

# Multi-Camera 3D Position Estimation using Conditional Random Field

Shusuke Matsuda<sup>1</sup>, Nattaon Techasartikul<sup>2</sup>, Hideyuki Shimonishi<sup>2</sup>

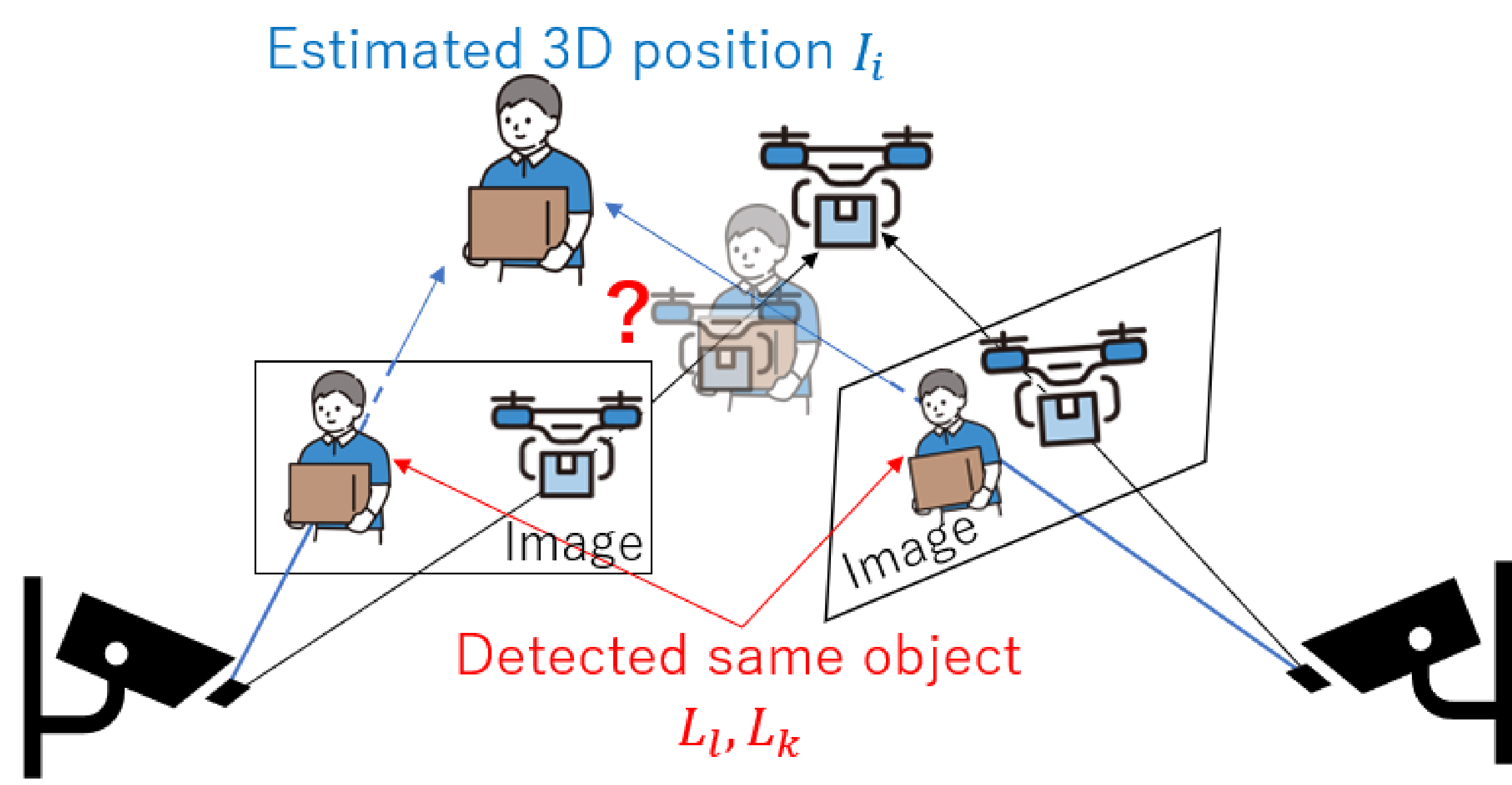
<sup>1</sup>Graduate School of Information Science and Technology, <sup>2</sup>Cybermedia Center, Osaka University

