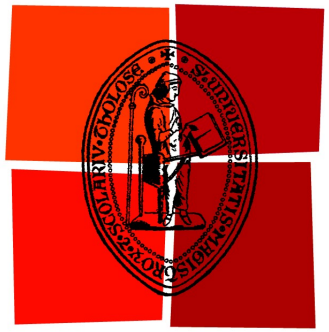


Université
de Toulouse

[simula]

NUS
National University
of Singapore

3D Interest Maps from Simultaneous Video Recordings

Axel Carlier
Université de Toulouse

Lilian Calvet
Simula Research Laboratory

Duong T. D. Nguyen and
Wei Tsang Ooi
National University of Singapore

Pierre Gurdjos and
Vincent Charvillat
Université de Toulouse

Regions Of Interest

- **ROI**: a region of a multimedia content that contains semantic information that a user or a group of users may find interesting.
- Highly subjective, dependent on
  - Users
  - Context
  - Etc.

→ Difficult to predict automatically

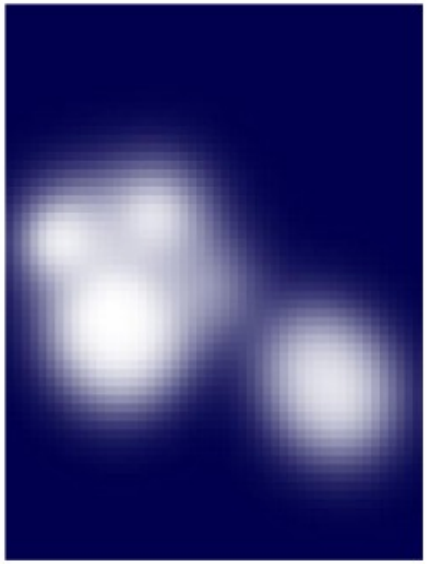
Related Work

- Saliency detection

Goferman et al, PAMI*'12*

- Usage-based approaches

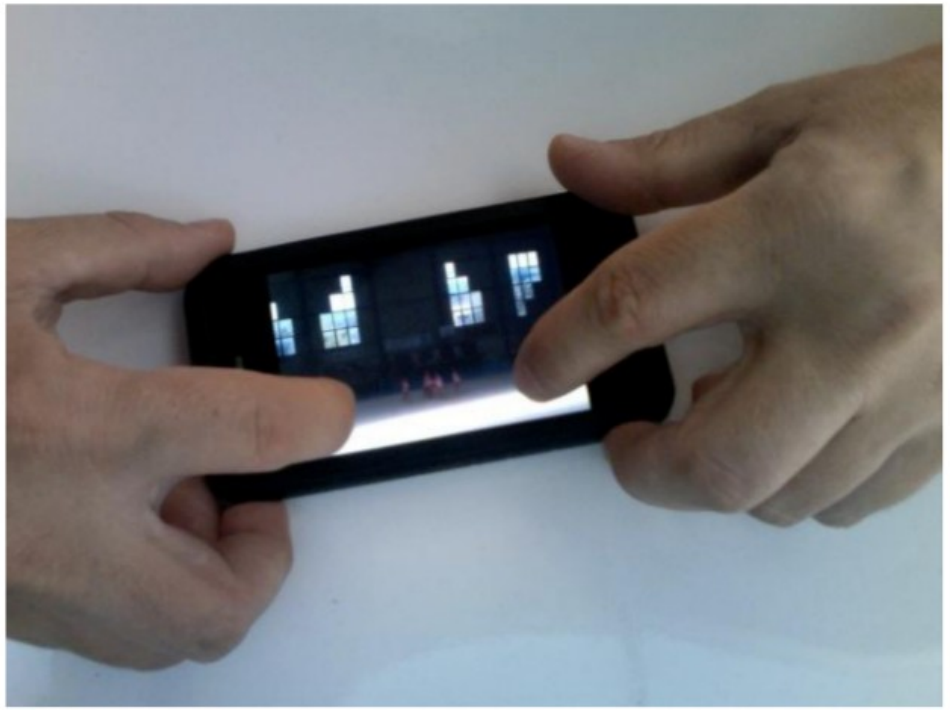
(a) Original image

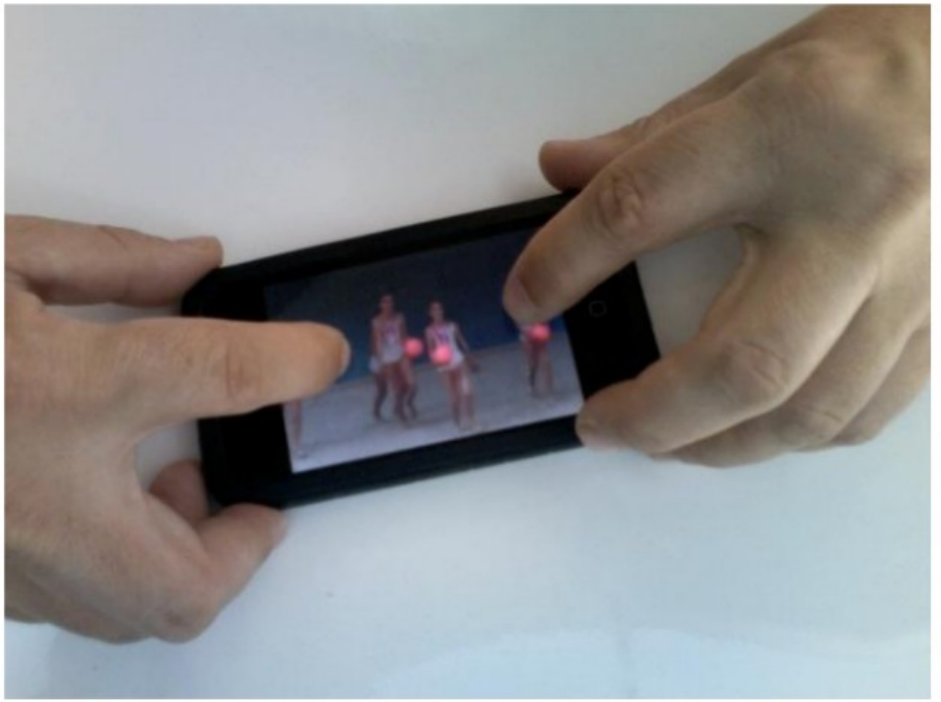


(b) User interest map

Xie et al, *CHI'05*

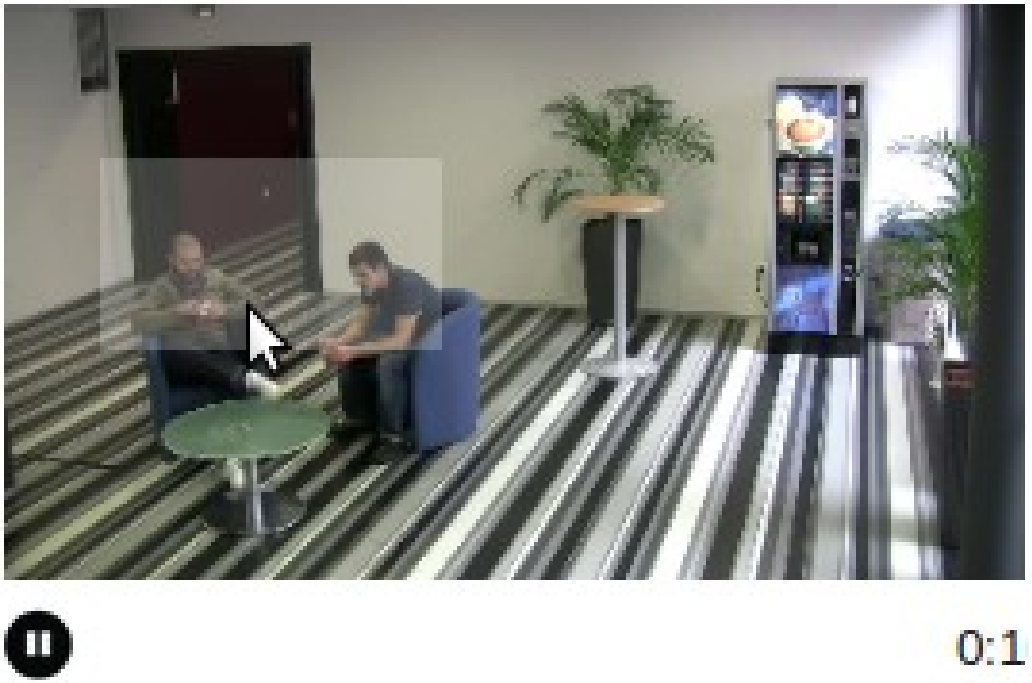
- What people zoom into is interesting





[Carlier et al] **Crowdsourced Automatic Zoom and Scroll for Video Retargeting**, MM *2010*

