

Nhóm 3:

Thành viên:

18521591 Đặng Nguyễn Anh Tuấn

18521503 Đặng Hữu Toàn

18521632 Nguyễn Văn

Đề tài: Augmented Reality with local feature

Mục lục

...

Mở

...

Phần 1: tổng quan về AR

Augmented Reality (AR) (Thực tế Tăng cường) được hiểu là công nghệ cho phép con người quan sát những vật trong thế giới thật thông qua một thiết bị điện tử,

Giúp cho phép lồng ghép thông tin ảo vào thế giới thực (và ngược lại),

Để người sử dụng tương tác với những nội dung số trong thực tại (như chạm vào, phủ vật thể lên trên - nói dễ hiểu là ghép ảnh theo dạng 3D),...

Tóm AR là công nghệ giúp bổ sung các chi tiết ảo vào thế giới thực tại thông qua các thiết bị số.

Ưu, Nhược, thách thức:...

Phần 2: các bước tiến hành

1. Tổng quan

Sử dụng python và openCV để chiếu lên màn hình mô hình 3d có vị trí và hướng ăn khớp với vị trí và hướng của mặt phẳng được xác định trước. Hơn nữa sẽ làm theo thời gian thực, để nếu có thay đổi về vị trí hoặc hướng của bề mặt, hướng và vị trí của mô hình cũng sẽ thay đổi theo

Để đạt được điều này, trước tiên phải xác định được bề mặt phẳng tham chiếu trong khung hình ảnh hoặc video.

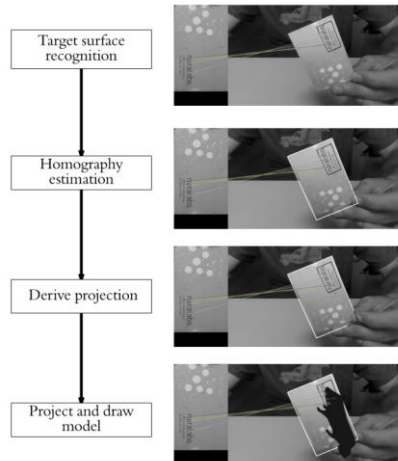
Sau khi xác định được, xác định sự chuyển đổi từ ảnh bề mặt tham chiếu (2D) sang ảnh mục tiêu (2D). Sự biến đổi này được gọi là phép đồng nhất.

(homography). Và sử dụng chút đại số để mở rộng phép biến đổi xử lý các trường hợp **chiều cao của điểm cần chiếu trong hệ tọa độ bề mặt tham chiếu khác 0.**

Cuối cùng, áp dụng sự biến đổi cho mô hình 3D và vẽ lên

Chia ra làm các phần sau:

- 1.1. Nhận biết bề mặt phẳng tham chiếu.(Recognize the reference flat surface)
- 1.2. Tính toán sự đồng nhất. (Estimate the homography.)
- 1.3. Từ phép đồng nhất mới tính, chuyển đổi từ hệ tọa độ bề mặt tham chiếu sang hệ tọa độ ảnh mục tiêu. (Derive from the homography the transformation from the reference surface coordinate system to the target image coordinate system)



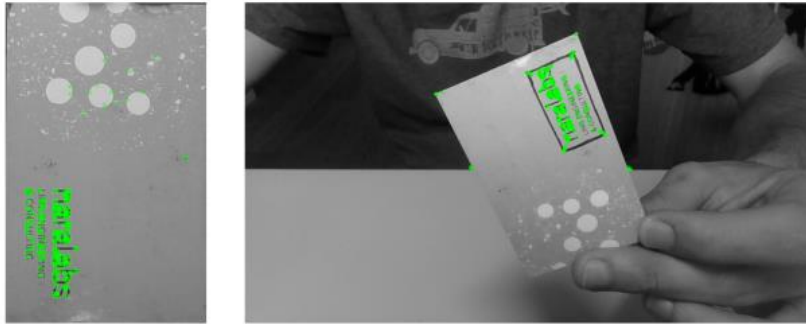
- 1.4. Chiếu mô hình 3D vào hình ảnh (không gian pixel) và vẽ nó.(Project our 3D model in the image (pixel space) and draw it)

2. Nhận biết bề mặt phẳng tham chiếu

2.1. nhận diện đặc trưng:

Commented [A.1]: Cái này ko hiểu lắm? như kiểu 2d thì hình nó flat height =0, còn 3d thì có cả dài rộng cao nên height nó sẽ khác 0, nên ms cần thêm các phép xử lý khác đúng ko?

