Understanding the Relationship Between Social Behavior and Social Attention

Hyun Soo Park, Yaser Sheikh, and Takeo Kanade
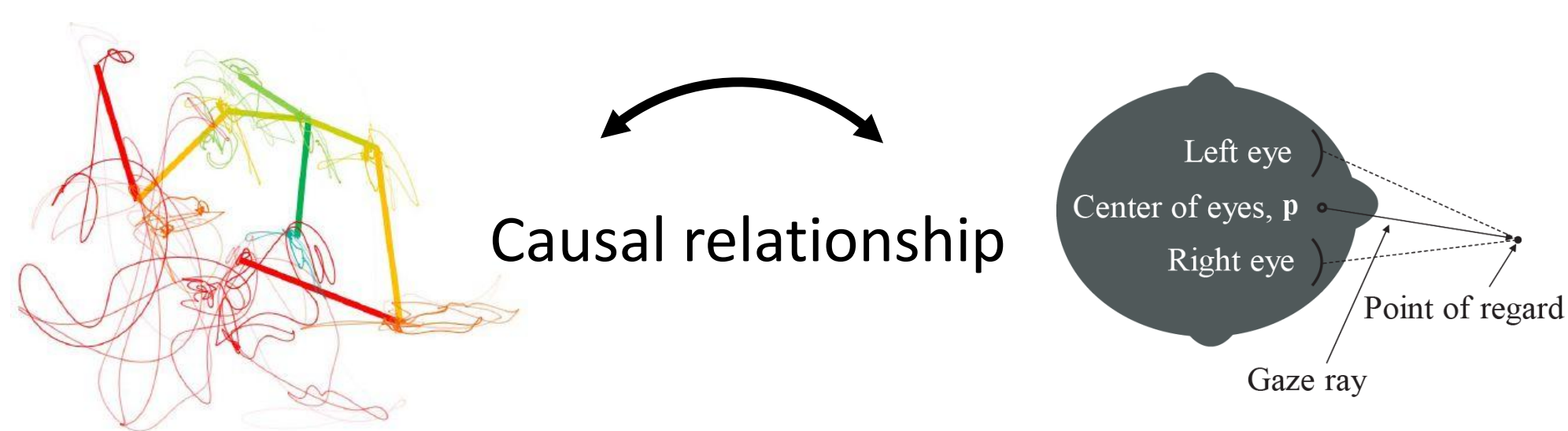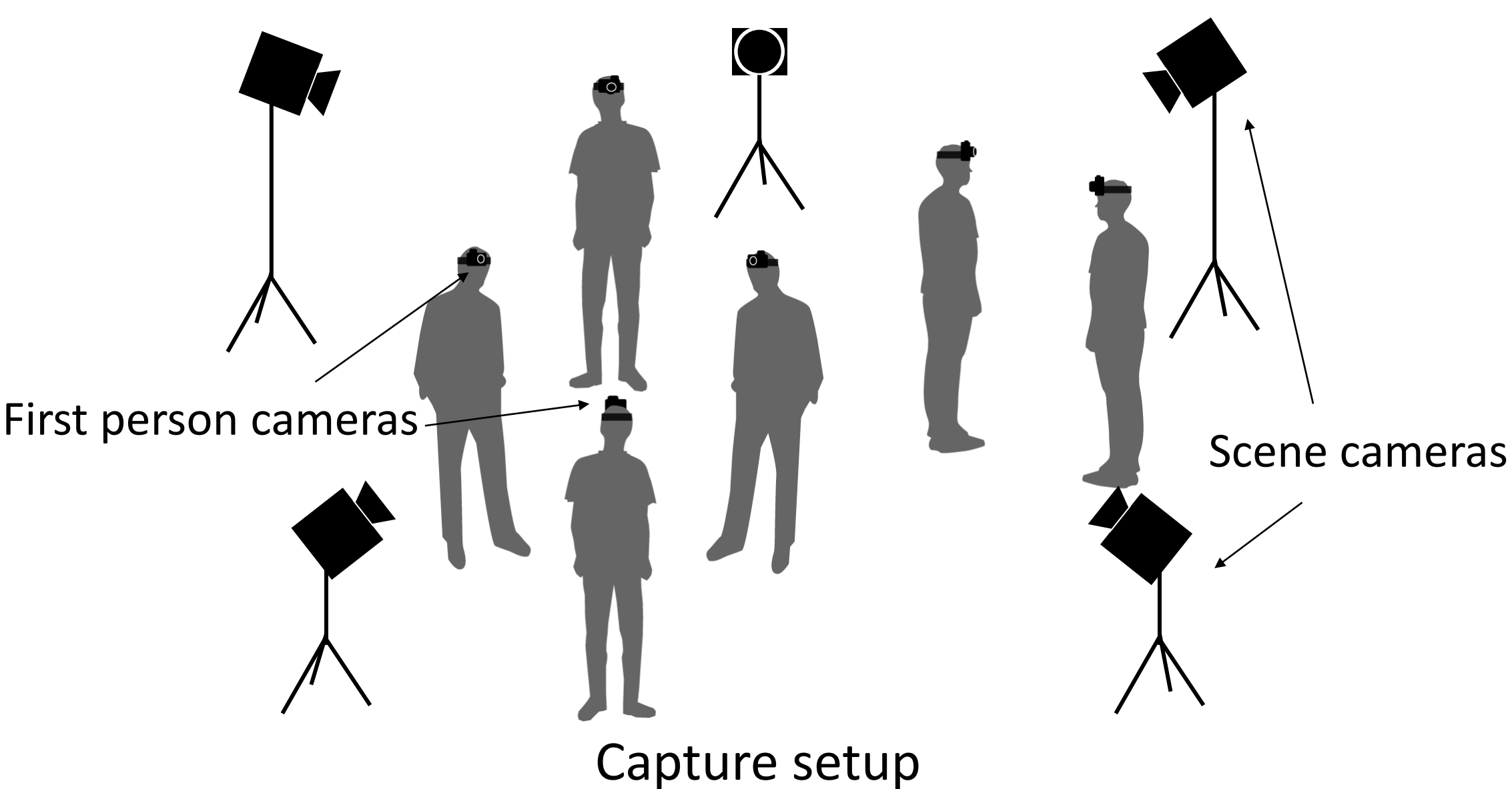Carnegie Mellon University