## Motivation

- Human-Robot Collaboration (HRC)
- Enhance efficiency and safety in HRC
- Create a digital twin to enable robot's perception of the physical world

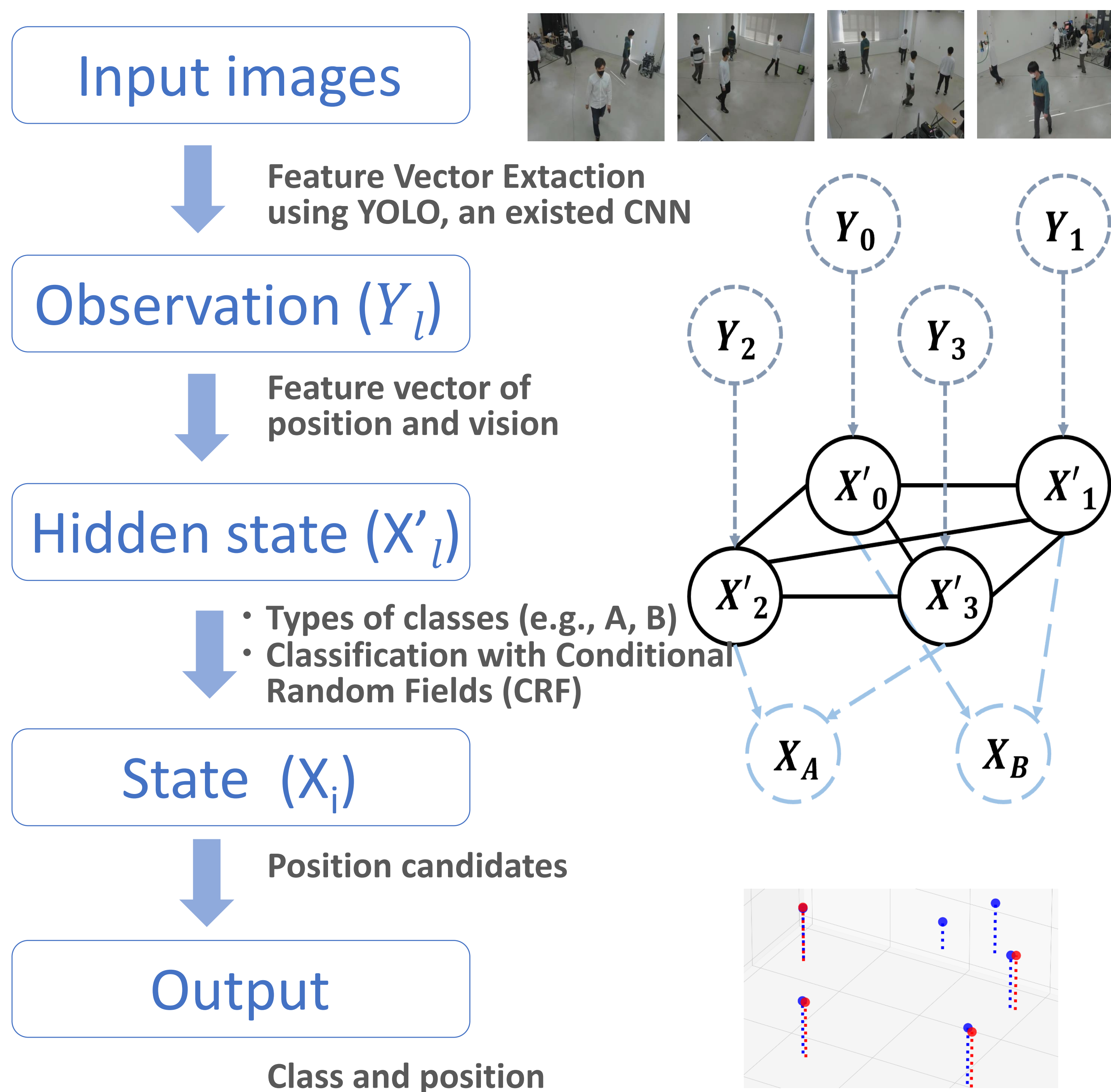
## Purpose

3D position estimation of the same objects using multiple cameras



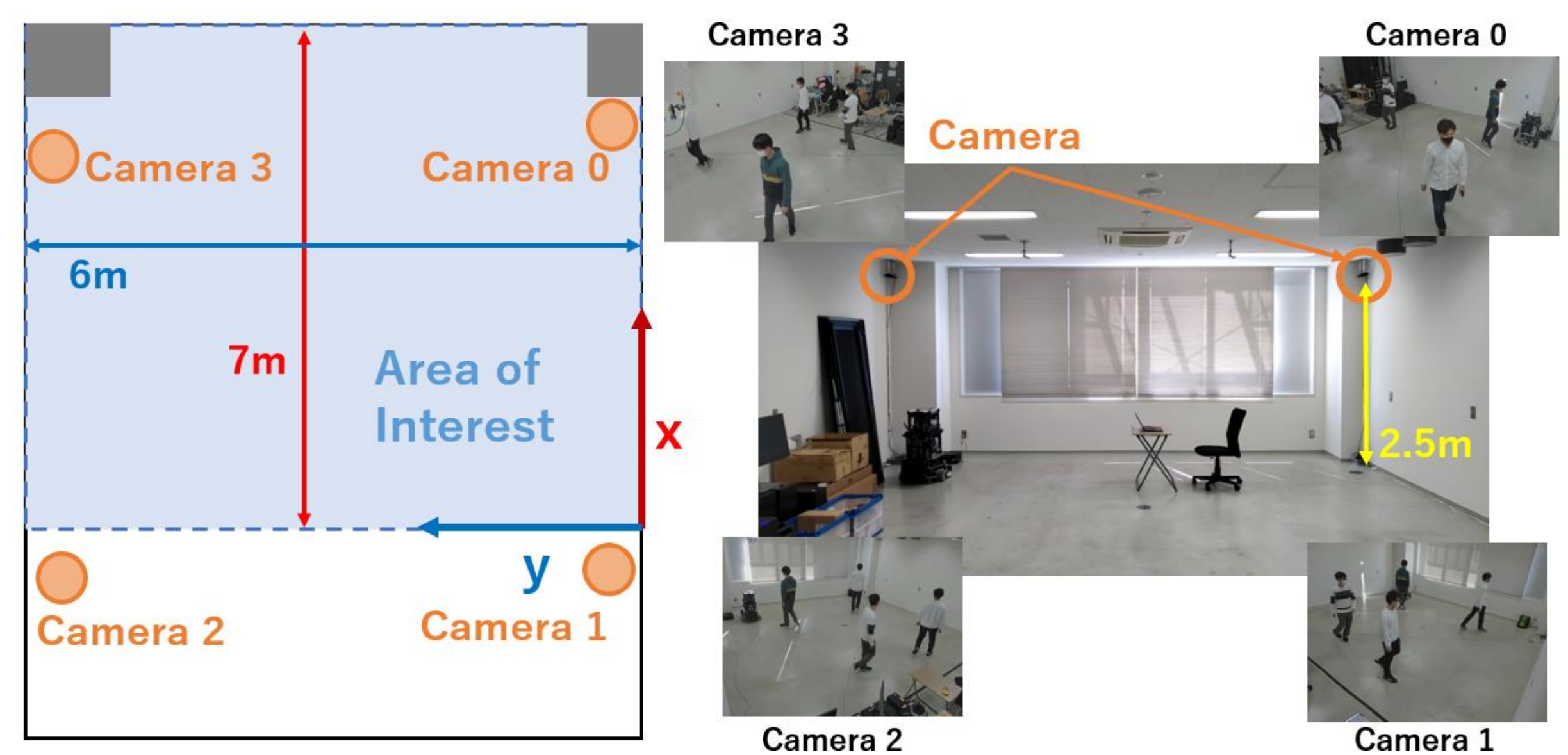
## Proposed Method

Detection of the same objects using Conditional Random Field (CRF), a graphical model, as a classification task



## Evaluation

### Environment setup



### Methods

- Unary: Only using hierarchical clustering
- CRF&p: CRF with position features
- CRF&p+v: CRF with position and vision features

### Results

#### 1. Detection of the same objects

	ARI	AMI	H	C	V
Unary	45.12	51.13	98.35	89.51	93.72
CRF & p	<b>78.86</b>	<b>81.46</b>	99.57	<b>94.80</b>	<b>97.12</b>
CRF & p + v	77.08	80.60	<b>99.72</b>	94.51	97.04

#### 2. 3D position estimation

	error Ave.	error S.D.	error Max
Unary	88	135	935
CRF & p	31	43	366
CRF & p + v	<b>24</b>	<b>30</b>	<b>317</b>