Zoomable Video



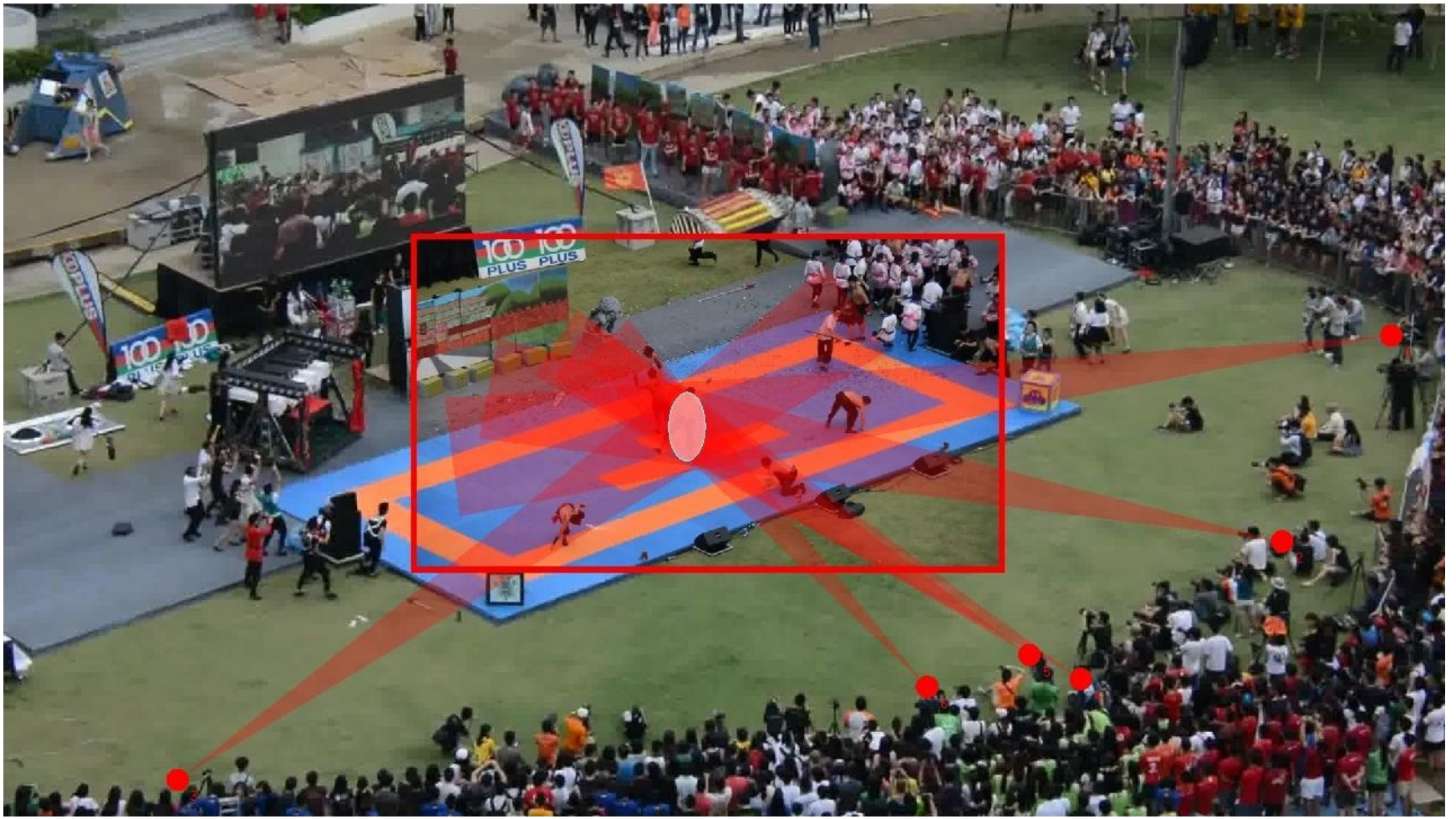
Viewports

User interest maps

From 2D to 3D

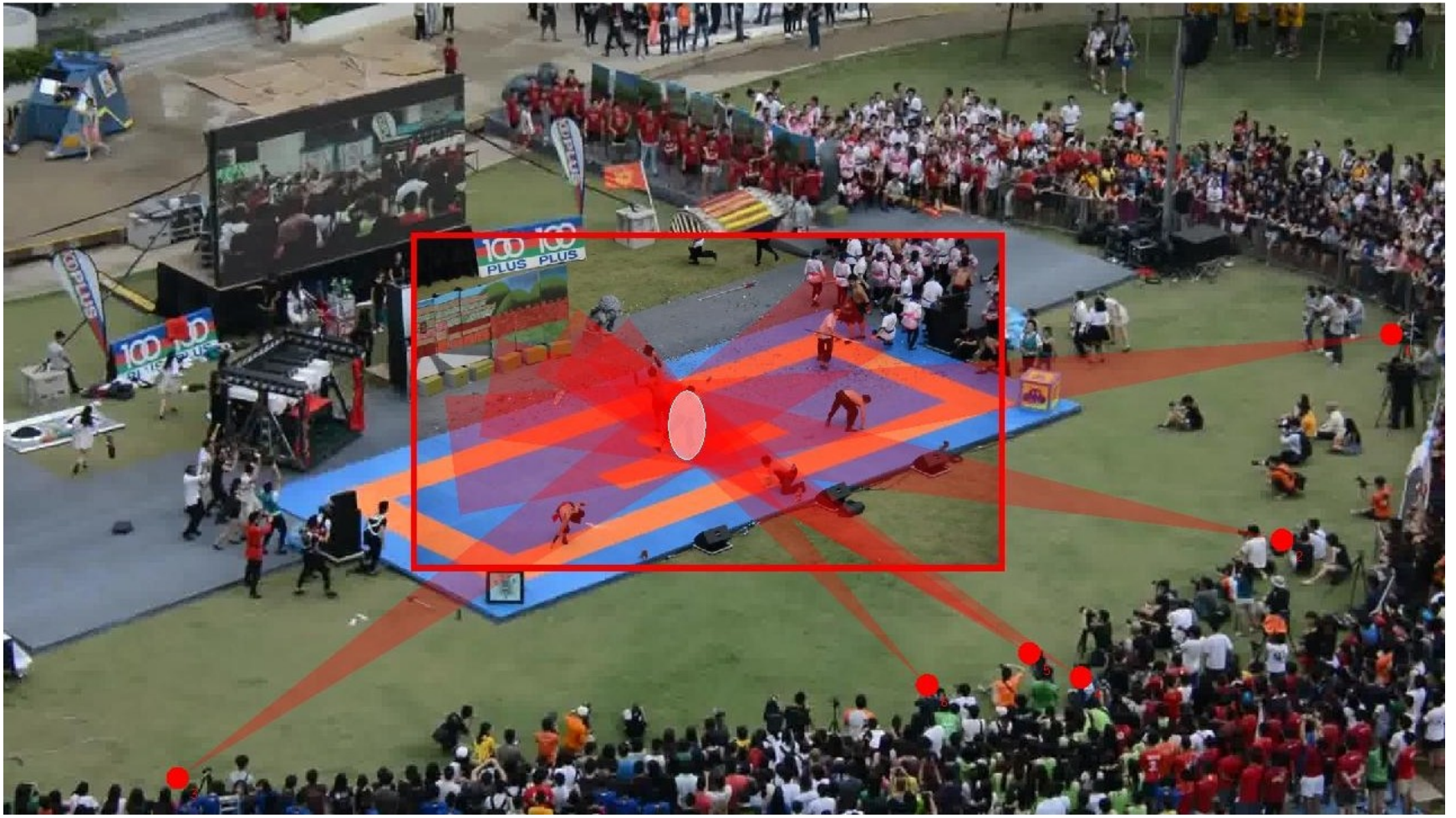
Idea

- What people choose to film is interesting



The setup

- Assumptions: calibration and synchronization to a certain extent

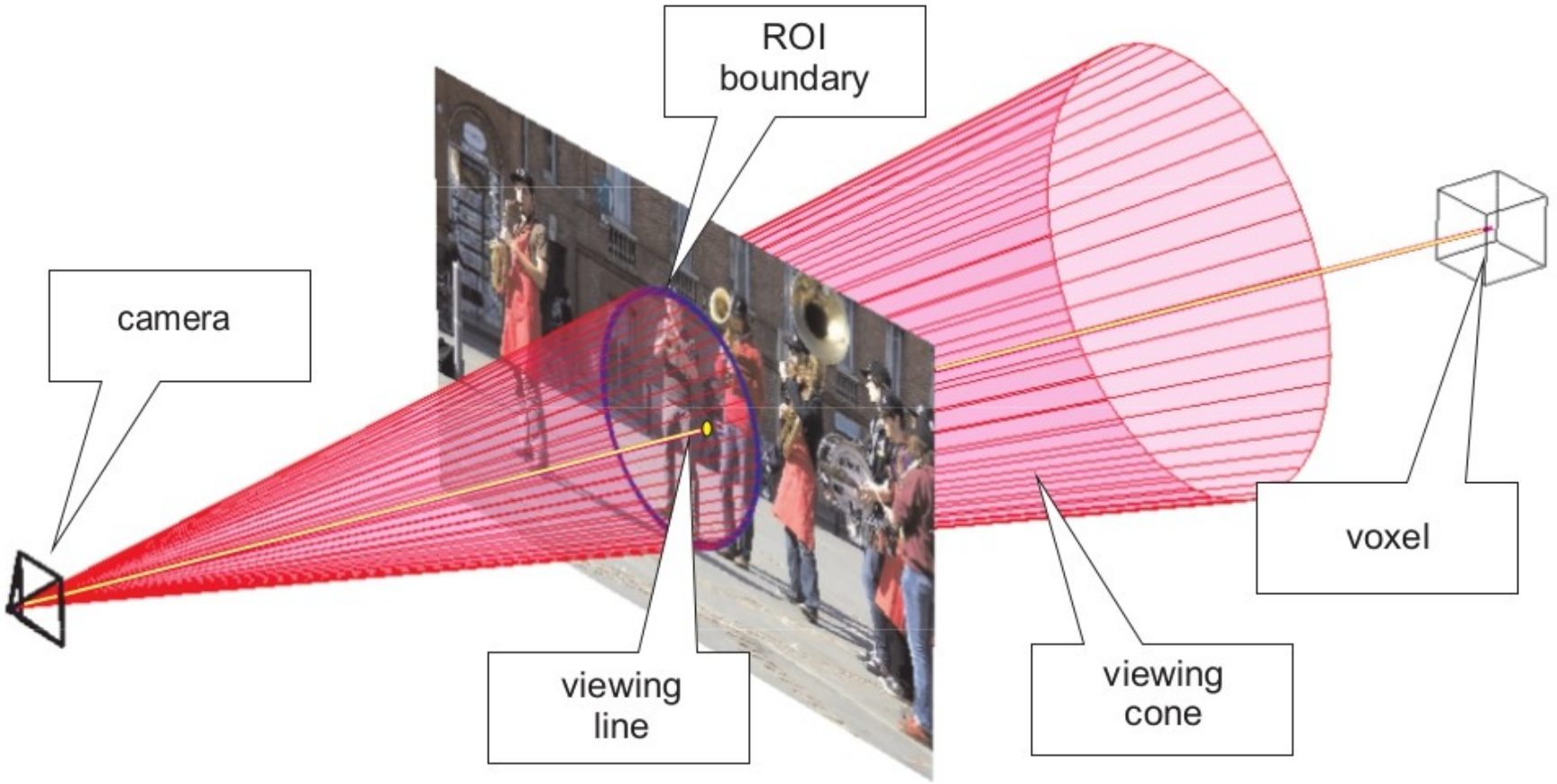


2D Regions Of Interest

With motion

Without motion

Back-Projection to 3D



