

MỘT KỸ THUẬT ĐỊNH VỊ ĐỐI TƯỢNG TRONG HỆ THỐNG CAMERA GIÁM SÁT PHỤC VỤ THEO DÕI TRỰC QUAN

Đỗ Năng Toàn¹, Hà Mạnh Toàn², Phạm Bá Mấy¹, Ngô Đức Vĩnh³

¹ Viện Công nghệ thông tin, Đại học Quốc gia Hà Nội

² Viện Công nghệ thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

³ Trường Đại học Công nghiệp Hà Nội

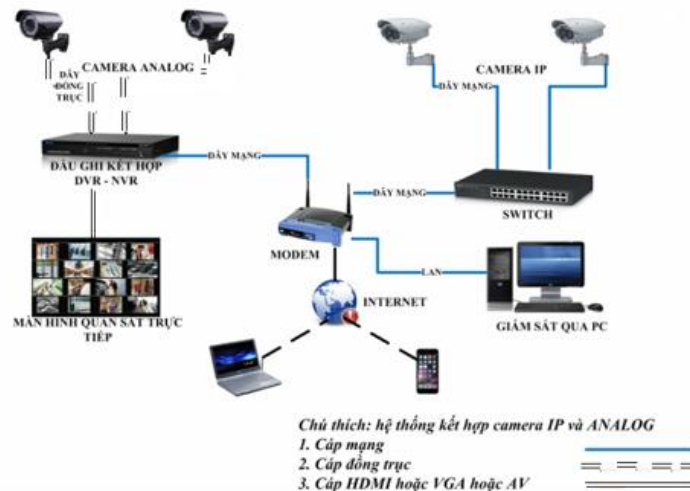
dntoan@vnu.edu.vn, hmtoan@ioit.ac.vn, mayioit@gmail.com, ngoducvinh@hau.edu.vn

TÓM TẮT: Hệ thống camera giám sát đang dần trở nên quen thuộc với sự phổ biến của các loại camera cùng với các dịch vụ lắp đặt camera. Thông thường những hệ thống này chỉ hỗ trợ việc quan sát từng màn hình camera riêng lẻ mà không thiết lập một cách nhìn toàn cảnh để có thể theo dõi, đánh giá vị trí của một đối tượng quan tâm trong khu vực. Khắc phục điều đó, bài báo trình bày một kỹ thuật định vị đối tượng phục vụ theo dõi trực quan trong hệ thống camera giám sát với việc kết hợp phát hiện vị trí đối tượng trên mỗi camera cụ thể và xây dựng ảnh xạ sang không gian địa lý đồng thời phân tích sai số dựa trên thông số lắp đặt của camera.

Từ khóa: object location, object detection, camera system...

I. GIỚI THIỆU

Hệ thống camera giám sát ngày nay đã và đang trở nên phổ biến rộng rãi và quen thuộc với cuộc sống con người, Hệ thống camera giám sát là một thành phần quan trọng trong việc đảm bảo an ninh tại các sân bay, ngân hàng, sòng bạc và các cơ quan cai quản. Gần đây hơn, các cơ quan chính phủ, doanh nghiệp và thậm chí cả các trường học đang hướng tới giám sát video như một phương tiện để tăng cường an ninh công cộng. Với sự gia tăng của các thiết bị và sự sẵn có của các mạng không dây băng thông rộng tốc độ cao, ngày càng nhiều camera giám sát an ninh đã được triển khai thực hiện và bước đầu cho thấy tính khả thi về mặt kinh tế và kỹ thuật.



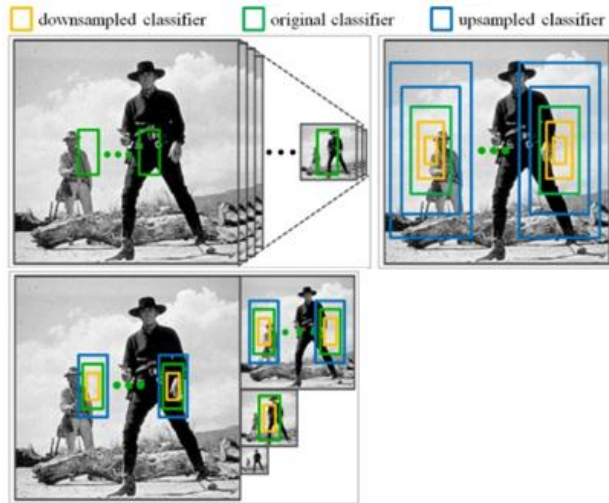
Hình 1. Một mô hình hệ thống camera giám sát

Các nghiên cứu trong hệ thống camera giám sát thường tập trung vào các thuật toán phân tích dữ liệu video để xác định những thông tin cần lưu ý. Một bài toán quan trọng trong đó là định vị đối tượng trong hệ thống camera.

