

ĐẠI HỌC QUỐC GIA TP.HCM
TRƯỜNG ĐẠI HỌC BÁCH KHOA



GÁN NHÃN ĐỐI TƯỢNG DI CHUYỂN QUA NHIỀU
CAMERA
(TO ASSIGN LABEL FOR MOVING OBJECTS IN
MULTIPLE CAMERAS)

GVHD: PGS TS. NGUYỄN THANH BÌNH
HVTH: VÕ HOÀNG AN (1670211)

NỘI DUNG

2



1. Giới thiệu đề tài

2. Cơ sở lý thuyết

3. Phương pháp đề xuất

4. Thí nghiệm và kết quả

5. Kết luận và hướng phát triển

NỘI DUNG

3

1. Giới thiệu đề tài

2. Cơ sở lý thuyết

3. Phương pháp đề xuất

4. Thí nghiệm và kết quả

5. Kết luận và hướng phát triển

GIỚI THIỆU ĐỀ TÀI

4

- ❑ Thị giác máy tính hướng tới việc xây dựng hệ thống có thể thay thế mắt người.
- ❑ Thị giác máy tính tập trung giải quyết hai bài toán cơ bản:
 - ❑ Nhận dạng đối tượng.
 - ❑ Theo dõi đối tượng.
- ❑ Ứng dụng của thị giác máy tính rất phổ biến : công nghiệp, dịch vụ, an ninh và đặc biệt là AI.

GIỚI THIỆU ĐỀ TÀI

5

- ❑ Đề tài tập trung vào việc *gán nhãn cho các đối tượng di chuyển qua nhiều camera.*
- ❑ Sử dụng hai camera được đặt theo hai trường hợp:
 - ❑ Hai camera song song có vùng không gian quan sát trùng lắp.
 - ❑ Hai camera chéo nhau có vùng không gian quan sát trùng lắp.

NỘI DUNG

6

1. Giới thiệu đề tài

2. Cơ sở lý thuyết

3. Phương pháp đề xuất

4. Thí nghiệm và kết quả

5. Kết luận và hướng phát triển

CƠ SỞ LÝ THUYẾT

7

- ❑ Cắt đoạn video để thu được danh sách các frame ảnh.
- ❑ Xác định vùng không gian trùng lặp.
- ❑ Phát hiện đối tượng di chuyển trong đoạn video.
- ❑ Rút trích đặc trưng của đối tượng.
- ❑ Gán nhãn cho đối tượng.

CƠ SỞ LÝ THUYẾT

8

- ❑ Cắt đoạn video để thu được danh sách các frame ảnh:
 - ❑ Matlab.
 - ❑ FFMPEG.
 - ❑ OpenCV.
 - ❑ EmguCV.

CƠ SỞ LÝ THUYẾT

9

- ❑ Xác định vùng không gian trùng lắp:
 - ❑ Hai camera đặt song song.
 - ❑ Hai camera đặt chéo nhau.

CƠ SỞ LÝ THUYẾT

10

- ❑ Phát hiện đối tượng di chuyển trong đoạn video:
 - ❑ Point detectors.
 - ❑ Background subtraction.
 - ❑ Segmentation.
 - ❑ Supervisioned learning.

CƠ SỞ LÝ THUYẾT

11

❑ Rút trích đặc trưng của đối tượng.

❑ Màu.

- ❑ Mô men màu (Moments color), biểu đồ màu (color histogram)...

❑ Hình dáng.

- ❑ Mô men bất biến của Hu (Hu's invariant moments)...

CƠ SỞ LÝ THUYẾT

12

- ❑ Gán nhãn cho đối tượng.
 - ❑ Phương pháp so trùng đặc trưng.
 - ❑ Phương pháp phân lớp và phân cụm.
 - ❑ Phương pháp chuỗi thời gian.

