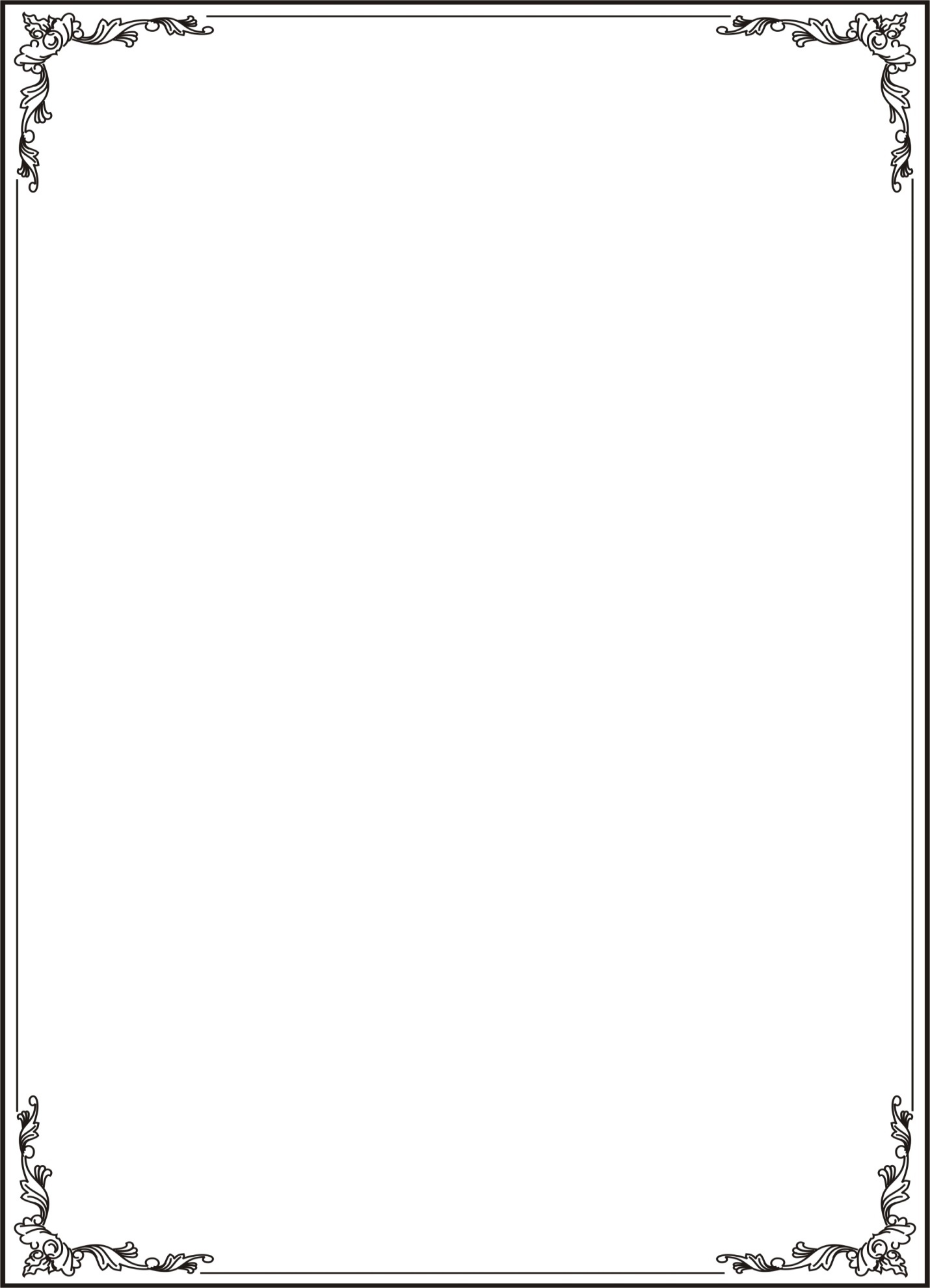
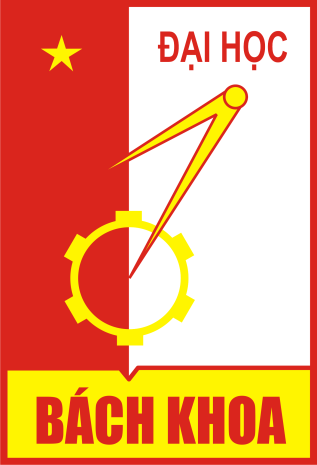
****

**ĐẠI HỌC BÁCH KHOA HÀ NỘI**

VIỆN CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG

----------



**Báo Cáo**

HỌC PHẦN: Xử Lý Ảnh

ĐỀ TÀI: **Đặc trưng SIFT và ứng dụng đối sánh logo**

*Nhóm sinh viên thực hiện:*

**Nguyễn Tiến Thiện (20153583)**

**Tăng Quang Huy (****2015)**

**Nguyễn Bình Minh (****2015)**

*Giảng viên hướng dẫn:* **TS.**

Mục lục

[1 Giới thiệu đặc trưng SIFT 3](#_Toc11444491)

[1.1 Giới thiệu 3](#_Toc11444492)

