**TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP. HCM**

A blue and white logo with a book and a square cap

Description automatically generated**KHOA CÔNG NGHỆ THÔNG TIN**

**BÁO CÁO MÔN: XỬ LÍ ẢNH**

**ĐỀ TÀI: RÚT TRÍCH ĐẶC TRƯNG ẢNH**

***THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024***

**TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP. HCM**

A blue and white logo with a book and a square cap

Description automatically generated**KHOA CÔNG NGHỆ THÔNG TIN**

**BÁO CÁO MÔN: XỬ LÍ ẢNH**

**ĐỀ TÀI: RÚT TRÍCH ĐẶC TRƯNG ẢNH**

**Lớp danh nghĩa: 12DHTH14**

**TKB chính thức: Thứ 2 – Tiết 4\_6**

**GVHD: Trần Đình Toàn**

**Sinh viên thực hiện:**

**Phạm Hồ Đăng Huy**

***THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MSSV** | **Họ Tên** | **Nhiệm Vụ Hàng Tuần** | **Kết Quả** | **GV đánh giá** |
| 2001215828 | Phạm Hồ Đăng Huy | Tìm hiểu và cài đặt | Hoàn Thành |  |

LỜI CẢM ƠN

Lời nói đầu tiên cho phép em – Phạm Hồ Đăng Huy cùng các cộng sự gửi lời cảm ơn đến thầy Trần Đình Toàn, hiện đang là giảng viên - người hướng dẫn chúng em để hoàn thiện đề tài báo cáo này.

Đề án này được hoàn thành dưới sự hướng dẫn một cách chi tiết và khoa học của thầy Trần Đình Toàn. Nhóm chúng em xin được bày tỏ lòng biết ơn sâu sắc đến thầy Trần Đình Toàn – người đã tận tình giúp nhóm chúng tôi nâng cao kiến thức và tác phong làm việc bằng tất cả sự mẫu mực của một người giảng viên và tinh thần trách nhiệm của người làm khoa học. Đồng thời, thầy cũng là người đã hỗ trợ nhóm chúng tôi và cho chúng tôi nhiều ý kiến đóng góp quý giá để giúp nhóm có thể hoàn thành luận án một cách hoàn thiện nhất.

Kế đến, nhóm chúng em cũng xin cảm ơn quý thầy cô giáo trường Đại Học Công Thương TP. Hồ Chí Minh đã giảng dạy và truyền đạt những kiến thức, kinh nghiệm và kỹ năng nền tảng cốt lõi vô cùng bổ ích để cho nhóm em làm đề án hiệu quả.

Hơn thế nữa, nhóm chúng em xin gửi lời cảm ơn chân thành đến Ban giám hiệu trường Đại Học Công Thương TP. Hồ Chí Minh đã giúp đỡ và tạo mọi điều kiện thuận lợi cho nhóm có việc học tập và nghiên cứu của nhóm.

Nhóm em xin chúc thầy luôn luôn mạnh khỏe và hạnh phúc!

Nhóm Trưởng

(Ký tên)

Phạm Hồ Đăng Huy

MỤC LỤC HÌNH ẢNH

[Hình 1. Phân loại Gradient 4](#_Toc162821651)

[Hình 2. Jump or Steal Discontinuity 5](#_Toc162821652)

[Hình 3. LoG và DoG 9](#_Toc162821653)

[Hình 4. Code mình họa rút trích màu dựa trên thuật toán K-Means 16](#_Toc162821654)

[Hình 5. Ảnh gốc và ảnh kết quả rút trích màu 17](#_Toc162821655)

[Hình 6. R-HOG (trên) và C HOG (dưới) 17](#_Toc162821656)

MỤC LỤC

[LỜI CẢM ƠN ii](#_Toc162821620)

[MỤC LỤC HÌNH ẢNH iii](#_Toc162821621)

[MỤC LỤC iv](#_Toc162821622)

[CHƯƠNG I. GIỚI THIỆU 1](#_Toc162821623)

[1. Lý do chọn đề tài 1](#_Toc162821624)

[2. Mục tiêu của báo cáo 1](#_Toc162821625)

[CHƯƠNG II. RÚT TRÍCH ĐẶC TRƯNG ẢNH 2](#_Toc162821626)

[1. Các yêu cầu của rút trích đặc trưng 2](#_Toc162821627)

[1.1. Trích chọn đặc trưng là gì? 2](#_Toc162821628)

[1.2. Đặc trưng tổng quát: 2](#_Toc162821629)

[1.2.1. Đặc điểm màu sắc: 2](#_Toc162821630)

[1.2.2. Đặc điểm kết cấu ảnh: 2](#_Toc162821631)

[1.2.3. Đặc điểm hình dạng: 2](#_Toc162821632)

[1.3. Đặc điểm trừu tượng: 2](#_Toc162821633)

[1.3.1. Đặc trưng các mức pixel: 3](#_Toc162821634)

[1.3.2. Đặc trưng cục bộ: 3](#_Toc162821635)

[1.3.3. Đặc trưng toàn cục: 3](#_Toc162821636)

[1.3.4. Đặc trưng đặc thù miền: 3](#_Toc162821637)

[2. Kỹ thuật rút trích đặc trưng hình ảnh 3](#_Toc162821638)

[2.1. Phát Hiện Biên và Phân Đoạn Ảnh 3](#_Toc162821639)

[2.1.1. Phát hiện biên 3](#_Toc162821640)

[2.1.2. Định nghĩa biên 4](#_Toc162821641)

[2.1.3. Mô tả toán học 5](#_Toc162821642)

[2.1.4. Các bước phát hiện biên 6](#_Toc162821643)

