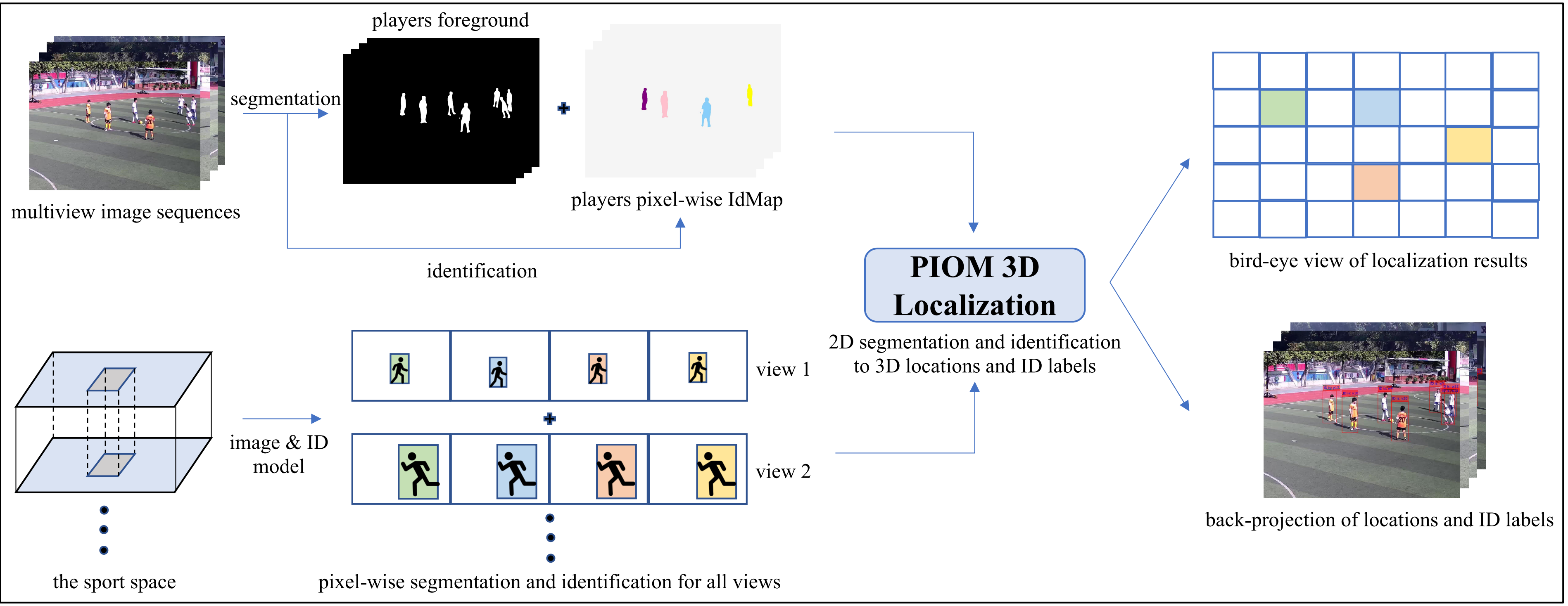
Multi-camera Sports Players 3D Localization with Identification Reasoning

Lead presenter: **Yukun Yang**

Co-authors: Ruiheng Zhang, Wanneng Wu, Yu Peng, Min Xu

Affiliation: University of Technology Sydney

Introduction



An overview of our proposed framework

- Firstly, we use 2D monocular segmentation and identification to extract sports players' segmentation masks and identification in pixel level.
- At the same time, with the 3D coordinate settings for the sports space, we introduce an Image&ID model and an image distance norm to fuse the multiview pixel-wise segmentation and ID labels with their 3D spatial relation.
- With multiple players PIOM localization algorithms, we then obtain sports players 3D location and their unique IDs. The results are given as probabilities of locations being occupied by specifically labeled players.

Contribution

We concluded our contribution as:

- We jointly apply object segmentation and one of our early research on player identification. With identification labels, we provide each segmentation mask with pixel-wise ID information.
- We introduce an Image&ID model to visually describe the position of a player standing at the sports court with a certain ID. Then we use an image distance norm to evaluate the overall errors between the model and multiview input evidence. This creates a tractable loss function to make iteration efficient.
- Different from existing methods which only construct one overall model for all players, we develop a Probabilistic and Identified Occupancy Map (PIOM) using a multiple dimensional Bayesian model. This model calculates the locations separately according to different ID labels, providing the results as probabilities of locations being occupied by multiple players with unique ID labels.
- We conduct experiments based on a youth football dataset and compare the performance with 3 previous baselines to show that our proposed method outperforms the previous localization approaches by a large margin.

Methods

1. Statistical modeling

We firstly define a 3D world coordinate system for the sport space and discretize the space into a certain number of 3D cubes, where each cube represents a typical human space that a sport player can occupy. We then construct players' locations into a set of conditional probabilities that are lately treated as the posterior probabilities when giving the likelihood probabilities and initializing the prior probabilities.

$$P\big((X,Y)|B\big) = P(X,Y)P\big(B|(X,Y)\big)$$

2. Image&ID model

We then introduce an Image&ID model to visually describe the relation between the likelihood probabilities and the players' segmentation and identities. The image distance norm calculates the errors between synthetic average images with ID and given evidence of pixel level segmentation and identification. After that we obtain the loss function for iteration.

$$R_k = arg\max_{\xi} \frac{\sum_c \left(\mathcal{R}_k^c \sum_{i=1}^{MN} r_i | r_i = \xi\right)}{\sum_c \left(\mathcal{R}_k^c \sum_{i=1}^{MN} r_i\right)}$$

3. Convergence and optimization

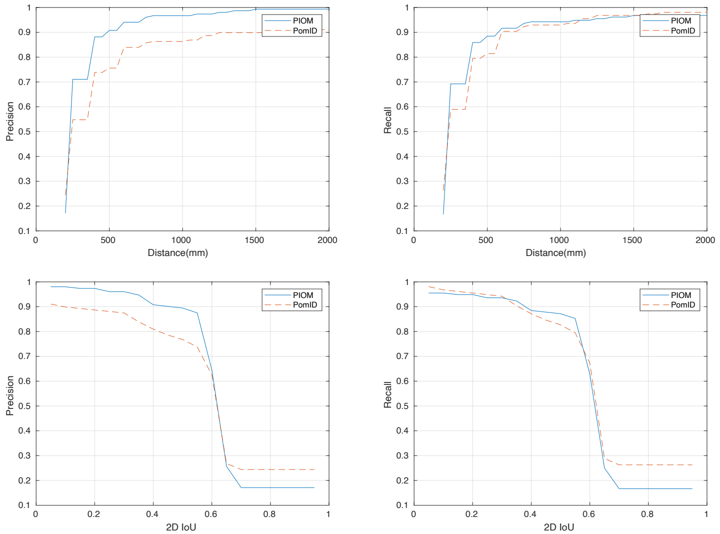
In order to compute the probabilities of occupancy and identification, we minimize the $\mathrm{K}-\mathrm{L}$ divergence between an estimated distribution and the true posterior probability that we are after. Afterwards, The likelihood probabilities are approximated by the normalized image distances between the synthetic average image with ID and the given 2D evidence.

Results

Back-projection of the results

Each player has a bounding box that is back-projected from the calculated 3D location and a unique ID label that can trace the specific player with location presented.

P-R curves



This shows the P-R curve on the metrices of bird-eye distance and 2D IoU, compared with our method PIOM and our previous method PomID.

Thank you!

Presenter: Yukun Yang