VNPT TECHNOLOGY

Logo

Description automatically generated

**BÁO CÁO THƯC TẬP**

**Project 1: TRACKING-USING-OPENCV**

Nhóm thực hiện: TTS-IPCam-1

Lê Minh Hưng

Lường Hữu Phong

Nguyễn Hoàng Anh

Mục lục

1. **Khái niệm cơ bản**
2. **Xử lý ảnh là gì.**

* Xử lý ảnh là quá trình chuyển đổi một hình ảnh sang dạng kỹ thuật số và thực hiện các thao tác nhất định để nhận được một số thông tin hữu ích từ hình ảnh đó. Hệ thống xử lý hình ảnh thường coi tất cả các hình ảnh là tín hiệu 2D khi áp dụng một số phương pháp xử lý tín hiệu đã xác định trước.
* Các loại xử lý hình ảnh chính:

+ Nhận diện – Phân biệt hoặc phát hiện các đối tượng trong hình

ảnh.

+ Làm sắc nét và phục hồi – Tạo hình ảnh nâng cao từ hình ảnh gốc.

+ Nhận dạng mẫu – Đo các mẫu khác nhau xung quanh các đối

tượng trong hình ảnh.

+ Truy xuất – Duyệt và tìm kiếm hình ảnh từ một cơ sở dữ liệu lớn gồm các hình ảnh kỹ thuật số tương tự như hình ảnh gốc.

1. **Ảnh số**
2. **Khái niệm**

* Ảnh số là dữ liệu máy tính lưu trữ thông tin về màu sắc để truyền tải đến người dùng. Ảnh số là dữ liệu do đó cũng tốn vùng nhớ, cũng có thể nén, có thể mất dữ liệu như những file khác. Và đặc biệt ảnh là loại dữ liệu quý giá giúp người dùng hiểu vấn đề nhanh hơn đọc chữ.
* Mỗi ảnh số là một ma trận, với một pixel là 1 giá trị trong ma trận. Với ma trận có thể cộng trừ nhân chia, hoán vị,… để biến đổi ma trận theo ý muốn. Và từ ma trận có thể chuyển đổi thành vecto và ngược lại (đại số tuyến tính và hình học)

1. **Các loại ảnh số**

* Ảnh số được chia làm 2 loại chính: Ảnh đen trắng và ảnh màu

+ Ảnh đen trắng: Là loại ảnh chỉ gồm 2 màu đen trắng, không chưa các màu khác. Ảnh đen trắng lại được chia thành 2 loại là ảnh nhị phân và ảnh xám. Người ta phân mức đen trắng của ảnh thành L mức. Nếu L bằng 2 nghĩa là chỉ có 2 mức 0 và 1, hay còn gọi là ảnh nhị phân. Nếu L lớn hơn 2 thì ta sẽ có ảnh xám

+ Ảnh màu: Tùy theo các trường hợp mà ảnh màu được biểu diễn bằng những không gian màu khác nhau. Không gian màu là những mô hình toán học dùng để mô tả các màu sắc trong thực tế dưới dạng số học.

1. **Kênh màu**

* Một hình ảnh màu được tạo từ các kênh màu khác nhau. Mỗi kênh màu là 1 ma trận.
* Ví dụ: RGB được tạo 24 bit phổ biến, hình ảnh về cơ bản được tạo từ ba tệp 8 bit nhỏ hơn. Các tệp nhỏ hơn này là các kênh của hình ảnh và chúng có phạm vi 256 giá trị, từ 0 – 255.