NỘI DUNG

13

1. Giới thiệu đề tài

2. Cơ sở lý thuyết

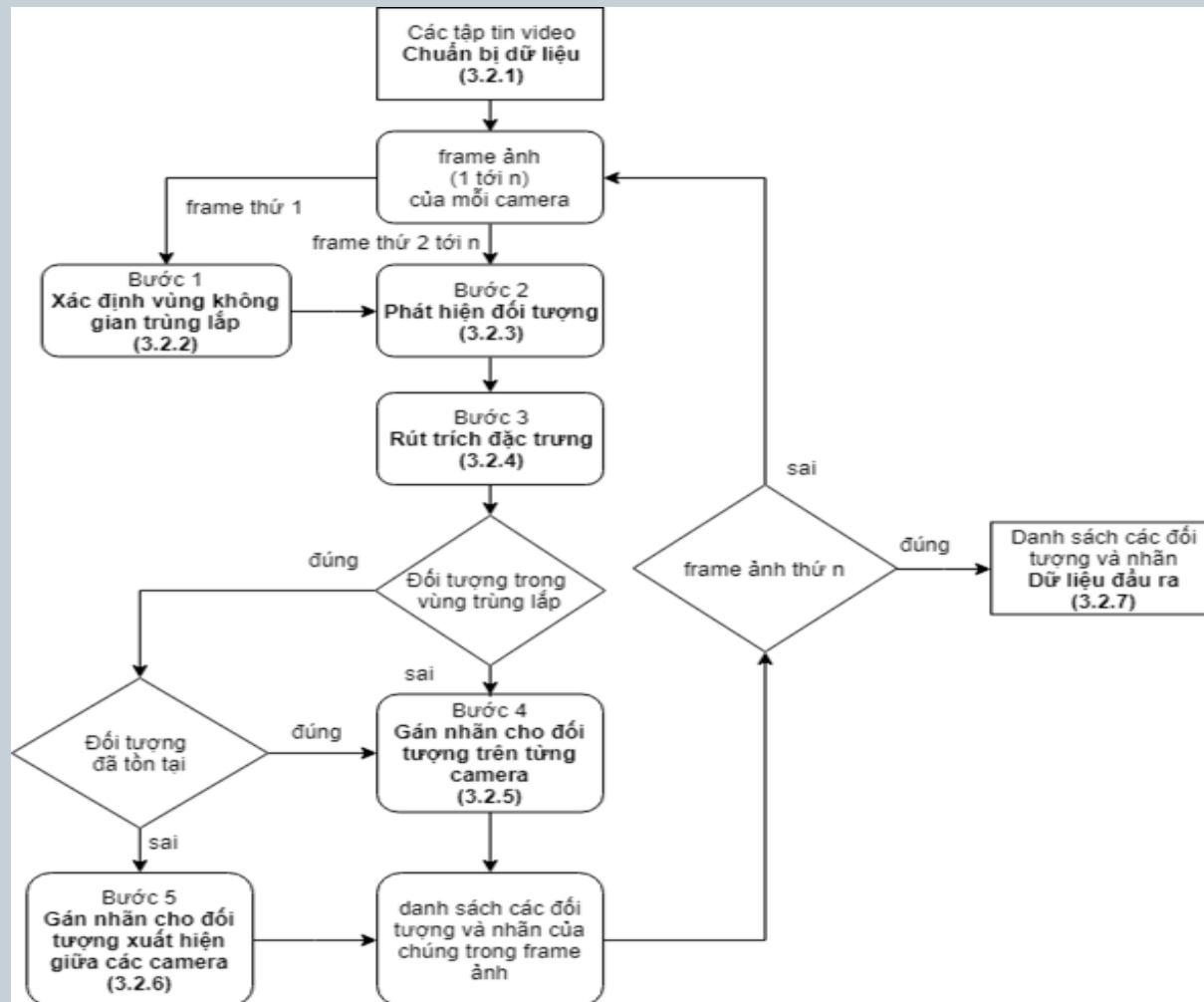
3. Phương pháp đề xuất

4. Thí nghiệm và kết quả

5. Kết luận và hướng phát triển

PHƯƠNG PHÁP ĐỀ XUẤT

14



PHƯƠNG PHÁP ĐỀ XUẤT

15

❑ Chuẩn bị dữ liệu:

- ❑ Thu thập các đoạn video từ camera quan sát được đặt song song hoặc chéo nhau.
- ❑ Cắt mỗi đoạn video thành một tập các frame ảnh theo mỗi 1s.

