Creating Visual Stories

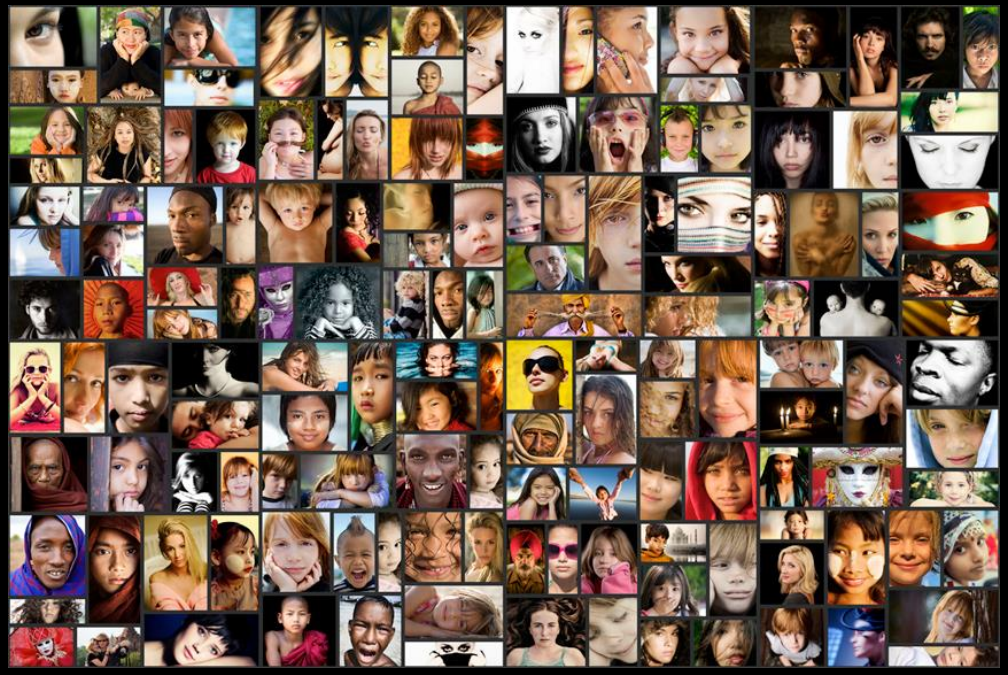
semantic of visual content

Prof. Ariel Shamir

Interdisciplinary Center, Herzliya

Data is Everywhere (and becomes visual)

- **350 million** new photos are posted on Facebook each day
- **100 hours** of video are uploaded to YouTube every minute



Data Gathering

Gordon Bell's
Life Bits

GoPro Cameras

Google Glass

Wearable Devices

The Semantic Gap

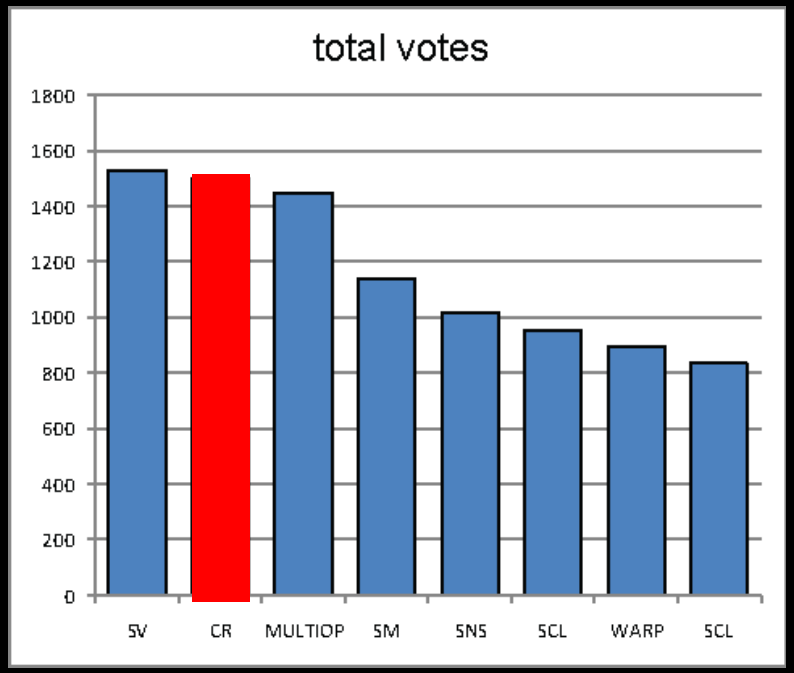
- The challenge is not to gather data, display it, store it, but… **understand** it!

- Goal: create a coherent visual story out of a set of images or videos
- Challenging because it involves **semantics**!

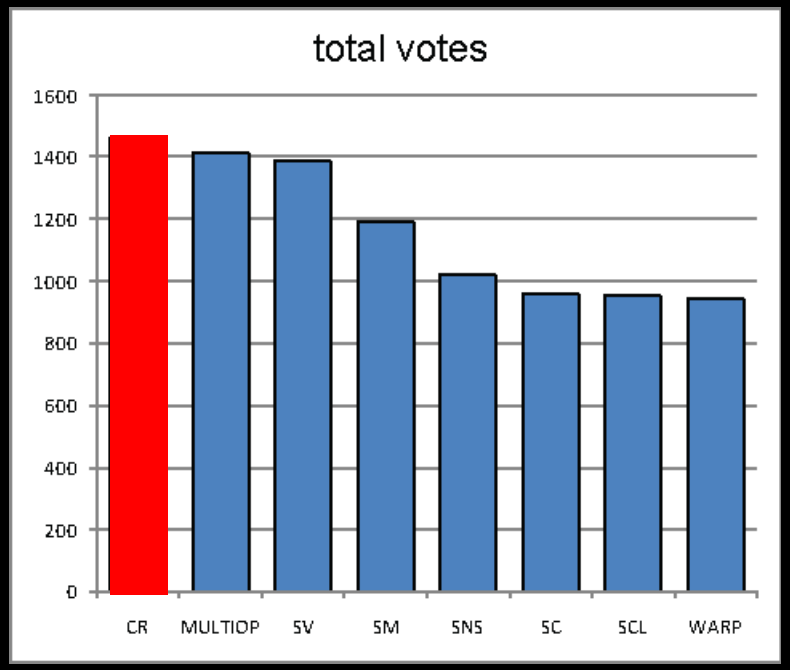
Video Retargeting

From: **Multi-operator Media Retargeting**
Michael Rubinstein · Ariel Shamir · Shai Avidan
ACM Transactions on Graphics, Volume 28, Number 3, SIGGRAPH 2009

Comparison of Operators



With original viewed



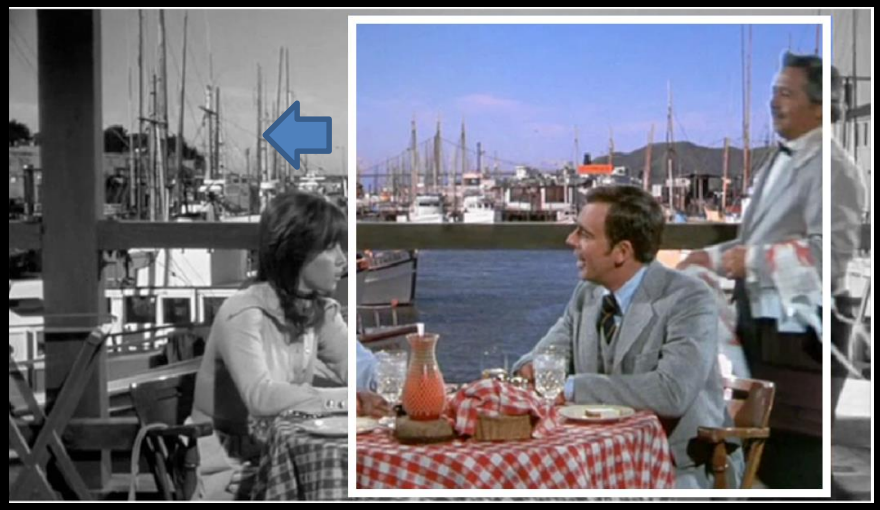
Without original viewed

From: **A Comparative Study of Image Retargeting**
Michael Rubinstein · Diego Gutierrez · Olga Sorkine · Ariel Shamir
ACM Transactions on Graphics, Volume 29, Number 5 (Proceedings SIGGRAPH Asia 2010)

Problem?