Đối với con người trong cuộc sống hàng ngày, định vị một đối tượng nào đó là công việc quen thuộc và hiển nhiên đến mức nhiều khi chúng ta không kịp nhận ra. Khi lái xe trên đường, ta liên tục phải đánh giá vị trí mỗi đối tượng khác trên đường để đưa ra các thao tác điều khiển xe hợp lý, khi chơi thể thao như đánh cầu lông hay đá bóng thì ta cũng luôn cần phải biết quả cầu hay quả bóng đang ở đâu... Nói chung, ta có thể hiểu đơn giản định vị một đối tượng là việc xác định vị trí của đối tượng đó trong một không gian chứa vật đó được định nghĩa trước.

Với cách hiểu như vậy, bài toán định vị đối tượng trên camera có thể được nhìn dưới nhiều góc nhìn khác nhau. Nhiều nhóm nghiên cứu tiếp cận việc định vị với không gian chứa vật được hiểu là không gian khung hình của từng camera riêng biệt. Một trong những công trình nổi tiếng nhất đó là của Paul Viola và Micheal Jones [1], trong đó các tác giả sử dụng tiếp cận boosting để kết hợp nhiều đặc trưng Haar lại thành 1 bộ phân lớp mạnh phục vụ phát hiện đối tượng trong ảnh. Phương pháp này đã rất thành công và trở thành một trong những phương pháp phổ biến cho việc phát hiện khuôn mặt. Dollar và các đồng nghiệp [2] đề xuất một phần mở rộng cho công trình của Viola-John, trong đó đặc trưng Haar được tính trên nhiều kênh của dữ liệu hình ảnh, bao gồm các kênh màu LUV, thang màu xám và một số

tính toán gradient. Từ đó các tác giả đưa ra một quy trình đơn giản và thống nhất cho việc tích hợp nhiều loại đặc trưng. Mở rộng cách tiếp cận này, với phân tích rằng các đặc trưng tính trên một mức có thể được sử dụng để mô tả gần đúng ở các mức tỉ lệ gần nó, công trình [3] đã trình bày một kỹ thuật phát hiện đối tượng ở nhiều mức tỉ lệ.



Hình 2. Tiếp cận tính xấp xỉ đặc trưng ở nhiều mức tỉ lệ [3]

Một cách tiếp cận khác đó là nhìn nhận không gian chứa vật là không gian chung mà các camera được thiết lập bên trong đó với các mối quan hệ không gian tương ứng. Chẳng hạn từ dữ liệu quan sát của các camera trong một hệ thống camera giám sát của một tòa nhà, ta có thể định vị đối tượng quan tâm trên cơ sở tương quan vị trí của camera quan sát trong tòa nhà, từ đó đưa ra ước lượng tương đối vị trí một người trong tòa nhà như đến gần cửa, cầu thang... hoặc gần camera khác. Tiếp cận này sử dụng việc định vị đối tượng trên một camera như một bài toán con, tức là từ dữ liệu video, cần thiết ước lượng được vị trí xuất hiện của đối tượng trong khung hình. Đây là một bước quan trọng để tính toán vị trí đối tượng này trong một không gian kết hợp giữa các camera. Một hướng sử dụng không gian chung này là trong các hệ thống cần đến sự chuyển tiếp camera. Công trình [4] có sử dụng đặc trưng đường trong việc kết hợp với các điểm được so khớp. Trong đó, mỗi đặc trưng đường được mô tả bởi 2 điểm và theo đó việc so khớp được ước lượng với các đặc trưng đường.



Hình 3. Ví dụ các đặc trưng đường trong [4], các đặc trưng khớp nhau được tô cùng màu

Thông thường các hệ thống camera giám sát chỉ cung cấp một giao diện quan sát theo nhóm của từng camera cụ thể chứ không đi sâu vào việc định vị trên không gian chung, thậm chí việc định vị đối tượng trên từng camera cụ thể nếu có cũng đã là tính năng mở rộng. Bài báo tiếp cận việc định vị đối tượng trong hệ thống camera giám sát theo hướng ước lượng tương đối vị trí đối tượng trong hệ tọa độ chung, chẳng hạn với camera giám sát trong một khu vực sảnh, vấn đề đặt ra đầu tiên là phát hiện vị trí của đối tượng trong tọa độ khung hình của camera, sau đó là tìm một cách thức ánh xạ những đối tượng đã được phát hiện trong khung hình camera lên không gian bản đồ 2D của khu vực sảnh quan sát. Đây cũng là cơ sở để tạo ra khung nhìn trên không gian bản đồ toàn cục mà cả hệ thống camera đang giám sát. Với quy trình được đặt ra, bài báo cũng tiến hành phân tích sai số dựa trên thông số được thiết lập của camera. Đây là một tiền đề quan trọng để đánh giá chất lượng của một thiết kế lắp đặt camera nếu phục vụ theo dõi trực quan trên bản đồ 2D.