đại khái, bước này trước tiên tìm kiếm trong cả hình ảnh tham chiếu và hình ảnh đích để tìm các đặc trưng nổi bật và mô tả một phần đối tượng được nhận dạng. Đặc trưng này sau đó có thể được sử dụng để tìm đối tượng tham chiếu trong hình ảnh đích. Sau



1 Ở bên trái, các đặc trưng được trích xuất từ bề mặt của mô hình. Ở bên phải, các đặc trưng được trích xuất từ ảnh mẫu. Chú ý cách mà các góc đã được phát hiện là điểm chú ý ở phía ngoài cùng trong hình bên phải.

đó sẽ cho rằng đã tìm thấy đối tượng khi phát hiện một số đặc trưng nhất định giữa ảnh mục tiêu và hình ảnh tham chiếu.

Đối với một vùng hoặc điểm của hình ảnh được gắn nhãn là đặc trưng, nó phải đáp ứng hai thuộc tính quan trọng:

- 2.1.1. Đầu tiên, phải thể hiện một số tính duy nhất ít nhất là **cục bộ**. Ví dụ các góc hoặc cạnh.
- 2.1.2. Thứ hai, vì không biết trước hướng, tỷ lệ hoặc độ sáng của một đối tượng trong hình ảnh muốn nhận diện, một đặc trưng lý tưởng phải bất biến với các phép biến đổi; tức là không thay đổi so với thay đổi tỷ lệ, xoay hoặc độ sáng.

Commented [A.2]: ???

2.2. Mô tả đặc trưng

Khi các đặc trưng đã được tìm thấy, nên tìm một biểu diễn phù hợp về thông tin mà chúng cung cấp. Điều này sẽ cho phép tìm kiếm chúng trong các hình ảnh khác và cũng để có được thước đo về mức độ tương tự của các đặc trưng được phát hiện giữa 2 hình khi so sánh.

Đây là lúc bộ mô tả được vào cuộc. Một bộ mô tả cung cấp sự biểu diễn thông tin được lấy bởi các đặc trưng và xung quanh nó. Sau khi các bộ mô tả đã được tính toán, đối tượng được nhận dạng sau đó có thể được trừu tượng hóa thành một

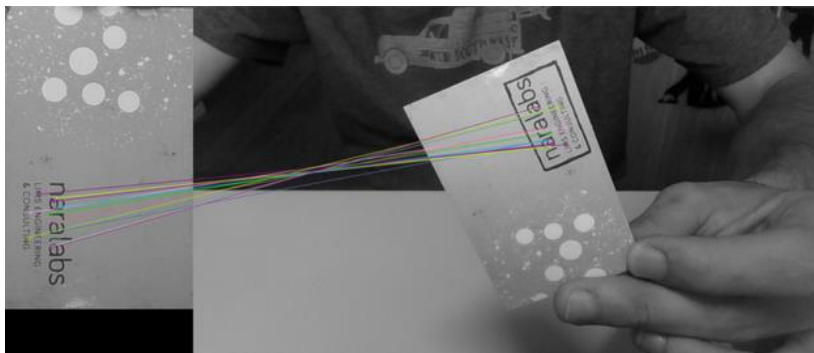
vector đặc trưng, hay là một vector có chứa các bộ mô tả của các điểm chính được tìm thấy trong hình ảnh với đối tượng tham chiếu

Có rất nhiều thuật toán trích xuất đặc trưng của hình ảnh và tính toán các bộ mô tả của nó như SIFT, SURF, hoặc Harris. Tuy vậy công cụ nhóm sẽ sử dụng được phát triển trong OpenCV được gọi là ORB (Oriented FAST and Rotated BRIEF: NHANH CHÓNG Định hướng và TÓM TẮT luân phiên). Hình dạng và giá trị của bộ mô tả phụ thuộc vào thuật toán được sử dụng và trong trường hợp này, các bộ mô tả thu được sẽ là các chuỗi nhị phân.