Abstract

A social scene typically includes a variety of human interaction. Human interactions arise in the form of motion such as body gestures and facial expressions. Among motion associated with human interactions, socially salient motion is motion that many people attend to. Therefore, this category of motion implicitly drives the holistic motion of the group of people.

In this project, we are developing a framework to understand the relationship between social behavior and social attention. We will reconstruct 3D human body motion and social saliency from scene cameras and first person cameras in the same 3D domain. This will allow us to identify socially salient motion and infer the causal relationship between the motion and people's attention, i.e., how the people respond to the socially salient motion.

Goal

Given videos from scene and first person cameras, infer the relationship between social behavior and social attention.

Capture setup

Input videos

Causal relationship

Left eye
Center of eyes, p
Right eye
Point of regard
Gaze ray

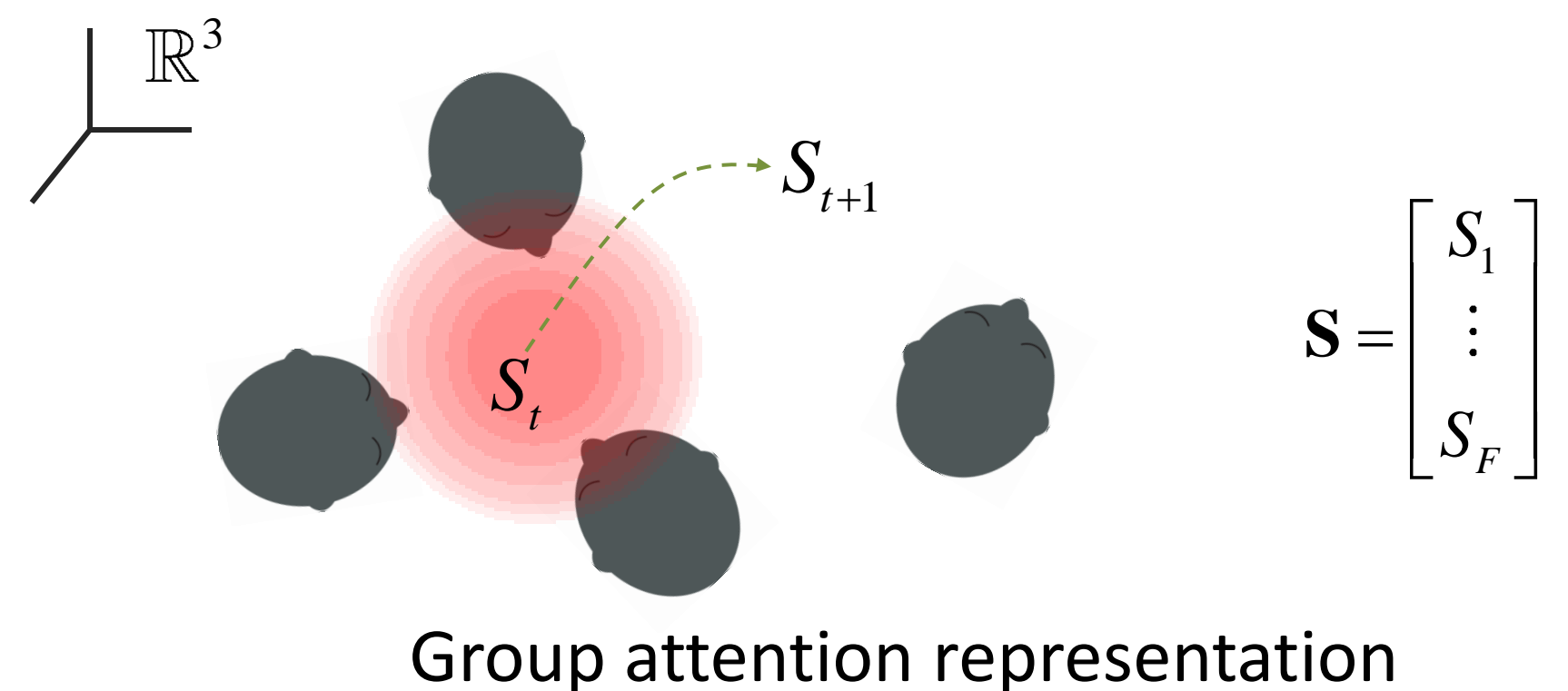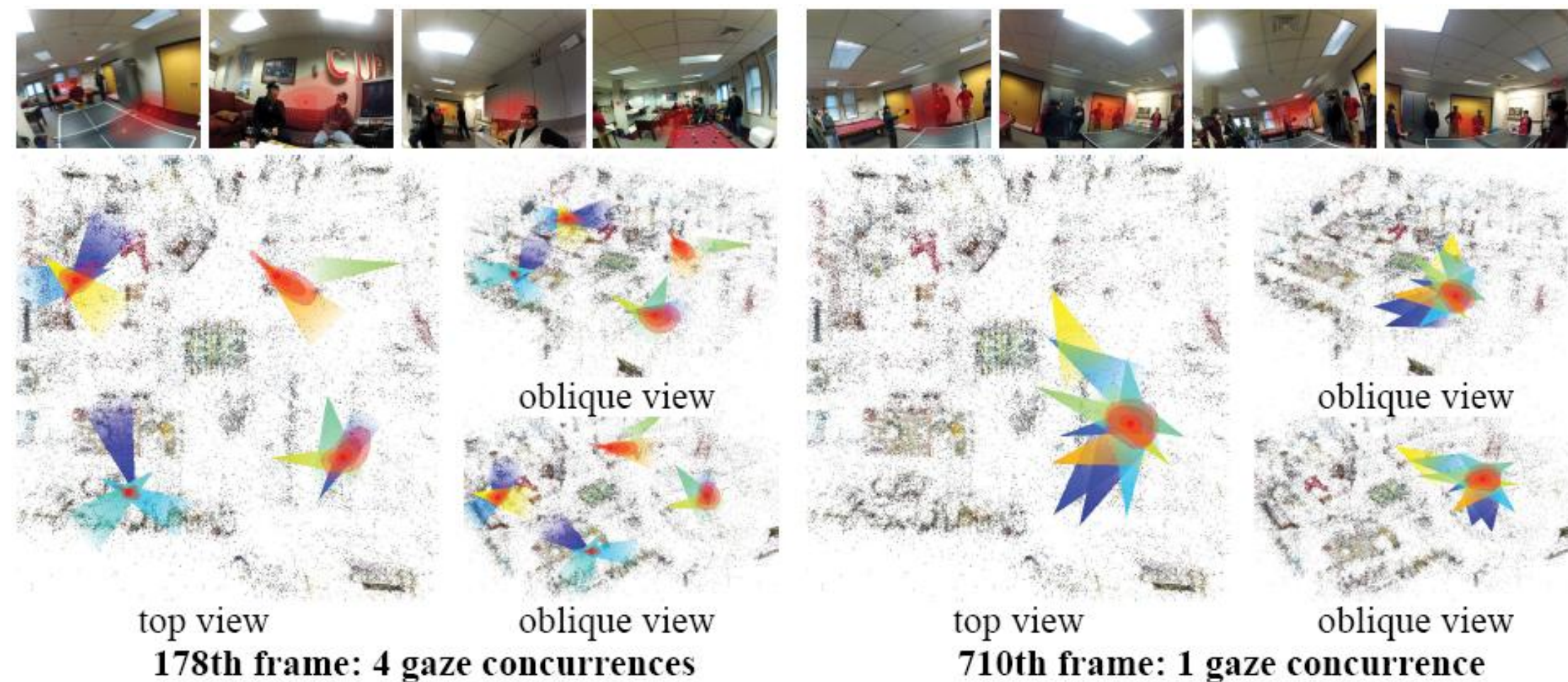## 3D Reconstruction of Human Motion