II. ĐỊNH VỊ ĐỐI TƯỢNG TRONG HỆ THỐNG CAMERA GIÁM SÁT PHỤC VỤ THEO DÕI TRỰC QUAN

Trong hệ thống camera giám sát, thông tin về đối tượng được xác định dựa trên những dữ liệu đầu vào thu nhận được từ từng camera giám sát. Mỗi camera sẽ thể hiện hình ảnh về một vùng không gian thật được chiếu lên cảm biến

của camera đó. Trong một hệ thống nhiều camera lắp đặt để giám sát một khu vực, bản chất là hệ thống sẽ thu thập và phân tích một tập các hình chiếu của từng vùng không gian nhỏ trong khu vực đó. Không gian của camera giám sát có thể hiểu trực tiếp là không gian ảnh tương ứng với khung hình thu nhận được của camera, cũng có thể hiểu theo ngữ cảnh liên hệ với vùng không gian thật hoặc theo nghĩa có sự liên kết với các camera khác trong hệ thống như một dạng liên kết giữa các nút trong một đồ thị.

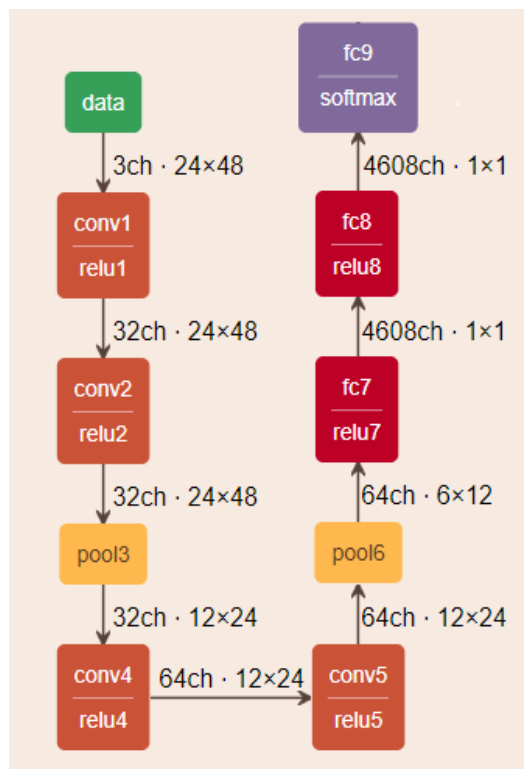
Vấn đề quan tâm của bài báo là ước lượng tương đối vị trí đối tượng trong hệ tọa độ chung, chẳng hạn với camera giám sát trong một khu vực sảnh, cần tìm một cách thức ánh xạ những đối tượng đã được phát hiện trong khung hình camera lên không gian bản đồ 2D của khu vực sảnh quan sát. Đây cũng là cơ sở nếu sau đó muốn đánh giá tương quan vị trí một người trong tòa nhà như đến gần cửa, cầu thang..., hoặc để ghép nối và tạo ra khung nhìn trên không gian bản đồ toàn cục mà cả hệ thống camera đang giám sát. Như vậy từ dữ liệu video thu được, chương trình sẽ xử lý để tính toán được vị trí xuất hiện của đối tượng trong khung hình. Trên cơ sở đó đưa ra ước lượng tương đối vị trí mỗi đối tượng trong không gian bản đồ 2D của khu vực quan sát. Quy trình thực hiện sẽ diễn ra theo 2 giai đoạn: xác định đối tượng trong hệ tọa độ của camera và xác định đối tượng trong hệ tọa độ chung; cụ thể hơn đó là định vị đối tượng trong một camera và sau đó ánh xạ vị trí đó từ không gian camera sang không gian chung.