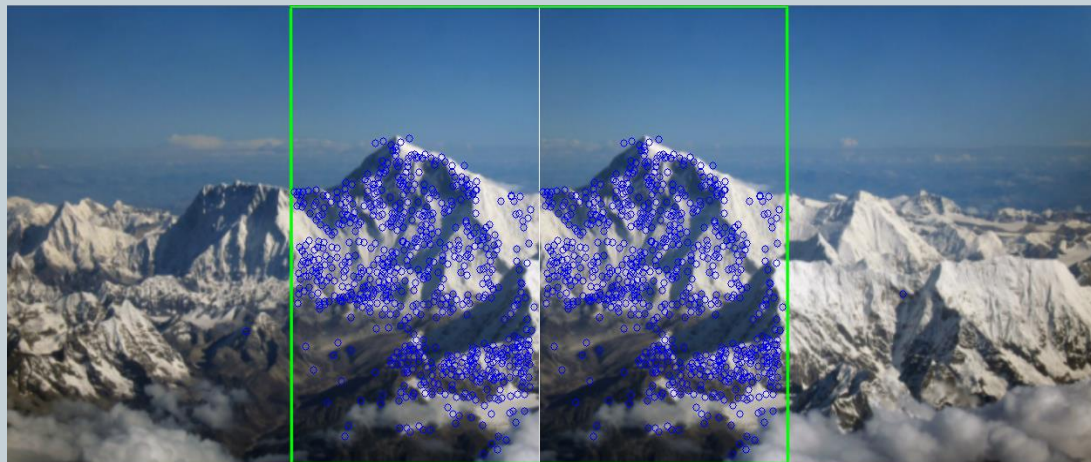
PHƯƠNG PHÁP ĐỀ XUẤT

16

❑ Xác định vùng không gian trùng lắp:

❑ Hai camera đặt song song.

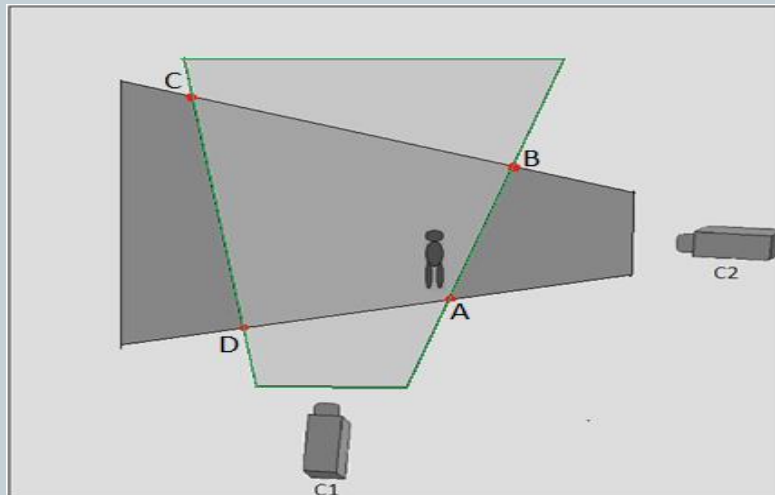
❑ Sử dụng đặc trưng SIFT để tìm điểm tương đồng giữa hai camera. Từ tập hợp các điểm tương đồng có thể xây dựng được vùng không gian trùng lắp.



PHƯƠNG PHÁP ĐỀ XUẤT

17

- ❑ Xác định vùng không gian trùng lắp:
 - ❑ Hai camera đặt chéo nhau.
 - ❑ Sử dụng phương pháp thủ công như hình dưới