[1.2 Thuật toán 3](#_Toc11444493)

[1.2.1 Phát hiện các điểm cực trị của Scale-space 4](#_Toc11444494)

[1.2.2 Định vị các điểm nổi bật 7](#_Toc11444495)

[1.2.3 Xác định hướng cho các điểm nổi bật 8](#_Toc11444496)

[1.2.4 Mô tả các điểm nổi bật 9](#_Toc11444497)

[2 Giới thiệu đặc trưng SURF 9](#_Toc11444498)

[2.1 Giới thiệu 9](#_Toc11444499)

[2.2 Thuật toán 10](#_Toc11444500)

[2.2.1 Xác định điểm nổi bật 11](#_Toc11444501)

[2.2.2 Gán hướng cho điểm nổi bật và mô tả đặc trưng SURF 12](#_Toc11444502)

[3 Ứng dụng 14](#_Toc11444503)

[3.1 Giới thiệu 14](#_Toc11444504)

[3.2 Phương pháp 14](#_Toc11444505)

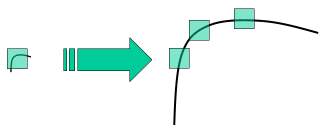
[3.3 Kết quả 14](#_Toc11444506)

[4 Tài liệu tham khảo 14](#_Toc11444507)

# Giới thiệu đặc trưng SIFT

## Giới thiệu

Trong quá trình nghiên cứu so sánh đối chiếu các điểm đặc trưng của ảnh số, người ta thường gặp những vấn đề liên quan đến việc các điểm đặc trưng tương đồng giữa các ảnh không có được sự so sánh đối chiếu chính xác nếu như các điểm đó không có cùng tỉ lệ. Điều này có nghĩa là nếu ta cố gắng tìm ra các điểm đặc trưng tương đồng từ hai ảnh được chụp lại từ hai góc độ khác nhau của một khung hình, ta sẽ không có được kết quả tốt nhất.



Hình 1. Mô tả sự không hoàn thiện của thuật toán Harris

Ví dụ: Trong hình trên, ta dễ dàng nhận ra ở bên trái là một góc đã được thuật toán tìm kiếm góc Harris phát hiện. Thế nhưng nếu như ta xem xét góc này với một tỉ lệ lớn hơn nhiều thì đó không còn lại một góc theo thuật toán Harris nữa, nhưng trên thực tế, đó hoàn toàn là một góc.

Để giải quyết vấn đề này, David G.Lowe đưa nghiên cứu và đưa ra đặc trưng cục bộ bất biến SIFT (Scale - Invariant Feature Transform) vào năm 2004 với các tính chất sau:

* Bất biến với các phép: thay đổi tỷ lệ ảnh, quay ảnh, thay đổi góc nhìn, thêm nhiễu ảnh, thay đổi cường độ chiếu sáng ảnh.
* Các đặc trưng cục bộ bất biến được trích rút từ các điểm nổi bật trên ảnh.

## Thuật toán

Phương pháp trích chọn đặc trưng cục bộ SIFT gồm các bước:

* Phát hiện các điểm cực trị Scale-Space (scale-space extrema detection): Áp dụng các hàm sai khác DoG để tìm các điểm có khả năng là điểm đặc trưng (candidate keypoints), đó là những điểm ít phụ thuộc vào sự phóng và xoay ảnh.
* Định vị các điểm nổi bật (Keypoint localization): Lọc và lấy ra các điểm tốt nhất từ tập các điểm nổi bật tiềm năng.
* Xác định hướng cho các điểm nổi bật (Oriented assignment): Mỗi keypoint được gán một hoặc nhiều hướng dựa trên dựa theo hướng gradient.
* Mô tả các điểm nổi bật (Keypoint description): Các điểm hấp dẫn sau khi được xác định hướng sẽ được mô tả dưới dạng các vector đặc trưng nhiều chiều.

### Phát hiện các điểm cực trị của Scale-space

Định nghĩa không gian tỉ lệ của một hình ảnh được mô tả bằng hàm dưới đây:

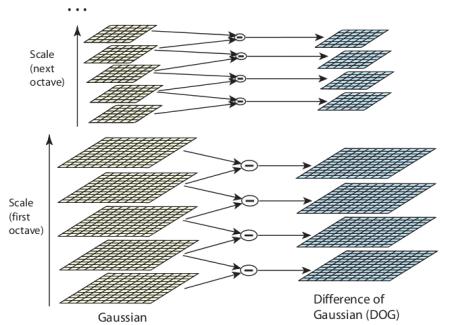
Với là Biến tỉ lệ Gaussian (Variable scale Gausisan)

là ảnh đầu vàolà hàm không gian tỉ lệ của ảnh

Để tìm những điểm nổi bật có tính bất biến cao, ta sẽ tìm cực trị cục bộ của hàm sai khác DoG (Difference of Gaussian), kí hiệu là . Hàm này được tính toán từ sự sai khác giữa 2 không gian đo cạnh nhau của một ảnh với các tham số đo lệch nhau một hằng số .

Với được tính bằng phép trừ giữa Hàm không gian tỉ lệ của ảnh và lần của chính nó.