A. Định vị đối tượng trong một camera

Bước đầu tiên của việc xử lý được tiến hành riêng rẽ trên dữ liệu của từng camera cụ thể. Với một khung hình đầu vào, chương trình cần đưa ra được vị trí cụ thể của các đối tượng quan tâm trong khung hình tương ứng là một danh sách các hình chữ nhật. Bước này được thực hiện dựa trên [5], trong đó, chương trình sẽ xét duyệt các vị trí có thể trên khung hình đầu vào; tại mỗi vị trí, một mạng nơ ron tích chập được áp dụng để trả lời câu hỏi khung hình đó có phải đối tượng quan tâm hay không. Trong [5], mạng nơ ron tích chập được thiết kế dựa trên công thức:

$$\text{INPUT} \rightarrow [(\text{CONV} \rightarrow \text{RELU})^*N \rightarrow \text{POOL?}]^*M \rightarrow [\text{FC} \rightarrow \text{RELU}]^*K \rightarrow \text{FC}$$

Trong đó, ký hiệu * là chỉ lặp lại, ký hiệu POOL? chỉ lựa chọn xem có sử dụng tầng POOL không. Ở đây, $N \geq 0$ và thường $N \leq 3$, $M \geq 0$, $K \geq 0$ và thường $K < 3$. Một trường hợp là $\text{INPUT} \rightarrow \text{FC}$, tương ứng mô hình phân lớp tuyến tính. Ở đây $N = M = K = 0$.

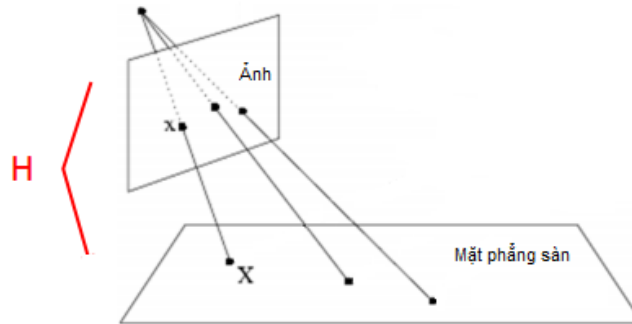


Hình 4. Cấu trúc mạng CNN trong [5]

B. Ánh xạ từ không gian camera sang không gian chung

Sau khi đã có được vị trí đối tượng trong khung hình camera, bước tiếp theo là thực hiện ánh xạ từ không gian camera sang không gian bản đồ 2D của khu vực quan sát. Từ những vị trí đối tượng đã được xác định trong không gian ảnh của khung hình camera, các vị trí này ta coi như các điểm 2D trên không gian ảnh, việc tính toán được thực hiện để

xây dựng một ánh xạ cho phép chiếu tương ứng các vị trí đối tượng đó lên mặt phẳng tương ứng với mô hình bản đồ 2D chứa không gian quan sát của camera. Quá trình này có thể quy về việc tính ma trận Homography, tức là xây dựng một ánh xạ từ mặt phẳng đối tượng đến mặt phẳng ảnh hoặc ngược lại.



Hình 5. Tương quan giữa điểm trên ảnh camera với điểm của đối tượng

Ma trận Homography liên quan đến việc biến đổi các điểm giữa hai mặt phẳng theo quan hệ như sau:

$$s \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Như vậy với hai tập điểm tương ứng $\{(x_i, y_i)\}$ và $\{(x'_i, y'_i)\}$, $i = 1..n$, việc tính ma trận Homography quy về việc cực tiểu hàm sai số:

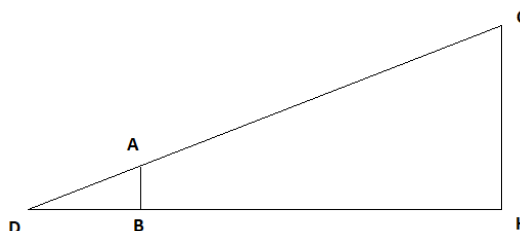
$$F = \sum_{i=1}^n \left(\left(x'_i - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 + \left(y'_i - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 \right)$$

Trong trường hợp có tồn tại một vài những cặp điểm ngoại lai, ta có thể áp dụng một số chiến lược khác chẳng hạn như chỉ ước lượng ma trận Homography chỉ với việc sử dụng ngẫu nhiên một số tập con các cặp điểm sau đó đánh giá sai số với từng ma trận Homography ước lượng được và chọn kết quả tốt nhất.