3D Interest Maps

- Measure of the interest of a voxel v_i

$$Int(v_i) = \frac{1}{|\mathcal{E}|} \sum_{E \in \mathcal{E}(v_i)} \frac{\mathrm{Vol}(\Lambda(E) \cap v_i)}{\mathrm{Vol}(v_i)}$$

$\mathcal{E}$	Set of 2D ROIs on all images
$\Lambda(E)$	Viewing cone of the ROI E

3D Interest Maps

- Probabilistic form of the interest

$$\widetilde{Int}(v_i) = Int(v_i) / \sum_{i \in I} Int(v_i)$$

A **3D interest map** is the limit form of the 3D histogram with voxels as bins with respect to this measure.

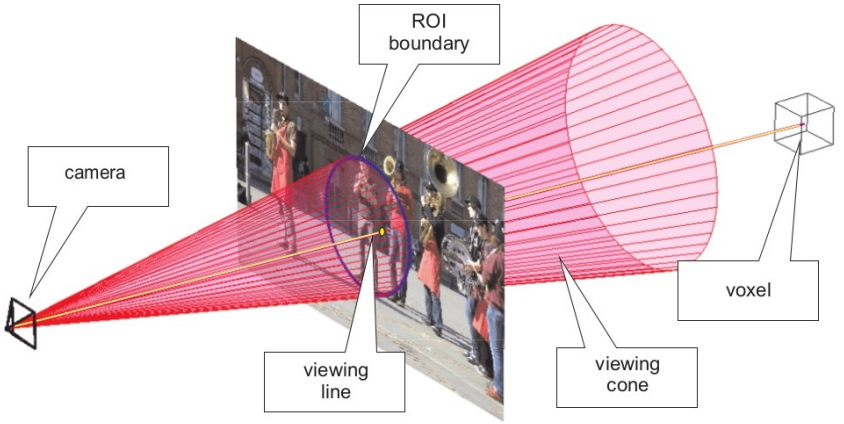
3D Interest Maps

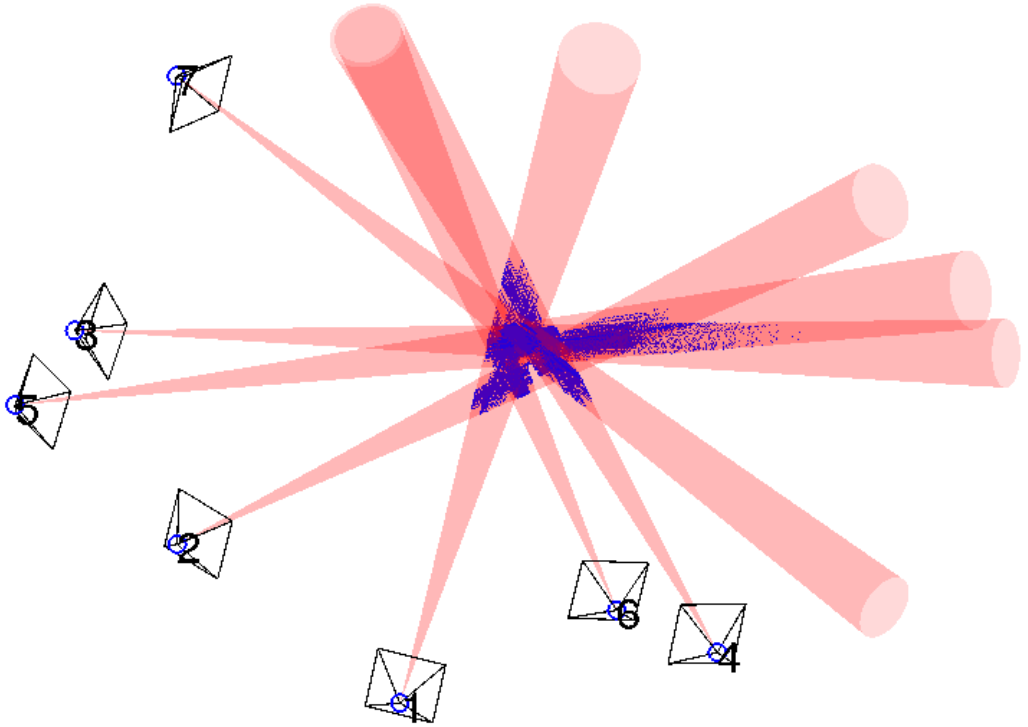
- We model our **3D interest map** with a Gaussian Mixture Model:

$$\widetilde{Int}(v) = \sum_{g=1}^{G} w_g \mathcal{N}(v; \boldsymbol{\mu}_g, \Sigma_g)$$

How to estimate the GMM parameters?