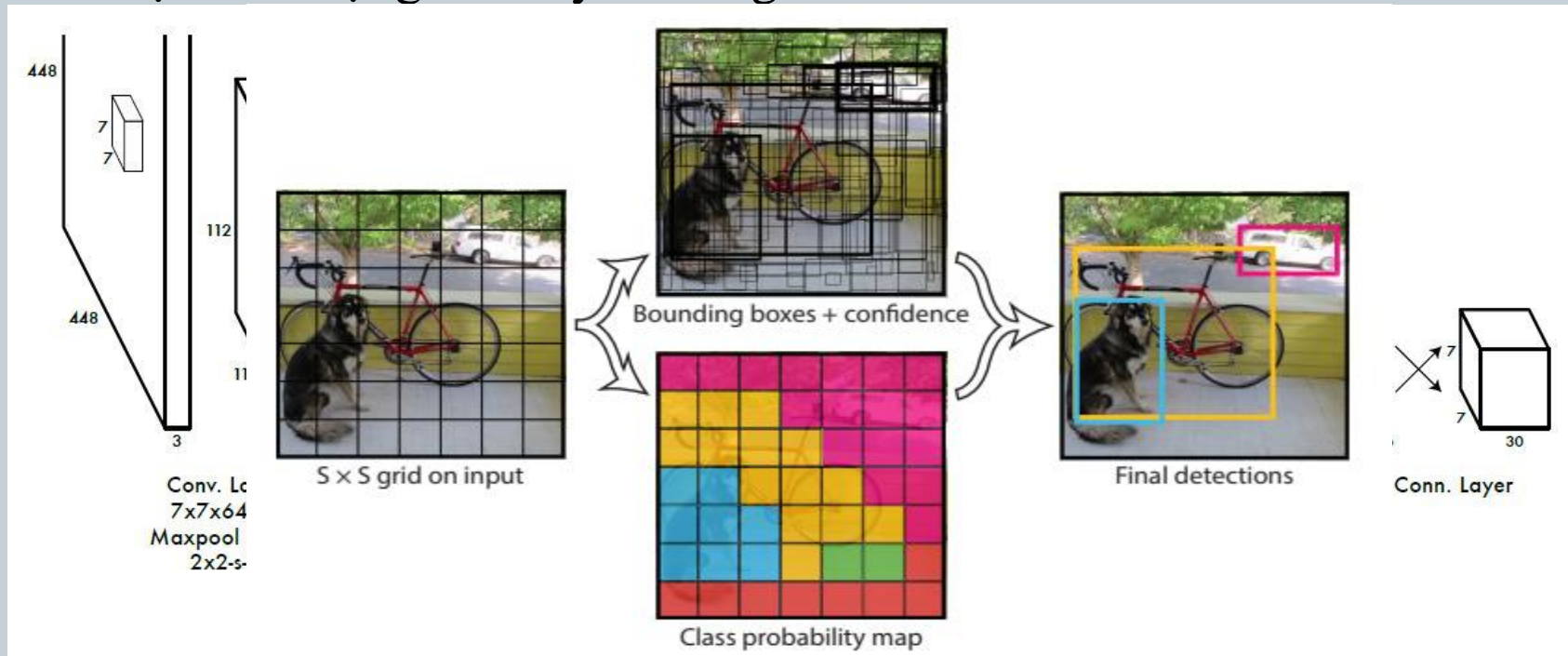


PHƯƠNG PHÁP ĐỀ XUẤT

18

❑ Phát hiện đối tượng:

- ❑ Sử dụng giải thuật YOLO thuộc lớp supervised learning để thu được đối tượng di chuyển trong mỗi frame ảnh.



PHƯƠNG PHÁP ĐỀ XUẤT

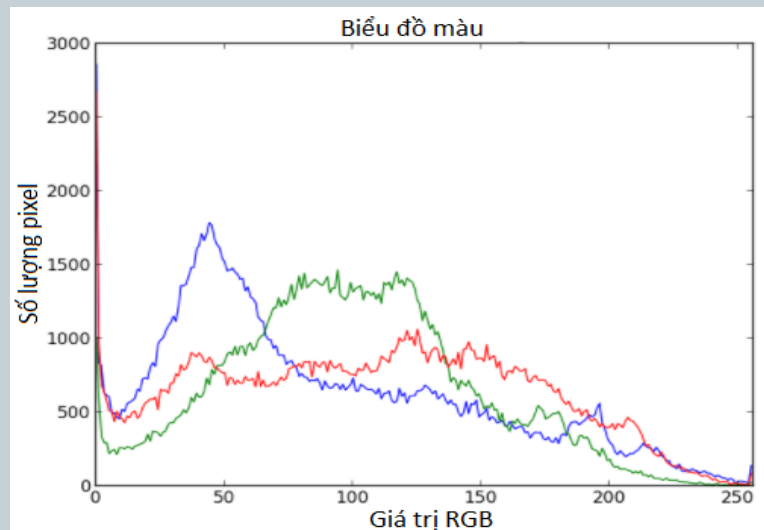
19

❑ Rút trích đặc trưng đối tượng:

❑ Sử dụng biểu đồ màu.



Ảnh của một chú chó



Biểu đồ màu của bức ảnh

PHƯƠNG PHÁP ĐỀ XUẤT

20

❑ Rút trích đặc trưng đối tượng:

❑ Sử dụng bảy moment bất biến của Hu.

