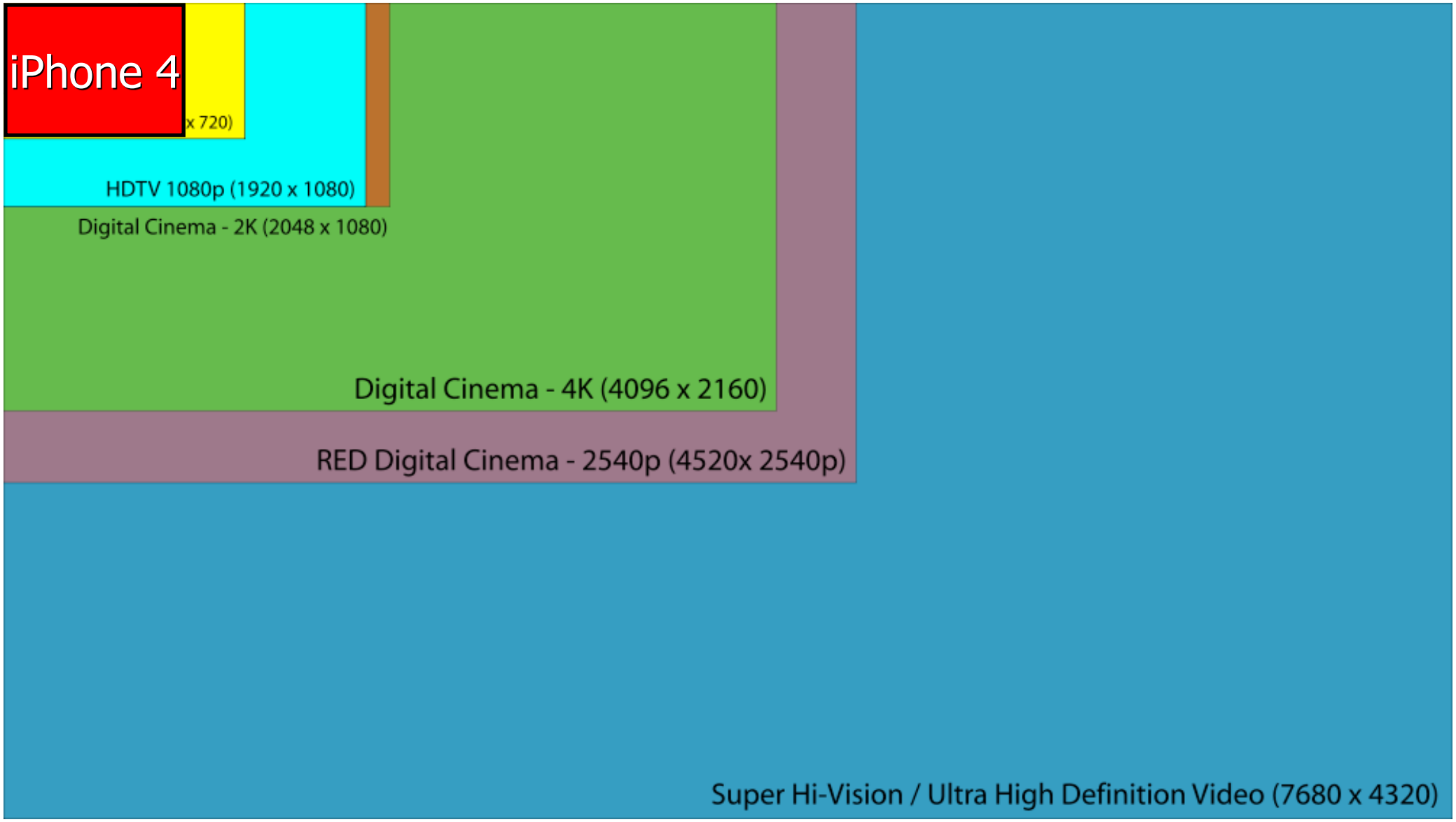
Overview

- Video retargeting
- Zoomable video
- Finding users' interests
- Creating shots
- Integrating reframing techniques
- Results validation

