**Xử lý ảnh**

**A. Đặc trưng về màu sắc**

Tính năng màu sắc là một đặc trưng rất cơ bản được sử dụng trong xử lý hình ảnh dựa trên sự tương tự màu sắc của hình ảnh.

Ưu điểm là đặc trưng về màu sắc tương đối mạnh mẽ khi xử lý các vấn đề nền và không phụ thuộc vào kích thước hình ảnh, hướng và thay đổi tỷ lệ. Hình ảnh chủ yếu được phân biệt dựa trên tính năng màu sắc của con người.

Nhược điểm là không phản ánh tính không gian và nhạy với phép thay đổi ảnh sáng

Các kỹ thuật khác nhau của kỹ thuật trích xuất tính năng màu được đưa ra như sau:

**1. Color histogram**

Color histogram là một tập hợp các bin, mỗi bin biểu thị xác xuất xuất hiện của một màu trong ảnh, một color histogram được định nghĩa là một vector:

H = {H[0], H[1], H[2],…, H[i], …, H[N]} (3.1)

trong đó i là màu trong biểu đồ màu tương ứng, H[i] là số pixel trong ảnh có màu i, N số bin, thường bằng với số màu của hệ màu của hình ảnh

A screenshot of a cell phone

Description automatically generated

1.1 Color histogram of an image

Để so sánh ảnh có kích thước khác nhau, color histogram có dạng chuẩn hóa có dạng:

H′ = {H′[0], H′[1], H′[2],…, H′[i], …, H′[N]} (3.2)

với *H* ′[*i*] = H[i]/P, P là tổng số pixels trong ảnh

Một color histogram lý tưởng khi các màu riêng biệt không nên được đặt chung một bin, và các màu tương tự phải được đặt chung một bin. Khi sử dụng ít màu sẽ làm giảm khả năng các màu tương tự bị gán cho các bin khác nhau, nhưng tăng khả năng các màu riêng biệt được gán cùng một bin và thông tin hình ảnh sẽ bị giảm theo một mức độ lớn hơn. Mặt khác, color histogram với số lượng bin lớn sẽ chứa nhiều thông tin hơn về nội dung hình ảnh, do đó làm giảm khả năng các màu riêng biệt được gán cùng một bin. Tuy nhiên chúng lại làm tăng khả năng các màu tương tự được gán cho các bin khác nhau, tăng dung lượng data phải lưu trữ và thời gian tích toán. Do đó phải có sự đánh đổi trong việc xác định có bao nhiêu bin nên được sử dụng trong color histogram. Số bin được lựa chọn phổ biến nhất là 64.

**1.1.1. Color histogram discrimination**

Các công thức khoảng cách màu là một thước đo tương tự giữa các hình ảnh dựa trên nhận thức về màu sắc. Ba công thức khoảng cách đã được sử dụng để truy xuất hình ảnh bao gồm histogram Euclidean distance (khoảng cách Euclide), histogram intersection (biểu đồ giao điểm) và histogram quadratic (cross) distance (khoảng cách bậc hai).

**Histogram Euclidean distance**

Giả sử có h, g là 2 color histogram, khoảng cách euclidean được tính bằng công thức:

**A picture containing object, drawing, clock

Description automatically generated**

**Histogram intersection distance**

A screenshot of a cell phone

Description automatically generated

**Histogram quadratic (cross) distance**

Công thức tính khoảng cách chéo:

A picture containing drawing

Description automatically generated

Khoảng cách chéo tính tương quan chéo giữa các bin của 2 histogram. Tập hợp tất cả các giá trị tương quan chéo được biểu dễ bằng ma trận A gọi là ma trận tương tự, phần từ (i,j) của ma trận A được tính bằng công thức:

*aij* =1−*dij* max(*dij*)

với dij là khoảng cách 2 màu

**2. Color moments**

Nếu coi giá trị màu tại mỗi điểm ảnh là biến ngẫu nhiên thì color histogram của ảnh là hàm mật độ phân bố xác suất của biến ngẫu nhiên. Mặt khác có thể biểu diễn một phân bố xác suất của biến ngẫu nhiên với các giá trị:

Kỳ vọng: A picture containing object, clock

Description automatically generated

Độ lệch chuẩn: A picture containing object, clock

Description automatically generated

Moment cấp 3: skewness (cho biết độ lệch về phía nào) A close up of a logo

Description automatically generated

**3. Vector liên kết màu**

Vector liên kết màu là sơ đồ tinh chỉnh của color histogram, phân chia mỗi bin thành các pixel liên kết và không liên kết. Mỗi pixel sẽ thuộc môt một vùng liên thông, dựa vào kích thước vùng liên thông nếu lớn hơn 1% kích thước ảnh thì pixel đó được tính là liên kết và ngược lại, nhỏ hơn sẽ tính là không.