- Tính: $\mu_{p,q} = \sum_{(x,y \in R)} (x - x_c)^p (y - y_c)^q$

Trong đó, (x_c, y_c) là trung tâm của đối tượng.

R là khung ảnh

$\mu_{p,q}$ là mô men trung tâm của thứ tự $(p+q)$

- Chuẩn hóa: $\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^\gamma}, \quad \gamma = \frac{p+q+2}{2}$

PHƯƠNG PHÁP ĐỀ XUẤT

21

❑ Rút trích đặc trưng đối tượng:

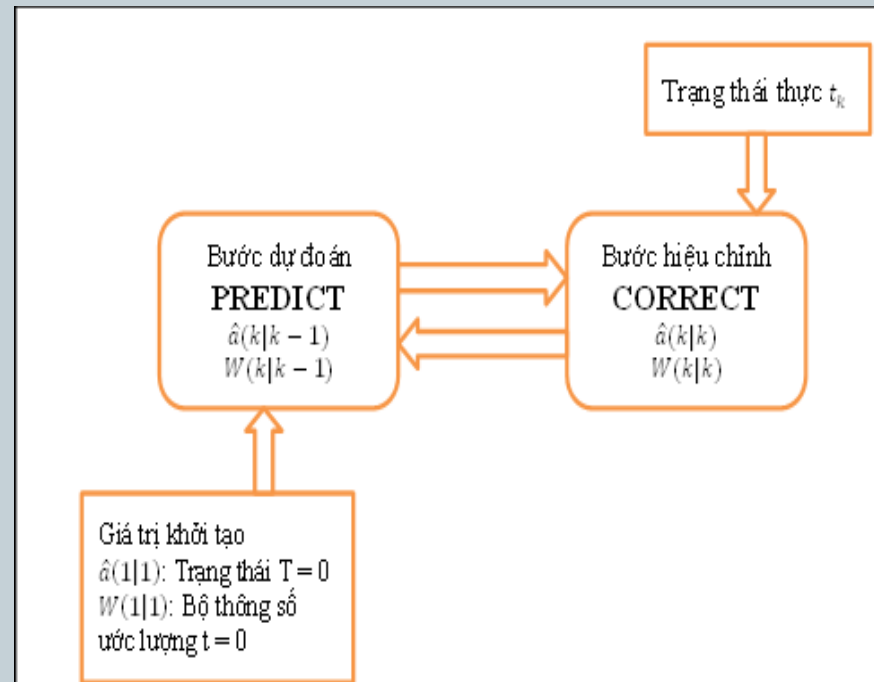
❑ Sử dụng bảy moment bất biến của Hu.

- $\phi_1 = \eta_{2,0} + \eta_{0,2}$
- $\phi_2 = (\eta_{2,0} - \eta_{0,2})^2 + 4\eta_{1,1}^2$
- $\phi_3 = (\eta_{3,0} - 3\eta_{1,2})^2 + (\eta_{0,3} - 3\eta_{2,1})^2$
- $\phi_4 = (\eta_{3,0} + \eta_{1,2})^2 + (\eta_{0,3} + \eta_{2,1})^2$
- $\phi_5 = (\eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2}) [(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{0,3} + \eta_{2,1})^2] + (\eta_{0,3} - 3\eta_{2,1})(\eta_{0,3} + \eta_{2,1}) [(\eta_{0,3} + \eta_{2,1})^2 - 3(\eta_{3,0} + \eta_{1,2})^2]$
- $\phi_6 = (\eta_{2,0} - \eta_{0,2}) [(\eta_{3,0} + \eta_{1,2})^2 (\eta_{0,3} + \eta_{2,1})^2 + 4\eta_{1,1}(\eta_{3,0} + \eta_{1,2})(\eta_{0,3} + \eta_{2,1})]$
- $\phi_7 = (3\eta_{2,1} - \eta_{0,3})(\eta_{3,0} + \eta_{1,2}) [(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{0,3} + \eta_{2,1})^2]$