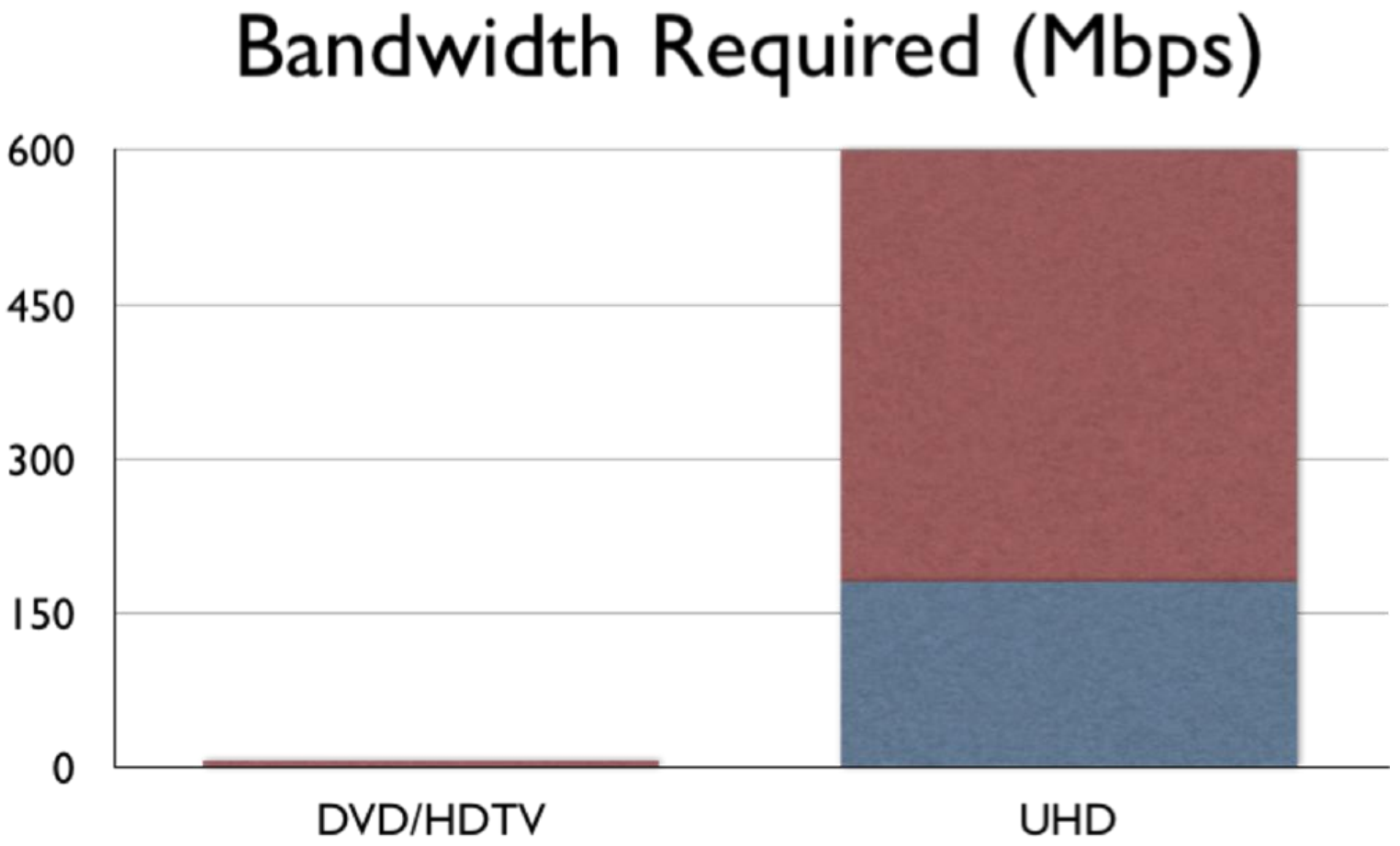


http://en.wikipedia.org/wiki/File:UHDV.svg

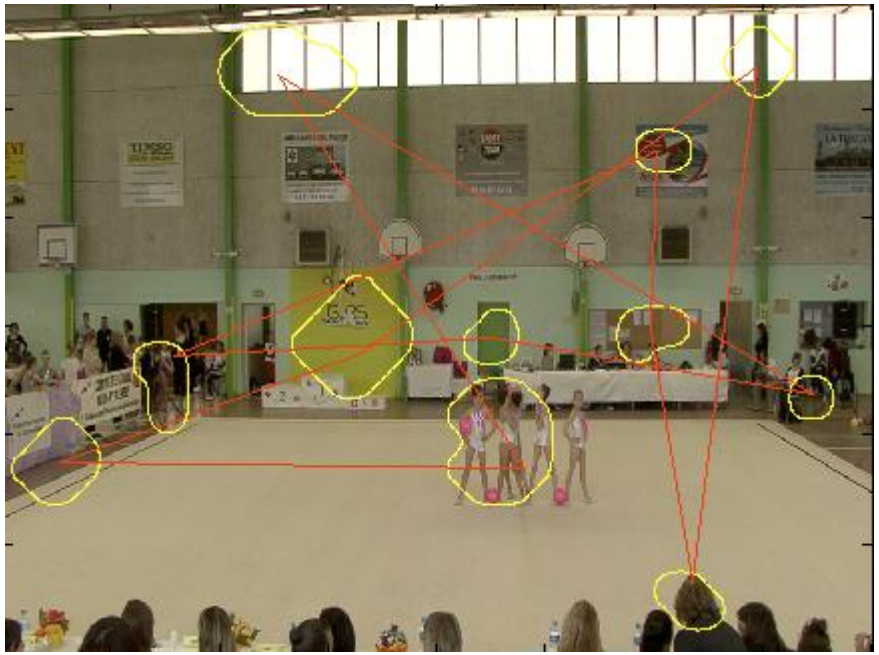
960 x 640



http://en.wikipedia.org/wiki/File:UHDV.svg


Bandwidth Required (Mbps)
600
450
300
150
0
DVD/HDTV
UHD

Approaches using Content Analysis



Saliency map

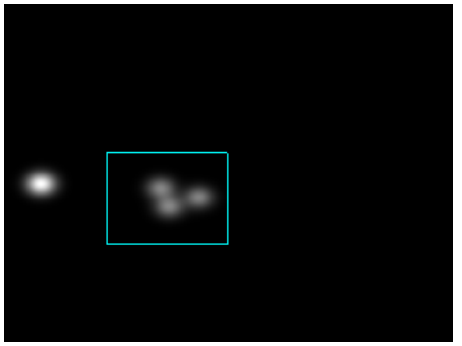
Motion detection

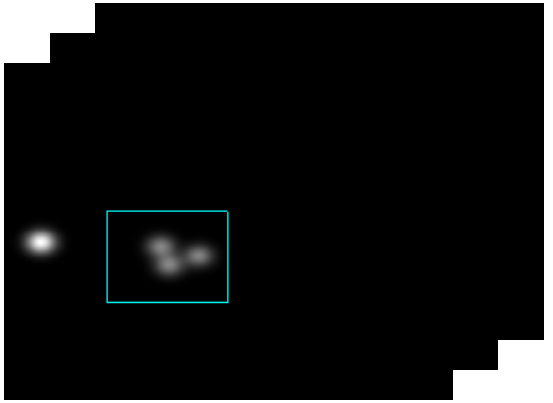
Liu, Gleicher MM 06
Avidan, Shamir Commun. ACM 09