Given a video with annotated landmarks, we reconstruct human body motion in 3D. The 3D reconstructed motion is a viewpoint-invariant representation and thus, it allows us to learn behavioral patterns across various scenes.
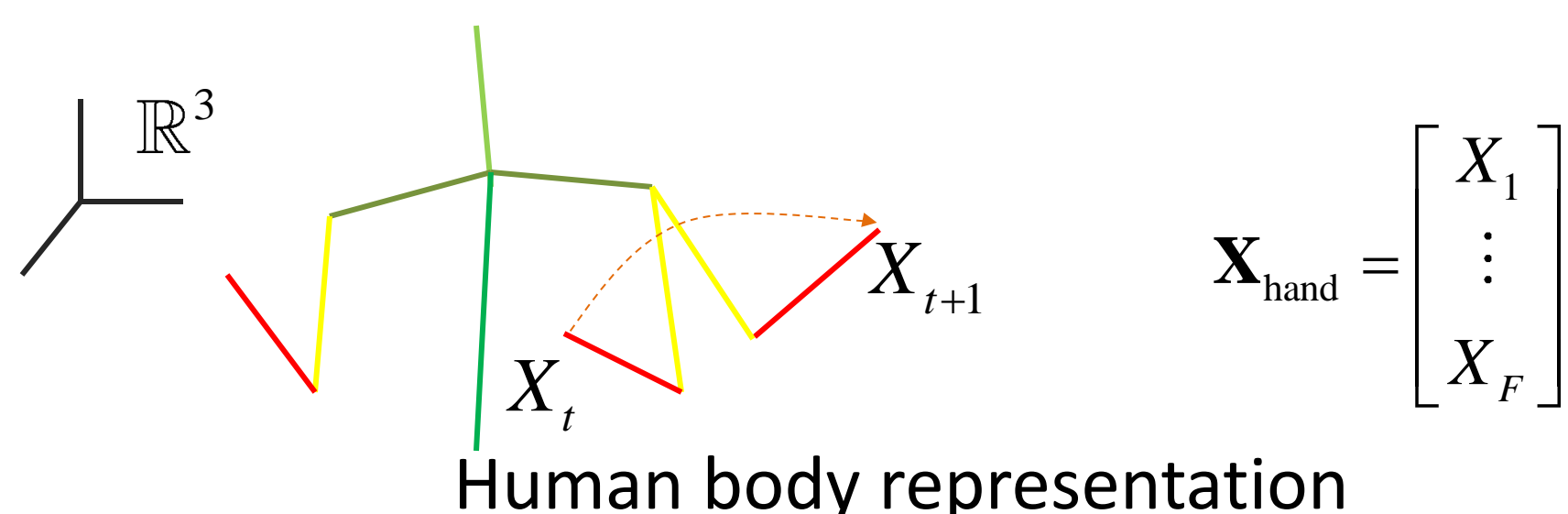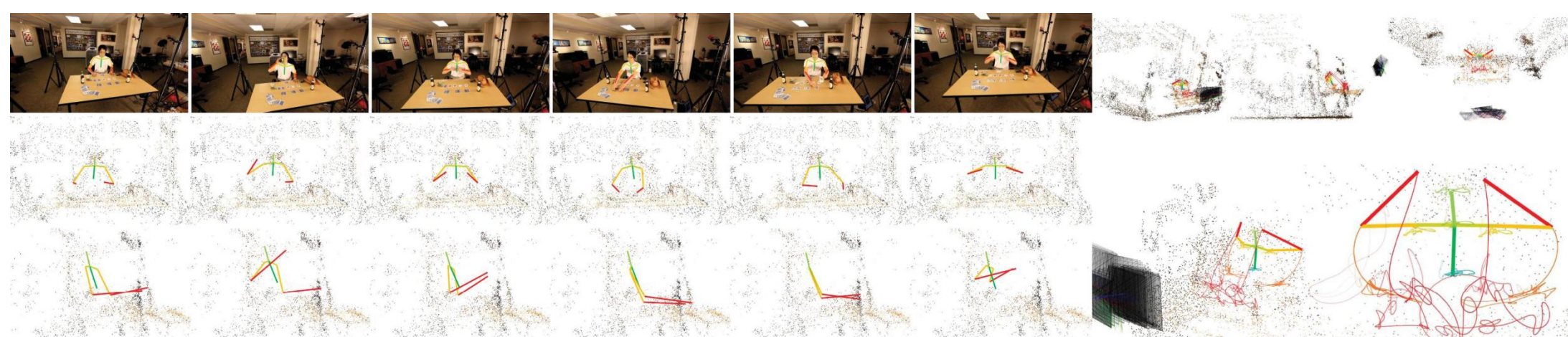
$$\mathbb{R}^3 \qquad X_{t+1} \qquad X_t \qquad \mathbf{X}_{hand} = \begin{bmatrix} X_1 \\ \vdots \\ X_F \end{bmatrix}$$
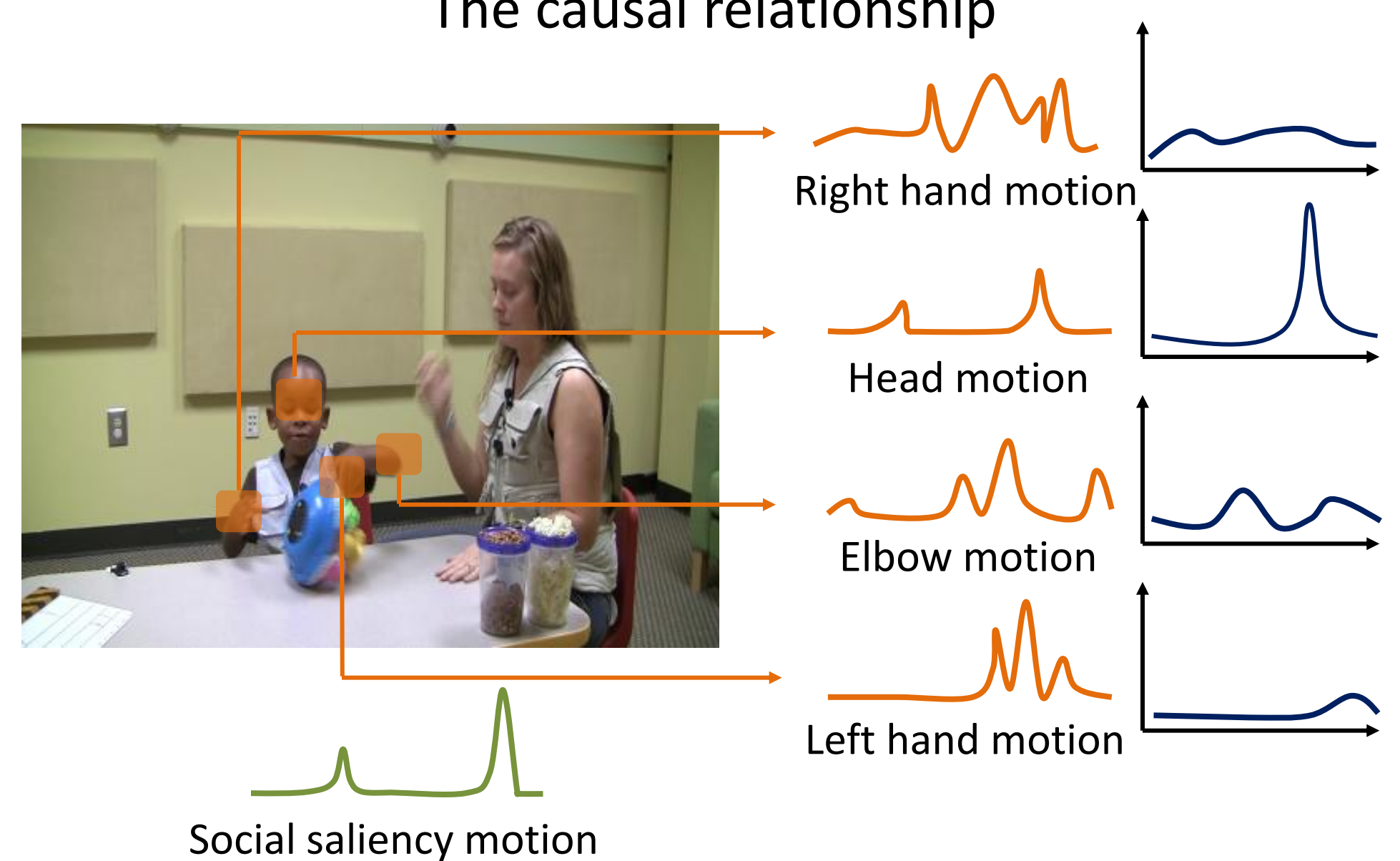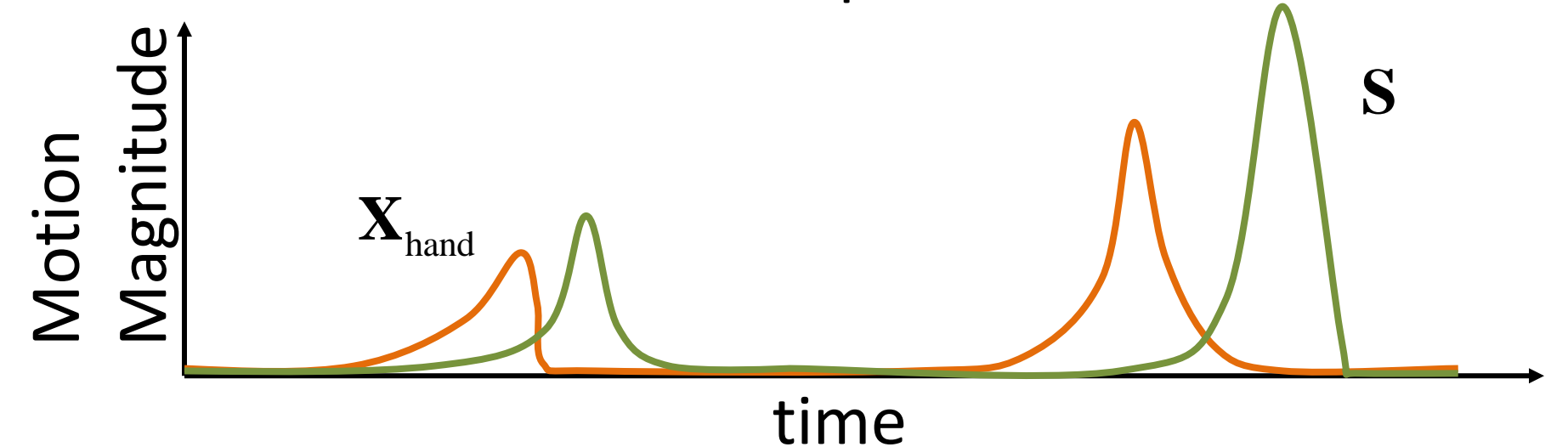
Human body representation

3D Reconstruction of Social Saliency

Social saliency occurs where people cognitively attend as a group. The 3D reconstructed social saliency enables us to represent social saliency in the same domain of human motion reconstruction.

top view   oblique view   top view   oblique view
178th frame: 4 gaze concurrences   **710th frame: 1 gaze concurrence**

$$\mathbb{R}^3 \qquad S_{t+1} \qquad S_t \qquad \mathbf{S} = \begin{bmatrix} S_1 \\ \vdots \\ S_F \end{bmatrix}$$

Group attention representation

Inference of the Causal Relationship

The hypothesis we wish to test is that a subset of behaviors, which we define as socially salient behavior, leads to characteristic change in social attention. Given 3D reconstruction of human body motion and social saliency, we infer the causal relationship between them.

Motion Magnitude | $\mathbf{X}_{hand}$ | $\mathbf{S}$ | time

The causal relationship

Right hand motion
Head motion
Elbow motion
Left hand motion

Social saliency motion

3D Reconstruction Interface

We are developing an interface to annotate and reconstruct human body motion in 3D.

3D reconstruction   Video input