GMM estimation





Mean Shift Clustering

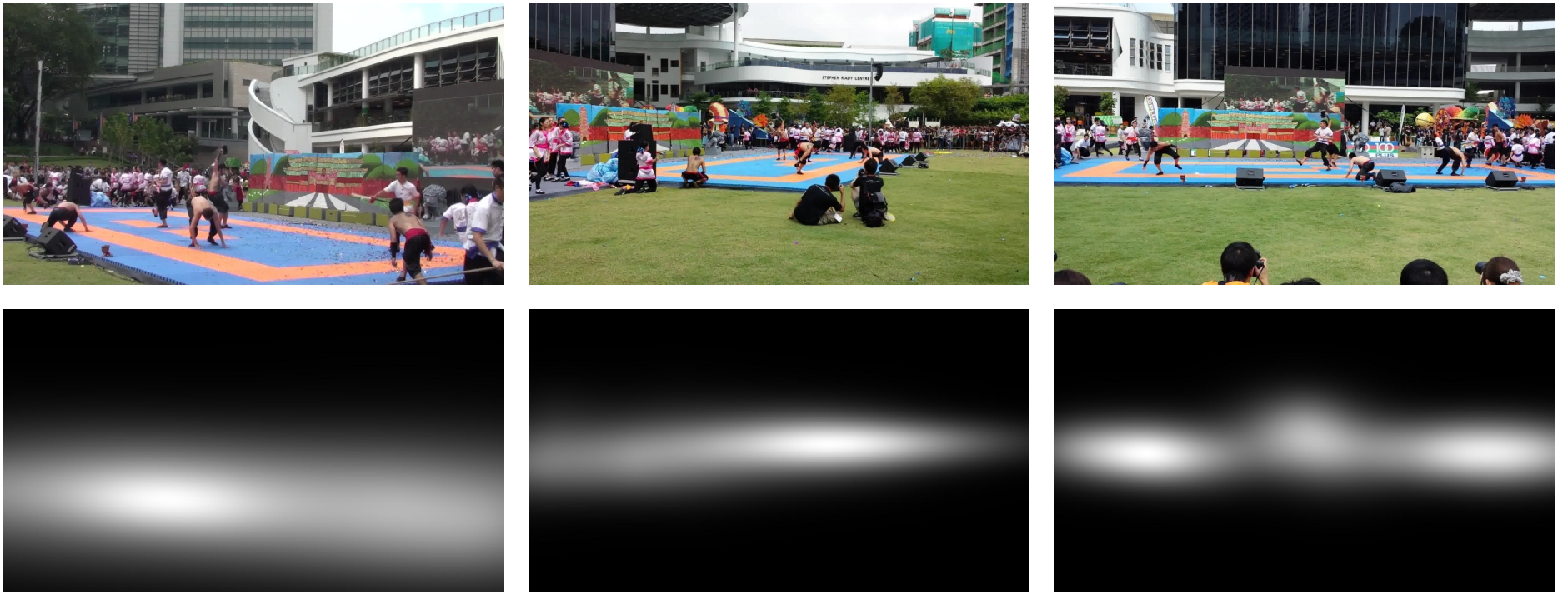
↓

1 Cluster

=

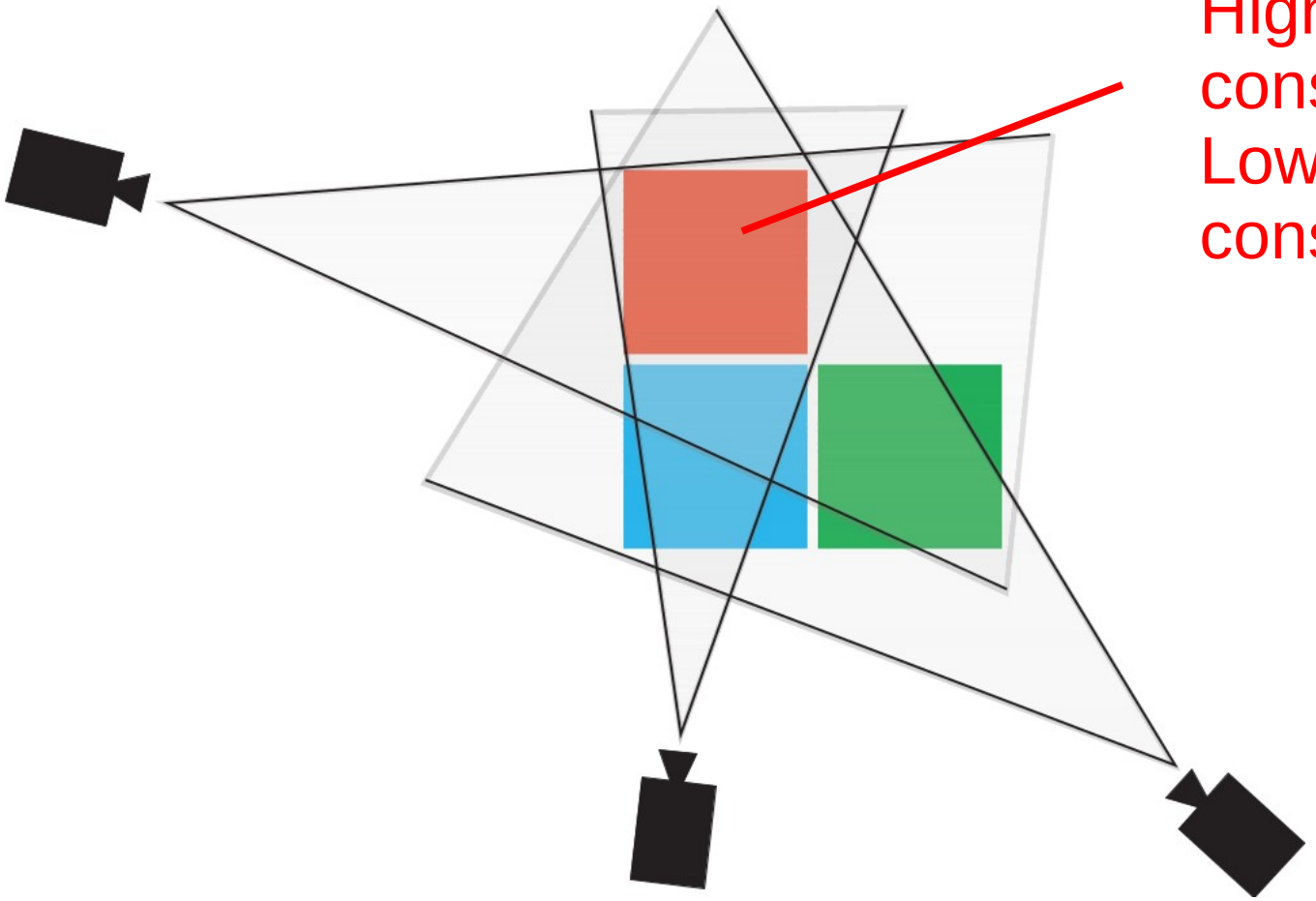
1 Gaussian

Results



Coarse results, because of the lack of photometric consistency

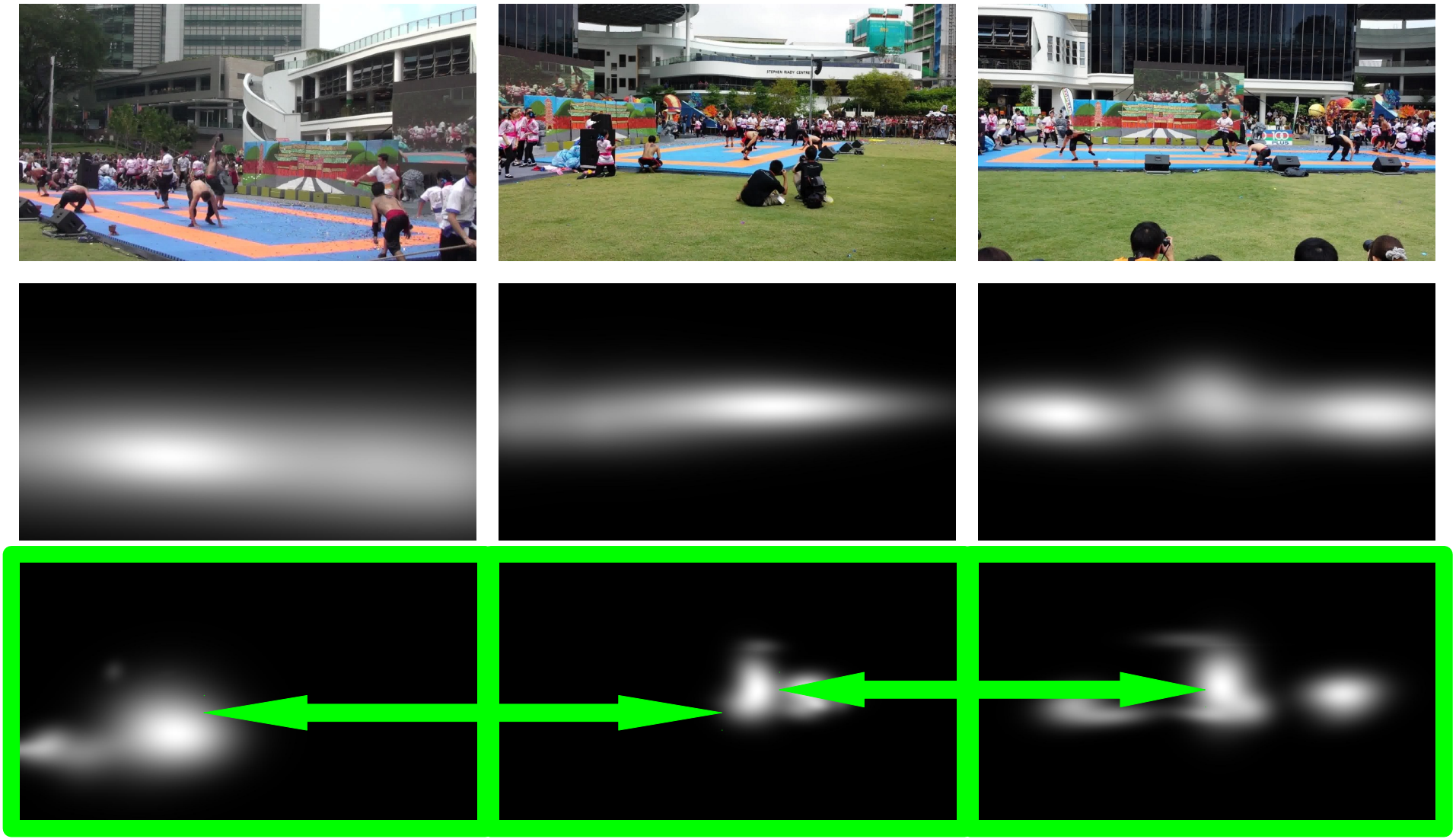