2.3. Đối xứng đặc trưng

Cách đơn giản nhất là lấy bộ mô tả của từng đặc trưng trong tập hợp đầu tiên, tính toán khoảng cách đến tất cả các bộ mô tả trong tập hợp thứ hai và trả về kết quả gần nhất là đối sánh tốt nhất (điều quan trọng là phải chọn một cách đo khoảng cách phù hợp với bộ mô tả đang được sử dụng. Vì bộ mô tả là chuỗi nhị phân nên sẽ sử dụng khoảng cách Hamming). Đây là một cách tiếp cận brute force và có nhiều phương pháp phức tạp hơn.

Ví dụ: có thể kiểm tra xem đối sánh được tìm thấy cũng là đối sánh tốt nhất khi tính khớp theo hướng ngược lại, từ các đặc trưng trong bộ thứ hai qua bộ đầu tiên. Điều này có nghĩa là cả hai đặc trưng phù hợp với nhau. Khi kết thúc đối sánh theo cả hai hướng, chỉ lấy đối sánh hợp lệ là đối sánh đã đáp ứng điều kiện trước đó. Hình 4 trình bày 15 đối sánh tốt nhất được tìm thấy bằng phương pháp này.



2 15 đối sánh brute force gần nhất được tìm thấy giữa bề mặt tham chiếu và ảnh đích

Cuối cùng, sau khi các kết quả phù hợp đã được tìm thấy, nên xác định một số tiêu chí để quyết định xem đối tượng đã được tìm thấy hay chưa. Đó là xác định một ngưỡng về số lượng đối sánh tối thiểu sẽ được tìm thấy. Nếu số lượng đối sánh trên ngưỡng,

Commented [A.3]: Có thể phải giải thích

Commented [A.4]: Có thể phải giải thích

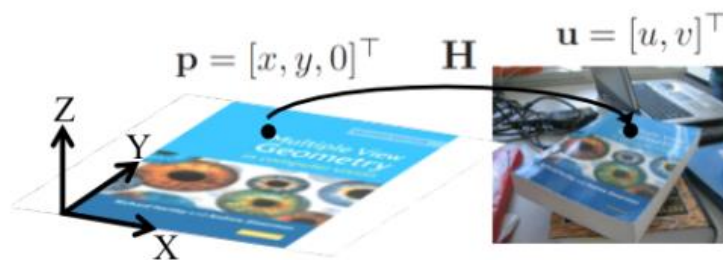
Commented [A.5]: Điều kiện gì?

thì sẽ giả sử đối tượng đã được tìm thấy. Nếu không, nó sẽ không có đủ bằng chứng để nói rằng việc nhận diện đã thành công.

Lưu ý cuối cùng và trước khi bước sang bước tiếp theo của quy trình, phải chỉ ra rằng, vì muốn có một ứng dụng thời gian thực, nên tốt hơn là triển khai kỹ thuật theo dõi (tracking technique) chứ không chỉ nhận dạng đơn thuần (plain recognition). Điều này là do việc nhận dạng đối tượng sẽ được thực hiện trong mỗi khung một cách độc lập mà không tính đến các thông tin có giá trị về vị trí của đối tượng tham chiếu có thể được bổ sung trong khung trước đó. Một điều khác cần lưu ý nữa là, bề mặt tham chiếu càng dễ tìm thấy thì khả năng phát hiện càng tốt.

3. Tính toán sự đồng nhất. (Estimate the homography.)

Khi đã xác định được bề mặt tham chiếu trong khung hiện tại và có một tập hợp các kết quả phù hợp, có thể tiến hành ước tính độ đồng nhất homography giữa cả hai hình ảnh. Như đã giải thích trước đây, chúng ta muốn tìm phép biến đổi các điểm



3 Homography giữa một mặt phẳng và một hình ảnh.

từ mặt phẳng bề mặt sang mặt phẳng ảnh (xem Hình 5). Sự chuyển đổi này sẽ phải được cập nhật mỗi khung mới.

Giải thích toán học

Homography là sự dịch chuyển sử dụng phép chiếu hình học, hay nói cách khác nó là sự kết hợp của cặp điểm trong phép chiếu phối cảnh. Ảnh thực trong không gian ba chiều có thể biến đổi về không gian ảnh bằng phép chiếu thông qua ma trận biến đổi Homography hay còn gọi là ma trận H. Các phép chiếu biến đổi thông qua ma trận Homography không đảm bảo về kích thước và góc của vật được chiếu, nhưng lại đảm bảo về tỉ lệ.