[2.1.5. Các phương pháp phát hiện biên 7](#_Toc162821644)

[2.2. Phân đoạn – Segmentation: 13](#_Toc162821645)

[2.2.1. Phân loại 14](#_Toc162821646)

[2.3. Rút trích màu trên ảnh 16](#_Toc162821647)

[2.4. Rút trích đặc trưng ảnh bằng HOG (Histogram of Orented Gradients) 17](#_Toc162821648)

[CHƯƠNG III. KẾT LUẬN 19](#_Toc162821649)

[TÀI LIỆU THAM KHẢO 20](#_Toc162821650)

1. GIỚI THIỆU

Trong thế giới ngày nay, rút trích đặc trưng ảnh là một phần quan trọng của xử lý ảnh. Quá trình này nhằm chuyển đổi dữ liệu hình ảnh từ không gian chiều cao sang không gian chiều thấp hơn, giữ lại những thông tin quan trọng và loại bỏ thông tin không cần thiết, từ đó tạo ra một biểu diễn đơn giản nhưng vẫn giữ được những đặc trưng quan trọng của hình ảnh.

* 1. Lý do chọn đề tài

Trong lĩnh vực xử lý ảnh, rút trích đặc trưng ảnh đóng vai trò quan trọng nhằm giúp cho việc phân tích và xử lý hình ảnh trở nên hiệu quả hơn. Hiểu biết sâu về các phương pháp và kỹ thuật rút trích đặc trưng ảnh không chỉ giúp cải thiện hiệu suất của các ứng dụng xử lý ảnh mà còn mở ra nhiều cơ hội mới trong nghiên cứu và phát triển công nghệ.

* 1. Mục tiêu của báo cáo

Mục tiêu của báo cáo này là cung cấp một cái nhìn tổng quan về rút trích đặc trưng ảnh, bao gồm các phương pháp, kỹ thuật và ứng dụng của chúng. Chúng tôi sẽ giải thích cách các phương pháp rút trích đặc trưng ảnh được áp dụng và biết cách chúng đóng vai trò quan trọng trong thực tiễn. Bằng cách này, tinh thần mà bài báo cáo này mang lại hy vọng sẽ giúp độc giả hiểu rõ hơn về tầm quan trọng của rút trích đặc trưng ảnh và ứng dụng của nó trong các lĩnh vực khác nhau.

1. RÚT TRÍCH ĐẶC TRƯNG ẢNH
   1. Các yêu cầu của rút trích đặc trưng

Trong lĩnh vực xử lý ảnh, rút trích đặc trưng là quá trình tìm ra các điểm đặc trưng của hình ảnh để hiểu và nhận dạng các đối tượng trong ảnh. Dưới đây là một số yêu cầu cũng như khái niệm cơ bản liên quan đến việc rút trích đặc trưng:

* + 1. Trích chọn đặc trưng là gì?

Trích chọn đặc trưng là quá trình tìm ra điểm đặc trưng của đối tượng so với đối tượng khác, tùy thuộc vào mục đích nhận dạng trong xử lý ảnh.

Ví dụ: Để phân biệt ảnh của cà chua chín, chúng ta có thể dựa vào màu sắc.

* + 1. Đặc trưng tổng quát:
       1. Đặc điểm màu sắc:

Biểu đồ màu (color histogram): Biểu diễn phân phối màu sắc trong ảnh.

Moments màu (color moments): Tính toán các giá trị thống kê liên quan đến màu sắc, chẳng hạn như trung bình, phương sai và độ lệch chuẩn.

* + - 1. Đặc điểm kết cấu ảnh:

Texture features: Liên quan đến cấu trúc và kết cấu của ảnh, ví dụ như đặc điểm của các vùng kết cấu như da, vải, gỗ, v.v.

HOG (Histogram of Oriented Gradients): Một phương pháp phổ biến để trích chọn đặc trưng từ hình ảnh, dựa vào phân phối hướng của gradient.

* + - 1. Đặc điểm hình dạng:

Biên và đường biên: Đặc trưng cho đường biên của đối tượng, rất hữu ích trong việc nhận dạng.

Biến đổi hình dạng: Đặc điểm này được trích chọn bằng việc thực hiện lọc vùng (zonal filtering), ví dụ như sử dụng các mặt nạ đặc điểm (feature mask) với hình dạng khác nhau.

* + 1. Đặc điểm trừu tượng:
       1. Đặc trưng các mức pixel:

Có thể tính toán tại mỗi điểm ảnh, ví dụ như màu sắc, vị trí, và các đạo hàm đầu tiên và thứ hai của giá trị mức xám tại mỗi điểm ảnh.

* + - 1. Đặc trưng cục bộ:

Có thể tính toán trên kết quả của các thuật toán phân đoạn hình ảnh và phát hiện biên, nghĩa là chúng dựa trên một phần của hình ảnh có các đặc tính đặc biệt. Hình dạng đối tượng là một ví dụ điển hình.

* + - 1. Đặc trưng toàn cục:

Nên được tính toán trên toàn bộ hình ảnh hoặc chỉ trên một khu vực cố định của hình ảnh. Thường thì chúng thực sự là các thuộc tính thống kê của một hình ảnh, ví dụ như biểu đồ tần số, trung bình, phương sai, và moment.

* + - 1. Đặc trưng đặc thù miền:

Các đặc điểm trong loại này phụ thuộc vào ứng dụng cụ thể, ví dụ như khuôn mặt người, dấu vân tay, và các đặc điểm khái niệm, là kết hợp của các đặc điểm cấp thấp cho một mục đích miền cụ thể.