PHƯƠNG PHÁP ĐỀ XUẤT

22

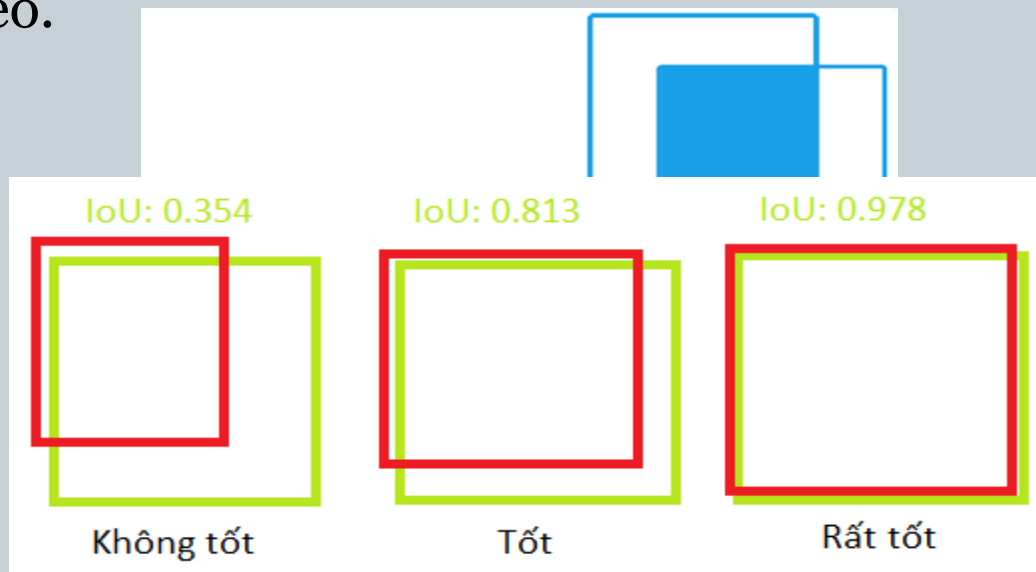
- ❑ Gán nhãn cho đối tượng trên từng camera:
 - ❑ Sử dụng giải thuật Kalman Filter.
 - ❑ **Bước dự đoán (Predict):** Từ bộ thông số, dự đoán vị trí đối tượng trong frame ảnh tiếp theo.
 - ❑ **Bước hiệu chỉnh (Correct):** Sử dụng các kết quả thu được để xác định lại bộ thông số cho bộ lọc



PHƯƠNG PHÁP ĐỀ XUẤT

23

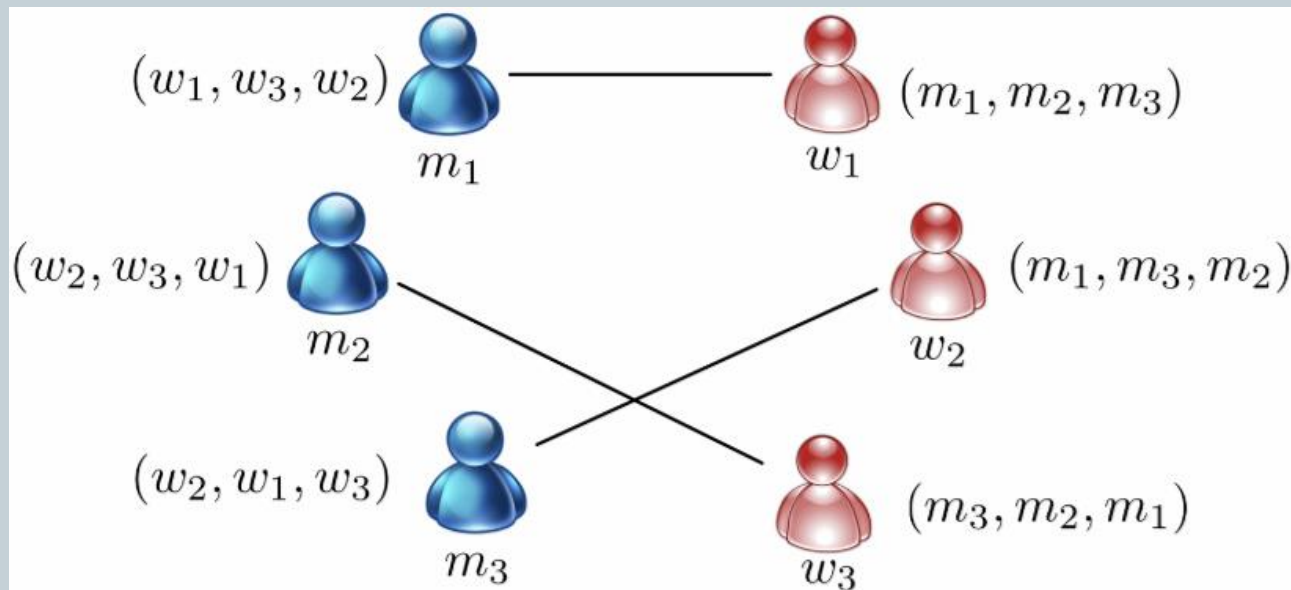
- ❑ Gán nhãn cho đối tượng trên từng camera:
 - ❑ Sử dụng giải thuật Kalman Filter.
 - ❑ Sử dụng công thức tính IoU để cải tiến việc ánh xạ giữa các giá trị ước lượng và các đối tượng phát hiện được trong frame ảnh tiếp theo.