Lý do để lựa chọn hàm Gaussian là vì kỹ thuật này rất hiệu quả để tính toán (cũng như làm tăng độ mịn của ảnh), mà thì luôn phải được tính rất nhiều để mô tả đặc trưng trong không gian đo, và sau đó, sẽ được tính bằng phép trừ ma trận điểm ảnh, một phép tính toán đơn giản.



Hình 2 Quá trình tính không gian đo (L) và hàm sai khác (D)

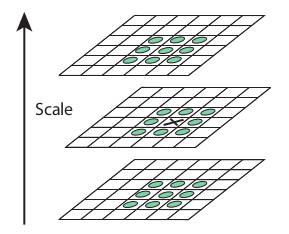
Hơn nữa, hàm sai khác DoG có thể được sử dụng để tạo ra một xấp xỉ gần với đạo hàm bậc hai Laplace có kích thước chuẩn của hàm Gaussiando tác giả Lindeberg đề xuất năm 1994[17]. Ông đã chỉ ra rằng việc chuẩn hóa đạo hàm bậc hai với hệ số là cần thiết cho bất biến đo trở nên đúng. Cụ thể, ông đã công bố rằng các giá trị cực đại và cực tiểu của chính là những giá trị có tính ổn định nhất (bất biến cao) so với một loạt các hàm đánh giá khác như: gradient, Hessian hay Harris.

Mối quan hệ giữa và được biểu diễn như sau:

Như vậy, có thể được tính thông qua việc xấp xỉ sự sai khác hữu hạn tại các tham số đo gần nhau và **:**

Từ công thức này, ta thấy khi mà hàm sai khác DoG được tính toán tại các tham số đo lệch nhau một hằng số , thì ta có thể sử dụng DoG để xấp xỉ đạo hàm bậc hai Laplace của Gaussian. Vì hệ số trong phương trình trên là hằng số trong mọi không gian đo nên nó sẽ không ảnh hưởng đến việc tìm các vị trí cực trị. Sai số trong việc xấp xỉ đạo hàm bậc 2 tiến về khi gần với . Tuy nhiên, các kết quả thử nghiệm của tác giả cho thấy quá trình xấp xỉ đạo hàm không ảnh hưởng đến việc dò tìm các vị trí cực trị thậm chí ngay cả khi chọn **k** khá xa, ví dụ .

Sau khi áp dụng hàm DoG ta thu được các lớp kết quả khác nhau từ ảnh gốc, bước tiếp theo là tìm các cực trị trong các lớp kết quả theo từng miền cục bộ. Cụ thể là tại mỗi điểm trên các lớp kết quả sẽ được so sánh với 8 điểm lân cận trên cùng lớp và 9 điểm lân cận trên mỗi lớp khác.



Hình 3 Quá trình tìm điểm cực trị trong các hàm sai khác DoG

Trong hình trên: điểm đánh dấu x sẽ được so sánh với 26 điểm lân cận (đánh dấu vòng tròn xanh). Điểm này sẽ được lấy làm điểm tiềm năng (điểm có thể làm điểm đặc biệt –candidate keypoint) nếu nó có giá trị lớn nhất hoặc nhỏ nhất so với 26 điểm lân cận như trên. Giải pháp cho việc tìm các điểm tiềm năng này là sử dụng thuật toán blob detection (dò tìm điểm) do Lindeberg đề xuất.

Vì số lượng các cực trị là rất lớn, vì vậy để tăng sự hiệu quả khi dò tìm các điểm cực trị (dò các điểm cực trị tốt nhất thay vì phải dò hết), ta cần xác định tần số lấy mẫu trong không gian đo và tần số lấy mẫu trong không gian quan sát (không gian ảnh). Thật không may là ta không thể xác định cả 2 loại tần số này một cách tự động trong mỗi tiến trình dò tìm.

Thay vì vậy, các tần số này sẽ được xác định thông qua phương pháp thử nghiệm.Sau khi thử nghiệm với nhiều nguồn dữ liệu ảnh khác nhau, tác giả đã chỉ ra tần số lấy mẫu trong không gian đo tốt nhất là 3 (giữ lại 3 lớp trong mỗi bộ 8 lớp), và tần số lấy mẫu .

### Định vị các điểm nổi bật

Sau bước một, ta sẽ thu được rất nhiều điểm tiềm năng có thể làm điểm đặc biệt, tuy nhiên một số điểm trong số đó là không cần thiết. Ở bước này, ta sẽ loại bỏ các điểm có độ tương phản kém (nhạy cảm với nhiễu) hoặc tính đặc trưng cục bộ ít hơn các điểm khác hoặc có xu hướng là đường biên đối tượng.Bước thực hiện này gồm 3 công đoạn:

* Phép nội suy lân cận cho vị trí đúng của điểm tiềm năng:

Phép nội suy lân cận sử dụng mở rộng Taylor (Taylor expansion) cho hàm Difference-of-Gaussian

Với là Độ dịch so với các điểm lân cận của điểm lấy mẫu.

Vùng chứa điểm nổi bật được xác định qua:

Nếu , điểm nổi bật nằm ở gần điểm lấy mẫu khác.