1. **Pixel**

* Pixel- một khối màu nhỏ hoặc một điểm ảnh raster, đây là đơn vị cơ bản nhất giúp tạo thành một bức ảnh có thể hiển thị trên màn hình LED, tương ứng với tọa độ ITS. Pixel được sản xuất dưới mạng lưới 2 chiều nhưng thực tế lại chi tương ứng với tỉ lệ quét của tần số màn hình hay thời gian. Pixel không có số liệu cụ thể để tính toán.
* Một pixel (điểm ảnh) tương ứng với một mảnh (mẩu) của tấm ảnh. Nhiều pixel bao nhiêu, hình ảnh hiển thị sẽ càng rõ và chính xác hơn so với bản gốc bấy nhiêu. Đây cũng là nguyên nhân mà cường độ điểm ảnh sẽ luôn thay đổi. Một hình ảnh hiển thị thông thường sẽ có các máu chủ đạo từ 3 đến 4 màu như xanh lá, đen, xanh lam, vàng, đỏ. Các chương trình có sử dụng đồ họa thường sẽ được tính bằng pixel dimensions- nghĩa là độ phân giải của hình ảnh.

1. **Các hệ không gian màu:**

* Có một số hệ không gian màu phổ biển: RGB, HSV, CMYK, CIE LAB,..

+ RGB: Đây là không gian màu được sử dụng phổ biến hiện nay bởi các thiết bị kĩ thuật số. Được tạo từ 3 màu cơ bản là R(Red), G(Green), B(Blue). Mỗi màu được mã hóa bởi 8 bit ứng với một pixel ảnh sẽ chiến 24 bit( 3byte). Với mô hình này số lượng màu tối đa đạt được là 16.777.216.

+ CMYK: dựa trên cơ sở trộn chất màu của các màu: Cyan(xanh lơ), Magenta( hồng sẫm), Yellow ( vàng), Key( đen). Nó hoạt động dựa trên nguyên lý hấp thụ ánh sáng, màu mà mắt người nhìn thấy là phần ánh sáng không bị hấp thụ. Trong CMYK, hồng sẫm cộng với màu vàng cho màu đỏ, xanh lơ với vàng sinh ra màu xanh lá cây, hồng sẫn với xanh lơ sinh ra màu xanh lam, tổ hợp xanh lơ-hồng sẫm- vàng sinh ra màu đen.

+ HSV: Đây là không gian màu mô tả màu sắc một cách tự nhiên hơn, dựa trên ba thuộc tính: Hue,Saturation và Bright (hoặc Value). Trong đó, Hue là bước sóng của ảnh sáng, Saturation là độ bão hòa và Bright( Value) là cường độ hay độ chói của ánh sáng.

+ CIE LAB: Tất cả các màu có cùng độ sáng sẽ nằm trên cùng của một mặt phẳng dạng hình tròn theo 2 trục a\* và b\_. Màu có giá trị a\_ dương thì ngả đỏ, còn màu có giá trị a\* thì ngủ lục. Tương tự nếu b\* dương thì ngả vàng và b\* âm thì ngả lam. Độ sáng của màu sẽ thay đổi theo trục dọc.

1. **Cách chuyển đổi giữa các hệ màu( Hay dùng)**

* Chuyển từ CMYK sang RGB:

Các giá trị R, G, B được cho trong khoảng 0..255.

Màu đỏ (R) được tính từ màu lục lam (C) và màu đen (K):

R = 255 × (1- C ) × (1- K )

Màu xanh lục (G) được tính từ màu đỏ tươi (M) và màu đen (K):

G = 255 × (1- M ) × (1- K )

Màu xanh lam (B) được tính từ màu vàng (Y) và đen (K):

B = 255 × (1- Y ) × (1- K )

* Chuyển từ RGB sang HSV:

Các giá trị R , G , B được chia cho 255 để thay đổi phạm vi từ 0..255 thành 0..1:

R '= R / 255

G '= G / 255

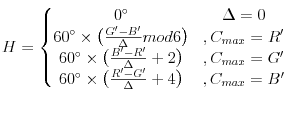
B '= B / 255

Cmax = max ( R ', G ', B ')

Cmin = tối thiểu ( R ', G ', B ')

Δ = Cmax - Cmin

Tính toán H :



Tính toán S :

Text

Description automatically generated with low confidence

Tính toán V:

Text

Description automatically generated with low confidence

* Chuyển từ HSV sang RGB:

Khi 0=< H=< 360, 0=<S=<1 và 0=<V=<1

Text, letter

Description automatically generated

1. **Video**
2. Khái niệm:

* Định nghĩa về video là 1 dạng hình ảnh động tập hợp của rất nhiều ảnh tĩnh xuất hiện liên tục nối tiếp nhau.