C. Vấn đề sai số

Qua 2 bước trên, ta đã có quy trình tính toán từ khung hình đầu vào của mỗi camera đến từng vị trí đối tượng trong bản đồ 2D của khu vực quan sát. Một vấn đề quan trọng cần đánh giá đó là mức độ sai số của mô hình này, cụ thể hơn trong trường hợp này ta xét sai số của bước ánh xạ từ không gian camera sang không gian chung. Có một bước quan trọng đó là cần tính toán vị trí đại diện là điểm 2D của mỗi đối tượng trong không gian ảnh để làm cơ sở thực hiện ánh xạ Homography. Với kết quả của việc định vị đối tượng trên một camera là các hình chữ nhật thể hiện vị trí đối tượng, có nhiều lựa chọn cho vị trí điểm 2D đại diện, có thể là điểm tâm của hình chữ nhật, có thể là trung điểm của cạnh dưới đáy... Như vậy điểm đại diện được chọn có thể sẽ không nằm trên mặt phẳng mà ta lựa chọn cho việc tính toán ma trận Homography. Chưa kể đến một thực tế là chính bước tính toán ra vị trí là hình chữ nhật cũng sẽ có sai số. Chẳng hạn với camera quan sát những người đi lại trong một sảnh hoặc một căn phòng, một lựa chọn tự nhiên cho mặt phẳng để tính ma trận Homography đó là mặt phẳng sàn, và như vậy điểm 2D cần thiết phải ước lượng được đó là điểm tâm của vùng người đó tiếp xúc với sàn. Với hình dung như vậy, ta dễ dàng thấy được khả năng tồn tại sai số khi ước lượng vị trí trong bản đồ 2D thông qua điểm 2D đại diện vị trí trên khung hình của camera.

Ta có thể phân tích một cách tương đối ảnh hưởng của sai số của quá trình tính toán như sau. Coi đối tượng là đoạn thẳng AB vuông góc với mặt sàn, điểm B là chân đối tượng nằm ở trên mặt sàn, camera C có hình chiếu là H trên mặt phẳng sàn, đường thẳng nối camera C và đi qua điểm A cắt mặt sàn tại điểm D. Như vậy, ta có thể xét các điểm A, B, C, D, H trên một phẳng phẳng đi qua camera C, qua đối tượng AB và vuông góc với mặt sàn.