Photometric Consistency



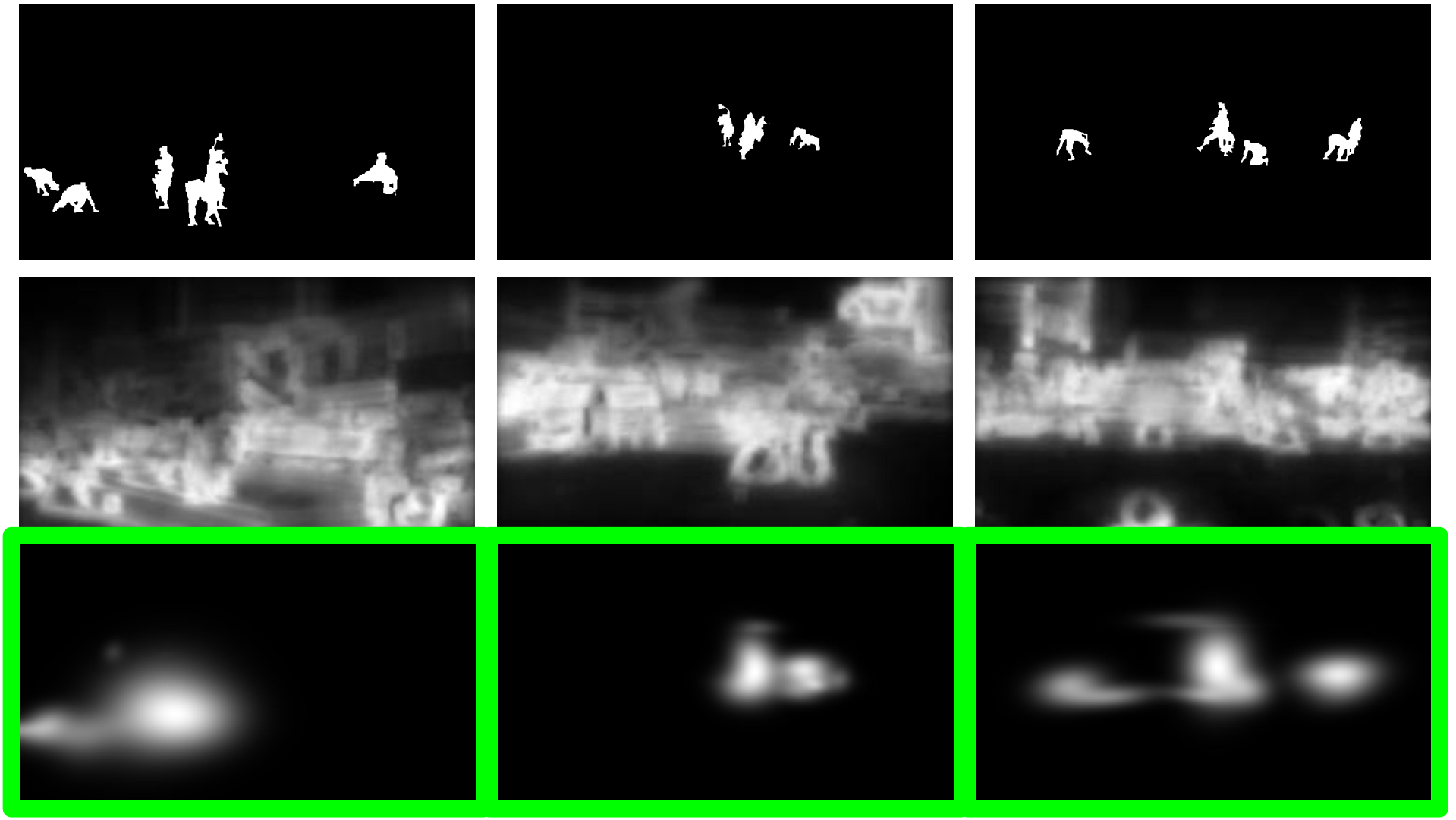
High geometric consistency
Low photometric consistency

[Furukawa et al] **Accurate, dense, and robust multiview stereopsis**, PAMI *2010* (a.k.a. PMVS software)