1. Chuẩn mã hóa:

* Chuẩn mã hóa (Codec): Video Codec là phần cứng hoặc phần mềm có nhiệm vụ mã hóa và giải mã video kỹ thuật số. Trong ngữ cảnh nén video, codec là bộ ghép nối giữa bộ mã hóa và bộ giải mã. Thiết bị nén video thường được gọi là bộ mã hóa và thiết bị giải nén video là bộ giải mã.
* Tại sao cần video codec: Để thu nhỏ video thành kích thước dễ quản lý, dễ đăng tải và download hơn, các nhà phân phối nội dung sử dụng công nghệ nén video được gọi là codec. Codec cho phép nén một video có dung lượng lớn để phân phối và lưu trữ dễ dàng.

1. **FPS:**

* Frame per second( Số khung hình trong 1 giây): Thường thì 1 giây video có 24 hình tĩnh và từ 24 hình trở lên thì mắt người của chúng ta sẽ có cảm giác là hành động đó xảy ra rất mượt mà. Con số 24 đã trở thành chuẩn mực và là lựa chọn hàng đầu của những người thực hiện video nói chung và video quảng cáo nói riêng.Tại sao là 24 khung hình/giây vì nó đồng nhất được âm thanh và hình ảnh.

1. **Thuật toán**
2. **Một số thuật toán cơ bản về xử lý ảnh**
3. Color Conversion

* Sử dụng một số công thức đã cung cấp ở trên áp đụng để tính toán các giá tri màu.
* Ngoài ra khi sử dụng python thì có thư viện Opencv có cung cấp sẵn 1 số hàm giúp chuyển đổi màu sắc một cách dễ dàng hơn( cv2.cvtColor(src,cv2.BRG2HSV)),…

1. Resize

* Sử dụng để thay đổi kích thước ảnh phù hợp với nhu cầu sử dụng
* OpenCV có cung cấp 1 hàm để có thể thay đổi kích là: cv2.resize(src,dsize())

1. Histogram

* Là dạng biểu đồ cột thống kê số lần xuất hiện của các mức sáng trong ảnh.
* Histogram thường được chuẩn hóa bằng cách lấy mỗi cột chia cho tổng giá trị 🡪 sau khi chuẩn hóa, tổng các giá trị trong histogram sẽ bằng 1.
* Histogram đã chuẩn hóa mang ý nghĩ là xác suất xuất hiện của các mức xám trong ảnh
* Trong opencv có cung cấp 1 số hàm về histogram để tìm và tính toán cv2. calHist(img, channels,mask,histSize,ranges)

1. Denoise

* Ảnh được số vì một số tác nhân nào đó sẽ gây ra nhiễu làm giảm chất lượng của ảnh. Chính vì thế các phương pháp denoise(khữ nhiễu) được ra đời. Nhằm tăng chất lượng ảnh khi xử lí chúng
* OpenCV cung cấp 1 số các hàm denoise: cv2.fastNlMeansDenoisingColored()

1. **Thuật toán Background Subtraction**

* Background subtraction là một phương pháp được sử dụng rộng rãi để phát hiện các đối tượng chuyển động trong một chuỗi khung hình từ máy ảnh tĩnh. Cơ sở của phương pháp này là phát hiện sử khác biệt các đối tượng hiện tại với background mà ta đã cung cấp. Việc trừ hậu cảnh này giúp chúng ta tách được đổi tượng ra khỏi nền nhằm nhận diện một cách dễ dàng hơn.
* **Frame difference**

+ Đây là phương pháp đơn giản nhất của giải thuật trừ nền. Ta sẽ tính hiệu giá trị pixel của khung hình hiện tại và khung hình trước đó nếu hiệu này lớn hơn ngưỡng T thì pixel tại vị trí đó thuộc về Foreground.