Nếu , những điểm này sẽ được đưa vào tập hợp mẫu tốt nhất để tiếp tục phân tích.

* Loại trừ các điểm có tính tương phản kém

Các điểm nhạy cảm với độ sáng và nhiễu thì không được trở thành điểm đặc biệt và cần loại bỏ khỏi danh sách điểm tiềm năng.

Dùng để loại những điểm cực trị không ổn định (có độ tương phản thấp).

Thay vào , ta có:

Nếu thì điểm lấy mẫu đó sẽ bị loại.

* Loại bỏ các điểm dư thừa theo biên

Sử dụng hàm DoG sẽ cho tác động mạnh đến biên khi vị trí của biên là khó xác định và vì vậy các điểm tiềm năng trên biên sẽ không bất biến và bị nhiễu. Và để tăng sự ổn định cho các điểm sẽ được chọn làm điểm đặc biệt ta sẽ loại trừ các điểm tiềm năng khó định vị (tức là vị trí dễ thay đổi khi có nhiễu do nằm ở biên).

Sau khi áp dụng hàm DoG sẽ làm đường biên ảnh không rõ ràng và độ cong chính sẽ có giá trị lớn hơn nhiều so với độ cong dọc theo biên vì vậy cần loại bỏ bớt các điểm đặc biệt dọc theo cùng một biên. Giải pháp cho việc này là sử dụng giá trị của ma trận Hessian cấp 2:

Các giá trị riêng của tỉ lệ thuận với độ cong của ,các giá trị riêng (giá trị nhỏ) và (giá trị lớn) có tỉ lệ sẽ được sử dụng. Các phần tử của là và .

### Xác định hướng cho các điểm nổi bật

Bằng việc gán một hướng cho mỗi điểm nổi bật dựa vào các thuộc tính ảnh cục bộ, bộ mô tả điểm nổi bật có thể được biểu diễn tương đối so với hướng này và do đó đạt được tính bất biến đối với các hiện tượng quay ảnh.

Độ đo của các điểm nổi bật được sử dụng để tìm ra một ảnh đã lọc Gaussian ***L*** với kích thước gần nhất sao cho mọi tính toán sẽ được thực hiện trong cùng một cách bất biến về độ đo.

Tính toán hướng và độ lớn cho điểm nổi bật:

Với:

là Độ lớn của vector định hướng

là Hướng của vector định hướng (biểu diễn qua góc )

là Ảnh Gaussian ở tỷ lệ nhỏ nhất

* Tiếp theo tính histogram của hướng tại điểm mẫu trong một vùng xung quanh keypoint thì được chia vào 36 bin để bao phủ 360 độ của hướng. Kích thước của vùng thu thập hướng xung quanh keypoint phụ thuộc vào scale của nó (scale càng lớn, khu vực thu thập càng lớn).
* Chọn trong histogram hướng để gán cho keypoint:
  + có histogram cao nhất hoặc
  + 80% của histogram cao nhất



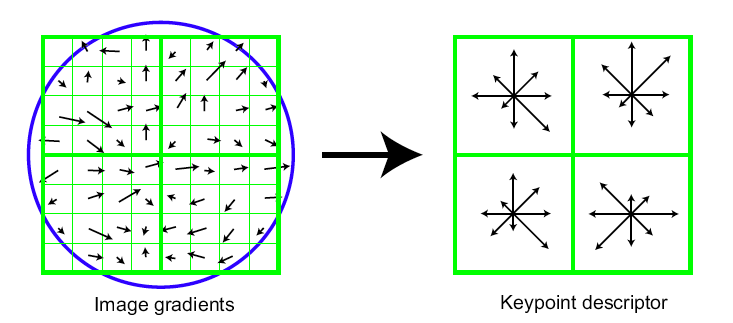
Hình 4 Histogram theo hướng

* Mỗi keypoint tại cùng scale có thể được gán 1 hoặc nhiều hướng

### Mô tả các điểm nổi bật

Các phép xử lý trên đây đã thực hiện dò tìm và gán tọa độ, kích thước, và hướng cho mỗi điểm nổi bật. Các tham số đó yêu cầu một hệ thống tọa độ cục bộ 2D có thể lặp lại được để mô tả vùng ảnh cục bộ và nhờ vậy tạo ra sự bất biến đối với các tham số đó. Bước này sẽ tính toán một bộ mô tả cho một vùng ảnh cục bộ mà có tính đặc trưng cao (bất biến với các thay đổi khác nhau về độ sáng, thu –phóng ảnh, xoay).

Cách tiếp cận này dựa trên một mô hình thị giác sinh học, cụ thể là mô hình noron phức tạp trong hệ thống não bộ. Các noron sẽ tương ứng với một gradient tại một hướng và tần số không gian cụ thể, nhưng vị trí của gradient trên võng mạc được phép trượt trên một phạm vi nhỏ của khung nhìn.