End giải thích ở đây, muốn thêm thì thêm

Có một số phương pháp ước tính các giá trị của ma trận đồng nhất homography, một trong những phương pháp đó, nhóm sẽ sử dụng RANdom SAMple Consensus (RANSAC). RANSAC là một thuật toán lặp lại được sử dụng để điều chỉnh mô hình

Commented [A.6]: Giải thích



(model fitting) khi có một số lượng lớn các giá trị ngoại lệ (Hình 12 đưa ra phác thảo chính của quy trình). Vì không thể đảm bảo rằng tất cả các kết quả đã tìm thấy thực sự là các kết quả phù hợp hợp lệ, nên phải xem xét rằng có thể có một số kết quả phù hợp sai (đó sẽ là giá trị ngoại lệ) và do đó, phải sử dụng phương pháp ước tính tốt hơn các giá trị ngoại lệ. Hình 11 minh họa những vấn đề mà có thể gặp phải khi ước lượng tính đồng nhất nếu coi rằng không có ngoại lệ.

Giải thích thuật toán RANSAC

...

Trong OpenCV đã có sẵn hàm xử lý ước lượng đồng nhất

Kết quả



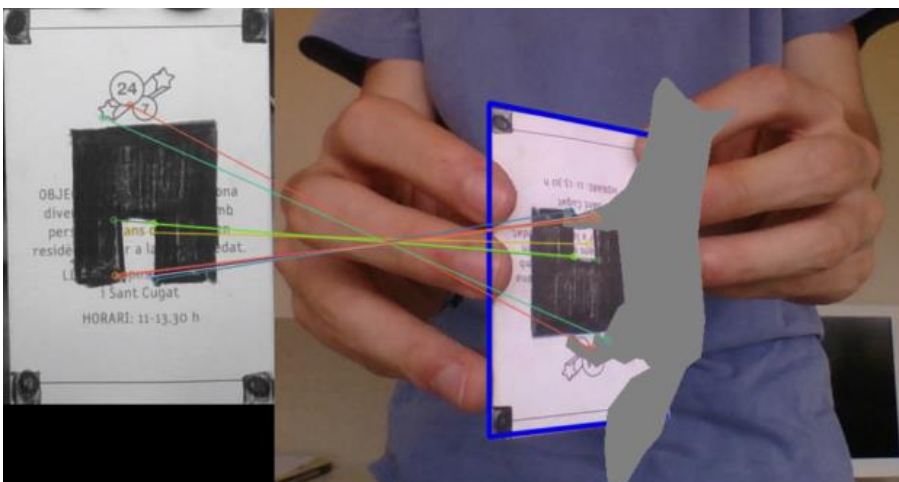
5 Các góc chiếu của bề mặt tham chiếu được tính bằng ước lượng đồng nhất (estimated homography)

4. chuyển đổi từ hệ tọa độ bề mặt tham chiếu sang hệ tọa độ ảnh mục tiêu (Derive projection) (projection Matrix)

Toán toán toán

...

5. Chiếu mô hình (Model projection)



giai đoạn cuối cùng:

sử dụng các mô hình đơn giản ở định dạng Wavefront .obj. Tại sao định dạng OBJ? Vì chúng dễ dàng xử lý và hiển thị trực tiếp bằng Python mà không cần phải sử dụng các thư viện khác như OpenGL. Vấn đề với các mô hình phức tạp là số lượng tiến trình mà chúng yêu cầu nhiều hơn những gì máy tính bình thường có thể xử lý. Vì ứng dụng hoạt động theo thời gian thực, điều này hạn chế độ phức tạp của các mô hình có thể kết xuất đồ họa.

Một số định dạng mô hình 3D (đa giác ít) được tải về từ clara.io.

Sử dụng code để tải hình học của mô hình lên. Khi mô hình được tải, chỉ cần triển khai chức năng đọc dữ liệu và chiếu nó lên trên khung hình video với ma trận chiếu mà đã thu được trong phần trước. Một khi đã xong, chỉ cần phải tô màu các mặt của mô hình.

Thêm hình kết quả

6. Thêm texture, giảm độ run (có thể có)

Phần 3: đánh giá

...

Kết luận

...

Tài liệu tham khảo