PHƯƠNG PHÁP ĐỀ XUẤT

24

- ❑ Gán nhãn cho đối tượng giữa các camera:
 - ❑ Sử dụng giải thuật GSA.
 - ❑ Các đối tượng xuất hiện trong vùng không gian trùng lắp sẽ được ánh xạ để tìm các cặp đối tượng tương ứng và gán nhãn.



PHƯƠNG PHÁP ĐỀ XUẤT

25

❑ Dữ liệu đầu ra:

- ❑ Đối tượng di chuyển đã được gán nhãn.
- ❑ Dữ liệu đặc trưng về màu sắc, hình dáng của đối tượng.

NỘI DUNG

26

1. Giới thiệu đề tài

2. Cơ sở lý thuyết

3. Phương pháp đề xuất

4. Thí nghiệm và kết quả

5. Kết luận và hướng phát triển

THÍ NGHIỆM

27

❑ Xây dựng tập dữ liệu:

❑ Tập dữ liệu được xây dựng sẵn:

- ❑ Thiết lập hai camera đặt chéo nhau ở độ cao 1.8m.
- ❑ Thông số kĩ thuật: độ phân giải 360x288, 25 frames/giây.

❑ Tập dữ liệu tự xây dựng:

- ❑ Thiết lập hai camera đặt song song ở độ cao 1.5m.
- ❑ Thông số kĩ thuật: độ phân giải 960x720, 30 frames/giây.

KẾT QUẢ DỰ KIẾN

28

- ❑ Tỉ lệ gán nhãn đúng trên từng camera:
 - ❑ Hai camera đặt song song: ~ 90%.
 - ❑ Hai camera đặt chéo nhau: ~76%.
- ❑ Tỉ lệ gán nhãn đúng trong cả hai camera:
 - ❑ Hai camera đặt song song: ~87.5%.
 - ❑ Hai camera đặt chéo nhau: ~76%.

NỘI DUNG

29



1. Giới thiệu đề tài

2. Cơ sở lý thuyết

3. Phương pháp đề xuất

4. Thí nghiệm và kết quả

5. Kết luận và hướng phát triển

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

30

❑ Ưu điểm của phương pháp:

- ❑ YOLO phát hiện đối tượng trong thời gian thực.
- ❑ Giải thuật Kalman Filter đơn giản, dễ hiện thực.
- ❑ Khả năng rút trích màu và hình dáng dễ dàng.
- ❑ Việc xác định vùng không gian trùng lấp thực tiễn và hiệu quả cao.
- ❑ Giải thuật GSA phù hợp với bài toán ánh xạ đối tượng trong vùng không gian trùng lấp.

❑ Nhược điểm của phương pháp:

- ❑ YOLO không mang lại tính chính xác cao như RCNN, Fast RCNN...
- ❑ Không giải quyết được trường hợp camera đặt không có vùng trùng lấp.
- ❑ Giải thuật Kalman Filter không giải quyết triệt để vấn đề phủ lấp đối tượng.
- ❑ Hiệu suất của phương pháp phụ thuộc nhiều vào chất lượng ghi hình của camera.

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

31

□ Hướng phát triển:

- Tìm phương pháp phát hiện đối tượng dựa trên phương pháp học sâu
- Cải thiện phương pháp rút trích đặc trưng đối tượng.
- Giải quyết được bài toán “*gán nhãn cho các đối tượng di chuyển qua nhiều camera không có vùng không gian trùng lắp*”.

Thank you!