Vector liên kết màu thường được dùng để phân biệt ảnh thiên nhiên với ảnh màu

**B. Đặc trưng về kết cầu ảnh**

Kết cấu hình ảnh còn gọi là vân ảnh chưa có định nghĩa tổng quát, thể hiện sự sắp xếp về mặt không gian của các giá trị độ chói, màu sắc. Kết cấu ảnh được tạo từ các phần từ kết cấu gọi là texel. Các đặc trưng kết cấu khác nhau được nhận ra bởi mắt người như tính đều đặn, tính định hướng, độ mịn, độ thô

Phân tích kết cấu hình ảnh sử dụng trong rất nhiều lĩnh vực và ứng dụng như phân loại kết cấu, tách vật thể đến tổng hợp hình ảnh hoặc nhận dạng mẫu.

Các phương pháp tiếp trích xuất đặc trưng kết câu ảnh: tiếp cận thống kê, tiếp cận cấu trúc, tiếp cận dựa trên biến đổi kết cấu, tiếp cận dựa trên model, tiếp cận dựa trên biểu đồ, học tập, cở sở dữ liệu.. Mỗi phương pháp có những ưu điểm và nhược điểm riêng.

**1. Tiếp cận cấu trúc**

Tiếp cận cấu trúc thương được áp dụng cho phân tích các kết cấu nhân tạo.

Ảnh kết cấu được tạo thành từ các phần tử kết cấu (texels) hay các mẫu (partern). Tiếp cận cấu trúc phân tích tương quan không gian giữa các texels hay các parterns

**2. Tiếp cận thống kê**

Tiếp cận thống kê sử dụng các thuộc tính thống kê của phân bố không gian các điểm ảnh làm mô tả kết cấu ảnh.

**2.1. Ma trận đồng hiện đa mức xám (co-occurrence matrix)**

Định nghĩa: Cho P là hàm xác định tương quan vị trí. Ma trận đồng hiện A: k x k phần tử, trong đó Aij là số lần xuất hiện của các điểm có mức xám Zi cùng với các điểm có mức xám Zj với các điểm đồng hiện này tuân theo hàm tương quan vị trí P.

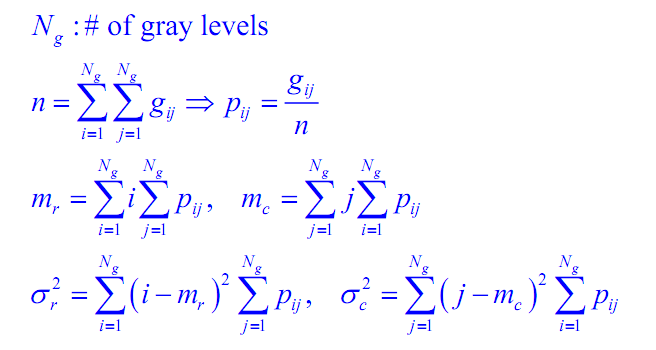
Ma trận đồng hiện cho phân bố của các cắp giá trị đồng xuất hiện tại những giá trị offset xác định

page9image36413152

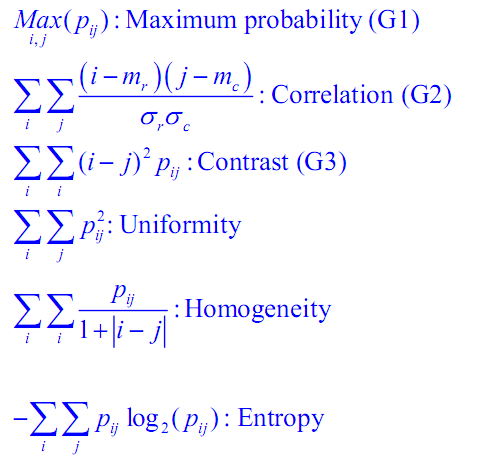
Ma trận đồng hiện cho thấy tương quan giữa các điểm ảnh. Vì ma trận phụ thuộc hàm vị trí P -> P sao cho phù hợp với pattern của texture. Các đặc trưng texture có thể rút ra từ ma trận đồng hiện:

* Giá trị xác suất lớn nhất
* Giá trị độ tương phản
* Tính đồng đều, đồng nhất
* Entropy

Các thông số thống kê của ma trận đồng hiện:



Với Ng là tổng mức xám, n là tổng pij là xác sất xuất hiện pij trong ma trận đồng hình, mr, mc là kỳ vọng theo hàng và cột, s^2 là phương sai theo hàng và cột.

Các đặc trưng có thể rút ra từ ma trận đồng hình: 

**Correlation**

Độ tương quan giữ hàng và cột

**Contrast**

Độ tương phản, là thước đo cường độ của một pixel và lân cận của nó. Trong nhận thức trực quan của thế giới thực, độ tương phản được xác định bởi sự khác biệt về màu sắc và độ sáng của vật thể và các vật thể khác trong cùng một trường nhìn.

**Uniformity**

Tính đồng nhất về kết cấu hình ảnh cả hình ảnh

**Homogeneity**

**Entropy**

Entropy (Ent) là thước đo tính ngẫu nhiên được sử dụng để mô tả kết cấu của hình ảnh đầu vào. Giá trị của nó sẽ là tối đa khi tất cả các yếu tố của ma trận cùng xảy ra là như nhau.

**Energy**

Năng lượng (E) có thể được định nghĩa là thước đo mức độ lặp lại của cặp pixel. Nó đo sự đồng nhất của một hình ảnh. Khi các pixel rất giống nhau, giá trị năng lượng sẽ lớn.

**2.2. Đặc trưng LBP (Local binary pattern)**

Là một trong những đặc trưng rất mạnh đề xuất bởi Ojala năm 1994, được sử dụng nhiều và rất tốt với bài toán phát hiện người (có người, không có người trong ảnh).

Đặc trưng LBP được trích xuất bằng cách so sánh giá trị mỗi pixel với các điểm lân cận của nó (lân cận 8, lân cận 25 ..) nếu giá trị tại điểm so sánh lơn hơn giá trị tại tâm thì gán nhãn 1, ngược lại gán nhãn 0, từ đó, mỗi pixel sẽ có một partern đại diện

Khi số parttern lớn, mà lại có nhiều parttern rất ít xuất hiện, nên các parttern này được chia làm uniform hoặc non-uniform parttern. Một parttern gọi là uniform nếu nó chưa nhiều nhất 2 lần đảo bit từ 0 sang 1 hoặc từ 1 sang 0 (thứ tự các bit trong parttern duyệt theo vòng tròn).

Sau đó mỗi uniform parttern được gán một nhãn. Tất cả các non-uniform được gán chung 1 nhãn.

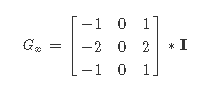
Ví dụ: Mã hóa (8, 1) lân cận sẽ có 2^8 = 256 parttern. Xây dựng histogram có vector đặc trưng có 256 chiều, trong đó có 58 uniform nên suy ra số chiều của LBP feature là 59.

**C. Đặc trưng HOG - Histograms of Oriented Gradients**

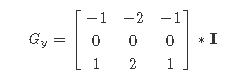
Phương pháp rút trích đặc trưng hình ảnh HOG xuất bản ở hội nghị CVPR 2005 được đề xuất bởi tác giả là Dalal và Triggs. Bài báo gốc HOG đề xuất phương pháp rút trích đặc trưng sử dụng các thống kê histogram về hướng trên ảnh gradient cho bài toán phát hiện người (human detection). CVPR là một trong những hội nghị thuộc hàng đỉnh của lĩnh vực thị giác máy tính. Do đó, bài báo HOG này xuất hiện ở đó quả thật là một điều gì đó không phải là ngẫu nhiên. Với ý tưởng cơ bản là hình dạng đối tượng cục bộ thường có thể được đặc trưng khá tốt bởi phân phối của gradient và hướng cạnh, thậm trí không có kiến thức chính xác về vị trí gradient hoặc cạnh tương ứng. Trong thực tế, điều này được thực hiện bằng cách chia cửa sổ hình ảnh thành các vùng không gian nhỏ (“cells”), vùng không gian lớn hơn (“blocks”). Hog được tính toán trên một lưới dày đặc các cell và chuẩn hóa sự tương phản giữa các block để nâng cao độ chính xác. Hog được sử dụng chủ yếu để mô tả hình dạng và sự xuất hiện của một object trong ảnh. Các bước tính toán HOG:

1. **Tình toán gradients**

* Sử dụng phương pháp phổ biến nhất là áp dụng một mặt nạ đạo hàm rời rạc theo 2 chiều dọc và ngang. Cụ thể, phương pháp sẽ lọc ma trận cường độ ảnh với các bộ lọc như Sobel mask hoặc Scharr. Để tính bộ lọc sobel, phép tích chập của kernel kích thước 3x3 được thực hiện với hình ảnh ban đầu. Nếu chúng ta kí hiệu I là ma trận ảnh gốc và Gx,Gy là 2 ma trận ảnh mà mỗi điểm trên nó lần lượt là đạo hàm theo trục x trục y. Chúng ta có thể tính toán được kernel như sau:
* Đạo hàm theo chiều ngang:



* Đạo hàm theo chiều dọc:



* Gía trị độ lớn gradient (gradient magnitude) và phương gradient (gradient direction) có thể được tạo ra từ 2 đạo hàm Gx và Gy theo công thức bên dưới:
* Độ lớn gradient:

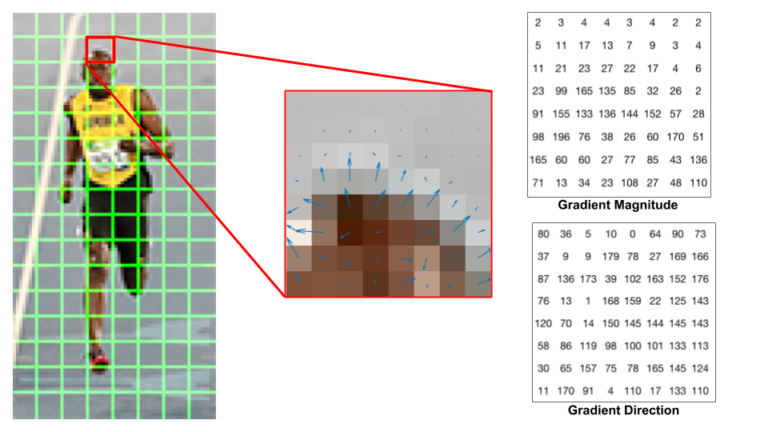


* Phương gradient:

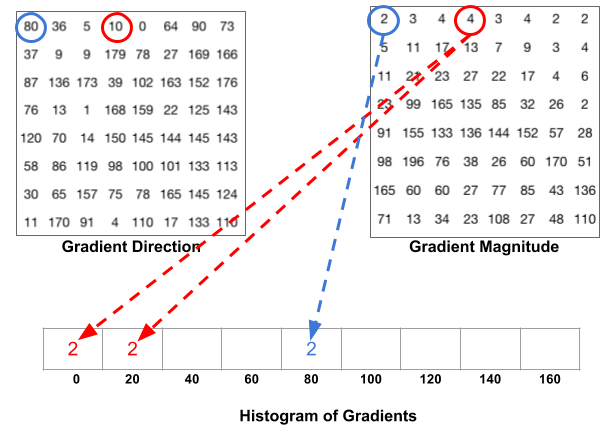


1. **Tình HOG**

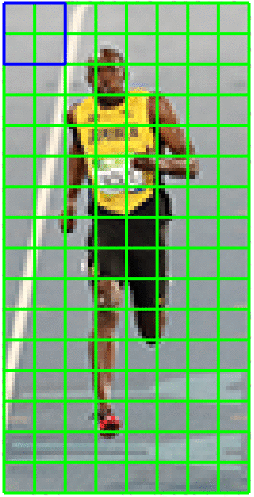
* Ta nhận thấy đặc trưng của mỗi bức ảnh được biểu diễn thông qua 2 thông số đó là mức độ thay đổi cường độ màu sắc (ma trận gradient magnitude) và hướng thay đổi cường độ màu sắc (ma trận gradient direction). Do đó chúng ta cần tạo ra được một bộ mô tả (feature descriptor) sao cho biến đổi bức ảnh thành một véc tơ mà thể hiện được cả 2 thông tin này.
* Hình ảnh được chia thành một lưới ô vuông mà mỗi một ô có kích thước 8x8 pixels. Như vậy chúng ta có tổng cộng 64 ô pixels tương ứng với mỗi ô. Trên mỗi một ô trong 64 pixels ta sẽ cần tính ra 2 tham số đó là độ lớn gradient (gradient magnitute) và phương gradient (gradient direction). Như vậy tổng cộng 8x8x2 = 128 giá trị cần tính bao gồm 64 giá trị gradient magnitute và 64 giá trị gradient direction như ma trận hình bên dưới:



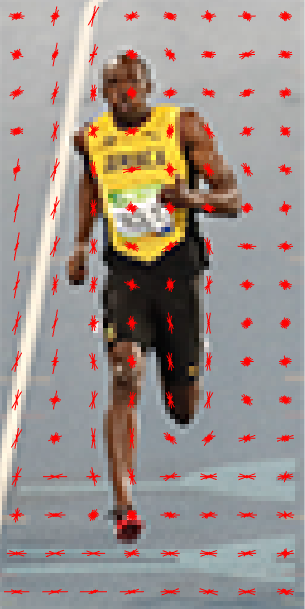
* Vector histogram sẽ được tạo ra như sau:
* Bước 1: Mapping độ lớn gradient vào các bins tương ứng của phương gradient
  + Các giá trị góc trong cần được chuẩn hoá về (0 – 180)
  + Sắp xếp các giá trị phương gradient theo thứ tự từ nhỏ đến lớn và chia chúng vào 9 bins. Độ lớn của phương gradient sẽ nằm trong khoảng [0, 180] nên mỗi bins sẽ có độ dài là 20 như hình bên dưới.



* Bước 2: Chuẩn hoá vector histogram theo block 16 \* 16
  + Chúng ta thấy rằng véc tơ histogram sẽ bị phụ thuộc vào cường độ các pixels của một bức ảnh. Với 2 bức ảnh có cùng nội dung nhưng bức ảnh biến thể tối hơn được tạo thành từ ma trận ảnh gốc nhân 1/2. Khi đó giá trị vector histogram của ảnh gốc cũng sẽ gấp đôi véc tơ histogram của ảnh biến thể. Chính vì thế cần chuẩn hóa vector histogram để cả 2 bức ảnh có cùng một véc tơ biểu diễn.
  + Một số dạng chuẩn hoá:
    - L1-norm: sau chuẩn hóa, tổng các giá trị của những phần tử trong vector bằng 1.
    - L2-norm: sau chuẩn hóa, độ dài vector bằng 1.
  + Qúa trình chuẩn hóa sẽ thực hiện trên một block kích thước 2x2 trên lưới ô vuông ban đầu (mỗi ô kích thước 8x8 pixel). Như vậy chúng ta sẽ có 4 véc tơ histogram kích thước 1x9, concatenate các véc tơ sẽ thu được véc tơ histogram tổng hợp kích thước là 1x36 và sau đó chuẩn hóa theo norm chuẩn bậc 2 trên véc tơ này. Việc di chuyển các window thực hiện tương tự như phép tích chập 2 chiều trong mạng CNN với step\_size = 8 pixels như hình ảnh bên dưới:



* Bước 3: Tính toán HOG feature vector
  + Sau khi chuẩn hóa các véc tơ histogram, chúng ta sẽ concatenate các véc tơ 1x36 này thành một véc tơ lớn. Đây chính là véc tơ HOG đại diện cho toàn bộ hình ảnh.
  + Biểu diễn phân phối HOG trên ảnh:
    - Đối với mỗi một ô trên lưới ô vuông, chúng ta biểu diễn phân phối HOG bao gồm nhóm 9 véc tơ chung gốc chiều dài bằng độ lớn gradient và góc bằng phương gradient. Khi đó chiều của nhóm các véc tơ sẽ tương đối giống với dáng của vận động viên trong ảnh, đặc biệt là tại các vị trí chân và tay. Cụ thể hãy xem hình bên dưới:



……………………..