Crowdsourcing

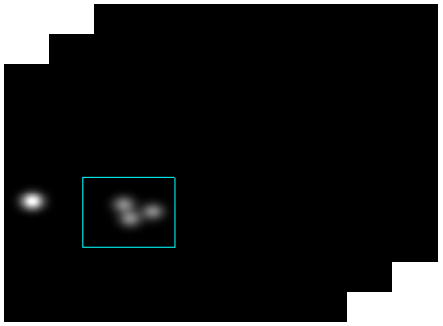
Heatmap

Hotspots

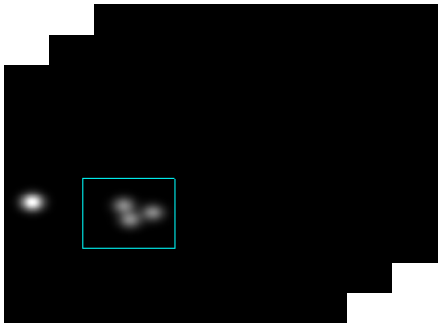




Shots



Shot 2

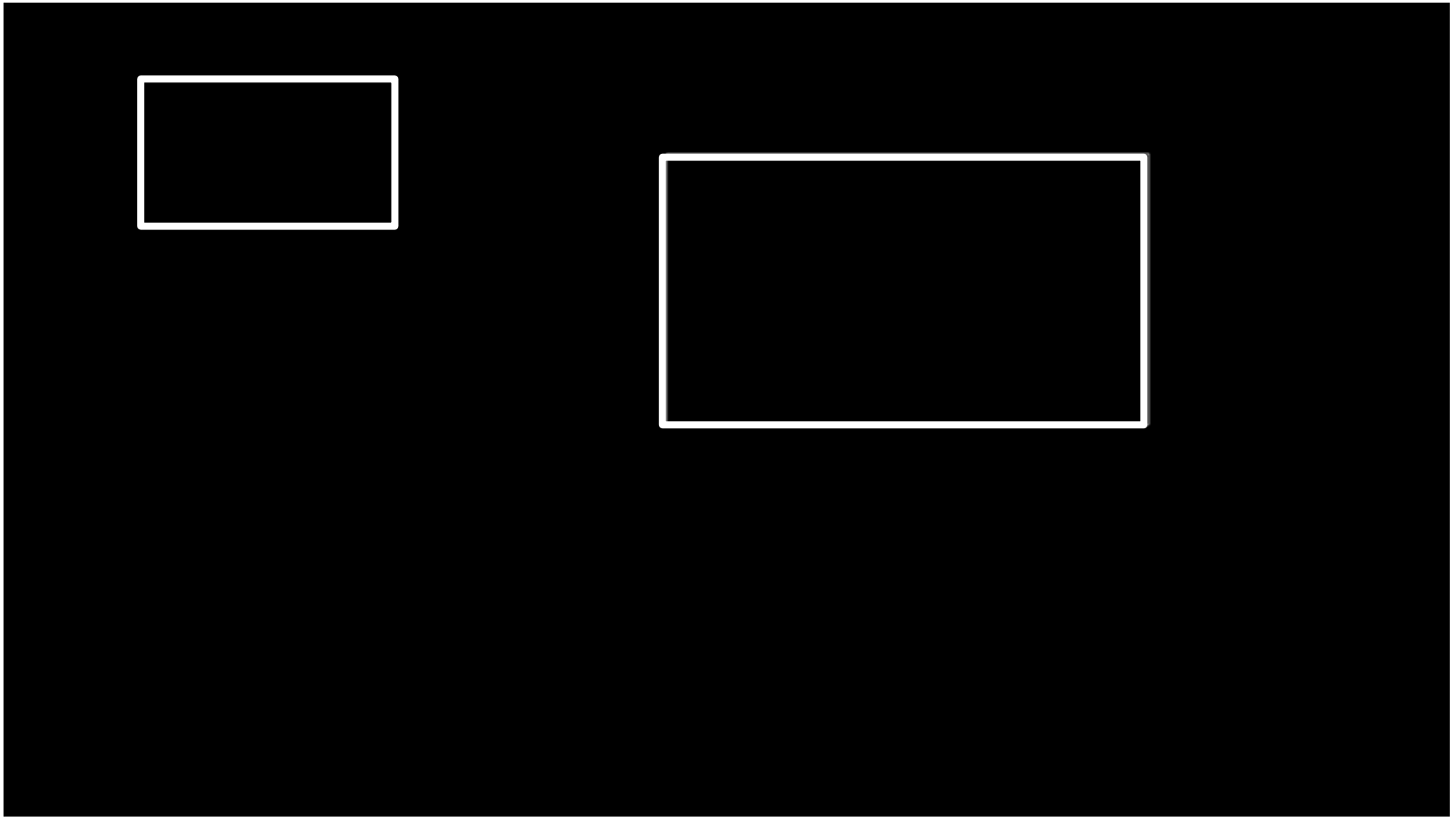