+ Ưu điểm : của phương pháp này là tốc độ tính toán nhanh do việc khởi tạo background image chỉ đơn giản là việc lấy khung hình trước đó.

+ Nhược điểm của phương pháp này là nó chỉ tốt với những đối tượng di chuyển liên tục nhưng khi có một đối tượng đứng yên trong khung hình quá lâu thì đối tượng này sẽ bị cho là background cùng với đó phương pháp này phụ thuộc rất nhiều vào ngưỡng T do đó với mỗi video ta đều phải chọn 1 ngưỡng phù hợp.

**Mean Filter**

Thay vì sử dụng một khung hình trước đó để làm background image thì Nean Filter sử dụng N khung hình phai trước khung hình đang xét để khởi tạo background image.

B ( x, y, t ) =

Trong đó B là background image tại thời điểm t, N là số khung hình trước thời điểm t để dùng cho việc tính toán ra background image ,

I (x,y,t ) là khung hình tại thời điểm t

* OpenCV cung cấp 1 số API để thực hiện điều này: cv2.createBackgroundSubtractorMOG2(), cv2.createBackgroundSubtractorKNN()

1. **Các thuật toán Morphological**
2. Erosion

* Đây là một phương pháp làm loại bỏ các pixel nhiễu cô lập, đồng thời loại bỏ các nhiễu xung quanh đối tượng đồng thời làm cho các cạnh của đối tượng mòng hơn và trờ nên mịn hơn.
* Trong opencv có cung cấp 1 hàm là: cv2.erode()

1. Dilation

* Với nhiễu xung quanh đối tượng thì giải thuật này sẽ biến các nhiễu đó thành cạnh của đối tượng. Đồng thời làm cho đối tượng nổi bật hơn
* OpenCV có cung cấp 1 hàm để làm điều này: cv2.dilate()

1. Opening

* Opening sẽ loại bỏ các đối tượng nhỏ khỏi nền trước (thường lấy làm điểm sáng) của hình ảnh và đặt chúng vào điểm sau. Loại bỏ các phần lồi mỏng của ảnh thu được
* Ta sẽ sử dụng Ersion trước sau đó sẽ sử dụng Dilation sau

1. Closing

* Closing thường được sử dụng để làm mịn đường viền và hợp nhất với các điểm đứt đoạn. Việc đóng sẽ loại bỏ các lỗ nhỏ khỏi hình ảnh thu được
* Ta sử dụng Dilation trước sau đó sẽ sử dung Ersion

1. **Edge and Contour**
2. Khái niệm

* Trong ảnh, một cạnh (edge) là một danh giới (boundary) hay một đường viền (contour) mà tại đây xuất hiện 1 sự thay đổi lớn về một vài khía cạnh vật lý của ảnh (cường độ, bề mặt phản chiếu,..)
* Edge detection là 1 kĩ thuật xử lí ảnh được sử dụng để tìm kiếm đường viền bao của các đối tượng trong ảnh

1. Các phương pháp

* Sử dụng mask filter đơn giản
* Sobel Operator

1. **Một số thuật toán tracking**

* Hiện tại trong OpenCV đang cung cấp 8 phương pháp object tracking khác nhau (OOSTING Tracker, MIL Tracker, KCF Tracker, CSRT Tracker, MedianFlow Tracker, TLD Tracker, MOSSE Tracker và GOTURN Tracker)
* Object tracking use Centroid tracking algorithms

step1. Tính toán centroid của các box

step2. Tính khoảng cách của mọi centroid có thể có

step3:Giả sử rằng khoảng cách nhỏ nhất của 2 centroid sẽ cho 1 đối tượng

step4 : Gán id cho đối tượng