New results

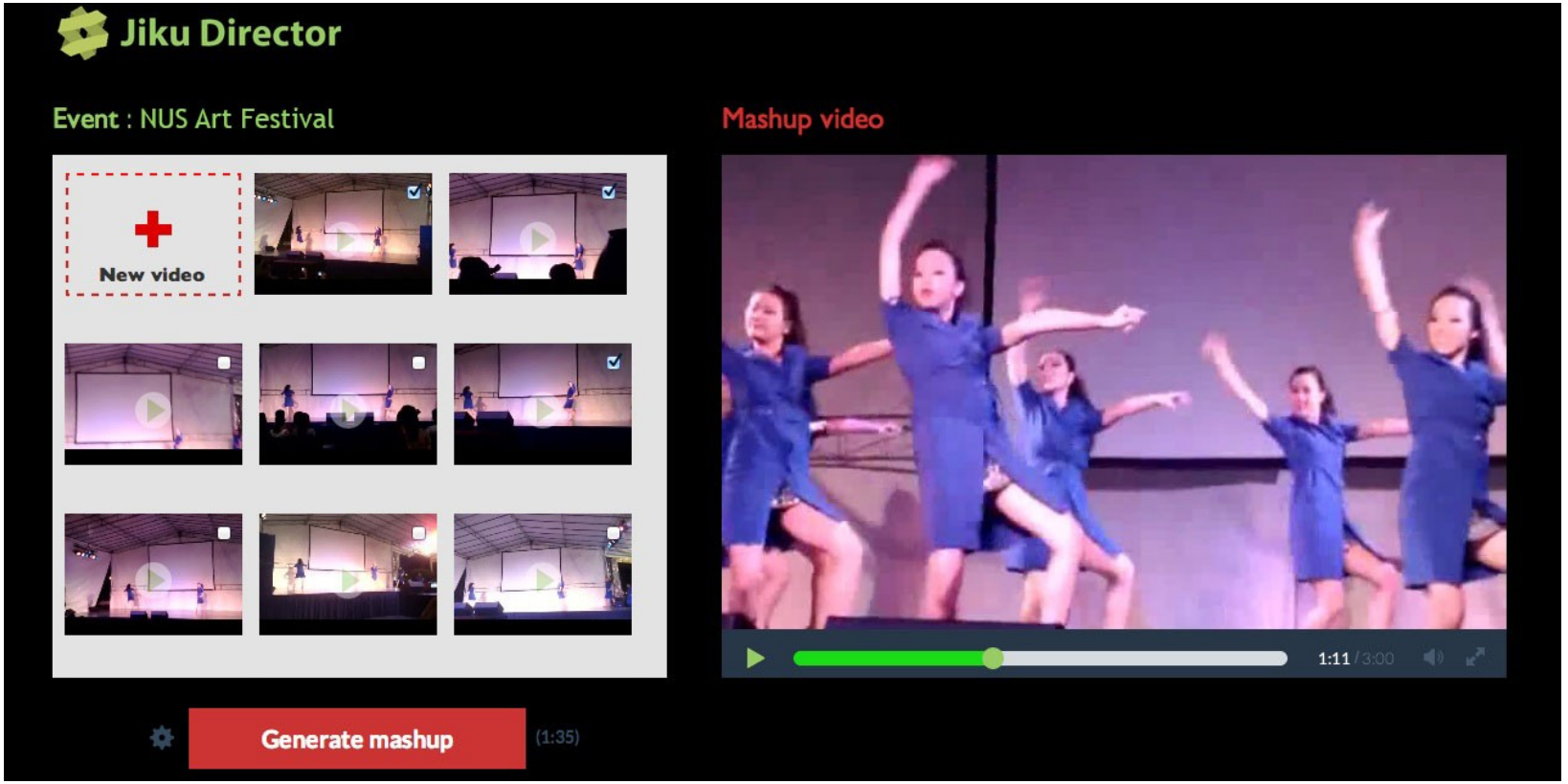


Comparison with Saliency



An application of 3D Interest Maps

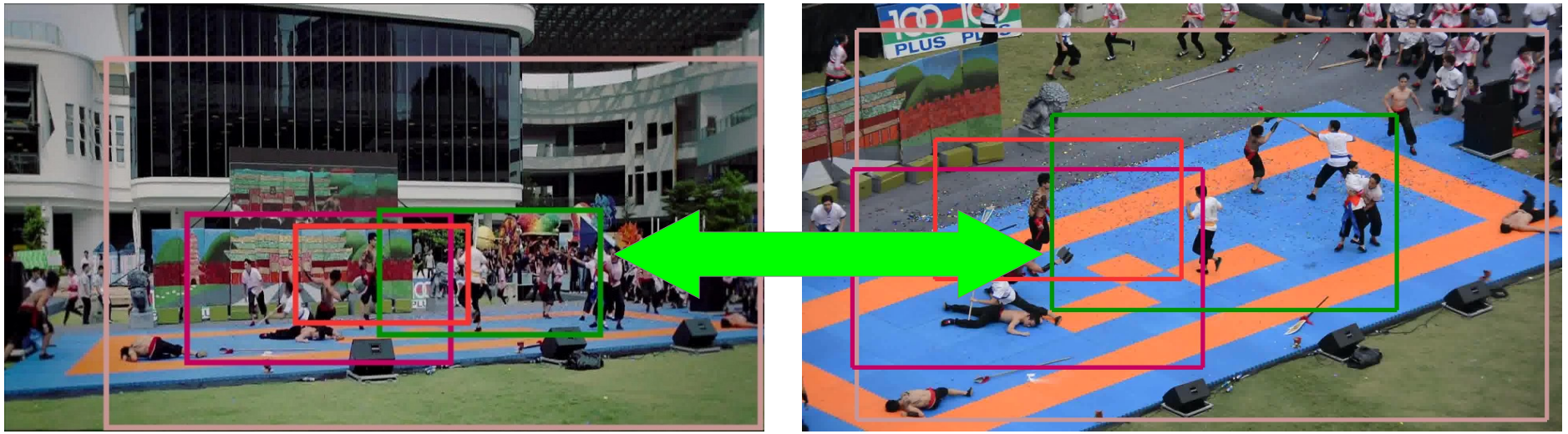
Mashup Video



[Saini et al] **Movimash: online mobile video mashup**, *MM 2012*

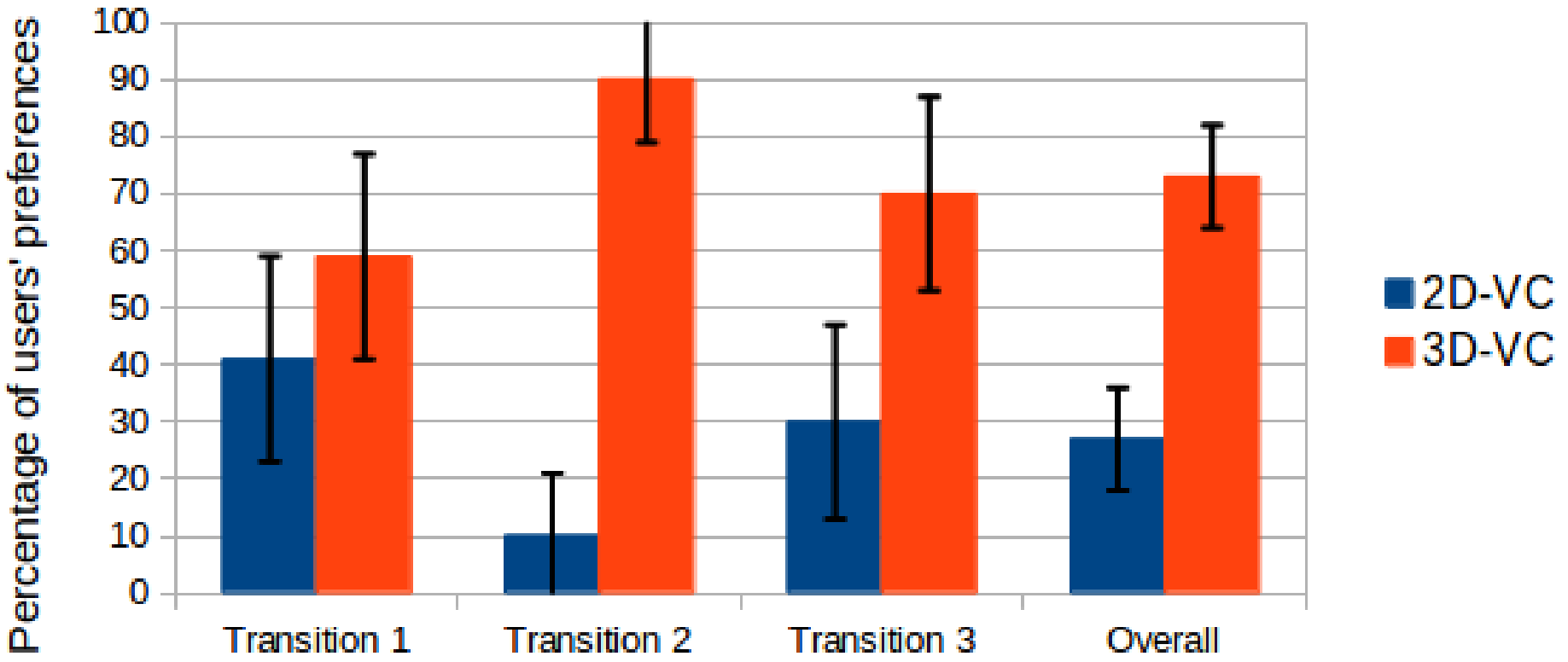
Automatic Video Edition

3D transition


PLUS

Result

Evaluation



2D-VC: JikuDirector 2.0 (demo)

3D-VC: this paper

Conclusion

- Formal definition of 3D interest maps
- A common space for representing interest in many simultaneously recorded videos
  - Thanks to our strong assumptions, this common space is the 3D space
- Many applications, including video mashup