Shot 1

Final Video

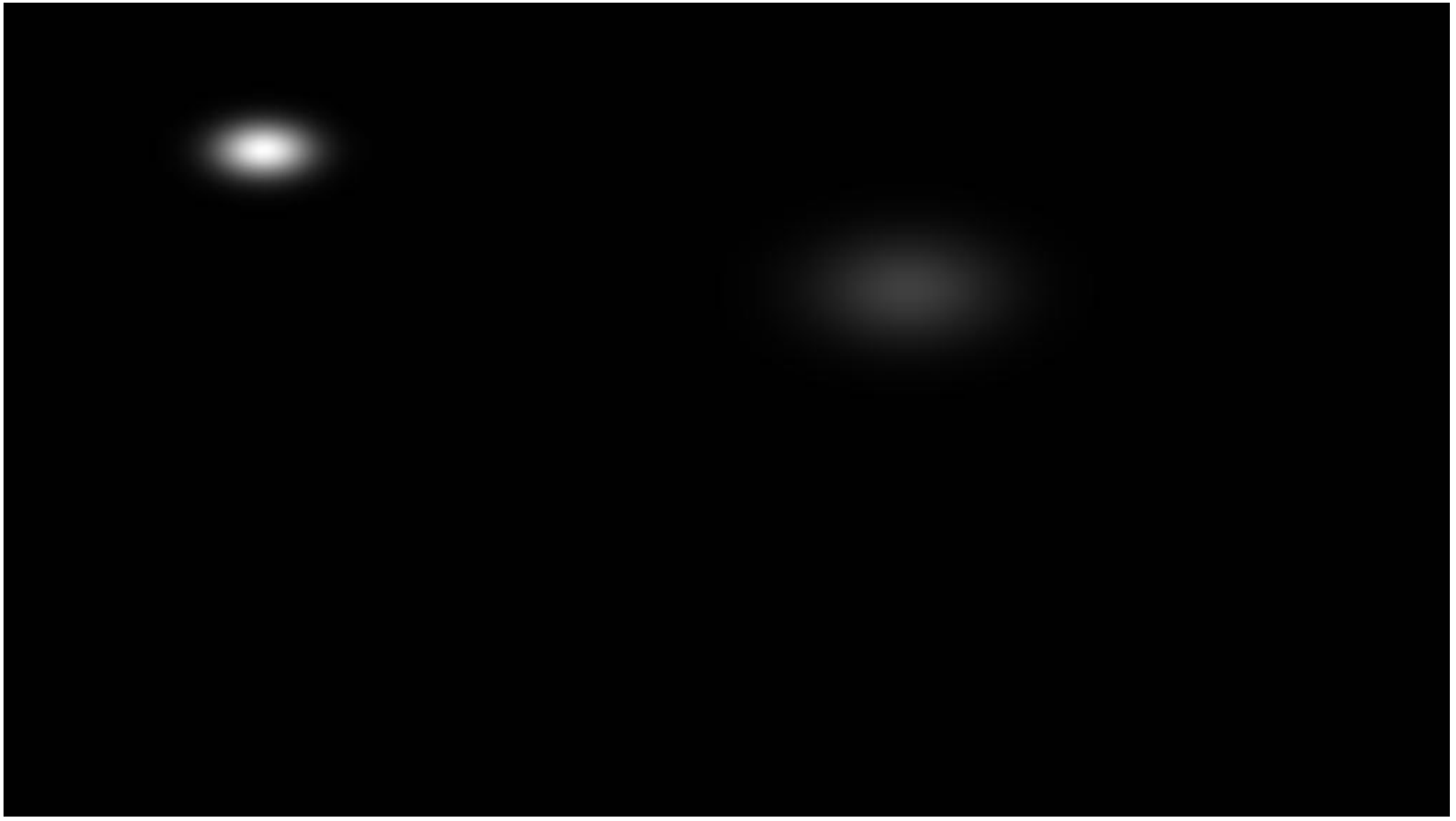
Creating Heatmaps

- Modelization of ROIs as a GMM (Gaussian Mixture Model)