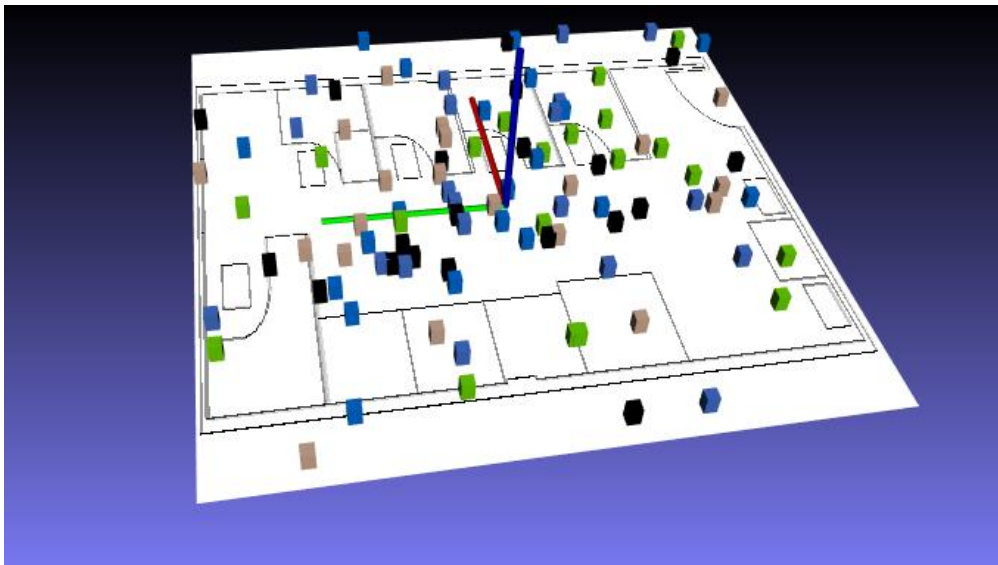


Hình 6. Mô phỏng mặt phẳng đi qua camera, đối tượng và vuông góc mặt sàn

Từ trên khung hình của camera, ta có thể phát hiện được vị trí của đối tượng AB và theo đó có thể xác định được điểm 2D đại diện trên khung hình và theo lý thuyết, điểm đại diện đó sẽ nằm trên đoạn thẳng là hình chiếu của đoạn thẳng AB trên khung hình camera. Như vậy, điểm đại diện sau khi thực hiện biến đổi Homography sẽ nằm trên đoạn thẳng DB và sai số của việc ước lượng sẽ là khoảng cách từ điểm đó đến điểm B. Dễ dàng nhận thấy, đoạn thẳng DB càng lớn thì giá trị sai số khả năng của việc ước lượng sẽ càng cao, và DB càng lớn khi mà giá trị góc \widehat{CDH} hoặc \widehat{CBH} càng nhỏ. Xét tương quan đến giá trị góc θ là góc giữa vector từ đối tượng đến camera so với mặt sàn, do ta không xét đối tượng là 1 điểm nên giá trị góc θ có thể hiểu là giá trị nằm trong $[\widehat{CDH}, \widehat{CBH}]$. Như vậy có thể nói rằng, sai số ước lượng sẽ càng lớn khi góc θ càng nhỏ, và tương tự, khi θ đạt $\pi/2$.

III. THỬ NGHIỆM

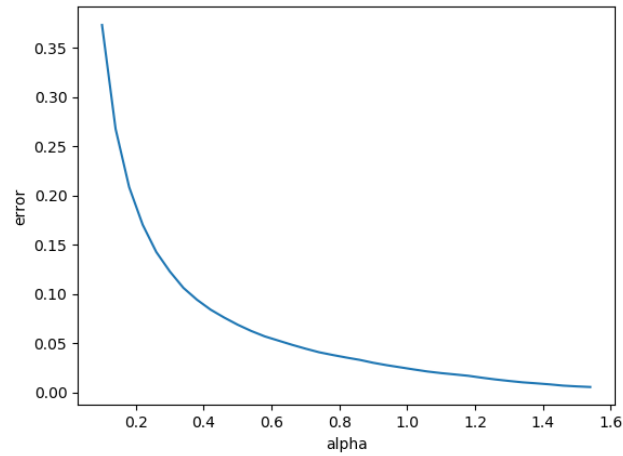
Quy trình được cài đặt thử nghiệm và đánh giá trên hai trường hợp: với dữ liệu giả lập và với dữ liệu thực từ camera giám sát. Với trường hợp dữ liệu giả lập, để phục vụ mục đích đánh giá sai số như đã được phân tích, bài báo sử dụng các kỹ thuật đồ họa máy tính để dựng mô phỏng các đối tượng trong một căn phòng. Trong thử nghiệm này, một mặt phẳng đại diện cho mặt phẳng sàn của căn phòng được thiết lập trên mặt phẳng Oxy với tâm mặt phẳng sàn là góc tọa độ và 100 đối tượng được sinh và cho di chuyển ngẫu nhiên trên mặt phẳng sàn.



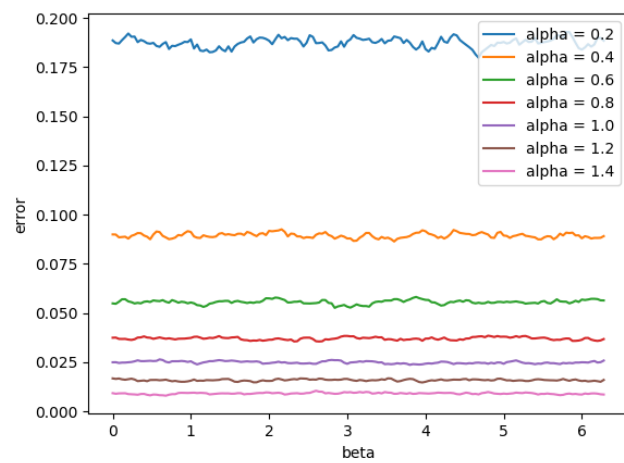
Hình 7. Dữ liệu được dựng giả lập: các đối tượng là các khối hộp có màu sắc được sinh và cho di chuyển ngẫu nhiên trên mặt phẳng sàn