Hình 5 Mô tả tạo bộ mô tả cục bộ

Ảnh trái là mô phỏng biên độ gradient và hướng tại mỗi mẫu ảnh trong một vùng lân cận với điểm nổi bật. Các giá trị đó tập trung trong một cửa sổ gaussian (nằm bên trong vòng tròn). Các mẫu này sau đó được gom lại thành một lược đồ hướng mô tả vắn tắt nội dung trong 4x4 vùng con như được mô tả ở bên phải với độ dài của mỗi hàng tương ứng với tổng biên độ gradient gần hướng đó bên trong một vùng.

# Giới thiệu đặc trưng SURF

## Giới thiệu

Phương pháp SIFT đã giải quyết được những hạn chế còn tồn tại ở thuật toán tìm kiếm góc Harris và trở thành một trong những thuật toán trích chọn đặc trưng mạnh mẽ nhất. Dù vậy, tốc độ xử lý của SIFT vẫn còn rất chậm và không phù hợp với các ứng dụng thời gian thực.

Để giải quyết bài toán này, người ta đã giới thiệu thuật toán trích chọn đặc trưng SURF (Speed Up Robust Features) có được sự cân bằng giữa yêu cầu tốc độ và sự chính xác.Đặc trưng tối ưu cả hai giai đoạn phát hiện đặc trưng (detector) và mô tả đặc trưng (descriptor) về mặt thời gian tính toán nhưng vẫn giữ được tính bền vững của đặc trưng.

Bộ phát hiện đặc trưng của SURF sử dụng phép xấp xỉ trên ma trận Hessian và ảnh tích hợp (Integral Image) để làm giảm thời gian tính toán một cách đáng kể. Bộ mô tả đặc trưng tương tự như đặc trưng SIFT, sử dụng vector 64 chiều chứa thông tin biến thiên trên ảnh dựa trên sự phân phối bậc nhất Haar wavelet tác động trên trục x và y, kết hợp với ảnh tích lũy làm tăng tốc độ tính toán.

SURF được mô tả bởi vector có số chiều ít hơn SIFT nên tốc độ so khớp nhanh hơn, tuy nhiên độ bền vững vẫn được đảm bảo. Hơn thế nữa, bằng việc đánh chỉ mục dựa trên dấu của Laplacian, đặc trưng SURF không chỉ giữ tính bền vững cho đặc trưng mà còn làm tăng tốc độ so khớp (tăng gấp 2 trong trường hợp tốt nhất).

Ảnh tích hợp (Integral Image) là một khái niệm do Viola và Jones đưa ra trong “Rapid object detection using a boosted cascade of simple Features”, là một mảng hai chiều với kích thước bằng kích thước của ảnh cần tìm các đặc trưng. Mỗi phần tử của mảng này được tính bằng cách tính tổng của điểm ảnh phía trên (dòng –1) và bên trái (cột -1) của nó. Bắt đầu từ vị trí trên, bên trái đến vị trí dưới, bên phải của ảnh, việc tính toán đơn thuần chỉ dựa trên phép cộng số nguyên đơn giản, do đó tốc độ thực hiện rất nhanh.

Ảnh tích hợp có công thức sau:

## Thuật toán

Thuật toán của kỹ thuật SURF gồm những bước dưới đây:

* Sử dụng bộ dò Fast-Hessian để xác định các điểm nổi bật
* Gán hướng cho các điểm nổi bật và mô tả đặc trưng SURF.
* So khớp đặc trưng.

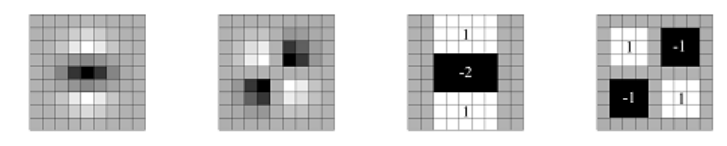
### Xác định điểm nổi bật

Bộ dò này được xây dựng dựa trên ma trận Hessian vì hiệu suất tốt của nó về thời gian tính toán và độ chính xác, kết hợp với việc sử dụng ảnh tích hợp để giảm thời gian tính toán đặc trưng.

Ma trận Hessian được định nghĩa như sau:

Với là tích của đạo hàm bậc hai của hàm Gaussian với ảnh tại điểm , có tỉ lệ .

Nếu như SIFT xấp xỉ việc tính Laplacian của hàm Gaussian (LoG) bằng việc tính sai khác của hàm Gaussian (DoG) thì SURF xấp xỉ việc tính đạo hàm cấp 2 của hàm Gaussian bằng các hộp lọc (box filters). Dưới đây là một ví dụ của việc tính xấp xỉ đạo hàm cấp hai của hàm Gaussian với hệ số tỉ lệ thấp nhất bằng hộp lọc:



Hình 6 Xấp xỉ đạo hàm cấp 2 hàm Gaussian bằng hộp lọc

Trong hình trên: Ảnh thứ nhất là đạo hàm ma trận đạo hàm cấp 2 Gaussian theo trục y, ảnh thứ hai theo trục x và trục y. Ảnh thứ ba và thứ tự lần lượt là các hộp lọc xấp xỉ với hai trường hợp của ảnh một và hai. Phép tích chập xoắn của ảnh ***I*** với các hộp lọc này được thực hiện rất nhanh bằng việc sử dụng kết hợp với ảnh tích lũy.

Ta xác định vị trí và hệ số tỉ lệ tương ứng của điểm đặc trưng dựa trên định thức của ma trận Hessian. Công thức tính xấp xỉ định thức ma trận Hessian:

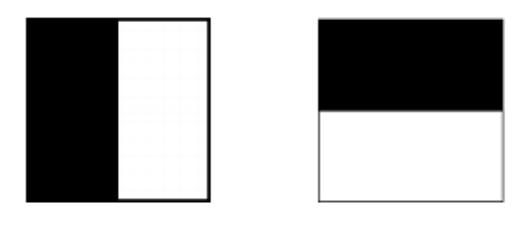
Trong đó là trọng số cân bằng của biểu thức định thức ma trận Hessian tùy thuộc vào hệ số tỉ lệ. là các hộp lọc xấp xỉ Gaussian như đã nói ở trên. Đối với tỉ lệ , kích thước hộp lọc là **9x9** thì giá trị trọng số **w** được tính bằng:

Vị trí, tỉ lệ và không gian rảnh mà điểm đặc trưng được xác định một phép loại trừ phi cực đại trong một vùng 3x3x3 sẽ được áp dụng.

### Gán hướng cho điểm nổi bật và mô tả đặc trưng SURF

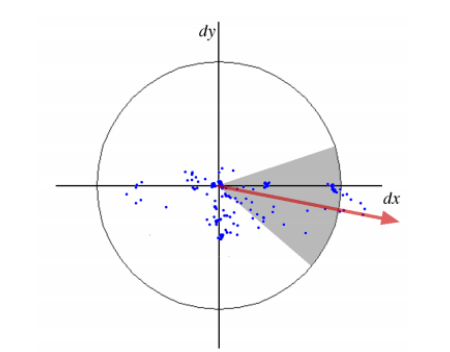
Đầu tiên, ta phải xác định vùng hình xung quanh điểm đặc trưng vừa tìm được, gán một giá trị hướng duy nhất cho điểm đặc trưng. Kích thước của hình tròn phụ thuộc và hệ số tỉ lệ tương ứng trong không gian ảnh mà điểm đặc trưng tìm được.

Ở đây các tác giả chọn bán kính của hình tròn là , trong đó là tỉ lệ mà tại đó điểm đặc trưng được tìm thấy. Hướng của đặc trưng được tính bằng Haar wavelet tác động theo hai hướng và . Trong đó, vùng tối có trọng số , vùng sáng có trọng số . Kích thước của wavelet cũng phụ thuộc vào hệ số tỉ lệ , ở đây các tác giả chọn .



Hình 7 Lọc Haar wavelet để tính sự ảnh hưởng trên hai hướng x và y

Haar wavelet có thể được tính một cách nhanh chóng bằng cách sử dụng ảnh tích lũy tương tự như hộp lọc xấp xỉ của đạo hàm cấp 2 hàm Gaussian. Vector hướng nào trội nhất sẽ được ước lượng và gắn vào thông tin của điểm đặc trưng. Hình dưới đây sẽ mô tả hướng và vùng ảnh hưởng của đặc trưng.



Hình 8 Vùng hình tròn xung quanh và hướng đại diện cho điểm đặc trưng

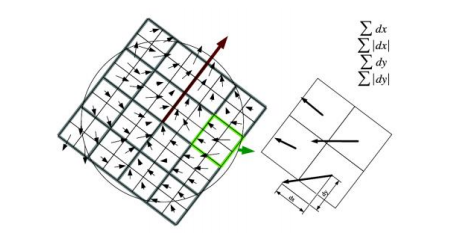
Tiếp theo, ta xây dựng các vùng hình vuông xung quanh điểm đặc trưng men theo vector hướng vừa ước lượng được ở bước trước đó. Vùng hình vuông này được chia nhỏ thành 4x4 hình vuông con để ghi nhận thông tin của trên miền không gian ảnh lân cận. Haar wavelet được rút trích trên toàn bộ không gian điểm ảnh. Wavelet tác động trên hai hướng ngang và dọc được cộng dồn các giá trị và trên mỗi hình vuông con.

Hơn thế nữa, các giá trị tuyệt đối và cũng được cộng dồn để lấy thông tin về độ lớn của sự thay đổi cường độ sáng trên ảnh.

Như vậy mỗi hình vuông con sẽ được mô tả bởi vector 4 chiều:

Như vậy vector mô tả cho tất cả 4x4 hình vuông con là một vector 64 chiều (4x4x4). Đây cũng chính là mô tả đặc trưng chuẩn của SURF (hay còn gọi là SURF-64). Ngoài ra còn có các phiên bản khác dựa trên cách chia hình vuông con như SURF – 36, SURF – 128…

Tuy nhiên thực nghiệm của các tác giả cho thấy rằng SURF – 64 cho tốc độ tính toán tốt nhất mà vẫn đảm bảo tính bền vững của đặc trưng. Haar wavelet bất biến với sự Trang 30 thay đổi của ánh sáng và sự tương phản khi ta chuẩn hóa vector mô tả đặc trưng về chiều dài đơn vị.



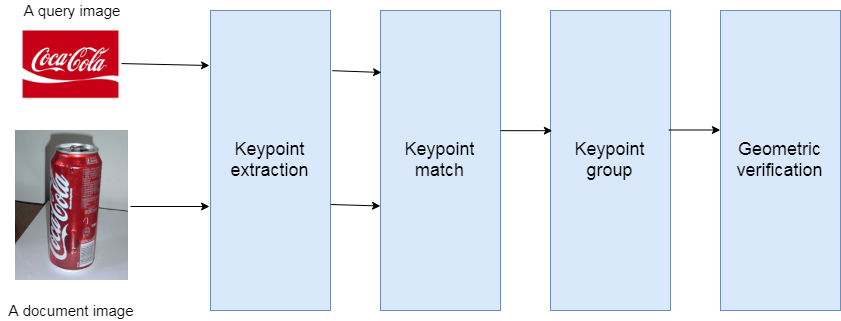
Hình 9 4x4 hình vuông xung quanh điểm đặc trưng

# Ứng dụng

## Giới thiệu

## Phương pháp

Phương pháp được thể hiện qua các bước sau:



* Xác định các điểm nổi bật: xác định các điểm nổi bật của ảnh logo và ảnh tài liệu. Mỗi điểm nổi bật được mô tả bởi một vector 128 chiều. Từ đó, logo và ảnh trong tài liệu được biểu diễn bởi tập các điểm nổi bật.
* Đối sánh các điểm nổi bật: tìm các ứng viên tốt nhất đối sánh với mỗi key-point từ ảnh logo trong tập key-points trong các ảnh tài liệu. Ứng viên tốt nhất đối sánh đơn giản là key-points gần nhất trên không gian khoảng cách Euclidean.
* Nhóm các điểm nổi bật: phương pháp phân cụm dựa trên mật độ (DBSCAN) được áp dụng để nhóm các điểm đã đối sánh. Tác dụng của bước này là phân đoạn vùng chứa logo trên ảnh tài liệu và loại bỏ các đối sánh không đúng (ngoại biên). Các điểm ngoại biên của phân cụm dựa trên mật độ là đối sánh đứng một mình hoặc nhóm nhỏ (có rất ít đối sánh).

### Xác định điểm nổi bật

Với logo đầu vào được biểu diễn bởi điểm nổi bật. Mỗi điểm nổi bật được mô tả bằng một vector đặc trưng.

Trong đó, là tọa độ x, y của điểm nổi bật thứ . là vector mô tả của điểm hấp dẫn với độ dài là 128 ( đặc trưng SIFT)

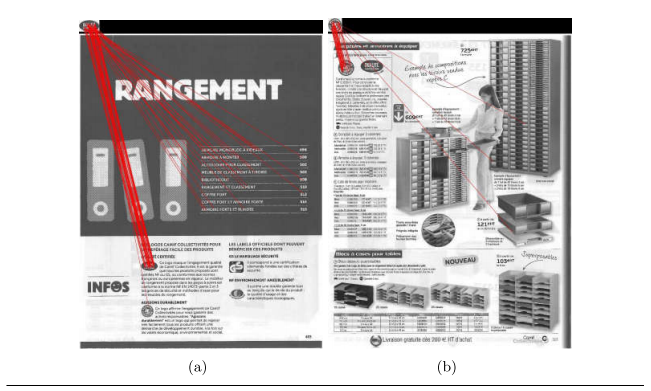
Tương tự, tài liệu được biễu diễn bởi tập các đặc trưng của điểm hấp dẫn của tài liệu thứ i:

### Đối sánh các điểm nổi bật

Sử dụng chiến lược tìm kiếm brute force để tìm tìm các ứng viên tốt nhất đối sánh với mỗi key-point từ ảnh logo trong tập key-points trong các ảnh tài liệu. Ứng viên tốt nhất được xác định dự trên độ đo khoảng cách của các cặp vector mô tả của các điểm nổi bật. Độ đo khoảng cách được sử dụng là khoảng cách Euclidean.

### Nhóm các điểm nổi bật

Sau khi đối sánh, thông thường, ta có thể dựa trên số cặp trùng khớp, so sánh ngưỡng và đưa ra kết quả truy vấn. Tuy nhiên, không phải tất cả các cặp đều trong một vùng logo. Một số trong số đó thuộc các vùng khác trong ảnh tài liệu.



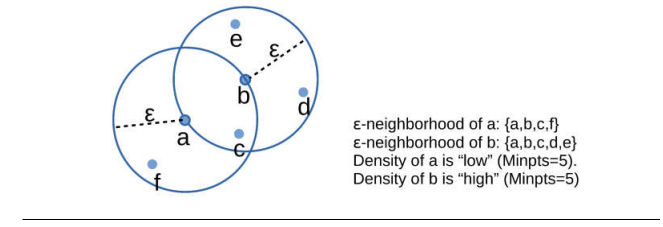
Hình 10 Có rất nhiều điểm trùng khớp nằm bên ngoài vùng logo.

Để giải quyết vấn đề trên, ta thêm bước phân cụm logo. Ta sử dụng giải thuật Density-based spatial clustering of application with noise (DBSCAN) để chọn các vùng logo.

DBSCAN được đề xuất bởi Ester et al., là một giải thuật phân cụm dựa trên mật độ. Tập các điểm hàng xóm của x là tập các điểm trong bán kính tính từ x:

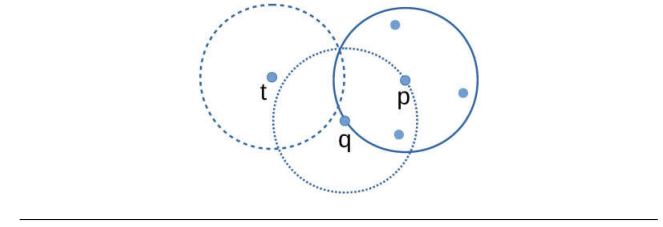
Trong đó, là khoảng cách từ điểm x đến y. Khoảng cách Euclidean thường được dùng .

Tập hàng xóm của x được coi là mật độ cao nếu nó chứa ít nhất điểm và mật độ thất nếu ngược lại. là ngưỡng mật độ do người dùng đặt.



Hình 11 Ví dụ tập điểm hàng xóm, các điểm mật độ cao và thấp

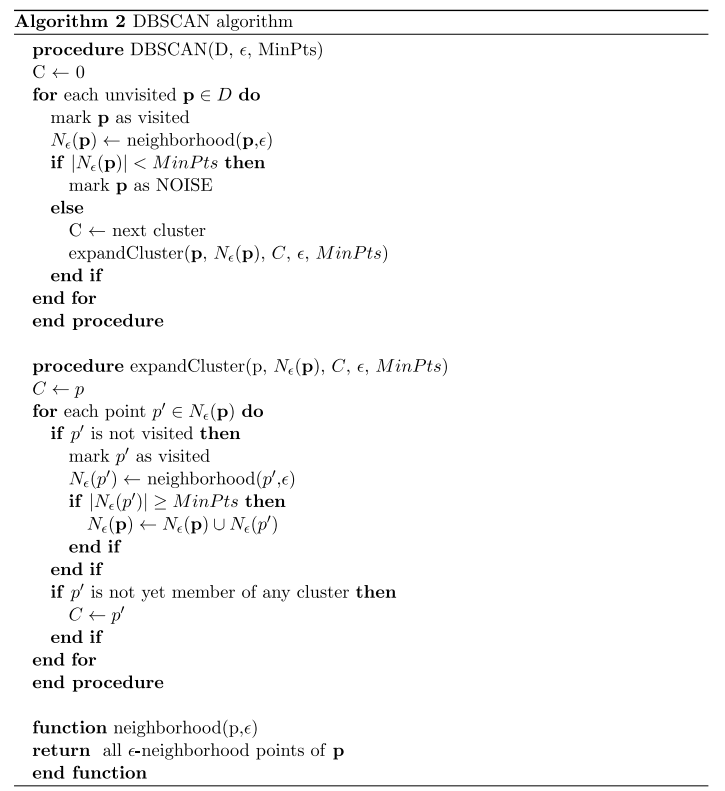
Điểm **p** là điểm lõi (core point) nếu nó có mật độ cao, có nghĩa là nó chứa ít nhất MinPts điểm trong tập hàng xóm của nó . Điểm biên là điểm có ít hơn MinPts điểm trong tập hàng xóm nhưng nằm trong tập hàng xóm của điểm lõi **p**.Cuối cùng, điểm không phải điểm lõi hoặc điểm biên được gọi là điểm nhiễu.



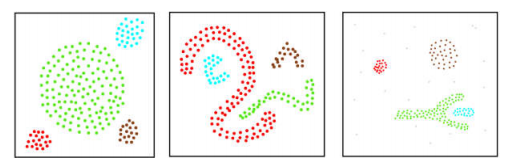
Hình 12 Ví dụ điểm lõi, điểm biên và điểm nhiễu

Ví dụ: , theo hình trên thì p là điểm lõi vì , q là điểm biên vì và nằm trong tập hàng xóm của p, t là điểm nhiễu.

a



Khởi tạo, tất cả các điểm được đánh dấu là chưa thăm. Thứ 1, tính tập hàng xóm của mỗi điểm với bán kính . Nếu tập hàng xóm không chứa ít nhất MinPts điểm, thì điểm đó được gán nhãn nhiễu. Ngược lại, bắt đầu một cụm C. Tiếp theo, tiến hành mở rộng cụm C để tìm tất cả các điểm có kết nối dày với x. Với mỗi kết nối dày với điểm , nếu là điểm lõi, thì tất cả các điểm hàng xóm của điểm kết nối dày được thêm vào cụm C. Ngược lại, nếu không là thành viên của bất kỳ cụm nào thì được thêm vào cụm C.



Hình 13 Ví dụ cụm được phân bởi DBSCAN

## Kết quả

# Tài liệu tham khảo

<http://kdientu.duytan.edu.vn/media/49682/le-dac-thinh-bao-cao-nckh.pdf>

Viet Phuong Le, logo detection, recognitn and spotting in context by matching logcal visual features