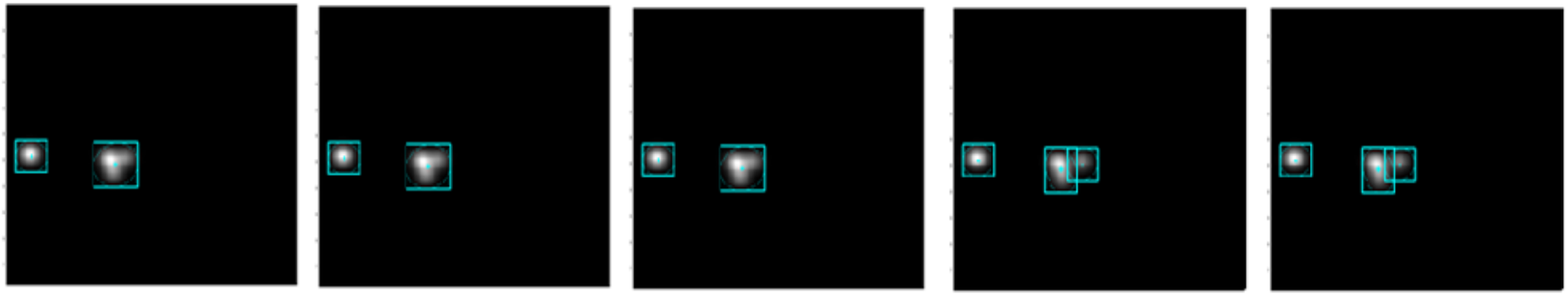


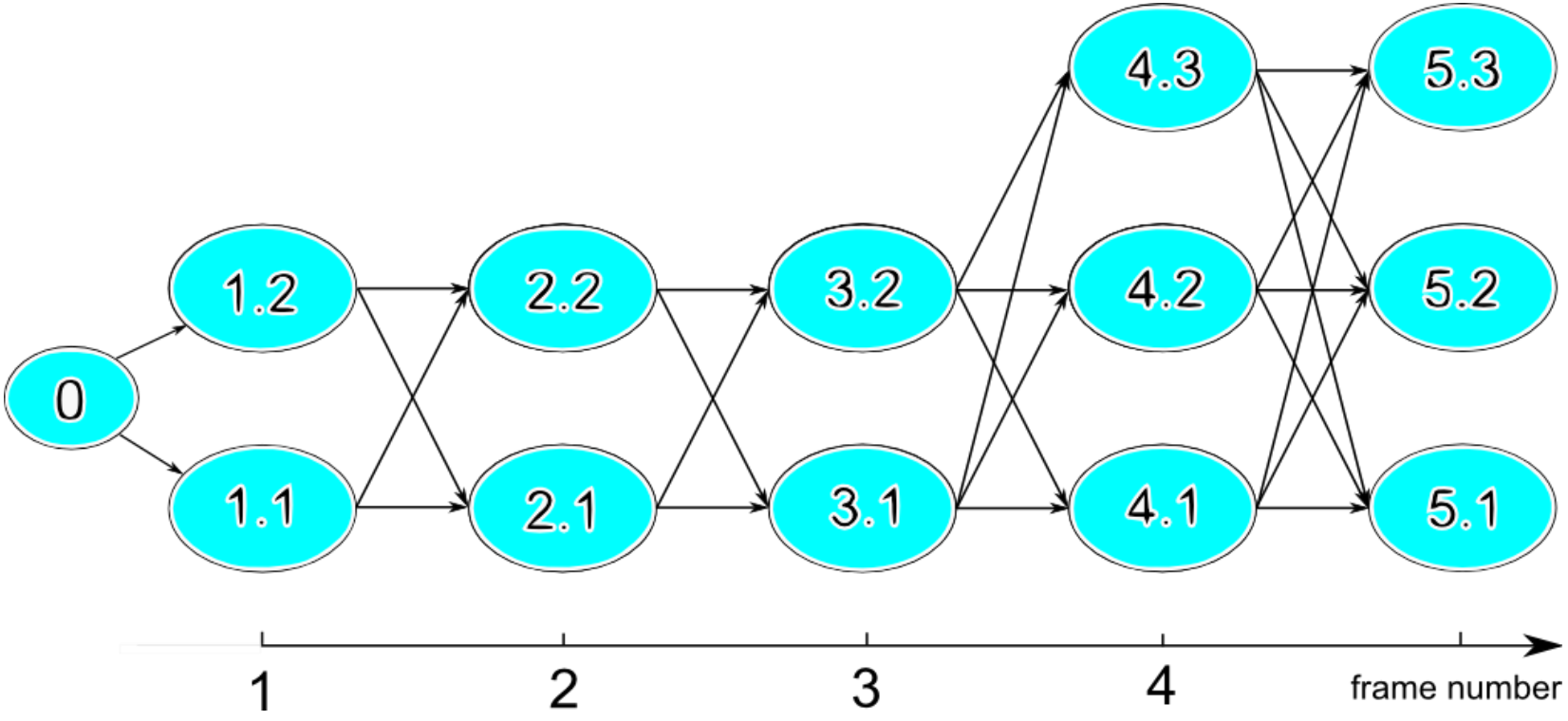
Creating Heatmaps

- Modelization of ROIs as a GMM (Gaussian Mixture Model)