Để tiến hành đánh giá sai số, camera được thiết lập đặt trên mặt cầu với tâm là gốc tọa độ $(0, 0, 0)$ và bán kính bằng 2 lần kích thước của căn phòng. Điểm 2D đại diện cho mỗi đối tượng được lấy bằng cách chọn ngẫu nhiên 1 điểm 3D trên trục thẳng đứng tại tâm của mỗi đối tượng với dao động trong đoạn $[-2/3, 2/3]$ của chiều cao đối tượng và hình chiếu của điểm 3D này trên mặt phẳng quan sát được chọn làm điểm 2D đại diện của đối tượng trên mặt phẳng ảnh. Để đánh giá sai số của ước lượng vị trí của mỗi đối tượng, ta thực hiện ánh xạ vị trí điểm 2D đại diện của đối tượng lên mặt phẳng bản đồ 2D và tính khoảng cách Euclid so với vị trí chính xác. Bằng cách tính trung bình trên 100 đối tượng được sinh và cho di chuyển ngẫu nhiên theo N khung hình liên tiếp, ta sẽ có một ước lượng tương đối về sai số của việc ước lượng, trong thử nghiệm, với mỗi lần đánh giá chúng tôi chọn $N=1000$. Chi tiết hơn trong việc đánh giá, chúng tôi tiến hành thử nghiệm với các trường hợp khác nhau của hai tham số góc lệch α và β . Trong đó α là góc lệch giữa vector từ gốc tọa độ đến camera so với mặt phẳng sàn và β là góc lệch giữa vector từ gốc tọa độ đến hình chiếu của camera trên mặt phẳng Oxy so với trục Ox. Như vậy, góc α sẽ nằm trong nửa đoạn $(0, \pi/2]$ và góc β sẽ nằm trong đoạn $[0, 2\pi]$. Ta đã phân tích sai số của việc ước lượng vị trí của một đối tượng phụ thuộc vào góc θ là góc giữa vector từ đối tượng đến camera so với mặt sàn, tuy nhiên các đối tượng luôn thay đổi vị trí trên vùng quan sát, nên việc sử dụng góc θ để đánh giá là không thực tế. Thay vào đó, vì trong thử nghiệm ta thiết lập mặt phẳng sàn trên Oxy với tâm trùng với gốc tọa độ nên ta sử dụng góc α như một xấp xỉ tương đối cho các góc θ của các đối tượng quan sát. Về mặt lý thuyết, ta có thể dự đoán rằng sai số sẽ không thay đổi khi ta giữ nguyên góc α và thay đổi góc β đồng thời khi cố định β , sai số sẽ lớn dần khi α dần về 0 và đạt cực tiểu khi α đạt $\pi/2$.

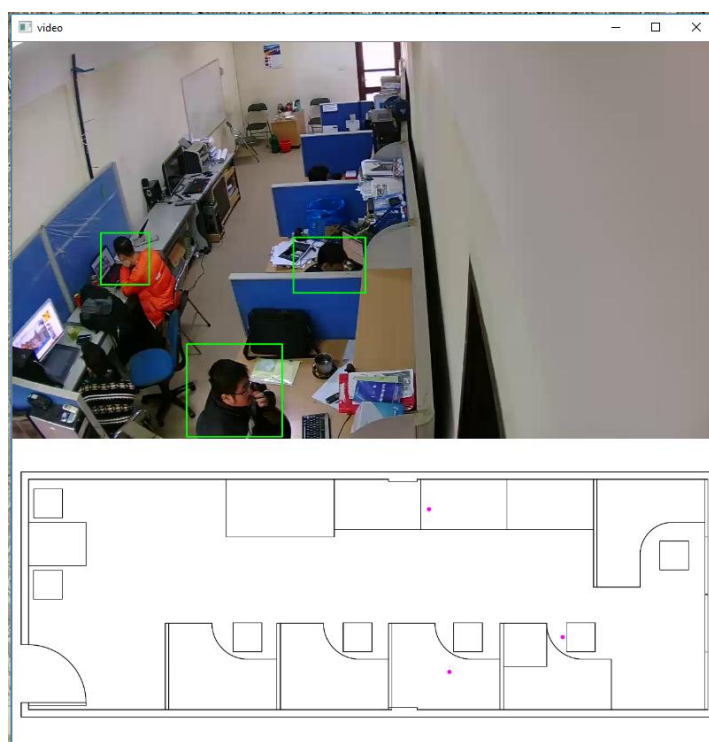
Trong thử nghiệm đánh giá tương quan giữa sai số hệ thống và giá trị α , dễ dàng nhận thấy dự đoán từ lý thuyết là chính xác khi sai số giảm dần khi giá trị α dần đến $\pi/2$ và sai số tăng khi α dần về 0. Điều này cũng càng được khẳng định với thử nghiệm khi ta cố định từng giá trị α cụ thể và cho β thay đổi. Có thể thấy, với mỗi mức α cụ thể, sai số của hệ thống tương đối ổn định khi ta thay đổi các giá trị β nhưng lại có sự cách biệt rõ ràng về mức sai số giữa các giá trị α khác nhau. Đây là những tiêu chí quan trọng để đánh giá chất lượng định vị của hệ thống với một thiết kế lắp đặt của các camera.