🡺 Việc hiệu quả trong việc rút trích đặc trưng giúp cải thiện độ chính xác của hệ thống nhận dạng ảnh và giảm tải tính toán.

* 1. Kỹ thuật rút trích đặc trưng hình ảnh
     1. Phát Hiện Biên và Phân Đoạn Ảnh

Đối với việc phát hiện đặc trưng ở mức cao, phân đoạn hình ảnh là cơ sở quan trọng. Đối với việc phân đoạn hình ảnh chính xác, phát hiện biên luôn đóng một vai trò rất quan trọng.

* + - 1. Phát hiện biên

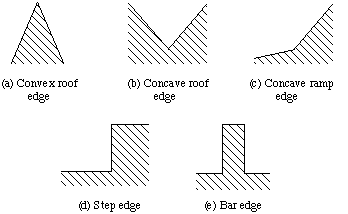
Đầu ra của quá trình phát hiện biên là một hình ảnh biên hoặc biểu đồ biên, trong đó giá trị của mỗi pixel phản ánh mức độ mạnh mẽ của pixel tương ứng trong hình ảnh gốc đáp ứng với yêu cầu là một pixel biên.

Các biên được sử dụng nhiều do tính đơn giản của việc tính toàn, bởi vì sự khác biệt trong giá trị pixel giữa các vùng là dễ tính toán bằng cách xem xét các độ dốc. Quan trọng hơn, chúng cung cấp các gợi ý hình ảnh mạnh mẽ có thể giúp quá trình nhận dạng.

* + - 1. Định nghĩa biên

Một biên có thể được coi là biên giữa hai vùng khác biệt trong một hình ảnh, có thể là các bề mặt khác nhau của đối tượng, hoặc có thể là biên giữa ánh sáng và bóng đổ trên một bề mặt duy nhất. Hầu hết các phương pháp phát hiện biên hoạt động dựa trên giả định rằng một biên xuất hiện khi có một sự không liên tục trong hàm độ sáng (hoặc độ sâu) hoặc một độ dốc hàm độ sáng (hoặc độ sâu) rất cao và thay đổi đột ngột trong hình ảnh.

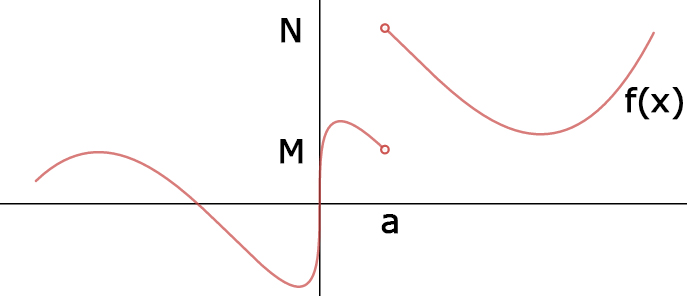
Sự không liên tục có thể được phân loại như sau:



Hình 1. Phân loại Gradient

* Gradient (A Gradient Discontinuity) - khi độ dốc (Gradient) của các giá trị pixel thay đổi qua một đường. Loại không liên tục này có thể được phân loại là:
  + Biên mái – Roof edges.
  + Biên dốc – Ramp edges.
  + Biên lồi – Convex edges.
  + Biên lõm – Cancave edges.

bằng cách ghi nhận dấu của thành phần độ dốc vuông góc với biên hai bên của biên. Các biên dốc có cùng dấu trong các thành phần độ dốc hai bên của không liên tục, trong khi các biên mái có dấu trái ngược trong các thành phần độ dốc.

* Bước hoặc nhảy (A Jump or Step Discontinuity) - khi các giá trị pixel thay đổi đột ngột qua một đường nào đó.

Hình 2. Jump or Steal Discontinuity

* Thanh (A Bar Discontinuity) - khi các giá trị pixel tăng nhanh sau đó lại giảm đi (hoặc ngược lại) qua một đường nào đó. Các biên có thể được mô tả bằng ba thông số cơ bản, bao gồm hướng, vị trí và cường độ.
  + - 1. Mô tả toán học

Trong hầu hết các hệ thống, một biên được mô hình hóa bằng đạo hàm. Theo mặt toán học, cho một hàm hình ảnh , hàm độ lớn của độ dốc và hàm có hướng của độ dốc có thể được xây dựng như sau:

Trong đó

Trong thực tế, thường được xấp xỉ bằng để đơn giản hóa tính toán. Lưu ý rằng ở đây chúng ta giả định rằng là liên tục, nhưng thực tế, thường được số hóa (hình ảnh kỹ thuật số) trong không gian. Sau đó, chúng ta phải tìm một phương pháp thích hợp để xấp xỉ và , trong đó đã có một số biến thể:

Thường thì, chúng ta để , từ đó thu được:

* + - 1. Các bước phát hiện biên

Các bước trong kỹ thuật trích biên có thể được chia thành ba bước phát hiện riêng biệt:

* Bước 1: Làm mờ hình ảnh. Bước này thường được sử dụng để loại bỏ nhiễu. Phương pháp thuận tiện và hiệu quả nhất là làm mờ Gaussian, được thực hiện bằng cách lồng ghép hình ảnh với một toán tử Gaussian. Hàm phân phối Gaussian hai chiều được định nghĩa như sau:

Nếu chúng ta viết lại công thức này trong hệ tọa độ, chúng ta có:

Hình dạng của phân phối và do đó lượng làm mờ có thể được điều khiển bằng cách thay đổi σ.

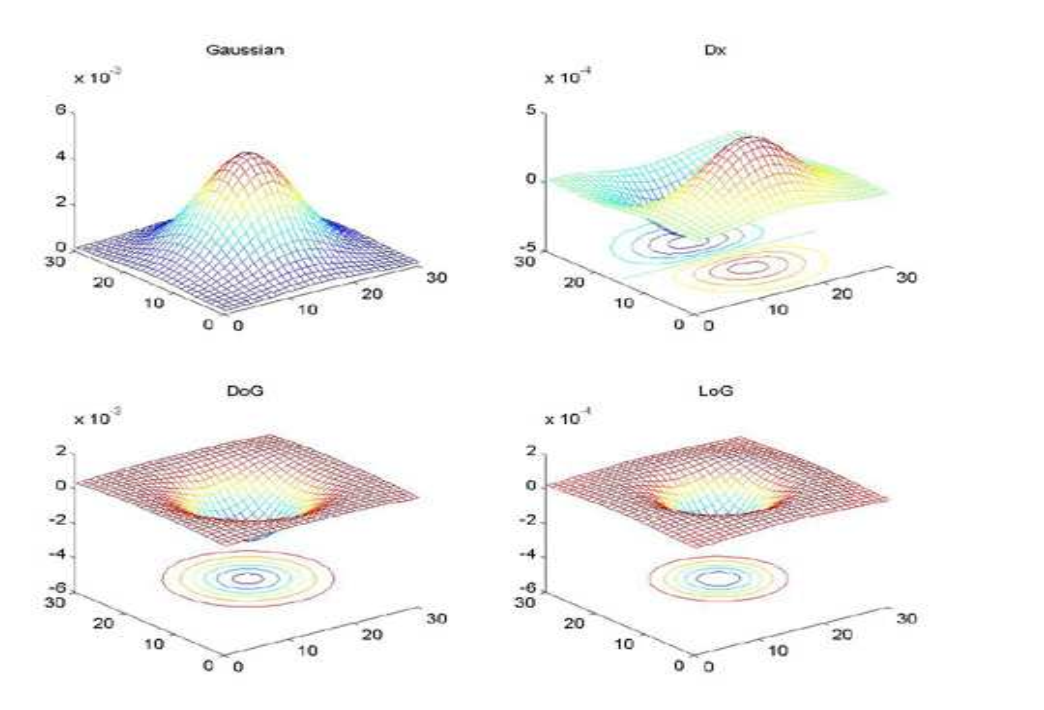
* Bước 2: Tìm các pixel trong hình ảnh nơi các biên có khả năng xuất hiện bằng cách tìm sự không liên tục trong các độ dốc. Các điểm ứng cử cho biên trong hình ảnh thường được gọi là điểm biên, pixel biên hoặc edgels.
* Bước 3: Liên kết các điểm biên này một cách nào đó để tạo ra mô tả của các biên dưới dạng đường, đường cong, vv.
  + - 1. Các phương pháp phát hiện biên

Tất cả các thuật toán phát hiện biên đều cố gắng định vị các biên trong một hình ảnh được biểu diễn bởi các yếu tố thực sự (thật) của cảnh nhưng không phải là nhiễu. Chúng có thể được phân loại vào bốn loại (ở đây chúng tôi xem xét thuật toán phát hiện biên Canny như một lớp có thể tách biệt do đóng góp đáng kể của nó vào việc phát hiện biên)

* + - * 1. Đạo hàm bậc nhất
* Ý tưởng cơ bản là xác định vị trí nơi đạo hàm bậc nhất là cực đại cục bộ như edgels.
* Một bộ lọc đối xứng lẻ sẽ xấp xỉ một đạo hàm bậc nhất, và các điểm cao điểm trong đầu ra lồng ghép sẽ tương ứng với các biên trong hình ảnh. Thường, đạo hàm bậc nhất của hình ảnh số được biểu thị dưới dạng một phép lồng ghép của hình ảnh số với một mặt nạ lồng ghép, cũng luôn được gọi là toán tử biên, sau đó các đầu ra kết quả được xử lý để tạo ra một bản đồ độ dốc. Độ lớn của bản đồ độ dốc được tính toán và phục vụ làm đầu vào cho một quá trình ức chế không cực đại. Cuối cùng, bản đồ kết quả của các cực đại cục bộ được ngưỡng để tạo ra bản đồ biên. Các toán tử biên thường được thiết kế để bao gồm một số lẻ pixel, ví dụ như kích thước mặt nạ 3\*3 hoặc lớn hơn. Một ưu điểm của việc sử dụng kích thước mặt nạ lớn là các lỗi do tác động của nhiễu được giảm bằng cách trung bình cục bộ trong hàng xóm của mặt nạ. Một ưu điểm của việc sử dụng mặt nạ kích thước lẻ là các toán tử được căn giữa và do đó có thể cung cấp một ước lượng mà có độ nghiêng về một pixel trung tâm (x, y).
  + Toán tử đơn giản
  + Toán tử Robert’s Cross: Sử dụng các hướng chéo để tính toán vector gradient.
  + Toán tử Prewitt: Phát hiện các biên dọc và ngang.
  + Toán tử Sobel: Dễ triển khai, đặc biệt là thông qua phương pháp pipeline
  + Toán tử Kirsch: Xét hướng cạnh biên một cách chính xác hơn, toán tử này sử dụng 8 mặt nạ.
* Ưu điểm:
  + Tưởng đối đơn giản.
* Nhược điểm:
  + Hướng.
  + Nhiễu, nhiều điểm cực đại cục bộ nhỏ sẽ được tạo ra bởi nhiễu.
  + Chỉ tìm được các biên giống với bộ lọc.
  + Cần xử lí phân mảnh.
  + Cần chọn ngưỡng thích hợp.
    - * 1. Đạo hàm bậc hai
* Ý tưởng cơ bản là tìm các pixel trong một hình ảnh mà các điểm giao nhau với giá trị bằng không được xem xét là các đoạn cạnh. Điều này dựa trên quan sát rằng một cực đại của đạo hàm đầu tiên xảy ra tại một điểm giao nhau của đạo hàm bậc hai.
* Một bộ lọc đối xứng chẵn sẽ xấp xỉ một toán tử đạo hàm bậc hai. Các điểm giao nhau trong đầu ra của việc tích chập với một bộ lọc đối xứng chẵn sẽ tương ứng với các cạnh, cực đại trong đầu ra của toán tử này sẽ tương ứng với sự gián đoạn tiếp tuyến, thường được gọi là thanh, hoặc đường thẳng.
  + Bộ lọc Laplacian: Tuyến tính và xoay đối xứng

Nó thường được số hóa bằng cách sử dụng mặt nạ tính chập

* + LoG (Laplacian of Gaussian) (Toán tử Marr – Hildreth): Trước khi sử dụng toán tử Laplacian, chúng ta làm mịn hình ảnh thông qua bộ lọc Gaussian.

Ở đây, để đơn giản hóa tính toán, chúng ta có thể xấp xí bằng cách sử dụng toán tử DoG (Difference of Gaussians) và toán tử Mexican Hat. DoG được xây dựng bằng cách lấy sự khác biệt của hai Gaussian có độ lệch chuẩn khác nhau. Thông qua các thực nghiệm, sẽ tìm thấy rằng tỷ lệ của các độ lệch chuẩn là 1:6 sẽ cho két quả xấp xỉ gần .

Hình 3. LoG và DoG

* Ưu điểm:
  + Thích hợp trong các lĩnh vực nhận thức tri giác của các loài động vật được phát hiện giống như LoG. Vì vậy, LoG có thể mô phỏng tầm nhìn của ác loài động rất tốt.
  + LoG là đối xứng. Do đó, các cạnh có thể được phát hiện ở tất cả các hướng.
  + Phát các điểm giao nhau có thể dễ dàng hơn chỉ bằng các tìm sự thay đổi dấu.
  + Đa phần các đường viền thu được có thể khác biệt so với các đường viền đúng được quan sát.
* Nhược điểm
  + Khó khăn với nhiễu, do tính chất của đạo hàm bậc hai.
  + Sẽ đánh dấu một số vị trí không phải là cạnh.
    - * 1. Canny Edge Detector

Các phiên bản Can 83 và Can 86 coi việc phát hiện cạnh như một vấn đề xử lý tín hiệu và cố gắng thiết kế một bộ phát hiện cạnh "tối ưu".

1. Mục Tiêu Tối Ưu

* Phát hiện tốt: Đánh dấu tất cả các cạnh thực sự.
* Vị trí tốt: Các cạnh phát hiện được phải gần nhất có thể với các cạnh thực sự.
* Phản hồi đơn giản: Đối với mỗi cạnh thực sự, chỉ có một phản hồi, nghĩa là các điểm không phải là cạnh không được phát hiện là cạnh.

1. Ràng buộc tối ưu:

Để đạt được các mục tiêu tối ưu, các ràng buộc tối ưu hóa toán học được định rõ, tương ứng là:

* Tối đa hóa Tỷ số Tín hiệu - Độ nhiễu (SNR).
* Tối thiểu hóa khoảng cách RMS của cạnh phát hiện được so với trung tâm của cạnh thực sự.
* Tối thiểu hóa khoảng cách trung bình giữa các cực đại cục bộ gây ra bởi nhiễu.

1. Thuật toán

* Làm mịn ảnh

Dữ liệu ảnh được làm mượt bằng hàm Gaussian hai chiều với độ rộng được xác định bởi một tham số người dùng. Trong thực tế, việc tích chập hai chiều với Gaussians lớn mất nhiều thời gian, nên trong thực tế, thường dùng phương pháp gần đúng này bằng cách sử dụng hai Gaussian một chiều, một căn với trục x và một căn với trục y.

* Tính toán đạo hàm

Tính toán độ dốc của bề mặt mượt của hàm hình ảnh được tích chập trong bất kỳ hướng nào từ độ dốc đã biết trong hai hướng. Để đơn giản, các giá trị trong mảng ảnh đã mượt theo hướng x được tích chập với đạo hàm bậc nhất của một Gaussian một chiều có σ giống nhau căn với y. Tương tự, các giá trị trong mảng ảnh đã mượt theo hướng y được tích chập với đạo hàm bậc nhất của một Gaussian một chiều có σ giống nhau căn với x. Sau đó, độ lớn và góc của độ dốc có thể được tính từ cạnh huyền và arctangent.

* Thoả mãn điểm cực đại

Mỗi pixel lần lượt tạo thành trung tâm của một khu vực lân cận chín pixel. Bằng cách nội suy các giá trị lưới rời rạc xung quanh, độ lớn độ dốc được tính toán tại biên của khu vực lân cận theo cả hai hướng vuông góc với pixel trung tâm. Nếu pixel đang xem xét không lớn hơn hai giá trị này (nghĩa là không tối đa), nó sẽ bị loại bỏ.

* Ngưỡng (Độ Nhoè)

Giới hạn giá trị cạnh trên và dưới được đặt. Nếu một giá trị nằm trên giới hạn ngưỡng trên thì nó sẽ được chấp nhận ngay lập tức. Nếu giá trị nằm dưới ngưỡng thấp nó sẽ bị từ chối ngay lập tức. Các điểm nằm giữa hai giới hạn sẽ được chấp nhận nếu chúng được kết nối với các pixel có tương tác mạnh.

* Tổng hợp (tổng hợp đặc điểm)

Nguyên tắc chung là bắt đầu từ phạm vi nhỏ nhất, sau đó tổng hợp các đầu ra phạm vi lớn hơn. Dữ liệu tổng hợp được trích xuất bằng cách tích chập một Gaussian có phạm vi lớn hơn với dữ liệu cạnh phạm vi nhỏ hơn. Quá trình được lặp lại để đánh dấu các cạnh ở phạm vi thứ hai không có mặt ở phạm vi đầu tiên, sau đó phạm vi thứ ba đến quy mô thứ hai và cứ tiếp tục. Canny quan sát "trong nhiều trường hợp, phần lớn các cạnh được nhận dạng bởi kênh màu nhỏ, và các kênh màu sau đó chủ yếu đánh dấu các cạnh bóng và cạnh đổ bóng, hoặc cạnh giữa các khu vực có kết cấu khác nhau ở Can 86. (Thường thì, bước này không được thực hiện như một phần của thuật toán phát hiện cạnh Canny.)

* + - * 1. Các mô hình phù hợp (Model Fitting)

Ý tưởng cơ bản là phát triển một mô hình biên và sau đó tính toán mức độ phù hợp của nó với hình ảnh. Một biên được mô hình bằng cách chỉ định bốn bậc tự do của nó: vị trí, hướng, và độ sáng không đổi ở hai bên của bước. Mô hình được đặt ở vị trí để phù hợp với dữ liệu bằng cách đơn giản là tìm kiếm sự phù hợp lỗi bình phương nhỏ nhất của mô hình tham số hóa với cửa sổ hình ảnh, một biên được khai báo có mặt khi lỗi bình phương nhỏ nhất được trọng số là nhỏ hơn một ngưỡng trước đặt.

[Chi 97] và [Chi 98] lập luận rằng tất cả các thuật toán phát hiện biên truyền thống đều gặp ba vấn đề chính:

* + - * Hiệu ứng bóng mờ không được tính đến.
      * Ước lượng hướng và vị trí (địa phương) rất không chính xác.
      * Chỉ có các biên bước được xem xét.

Trong [Her 70], biên được phân loại thành ba loại:

Biên Delta: A black hand and a black line

Description automatically generated Biên delta có sự gián đoạn bậc không. Chúng có thể được biểu diễn dưới dạng hàm delta.

Biên bước: A black square with a black line

Description automatically generated Biên bước cũng có sự gián đoạn bậc không. Chúng có thể được biểu diễn dưới dạng hàm bước.

Biên nếp gấp: A black triangle with a white background

Description automatically generated Biên nếp gấp có sự gián đoạn bậc một. Chúng có thể được biểu diễn dưới dạng hàm tam giác.

Trong thực tế, ba loại biên này có mối liên quan: Nếu biên delta có thể được xác định bằng cách gấp hình ảnh với một mặt nạ, biên bước có thể được xác định bằng cách gấp hình ảnh với đạo hàm bậc một của cùng một mặt nạ và cũng như biên nếp gấp được xác định bằng đạo hàm bậc hai của mặt nạ. Điều này là bởi vì hàm delta là đạo hàm bậc một của hàm bước, mà lại là đạo hàm bậc một của hàm chéo.

[Chi 97] và [Chi 98] mô hình edgel qua ba tham số: vị trí, hướng và biên độ.

* Thuật toán:
  + Bước 1: Lấy mẫu hình ảnh: Gấp hình ảnh với một mặt nạ. Và sau đó lấy mẫu hình ảnh gấp tại mỗi điểm ảnh ở bốn hướng để xây dựng các khối .
  + Bước 2: Ước lượng tham số: Từ các khối được tạo ra, ba tham số edgel được ước lượng bằng phương pháp Least Squares (LS).
  + Bước 3: Liên kết biên
* Nhược điểm: Tính toán tốn kém.
  + 1. Phân đoạn – Segmentation:

Theo nghĩa tổng quát, phân đoạn hình ảnh được định nghĩa là quá trình chia một hình ảnh kỹ thuật số thành các vùng không giao nhau (không chồng lấn). Nó cố gắng nhóm các pixel lại với nhau thành các vùng tương tự nhau.

* + - 1. Phân loại
* **Phân đoạn hình ảnh thông thường (Generic Image Segmentation):**

Các thuật toán cho mục tiêu này cố gắng tách các vùng dựa trên một số sự gián đoạn nào đó trong hình ảnh (hơi giống như phát hiện cạnh).

* **Phân đoạn về cấu trúc (Texture Segmentation):**

Ở đây chủ yếu là cấu trúc vật liệu được sử dụng làm tiêu chí cho việc phân đoạn. Các thuật toán sử dụng ý tưởng này xác định vị trí cấu trúc và vùng không có cấu trúc hoặc các vùng chứa các loại cấu trúc khác nhau.

* **Phân đoạn theo ứng dụng cụ thể (Application-Specific Segmentation):**

Loại phân đoạn này không thuộc về phân đoạn hình ảnh. Tuy nhiên, để hoàn chỉnh về mặt thảo luận, chúng tôi đề xuất chú ý đặc biệt đến nhánh này. Chúng tôi tin rằng ngày càng nhiều loại phân đoạn này sẽ xuất hiện hữu ích cho các ứng dụng đặc biệt, như phân đoạn đường viền để truy xuất hình ảnh.

Trước hết, chúng tôi sẽ đưa ra một cái nhìn tổng quan ngắn gọn về phân đoạn hình ảnh thông thường. Phân đoạn cấu trúc sẽ được thảo luận khi nói về các đặc trưng cấu trúc. Và phân đoạn theo ứng dụng cụ thể sẽ được bàn luận trong phần này về các đặc trưng cụ thể của miền.

* + - * 1. Phân đoạn hình ảnh thông thường (Generic Image Segmentation)

Trong quá trình phân đoạn, trước tiên chúng ta phải định nghĩa thuật ngữ "kết nối", quyết định sự hình thành của các vùng. Hiện nay, các phương pháp ước lượng kết nối thông thường được sử dụng là bốn vùng lân cận và tám vùng lân cận. Đã được phát hiện rằng kết quả dựa trên tám vùng lân cận gần với nhận thức của con người hơn.

Các phương pháp phân đoạn thường được phân loại như sau:

1. **Phân loại dựa trên vùng**: Gán mỗi pixel vào một đối tượng hoặc vùng cụ thể.

**Dựa trên ngưỡng** **(Thresholding – based):** Phương pháp dựa trên vùng thường được gọi là dựa trên ngưỡng, sử dụng ngưỡng toàn cầu hoặc thích nghi để phân đoạn hình ảnh. Ví dụ, chúng ta có một hình ảnh và một ngưỡng t, và chúng ta muốn có một biểu diễn phân đoạn sau đó chúng ta lập công thức như sau:

Nếu thì , ngược lại .

Điểm khó khăn của phương pháp này là chúng ta không thể sử dụng một ngưỡng tiêu chuẩn cho tất cả các hình ảnh, hoặc thậm chí, chúng ta không thể sử dụng một ngưỡng đơn lẻ cho một hình ảnh. Điều này làm cho việc lựa chọn ngưỡng trở nên phức tạp. Do đó, sự tập trung trong phương pháp này là tìm kiếm các ngưỡng từ hình ảnh một cách tự động.

* Ngưỡng cố định
* Ngưỡng dựa trên histogram
  + - Thuật toán Isodata
    - Thuật toán đối xứng nền
    - Thuật toán tam giác

**Phát triển vùng (Region Growing):** Một nhánh mới trong phương pháp dựa trên vùng và phát triển dựa trên toán tử hình thái (Đóng – Mở) trong phân đoạn hình ảnh. Nó hiệu quả trong việc xử lý phân đoạn hình ảnh khoa học.

1. **Dựa trên ranh giới (Boundary – based):** Cố gắng chỉ định vị các biên giữa các vùng.

**Theo dõi ranh giới (Boudary tracking)**: Ở đây, từ ban đầu, các thuật toán cố gắng định vị đường viền gần bằng cách theo dõi từng pixel. Phương pháp này chỉ hữu ích cho các hình ảnh không có nhiễu hoặc trong tình huống mà sự can thiệp của con người có thể ngăn chặn sự trượt đường ray.

**Ngưỡng hình ảnh độ dốc (Gradient image thresholding)**: Dựa trên ý tưởng sử dụng một mức xám vừa phải làm ngưỡng để phân tách các đối tượng và nền.

1. **Dựa trên cạnh (Edge – based):** Cố gắng xác định các pixel biên trước và sau đó kết nối chúng lại để tạo ra các biên cần thiết. Do đó, thuật toán theo phương pháp này có thể được chia thành hai bước: tìm biên và liên kết biên. Liên kết biên có thể được thực hiện qua một số kỹ thuật:

* **Tìm kiếm Heuristic**: Sử dụng một hàm đánh giá cạnh.
* Tìm đường cong phù hợp **(Curve fitting).**

Cuối cùng nhưng không kém phần quan trọng, đối với hình ảnh 2D và hình ảnh 3D, có các phương pháp khác nhau:

* **2D**: Nhóm các pixel lại với nhau theo tỷ lệ thay đổi của độ sáng của chúng qua vùng.
* **3D**: Nhóm các pixel lại với nhau theo tỷ lệ thay đổi của độ sâu trong hình ảnh, tương ứng với các pixel nằm trên cùng một bề mặt như một mặt phẳng, hình trụ, hình cầu, vv.
  + 1. Rút trích màu trên ảnh

Để gia tăng độ chính xác khi đối sánh hình ảnh qua đặc trưng cấp thấp, các màu sắc chủ đạo của hình ảnh được trích xuất. Đối với đặc trưng màu sắc, mỗi điểm ảnh được gom cụm theo các màu sắc cơ bản của dãy màu Newton (Red, Green, Blue, Yellow, Orange, Purple). Vì vậy, mỗi điểm ảnh là một véc-tơ dữ liệu trong không gian RGB và thực hiện gom cụm theo thuật toán k-Means dựa trên 6 cụm của các màu Newton. Thuật toán k-Means có trong openCV áp dụng trong trường hợp này được mô tả như sau:

Hình 4. Code mình họa rút trích màu dựa trên thuật toán K-Means

A close-up of a color palette

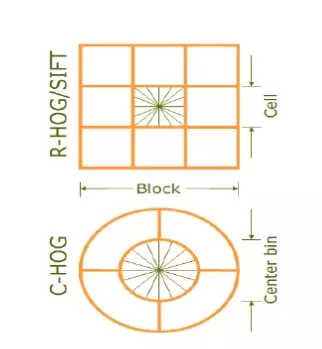
Description automatically generated

Hình 5. Ảnh gốc và ảnh kết quả rút trích màu

* + 1. Rút trích đặc trưng ảnh bằng HOG (Histogram of Orented Gradients)

Histogram of gradient (HOG) là đặc trưng được dùng nhiều trong lĩnh vực phát hiện đối tượng, được đề xuất bởi Bill Triggs và Navel Dalal vào năm 2005 tại viện nghiên cứu INRIA.

Ý tưởng chính của phương pháp là hình dạng và trạng thái xuất hiện của vật có thể được đặc trưng bằng sự phân bố về cường độ và hướng của cạnh. Đặc trưng này được phát triển dựa trên SIFT, đặc trưng HOG được tính trên cả một vùng. Do sự biến thiên màu sắc trong các vùng là khác nhau, kết quả là các vùng sẽ cho ta vector đặc trưng của nó. Vì vậy, để có được đặc trưng của toàn bộ cửa sổ (window) ta phải kết hợp nhiều vùng liên tiếp lại với nhau.

Đặc trưng HOG có một vài biến thể thường gặp như: R-HOG, R2-HOG, C-HOG, …Các đặc trưng này khác nhau ở các phân bố và hình dạng của các ô như trong hình:

Hình 6. R-HOG (trên) và C HOG (dưới)

Các bước thực hiện để rút trích đặc trưng ảnh bằng HOG, hoạt động cơ bản của các giải thuật phát hiện đối tượng:

1. Đọc ảnh
2. Tiền xử lý (preprocessing): hình ảnh được đưa qua bước tiền xử lý để thực hiện các thao tác như cân bằng sáng, làm mờ, …
3. Trích đặc trưng ảnh (feature extraction): bằng cách sử dụng các phương pháp rút trích đặc trưng ảnh ta sẽ thu được vector đặc trưng của ảnh. Nói một cách nôm na thân quen đó chính là bạn mã hóa hình ảnh thành một vector, và vector này mang những đặc trưng (các số thực) đại diện cho ảnh đó.
4. Huấn luyện mô hình học máy (training): với phương pháp truyền thống, ta thường sử dụng mô hình SVM trong machine learning để phân tách các vector đặc trưng thành các lớp cần phân loại.
5. Kiểm thử (validation): sau khi huấn luyện xong mô hình học máy bạn cần phải đánh giá mô hình mình đã huấn luyện đạt độ chính xác là bao nhiêu phần trăm trên tập kiểm thử này. Khi bạn đã hài lòng với kết quả kiểm thử, ta có thể dừng quá trình huấn luyện.
6. KẾT LUẬN

Trong bài báo cáo này, chúng ta đã thảo luận về các phương pháp trích xuất đặc trưng từ hình ảnh, tập trung vào việc phân tích các phương pháp phát hiện cạnh và phân đoạn hình ảnh. Đầu tiên, chúng ta đã đề cập đến các phương pháp phát hiện cạnh, bao gồm phương pháp dựa trên đạo hàm bậc nhất, đạo hàm bậc hai và phát hiện cạnh Canny. Mỗi phương pháp này đều có ưu nhược điểm riêng và có ứng dụng khác nhau trong thực tế.

Tiếp theo, chúng ta đã xem xét về phân đoạn hình ảnh, bao gồm phân loại dựa trên ngưỡng, dựa trên vùng và dựa trên biên. Mỗi phương pháp phân đoạn này đều có cách tiếp cận và ứng dụng riêng biệt, đáp ứng nhu cầu phân tích hình ảnh cho các mục đích khác nhau.

Tổng kết, bài báo cáo này đã cung cấp một cái nhìn tổng quan về các phương pháp trích xuất đặc trưng từ hình ảnh và sự quan trọng của chúng trong phân tích hình ảnh và thị giác máy tính. Việc hiểu và áp dụng các phương pháp này có thể cung cấp những thông tin quý giá để giải quyết các vấn đề thực tế trong lĩnh vực xử lý hình ảnh và trí tuệ nhân tạo.

TÀI LIỆU THAM KHẢO

Dr. Dafda. (n.d.). *Image Segmentation.* India.

Geeksforgeeks. (n.d.). *OpenCV - Python Tutorials*. Retrieved from https://www.geeksforgeeks.org/opencv-python-tutorial/

[Image Segmentation using K Means Clustering - GeeksforGeeks](https://www.geeksforgeeks.org/image-segmentation-using-k-means-clustering/)

Khanh's Blog. (n.d.). *Image Segmetation*. Retrieved from https://phamdinhkhanh.github.io/2020/06/10/ImageSegmention.html#12-b%C3%A0i-to%C3%A1n-image-segmentation

MathWork. (n.d.). *Structuring Elements*. Retrieved from https://www.mathworks.com/help/images/structuring-elements.html

Nero Phung. (n.d.). *Otsu Threshold*. Retrieved from https://nerophung.github.io/2019/09/26/otsu-threshold

PGS. TS. Đỗ Toàn Năng. (n.d.). *Bài Giảng Xử Lí Ảnh.*

TAPIT. (n.d.). *Morphological Image Processing*. Retrieved from https://tapit.vn/xu-ly-anh-hinh-thai-hoc-morphological-image-processing/

[Can 83] J. F. Canny. Finding edges and lines in images. Master's thesis, MIT. AI Lab. TR-720, 1983.

[Can 86] J. F. Canny. A computational approach to edge detection. IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679-698, 1986.

[Chi 97] Pei-Chun Chiang and Thomas O. Binford. Generic, Model-Based Edge Estimation in the Image Surface. In Proceedings of Image Understanding Workshop, Vol. 2, pp. 1237-1246, May 1997.

[Chi 98] Pei-Chun Chiang and Thomas O. Binford. Edge Detection: Modeling, Estimation, and Aggregation. submitted to IEEE International Conference on Image Processing, 1998

[Her 70] A. Herskovits and T. O. Binford. On Boundary Detection. MIT Artificial Intelligence Lab., AI Memo 183, Cambridge, Mass., 1970