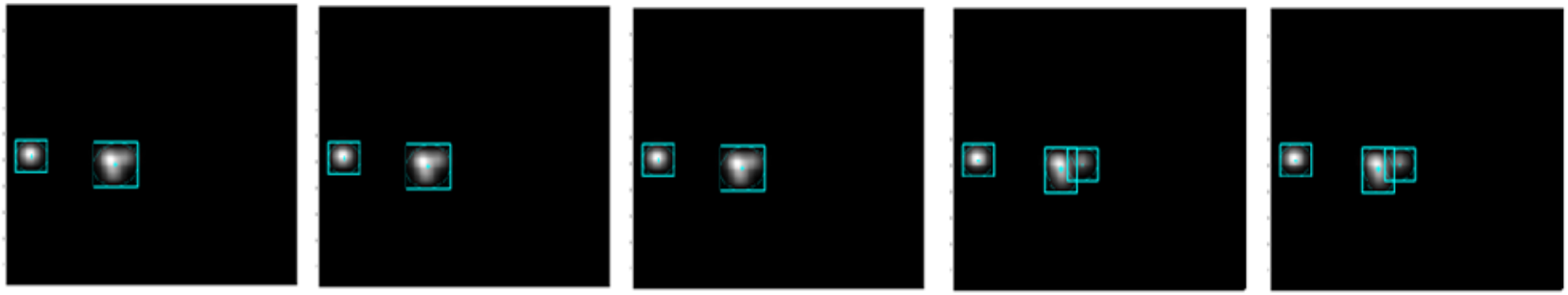


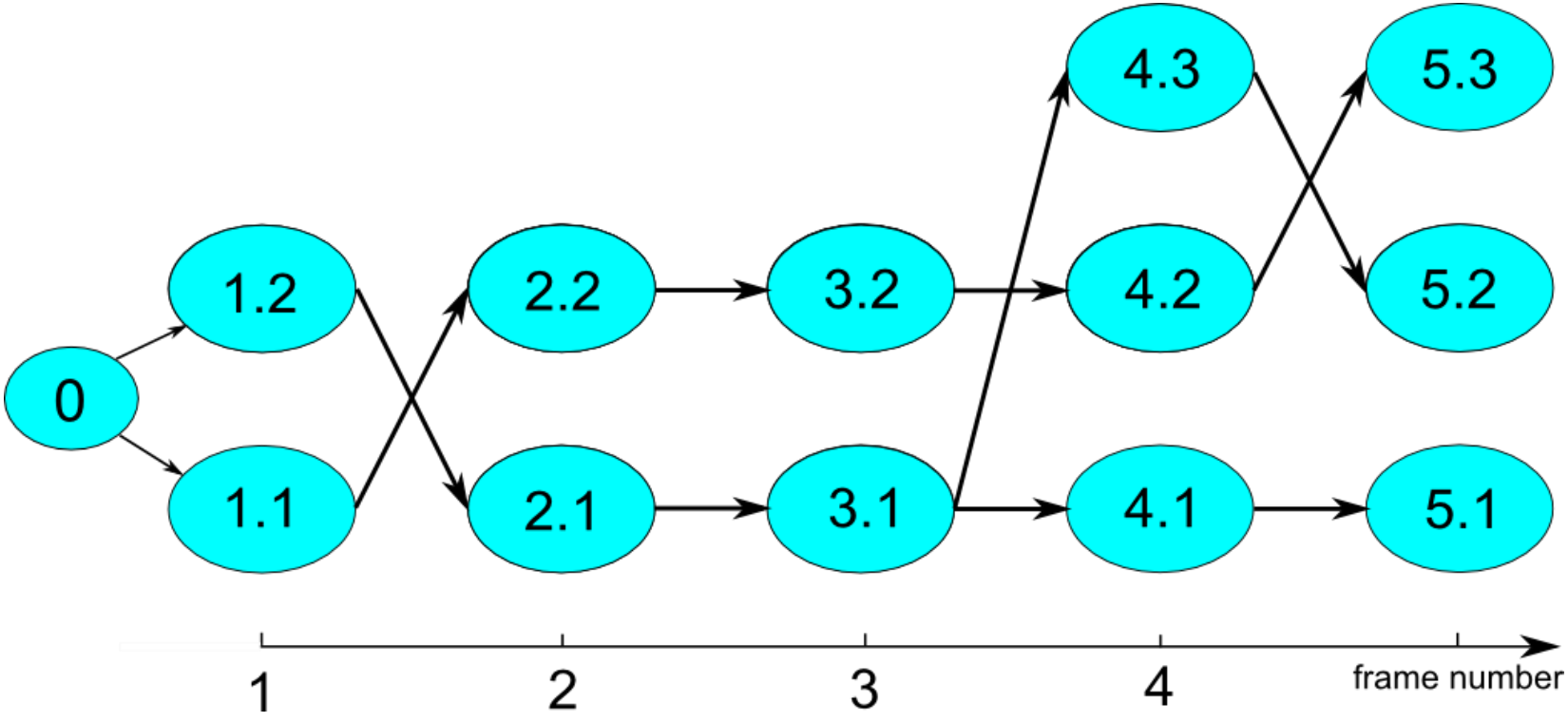
Building a ROI dynamics graph