Hình 8. Thử nghiệm đánh giá sai số của hệ thống khi cho thay đổi α trong $(0, \pi/2]$ với $\beta=0$



Hình 9. Thử nghiệm đánh giá sai số của hệ thống khi cho thay đổi β trong $[0, 2\pi]$ với từng mức giá trị α cụ thể



Hình 10. Triển khai áp dụng trên hệ thống camera

Bên cạnh việc thử nghiệm với dữ liệu mô phỏng để đánh giá sai số, chúng tôi cũng tiến hành triển khai thử nghiệm trong hệ thống camera giám sát với chức năng định vị người. Để thực hiện, chúng tôi tiến hành vẽ lại sơ đồ phòng để làm dữ liệu bản đồ 2D của khu vực quan sát. Do góc nhìn thường bị khuất mất một phần thân dưới, chúng tôi hướng đến việc phát hiện người dựa trên việc phát hiện vùng đầu người, theo đó mạng CNN được tiến hành học với dữ liệu đầu người trong bộ dữ liệu SCUT-HEAD [6] với vùng đầu người được mở rộng. Do vị trí bất đối tượng sẽ ở cao hơn so với sàn, chúng tôi thực hiện xây dựng ánh xạ cặp điểm cho việc tính ma trận Homography với các điểm đặt trên mặt bàn thay vì đặt trên sàn. Cũng như phân tích trong dữ liệu mô phỏng, chất lượng định vị trên bản đồ sẽ tốt với với những đối tượng ở dưới, tức là tương ứng với những vị trí có góc α lớn, và sẽ kém hơn với những vị trí ở xa, chẳng hạn như ở vùng bàn uống nước, khi đó việc định vị có thể nhầm với vùng nằm ngoài lề phòng.

IV. KẾT LUẬN

Định vị đối tượng trong hệ thống camera giám sát là vấn đề quan trọng trong nghiên cứu lý thuyết xử lý ảnh cũng như trong việc ứng dụng thực tế trên các hệ thống camera giám sát. Thông thường các hệ thống camera giám sát chỉ cung cấp một giao diện quan sát theo nhóm của từng camera cụ thể chứ không đi sâu vào việc định vị trên không gian chung. Bài báo đã trình bày một kỹ thuật định vị đối tượng và thể hiện trên không gian bản đồ 2D phục vụ theo dõi trực quan đồng thời tiến hành những thử nghiệm cho phép đánh giá sai số với những ảnh hưởng của góc lắp đặt camera. Đây là cơ sở quan trọng khi ta thiết kế lắp đặt các camera quan sát phục vụ việc định vị để theo dõi trực quan trên không gian bản đồ 2D.

TÀI LIỆU THAM KHẢO

- [1] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." CVPR (1) 1 (2001): 511-518.
- [2] P. Dollar, Z. Tu, P. Perona, and S. Belongie, "Integral Channel Features," Proc. British Machine Vision Conf., 2009.
- [3] P. Dollar, S. Belongie, and P. Perona, "The Fastest Pedestrian Detector in the West," Proc. British Machine Vision Conf., 2010.
- [4] Zhang, Zhong, et al. "Video surveillance using a multi-camera tracking and fusion system." Multi-Camera Networks: Principles and Applications (2009): 435-456.
- [5] Hà Mạnh Toàn, Nguyễn Văn Năng, Trịnh Hiền Anh, Nguyễn Thị Lan Phương, Trịnh Xuân Hùng, "Một kỹ thuật phân lớp người sử dụng mạng nơron tích chập", tr.182-185, Kỷ yếu Hội thảo Quốc gia, Một số vấn đề chọn lọc của Công nghệ thông tin và truyền thông, @ lần thứ XXI, Thanh Hóa, 27-28/07/2018.
- [6] Peng, Dezhi, et al. "Detecting Heads using Feature Refine Net and Cascaded Multi-scale Architecture." 2018 24th International Conference on Pattern Recognition (ICPR). IEEE, 2018.

AN OBJECT LOCATING TECHNIQUE IN CAMERA SYSTEM FOR VISUAL MONITORING

Do Năng Toàn, Hà Mạnh Toàn, Phạm Bá Máy, Ngô Đức Vinh

ABSTRACT: Surveillance camera system is gradually becoming familiar with the popularity of cameras with camera installation services. Usually these systems only support the observation of individual camera screens without establishing a panoramic view to be able to track and assess the location of an interested area. Overcoming that, the paper presents an object locating technique for visual monitoring in surveillance camera system with the combination of detecting the object position on each specific camera and building the map to geospatial space, then analyzing errors based on camera settings.