Enhancing Crop

- Panning



- Cutting

Re-Editing of a Movie

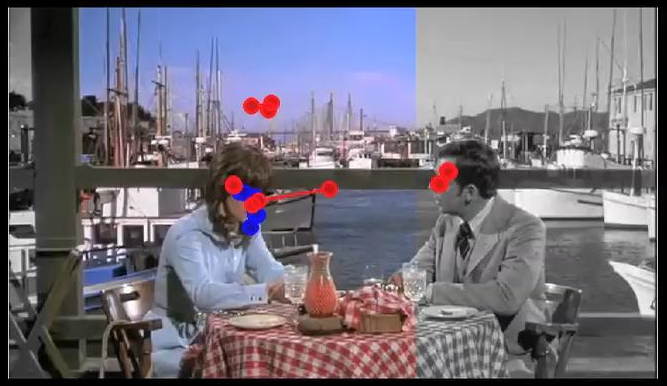
Our Result 1:1

But… how does the algorithm know what is important?

Saliency Measures

Semantics: Eye Tracking

- Viewers look at what is important:



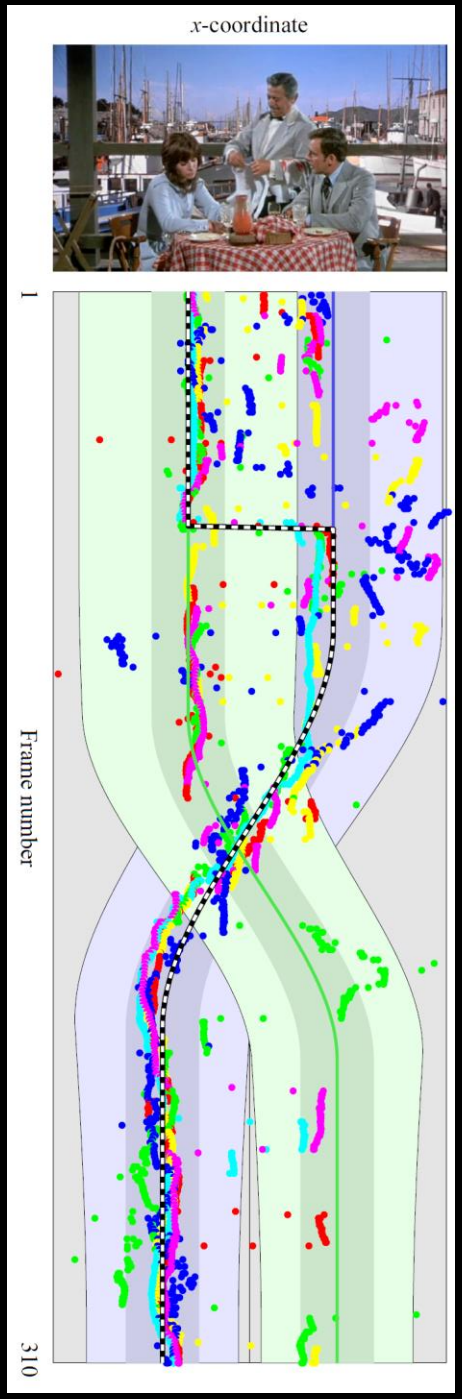
Red: Gaze Data on Original Widescreen

Blue: Gaze Data on Our Result (1:1)

Two Challenges

1. What is important?
2. How to edit to follow the story?

Fitting Curves & Cutting



From: Gaze-driven Video Re-editing
Eakta Jain, Yaser Sheikh, Ariel Shamir, Jessica Hodgins
ACM Transactions on Graphics, Volume 34, Issue 2, February 2015
Article No. 21

Sports

- A human filming crew is very expensive
- Static cameras capturing all field are boring

- We want to know where to point them and when to zoom in/out?

Same Two Challenges

1. What is important?
2. How to edit to follow the story?

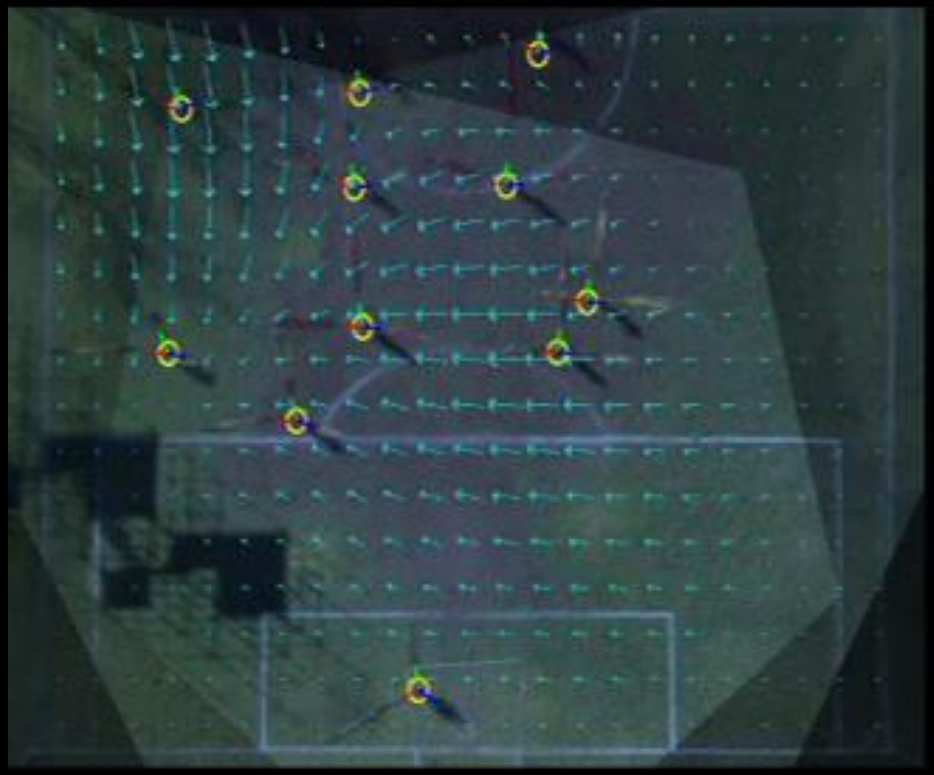
Key Idea

- You cannot track the eyes of viewers anymore
- But… you can track the players

1. Track Movement of Individual Players

Rectification

2. Build a Motion Field



3. Find Points of Convergence

Semantics: players movement

Soccer

Ball

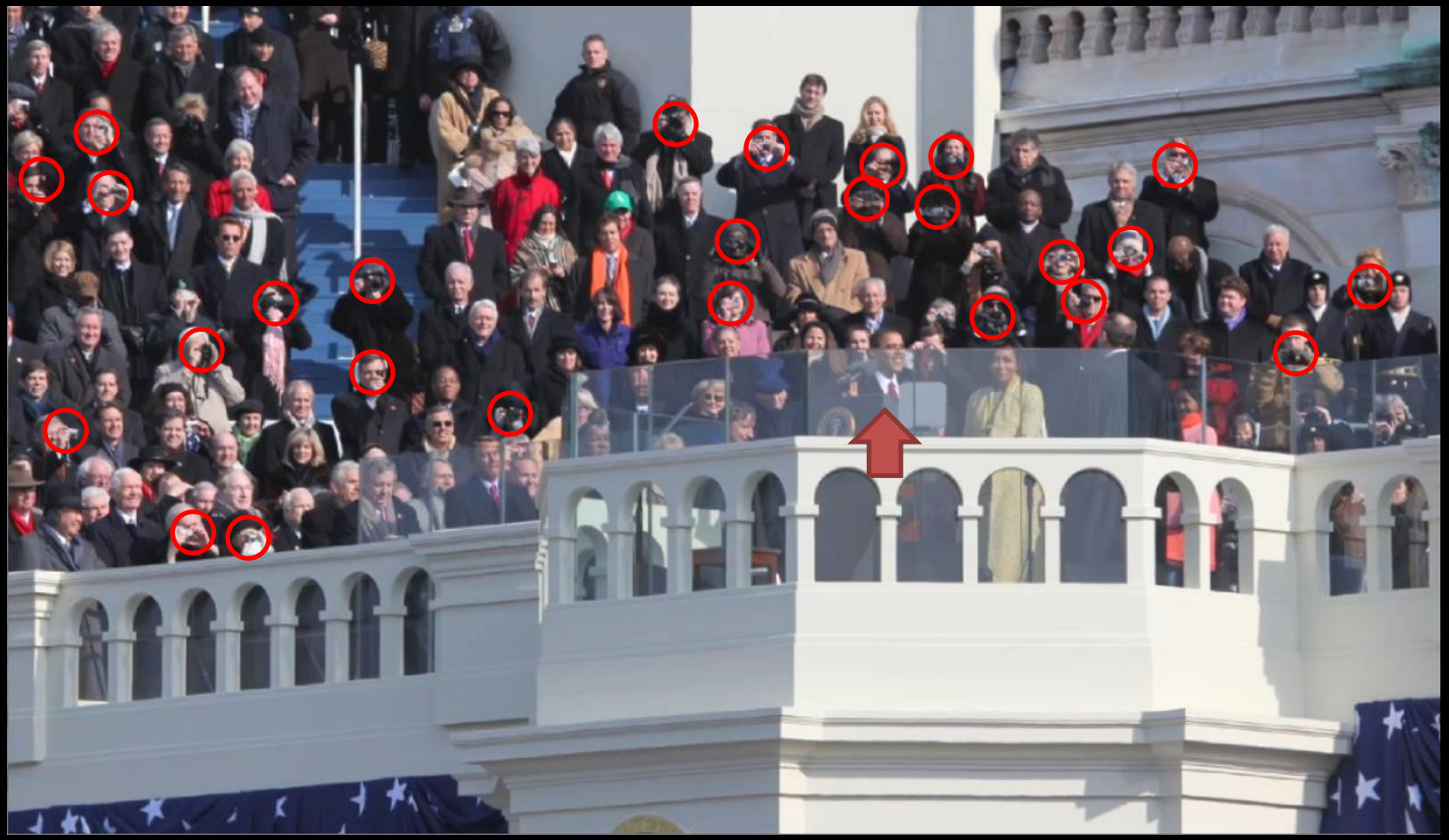
Ball

Basket Ball

Automatic Camera Control

Motion Fields to Predict Play Evolution in Dynamic Sport Scenes
Kihwan Kim · Matthias Grundmann · Ariel Shamir · Iain Matthews · Jessica Hodgins · Irfan Essa
Proc. Computer Vision and Pattern Recognition 2010

Many (Social) Cameras




Input: Synchronized Videos


Output: Coherent Video of Event













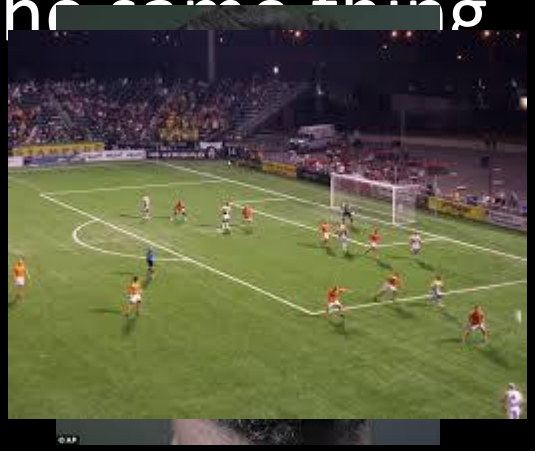


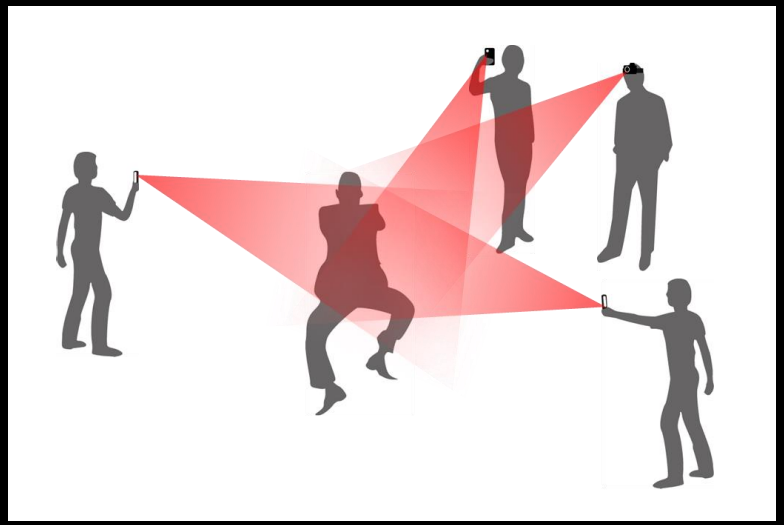

Challenges Again

1. What is important?
2. How to edit to follow the story?

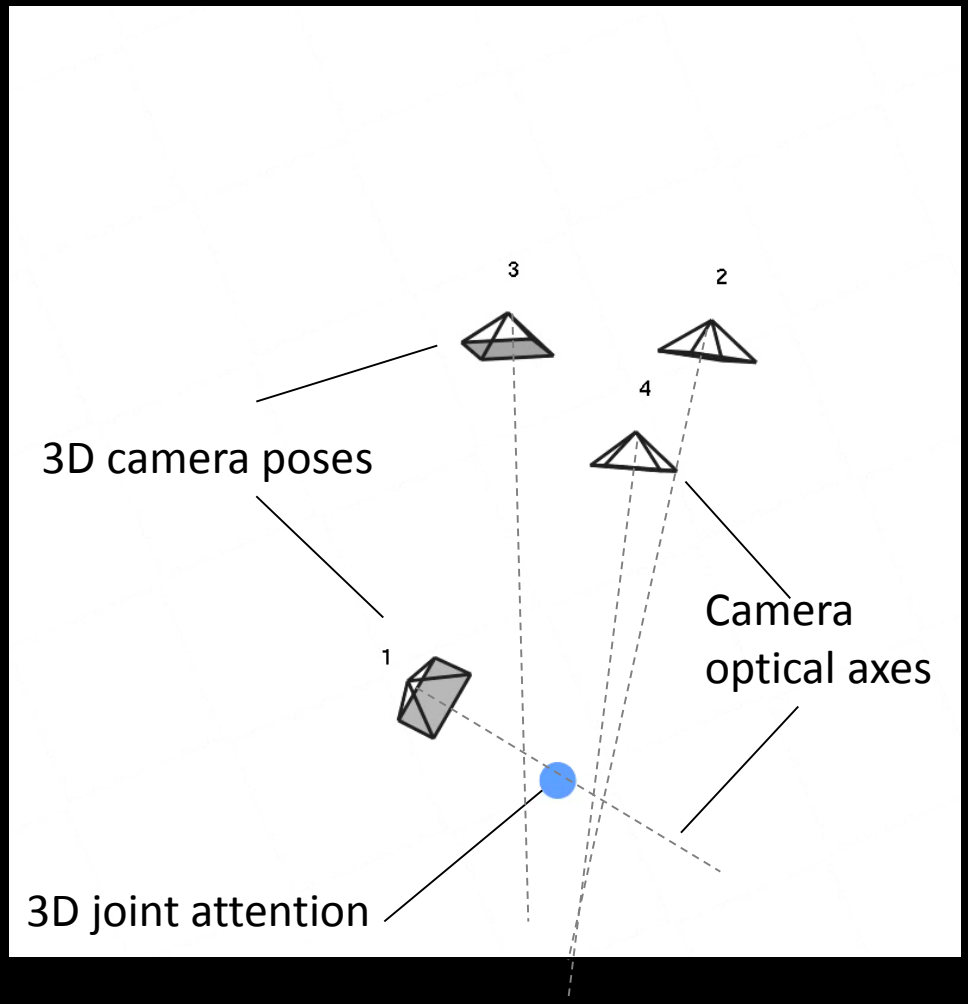
Semantics: Social Cues

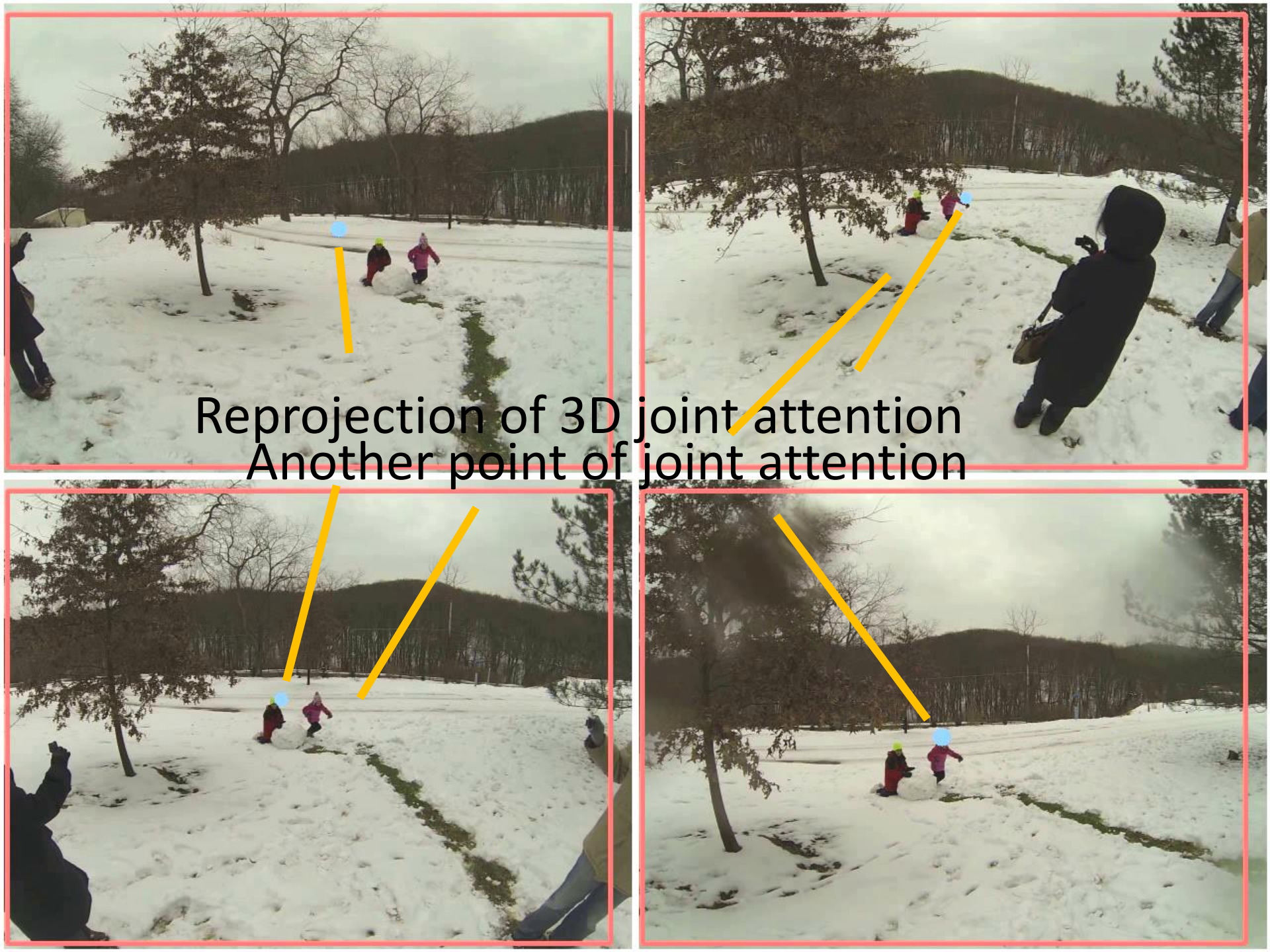
- It is difficult to know from the perspective of a single person
- And we don't have a group of players
- But… we have a group of viewers looking at the same thing





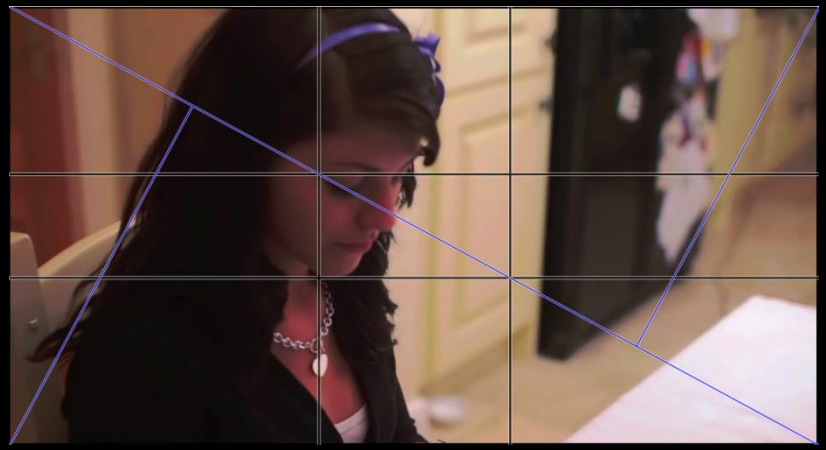
Structure from Motion
→ 3D Joint Attention



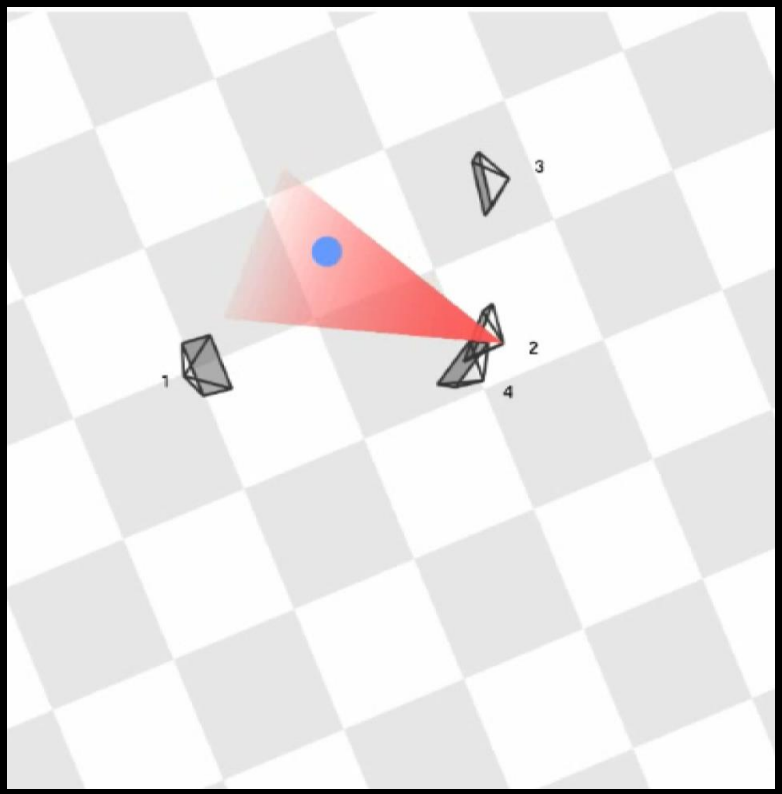

Reprojection of 3D joint attention
Another point of joint attention

Editing: Cinematography Rules

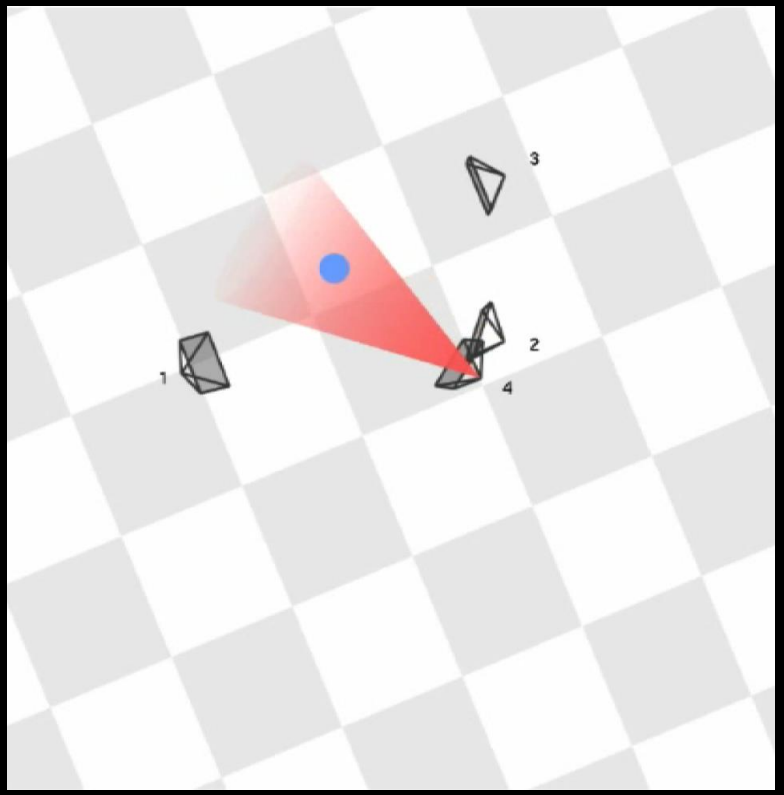
- Jump cut avoidance
- 180 degree rule
- Length of shot
- Variety
- Composition
- More…



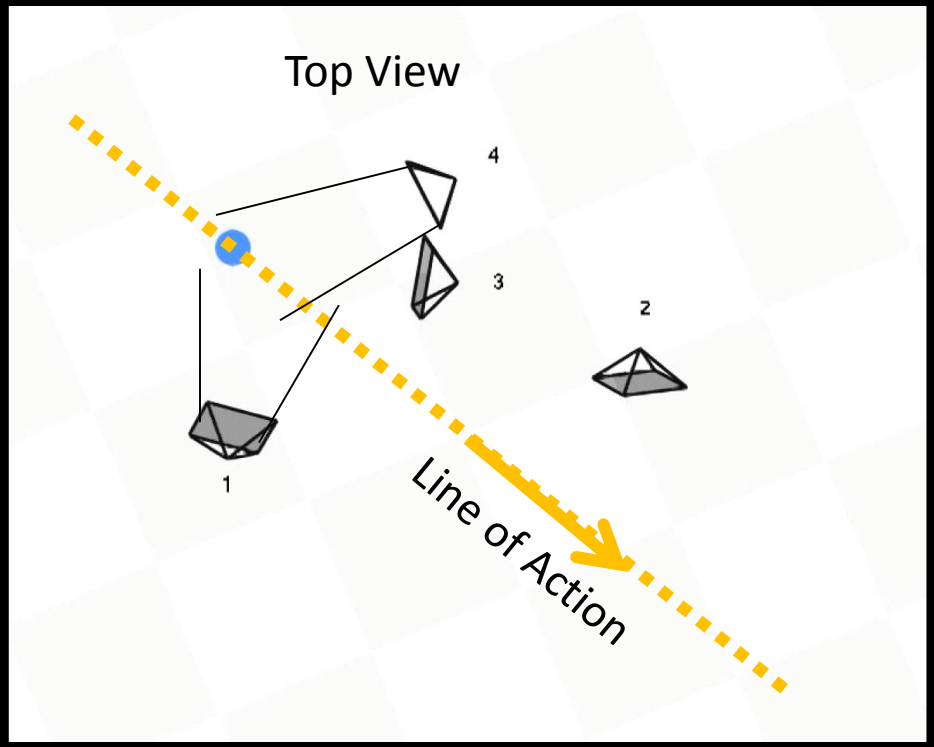
Jump Cut Avoidance



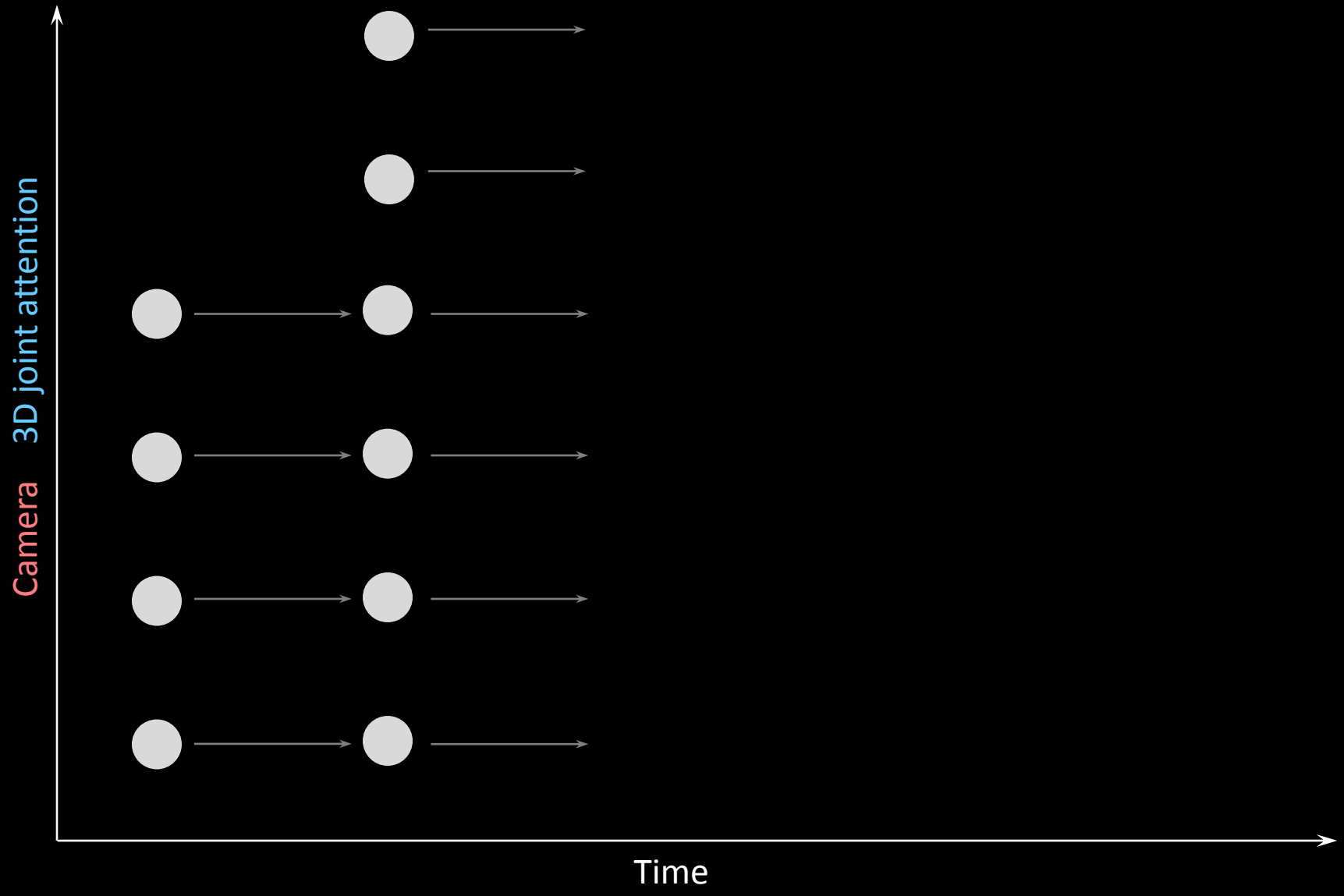
Jump Cut Avoidance



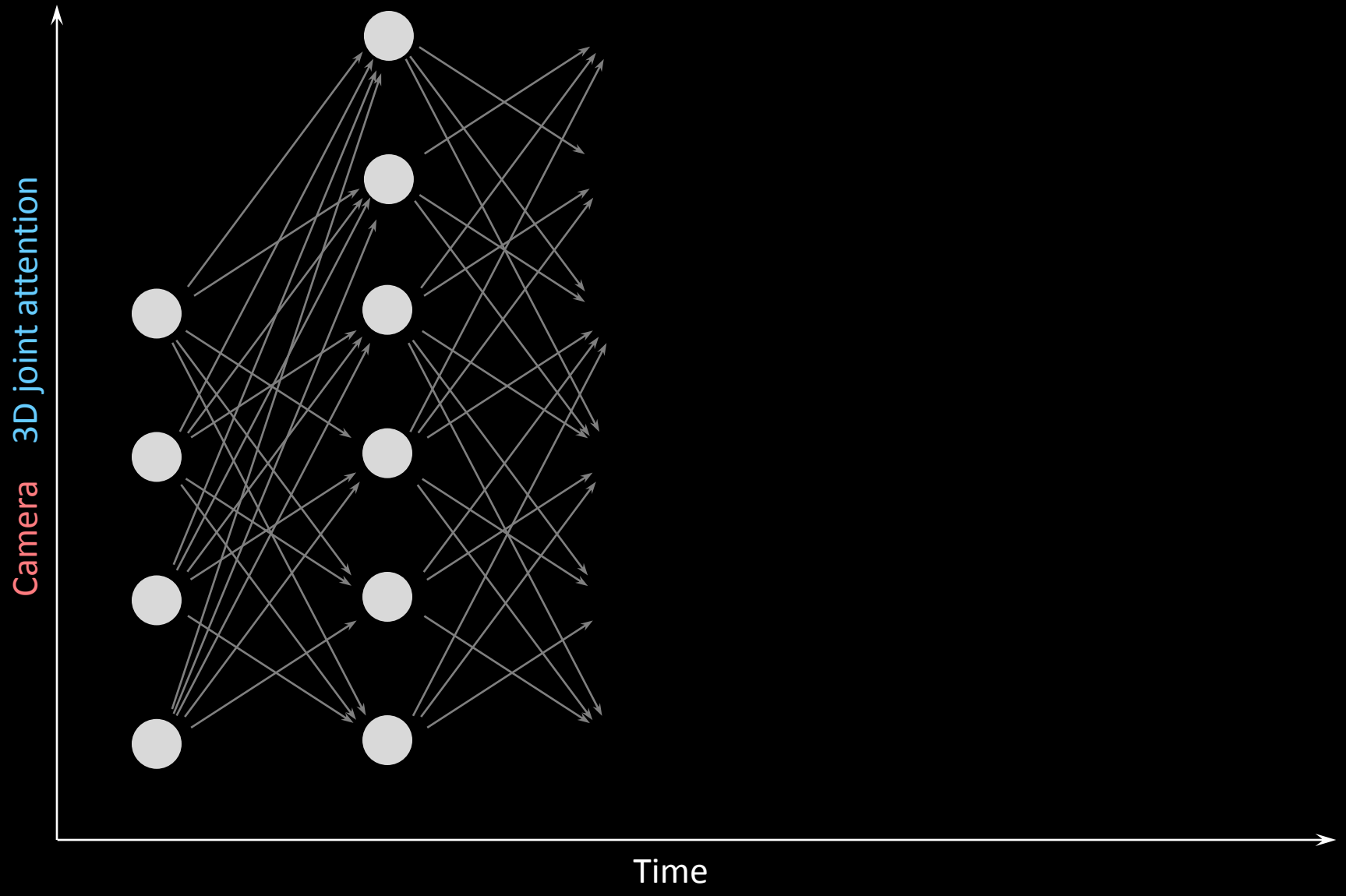
180 Degrees Rule



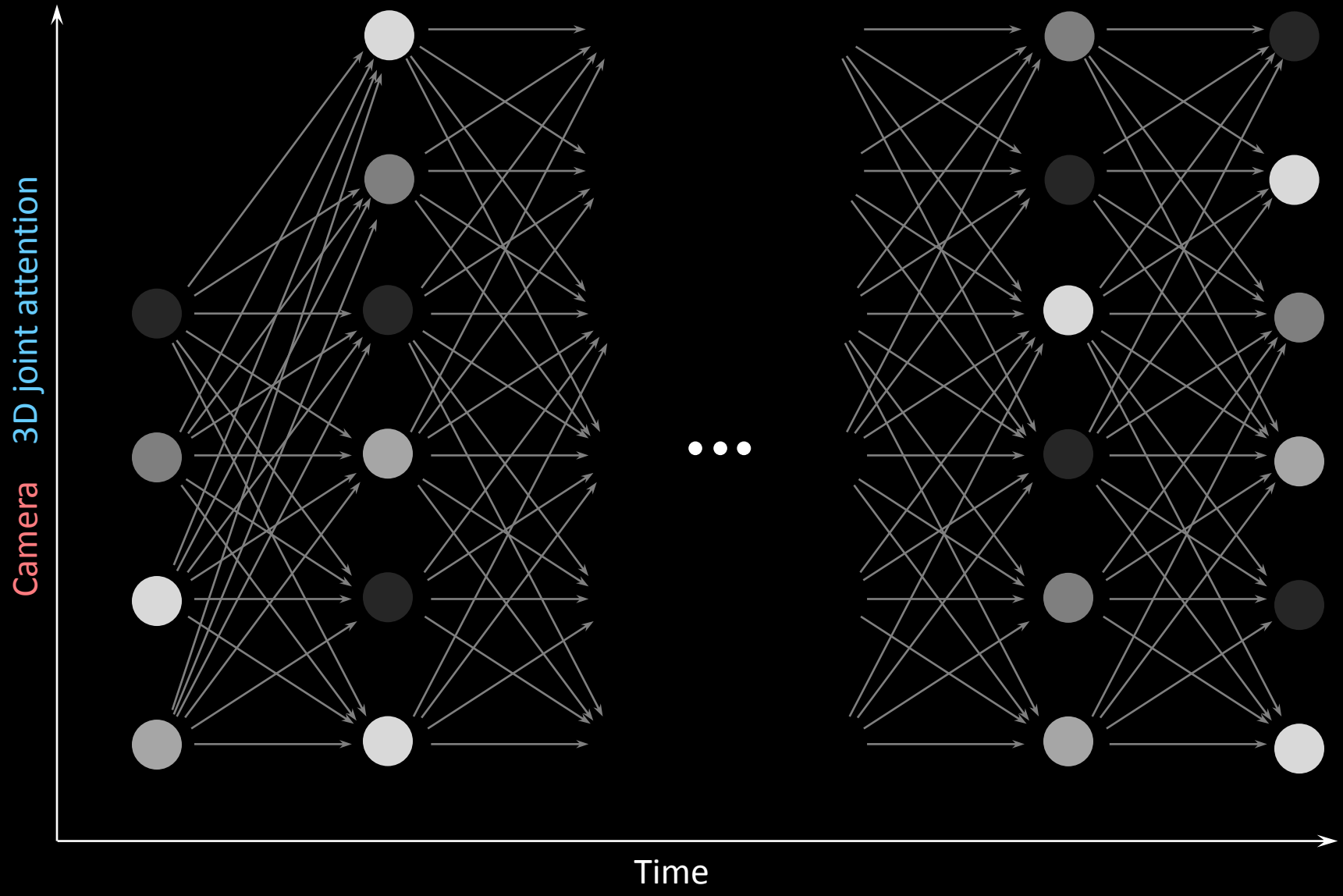
Graph Representation



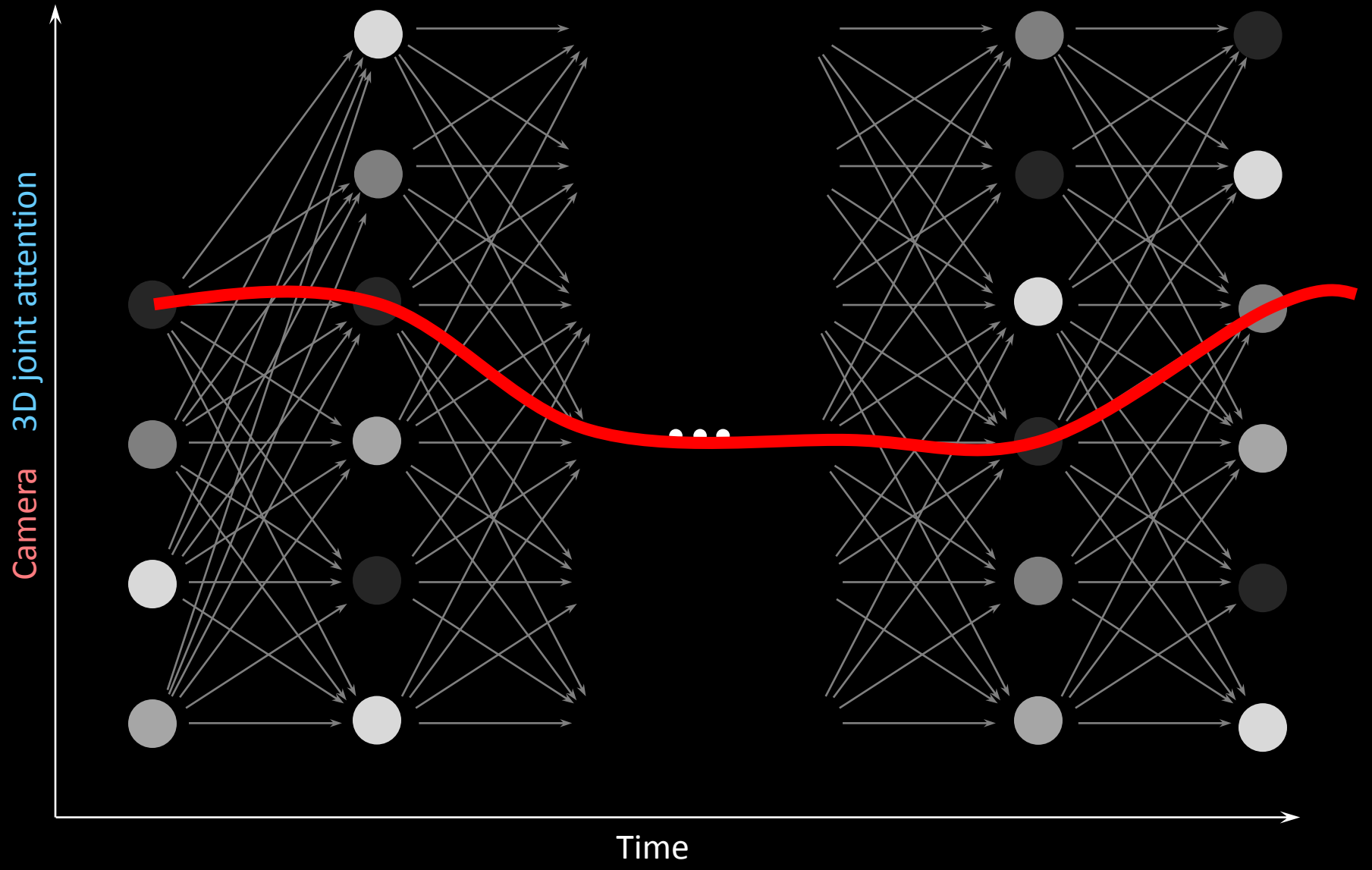
Graph Representation

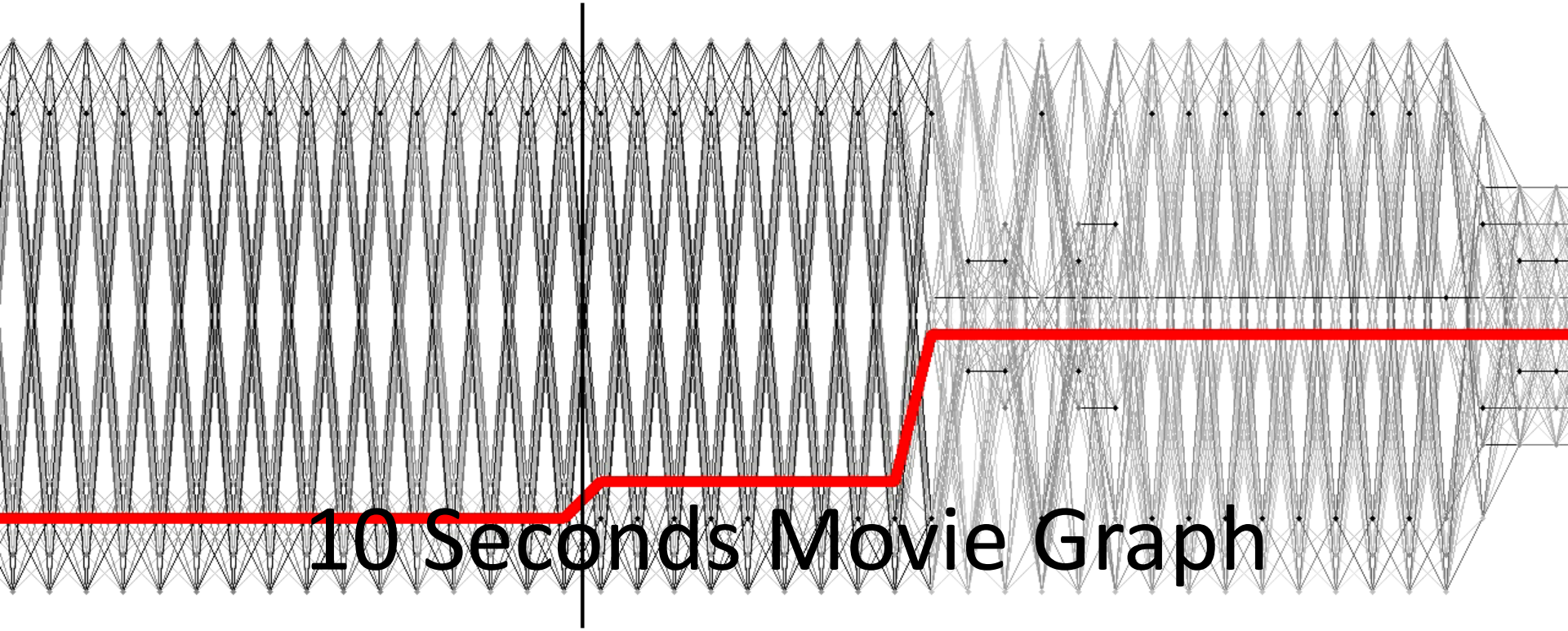


Graph Representation

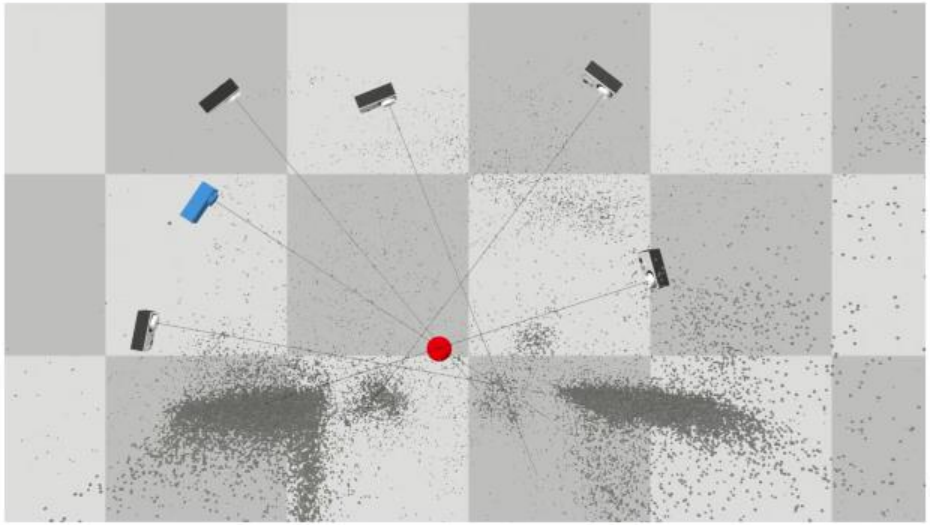


Graph Path = Movie

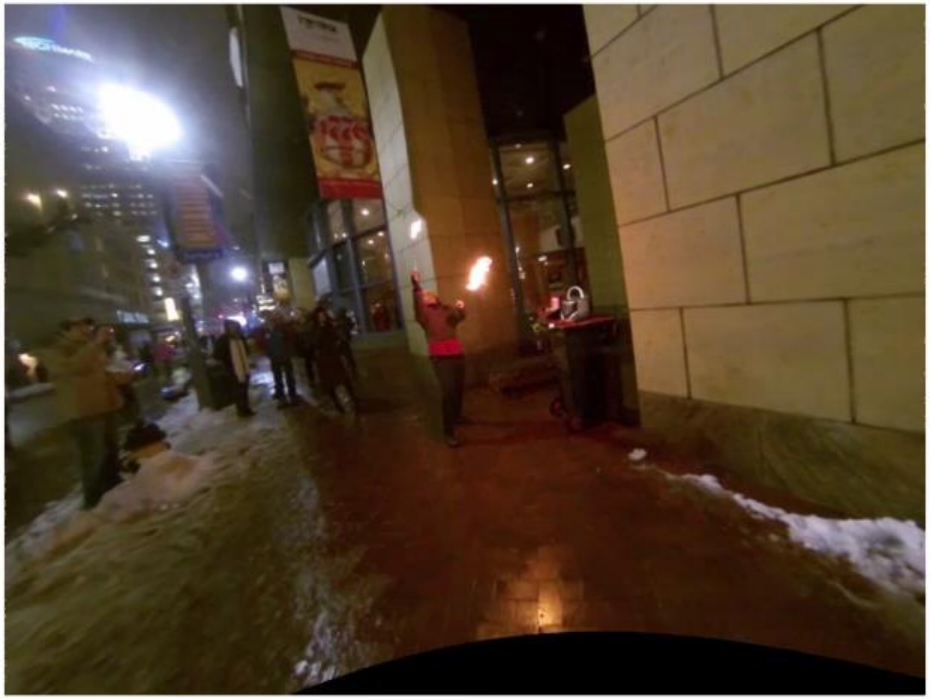



10 Seconds Movie Graph

Busker Performance



3D view



Output

Wearable Devices

- A new point of view
- A new medium

כלכליסט

שני 16 מרץ

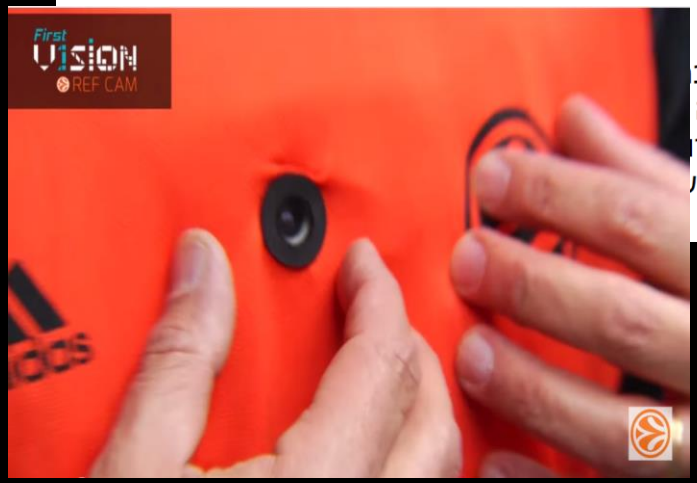
מצלמות על חולצות הכדורגלנים. זה כבר ההווה

אנדרס אינייסטה הוא אחד מהמשקיעים בחולצות שיכולות לשדר בלייב את נקודת המבט של השחקנים או השופטים במהלך משחק

שירות כלכליסט

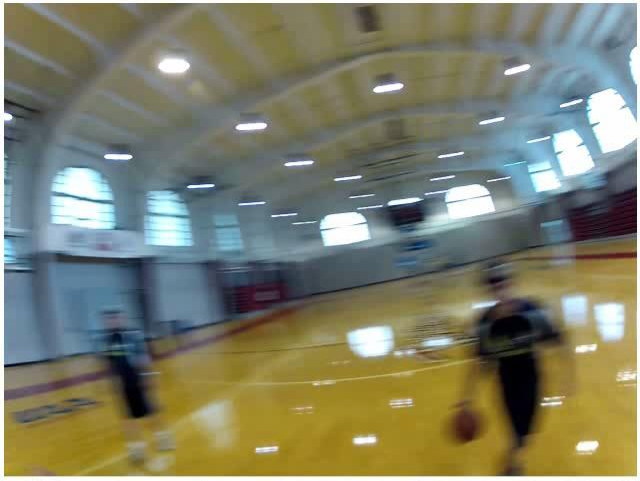
15.03.15, 10:37

חברת פירסט ויז'יון (First V1sion) הספרדית יצרה חולצה לשחקני הכדורגל שמותקנת בה מצלמה בחזה ותספק שידור וידיאו סטנדרטי. המכשיר נוסה באימונים על ידי שחקני ברצלונה ועל ידי שופטים במשחק הכדורסל בין ברצלונה וריאל מדריד. התכנים שנוצרו הופצו על ידי 150 משדרים.



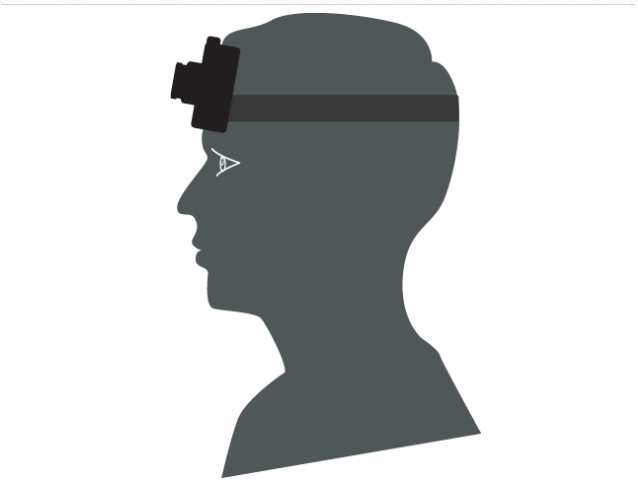
אנדרס אינייסטה, קשר ברצלונה ונבחרת ספרד, מאמין כל כך ב
שהוא השקיע עשרות אלפי יורו בטכנולוגיה. טלפוניקה השקיעה
אחזקה של 7% בפירסט ויזיון. החברה מקווה לגייס 5 מיליון יור
ההמונים BankToTheFuture, על מנת לממן את השלב הבא
יהיה גם פרזנטור של החברה.











Basketball Scene: Random

Basketball Scene: Ours

Summary

- Huge data … becomes visual!
- The challenge is semantics not technological
- Using humans as “agents” systems can learn:
  - By following eye movement
  - By following body movement
  - By following direction of view

Eakta Jain, Yaser Sheikh, Ariel Shamir, Jessica Hodgins

Gaze-driven Video Re-editing

ACM Transactions on Graphics, Volume 33, to appear, 2014

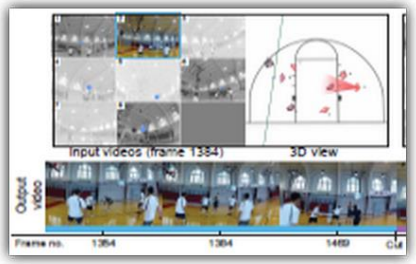
BibTeX More »

Kihwan Kim · Matthias Grundmann · Ariel Shamir · Iain Matthews · Jessica Hodgins · Irfan Essa

Motion Fields to Predict Play Evolution in Dynamic Sport Scenes

Proc. Computer Vision and Pattern Recognition, Pages 840 - 847, CVPR 2010

BibTeX



Ido Arev, Hyun-Soo Park, Yaser Sheikh, Jessica Hodgins, Ariel Shamir

Automatic Editing of Footage from Multiple Social Cameras

ACM Transactions on Graphics, Volume 33, Number 4, (SIGGRAPH Conference Proceedings), Article 81, 2014

BibTeX More »

Thank You