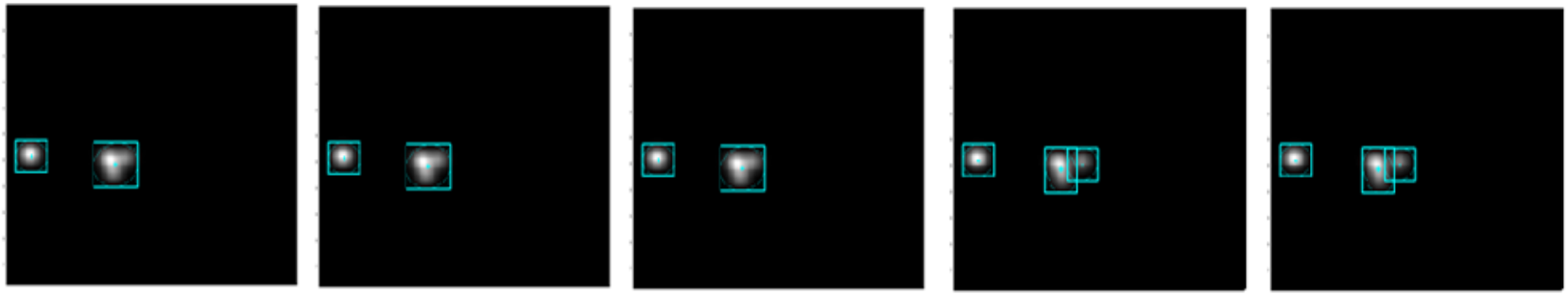


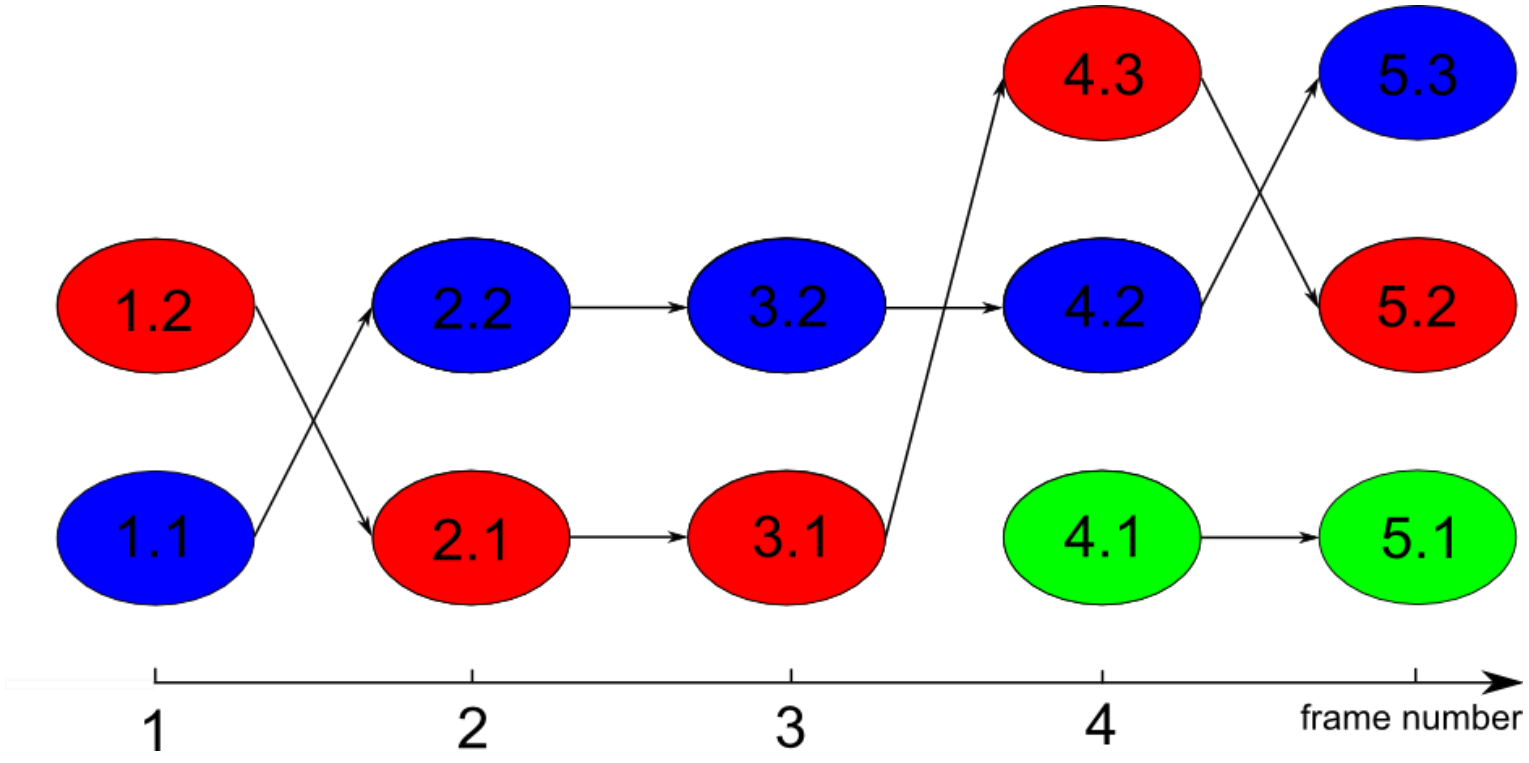
Minimal spanning tree

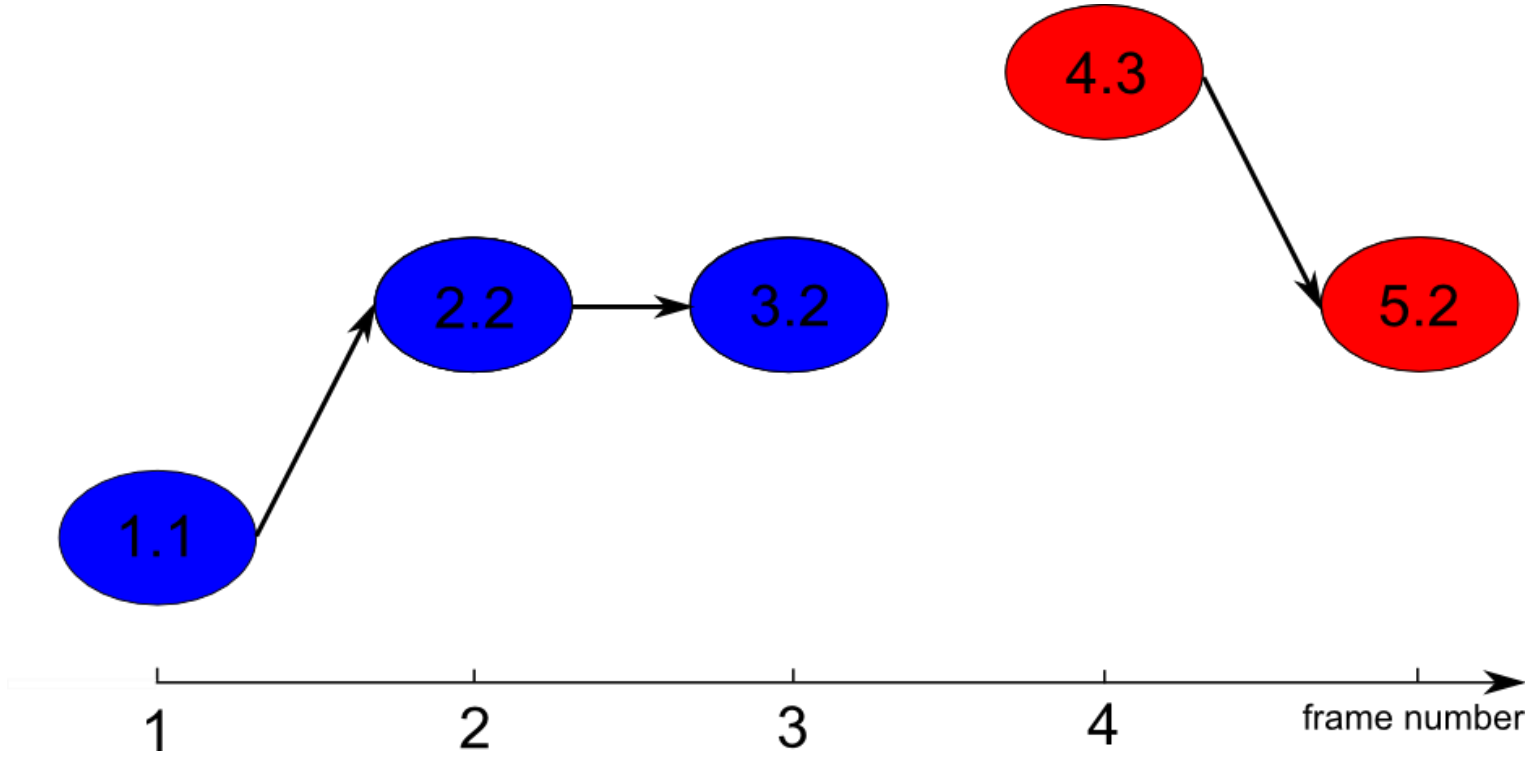




Cutting the tree into shots





Shots selection

Shots are selected according to their popularity: