Lời cảm ơn

Lời giới thiệu

CHƯƠNG I

GIỚI THIỆU

Chữ Nôm là một loại hình văn tự do người Việt sáng tạo ra, lấy chất liệu là chữ Hán làm cơ sở để tạo thành chữ Nôm. Chữ Nôm được bắt đầu hình thành từ những năm đầu khi người Trung Hoa chinh phục đất Giao Chỉ, và được phát triển cùng thời gian. Nhưng đến đầu thế kỉ XX khi chữ quốc ngữ được phổ biến thì chữ nôm dần dần mai một. Ngày nay, số lượng người biết và có thể đọc được các tài liệu lịch sử, văn học bằng chữ Nôm còn rất ít.

Bên cạnh đó, với sự phát triển nhanh chóng của công nghệ thông tin trong các lĩnh vực: công nghệ phần cứng, công nghệ phần mềm, công nghệ nhúng, việc ứng dụng công nghệ thông tin vào việc bảo tồn di sản văn hóa chữ Nôm và Hán – Nôm là lĩnh vực nghiên cứu quan trọng.

Khái niệm

Thuật ngữ panorama xuất hiện trước khi con người có máy ảnh panorama. Nguyên gốc của thuật ngữ này là do họa sĩ người Ireland – Robert Baker – dùng để mô tả những bức tranh diện rộng ở Edinburgh (Đức).

Năm 1881, họa sĩ người Hà Lan Hendrik Willem Mesdag đã tạo nên trường pháo panorama Mesdag với những ống hình trụ cuốn các bức tranh toàn cảnh với những kích cỡ khổn lồ, cao 14m và dài có thể từ 40 – 120 m. Thế kỉ XIX, cí hai bức tranh panorama được coi là lớn nhất thời kì này, đó là bức tranh mô tả trận chiến tại Atlanta với chiều cao gần 13m, dài 110m và bức tranh ở Wroclaw(Ba Lan) với chiều cao là 15m và chiều dài là 120m.

Cùng với sự phát triển công nghệ thông tin, người ta đã tạo ra ảnh panorama từ nhiều bức ảnh nhỏ. Ảnh chụp panorama sẽ cho chúng ta những tấm ảnh với góc rộng lý tưởng.

*Ảnh panorama là ảnh được ghép từ những ảnh số chụp từng phần của một phong cảnh (những ảnh này có phần gối nhau)lại thành một ảnh toàn cảnh hoàn chỉnh.*

Ta có thể hiểu một cách đơn giản Panorama là chế độ chụp ảnh khổ rộng bằng cách chụp nhiều tấm ảnh liên tiếp, với thông tin của tấm ảnh trước được thể hiện một phần trong tấm ảnh sau, nhằm hỗ trợ người dùng. Sau đó với sự trợ giúp của phần mềm xử lý ảnh thì ta sẽ có được 1 tấm ảnh khổ rộng.

Hay nói một cách tổng quát panorama là một cách nhìn rộng của một vật chất trong không gian. Nó cho phép biểu thị một góc nhìn rộng của các bức tranh, bản vẽ đồ họa, nghệ thuật nhiếp ảnh, phim hoặc video, hay mô hình 3 chiều.

Panorama được chia làm nhiều loại, trong đó có các loại phổ biến như sau:

ảnh phẳng (flat): có chiều rộng lớn hơn nhiều lần chiều cao.

ảnh vertorama: ghép ảnh với nhau theo chiều dọc.

ảnh trụ (cylinder): chụp một vòng 360 độ để tạo ảnh panorama, sau đó dán nó vào mặt hình trụ.

ảnh cầu (sphere ): Ảnh được phủ lên khắp quả cầu tròn.

ảnh hộp (cube): Ảnh được phủ lên một hình hộp.

Các bước để tạo một ảnh panorama từ nhiều ảnh khác được thể hiện trong sơ đồ sau:

Trích chọn đặc trưng

So sánh điểm đặc trưng

Tính ma trận Homography

Hòa trộn ảnh

Mục đích của khóa luận này là tìm hiểu các kĩ thuật xử lý ảnh và xây dựng phương pháp tạo ảnh panorama để phục vụ cho việc nhận dạng chữ Nôm. Hai ảnh cùng đối tượng hay cùng cảnh được chụp bởi camera cầm tay, sau khi được xử lý như sơ đồ trên sẽ tạo thành ảnh panorama.

Chương 2 của khóa luận giới thiệu về các vấn đề liên quan đến tìm kiếm điểm đặc trưng của ảnh số, so sánh các điểm đặc trưng. Kĩ thuật và thuật toán được sử dụng trong quá trình xây dựng module tạo ảnh panorama.

Trong chương 3, tôi xin giới thiệu về ma trận homography, cách tính ma trận homography bằng thuật toán RANSAC, và kĩ thuật hòa trộn ảnh để tạo nên ảnh panorama.

Chương 4 của khóa luận là một số kết quả thử nghiệm của module tạo ảnh panorama.

Chương 5 là phần kết luận. Trong phần này, tôi xin tổng kết những gì đã đạt được, nêu lên những hạn chế của khóa luận và định hướng phát triển tiếp đề tài.

CHƯƠNG II

TRÍCH CHỌN ĐẶC TRƯNG

*Chương này trình bày một số phương pháp được sử dụng hiệu quả trong cách tiếp cận và trích chọn các điểm đặc trưng của đối tượng và sử dụng đặc trưng này trong việc so sánh các đối tượng và tạo ảnh panorama.*

1. Bài toán trích chọn đặc trưng

Không giống như con người, khả năng nhìn nhận ảnh của máy tính là rất hạn chế. Con người có thể dễ dàng hiểu được nội dung của mỗi bức ảnh còn với máy tính, mỗi bức ảnh là một ma trận các pixel. Việc tìm kiếm các đặc trưng và biểu diễn ảnh thông qua các điểm đặc trưng là một phương pháp giúp cho máy tính nhìn nhận tốt hơn. Bài toán trích chọn đặc trưng là vấn đề nghiên cứu cơ bản trong các lĩnh vực khác nhau của thị giác máy như tìm cặp điểm đối sánh, nhận dạng đối tượng, tính độ sâu… Đối với một ảnh đặc trưng có thể ở dưới dạng khác nhau như: điểm ảnh, đường thẳng, đường cong, biên … tùy thuộc vào đặc điểm và độ phức tạp của bài toán để lựa chọn các đặc trưng cho phù hợp.

Trong xử lý ảnh và thị giác máy, trích chọn đặc trưng là quá trình xử lý làm cho mẫu ban đầu được biến đổi thành các mẫu mới dễ dàng nhận dạng hơn. Các đặc trưng thường là các điểm đặc trưng có chứa thông tin về mức xám, kết cấu, đường nét, hoặc nội dung ảnh.

*điểm đặc trưng trong ảnh là một điểm ảnh có chứa nhiều thông tin hơn các điểm ảnh lân cận. Biểu diễn ảnh theo điểm đặc trưng sẽ cô động hơn, giảm được không gian tìm kiếm trong các bài toán ứng dụng.*

trích chọn đặc trưng có từ rất sớm trong lĩnh vực nghiên cứu thị giác máy và có rất nhiều ứng dụng. Trong đó có ba hướng ứng dụng quan trọng nhất đó là: đăng kí ảnh (image registration), nhận dạng khuôn mặt và theo dấu chuyển động. Đăng kí ảnh là bài toán liên quan đến sự tìm sự tương ứng giữa hai hay nhiều ảnh ở các khung nhìn khác nhau hoặc ở các thời điểm khác nhau và là bước đầu tiên trong nhiều kĩ thuật xử lý ảnh như lập thể hình học, trộn ảnh và các ứng dụng liên quan đến đối sánh mẫu. Các đặc trưng nổi bật của ảnh là yếu tố quan trọng để đăng kí ảnh thành công. Trong lĩnh vực nhận dạng khuôn mặt các đặc trưng là các vị trí mô tả các vị trí nối bật của đường nét khuôn mặt như mắt, mũi, miệng … Theo dấu ảnh là hướng ứng dụng thứ ba. Khi đối tượng chuyển động cần xác định các thành phần của đối tượng tại vị trí mới như thế nào so với vị trí cũ.

Các đặc trưng của ảnh được phân thành một số dạng như: đường biên, đường thẳng, viền tròn và các điểm góc.

Đường biên là chuỗi các vị trí mà tại đó có sự thay đổi mạnh mẽ về cường độ điểm ảnh. Đường biên đặc trưng không nhất thiết phải trùng với đường biên vật lý.

Đường thẳng là dải hẹp dọc theo đối tượng và có cùng cường độ điểm ảnh. Có thể coi đường thẳng này là hai đường biên song song.

Viền tròn là các đối tượng đặc có cường độ các điểm ảnh xấp xỉ nhau.

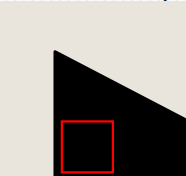
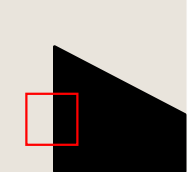
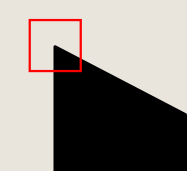
Điểm góc là vị trí gấp khúc của đường biên. Trong một cảnh, điểm góc có thể hiểu là điểm cuối của đường biên.

Trong bài toán tạo ảnh panorama này, mối liên hệ giữa các ảnh phải được chỉ rõ để biết được một điểm thuộc ảnh bên này sẽ tương ứng với điểm nào thuộc ảnh còn lại. Tuy nhiên, việc so sánh một điểm ảnh với tất cả các điểm ảnh của ảnh còn lại để tìm ra mối liên hệ giữa các ảnh là điểu không thể thực hiện được bởi độ phức tạp quá lớn. Vì thế việc tìm kiếm những điểm đặc trưng của ảnh, từ đó đối sánh các điểm đặc trưng là điều cần thiết trong bước đăng kí ảnh của bài toán tạo ảnh panorama.

Trong bài toán tạo ảnh panorama từ các ảnh, các điểm góc của các đối tượng trong ảnh là những điểm có nhiều khả năng tìm được điểm tương ứng với chúng trong ảnh còn lại.

Đã có rất nhiều nghiên cứu được đưa ra về vấn đề phát hiện điểm góc. Ban đầu các góc được phát hiện nhờ vào “độ nhọn” của đường biên: biên của đối tượng được lưu dưới dạng mã xích, góc được phát hiện thông qua việc tìm kiếm những vị trí trên biên bị “uốn” một cách đáng kể. Kĩ thuật phát hiện góc này rất phức tạp và phải triển khai trên nhiều bước. Sau đó, L.Kitchen đã đề xuất một phương pháp để cô lập điểm góc dựa vào một số phép toán trên Gradient, phương pháp này không phức tạp như dựa vào độ nhọn nêu trên.

Moravec đã đề xuất hàm phát hiện góc dựa trên việc dùng một cửa sổ đặt trên ảnh và xác định độ thay đổi trung bình cường độ xám khi dịch chuyển cửa sổ này theo bốn hướng. Nếu cường độ xám chỉ thay đổi theo hai hướng thì ta xác định đó là biên, còn nếu cường độ xám thay đổi theo bất cứ hướng nào thì ta xác định được đó là góc. Phương pháp này có một số nhược điểm như: cửa sổ trượt là hàm nhị phân, chỉ trượt được theo số hướng nhất định (bốn hướng). Harris và Stephens đã cải tiến phương pháp của Moravec bằng cách sử dụng hàm cửa sổ trượt là hàm Gaussian và có thể trượt theo bất kì hướng nào bằng cách áp dụng khai triển Taylor.

Trong hình 2.1(a), cửa sổ trượt nằm trên vùng hình ảnh không có góc cạnh, sau khi di chuyển cửa sổ trượt sẽ không có sự thay đổi về cường độ xám.

Trong hình 2.1(b), cửa sổ trượt nằm trên vùng hình ảnh có chứa cạnh, khi ta di chuyển cửa sổ trượt theo hai hướng sẽ có sự thay đổi về cường độ xám.

Trong hình 2.1(c), cửa sổ trượt nằm trong vùng hình ảnh có chứa cạnh, khi ta di chuyển theo bất cứ hướng nào đều có sự thay đổi về cường độ xám.

Dựa vào điều này ta có thể phát hiện ra điểm nào là góc, điểm nào không phải là góc.

1. Thuật toán tìm kiếm góc Harris

Giả sử ta có ảnh xám I với mỗi điểm (u, v) có độ dịch chuyển là (x, y) ta có thể tính toán sai lệch khi cửa sổ khung hình dịch chuyển từ (u, v) sang (u+x, v+y) theo công thức sau:

(2.1)



S(x, y) *là tổng bình phương độ lệch.*

w(u, v) *là cửa sổ trượt của khung hình tại vị trí (u, v).*

I(u, v) *và* I(u+x, v+y) *là cường độ xám của pixel tại các vị trí (u, v) và (u+x, v+y)*

Hàm cửa sổ trượt w(u, v) thay vì là hàm nhị phân và chỉ trượt được theo bốn hướng như trong thuật toán của Moravec, Harris đã thay hàm cửa sổ bằng hàm Gaussian. Để cửa sổ khung hình có thể trượt theo nhiều hướng khác nhau ta khai triển Taylor I(u+x, v+y), và kết quả là:



Ix, Iy *là các đạo hàm của I theo các thành phần x, y*

Khi đó công thứ (2.1) có thể viết lại như sau:



Nếu sự dịch chuyển giữa các cửa sổ khung hình là nhỏ thì có thể biểu diễn xấp xỉ dưới dạng ma trận là:



Trong đó, A là một cấu trúc Tensor:



Gọi λ1 và λ2 là các giá trị riêng của A. Khi đó:



CHƯƠNG III

TÍNH MA TRẬN HOMOGRAPHY BẰNG RANSAC

1. Homography

Homography là một định nghĩa của toán học. Đó 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 3 chiều có thể biến đổi về không gian ảnh bằng phép chiếu thông qua ma trận biến đổi Homography H. Các phép chiếu biến đổi thông qua ma trận Homography không đảm bảo về kích thước và góc của vật được chiếu, nhưng bảo đảm về tỷ lệ.

Trong lĩnh vực thị giác máy, Homography là một ánh xạ từ mặt phẳng đối tượng đến mặt phẳng ảnh. Hầu hết việc xử lý liên quan đến hai ảnh bất kì có cùng một phần chung đều liên quan đến ma trận Homography. Nó có rất nhiều ứng dụng như: sửa ảnh, tính toán sự chuyển động, xoay, dịch chuyển giữa hai ảnh.

Một đối tượng có thể được ánh xạ đến một mặt phẳng ảnh thông qua ma trận biến đổi Homography H.

**cX’ = HX**

c: *là tỉ lệ của phép chiếu và là hằng số khác 0.*

X’: *ảnh kết qua của phép ánh xạ.*

X: *đối tượng được ánh xạ.*

H: *ma trận Homography.*

Ma trận H là ma trận đã được chứng minh là khả nghịch. Nghĩa là, từ ảnh X’ muốn tái tạo ảnh X ta cần xác định ma trận Homography. Dưới đây em xin trình bày thuật toán DLT(*Direct Linear Transform*) dùng để tính ma trận Homography và cách chuẩn hóa thuật toán DLT.

Thuật toán DLT

Thuật toán DLT là một thuật toán đơn giản dùng để tính các phần tử của ma trận Homography từ một tập các cặp điểm tương ứng.

Trong tọa độ không đồng nhất, công thức (3.1) có thể được viết lại như sau:

Lần lượt chia dòng thứ nhất của công thức (3.2) cho dòng thứ ba và dòng thứ hai cho dòng thứ ba, ta được biểu thức sau:

Hai công thức (3.3) và (3.4) có thể được viết dưới dạng ma trận như sau:

Trong đó,

Và h = ( h1 h2 h3 h4 h5 h6 h7 h8 h9 )T

Với mỗi cặp điểm tương ứng ta có hai biểu thức, mặt khác do ma trận H là ma trận có D.O.F là 8, nên chỉ cần 4 cặp điểm tương ứng là ta có thể xác định được ma trận H.

Áp dụng công thức SVD(*Singular Value Decomposition*) cho ma trận A ta được:

Trong đó, là giá trị đặc biệt. sẽ được sắp xếp theo thứ tự giảm dần. Nghĩa là sẽ có giá trị nhỏ nhất.

Khi đó giá trị của hi bằng giá trị nhỏ nhất của cột vi.

Dưới đây là giải thuật của thuật toán DLT tính toán ma trận Homography cho ánh xạ trong không gian hai chiều:

Input:

n cặp điểm { } với

Output:

Ma trận Homography H, sao cho

Thuật toán:

* Với mỗi cặp điểm, tính toán ma trận như trong công thức (3.5)
* Từ n ma trận nhóm vào thành ma trận A có kích thước 2n x 9
* Tính SVD của ma trận A, khi đó hi bằng giá trị nhỏ nhất của vi
* Các phần tử của ma trận H là hi

Theo lý thuyết, với 4 cặp điểm tương ứng sẽ tìm được một ma trận H với D.O.F bằng 8. Tuy nhiên, trong thực tế không phải lúc nào các cặp điểm tương ứng cũng là chính xác, khi đó giá trị của ma trận H sẽ trở nên thiếu chính xác. Vì vậy, quá trình tính toán ma trận H đòi hỏi một bước chuẩn hóa.

Với 4 điểm tương ứng trong thuật toán DLT, Hartley và Zisserman đã đưa ra bước chuẩn hóa để đảm bảo rằng kết quả của thuật toán sẽ cho ra kết quả chính xác. Dưới đây là thuật toán DLT chuẩn hóa:

Input:

n cặp điểm { } với

Output:

Ma trận Homography H, sao cho

Thuật toán:

* Chuẩn hóa x:
* Chuẩn hóa x’:
* Áp dụng thuật toán DLT với các cặp điểm tương đồng đã chuẩn hóa.
* Ma trận H được tính bằng công thức

1. RANSAC

RANSAC là viết tắt của cụm từ “*RANdom SAmple Consensus*” có thể dịch là “đồng thuận mẫu ngẫu nhiên”. Được công bố bởi Fischler và Bolles vào năm 1981. Đây là thuật toán khử nhiễu khá hiệu quả.

Ý tưởng của thuật toán là: Từ tập dữ liệu đầu vào có hai loại dữ liệu là “*inlier*” và “*outlier*”, trong đó, “*inlier*” là các dữ liệu không phải nhiễu và “*outlier*” là các dữ liệu nhiễu, ta sẽ thu được một mô hình cần xây dựng phù hợp nhất với dữ liệu đầu vào, tập các dữ liệu là “*inlier*” và tập các dữ liệu là “*outlier*”.

Quá trình thực hiện của thuật toán RANSAC được môt tả như sau:

Từ tập dữ liệu đầu vào gồm có “*inlier*” và “*outlier*” ta chọn n dữ liệu ngẫu nhiên, tối thiểu để xây dựng mô hình.

Ta tiến hành xây dựng mô hình với n dữ liệu đó, sau đó đặt ra một ngưỡng dùng để kiểm chứng mô hình.

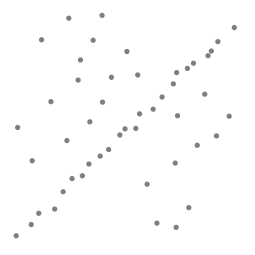
Ta gọi tập dữ liệu ban đâu trừ đi tập n dữ liệu để xây dựng mô hình là tập dữ liệu kiểm chứng. Ta sẽ tiến hành kiểm chứng mô hình đã xây dựng bằng tập dữ liệu kiểm chứng. Nếu kết quả thu được từ mô hình vượt quá ngưỡng, thì điểm đó là “outlier”, còn không đó sẽ là “inlier”.

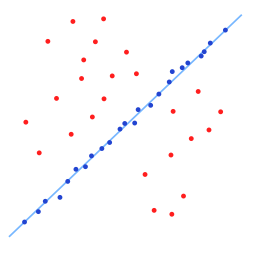
Quá trình này sẽ được lặp đi lặp lại trong k lần, tùy theo độ phức tạp của mô hình và yêu cầu chính xác của kết quả mà số k sẽ được thay đổi. Và kết quả là mô hình nào có số “inlier” nhiều nhất sẽ được chọn là mô hình tốt nhất.

Để làm rõ hơn thuật toán RANSAC em xin đưa ra một ví dụ được đăng trên Wikipedia như sau:

Cho tọa độ của N điểm trên mặt phẳng. Giả sử có một số điểm nằm trên một đường thẳng, còn lại là nhiễu. Yêu cầu đặt ra là xác định phương trình đường thẳng, tập các điểm được coi là nằm trên đường thẳng và tập các điểm là nhiễu.

Trong bài toán này, tập dữ liệu đầu vào là tọa độ của N điểm trên mặt phẳng. Dữ liệu đầu ra sẽ là một mô hình, mô hình này là một phương trình đường thẳng, tập các dữ liệu “*inlier*”, là tập các điểm được coi là nằm trên đường thẳng, và tập các dữ liệu nhiễu “outlier” không nằm trên đường thẳng.





Quá trình giải bài toán trên bằng thuật toán RANSAC như sau:

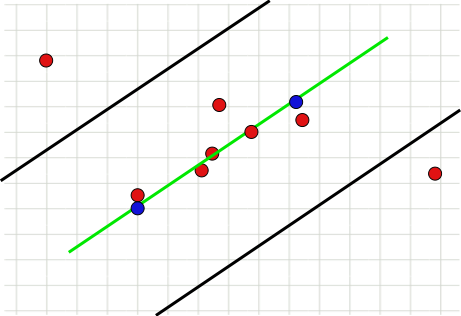
Chọn số lần lặp là k, tùy vào yêu cầu thực tế mà đặt k cho phù hợp.

Bước 1: Chọn hai điểm ngẫu nhiên từ tập N điểm trên mặt phẳng để xây dựng mô hình. Vì chỉ cần hai điểm là có thể xây dựng mô hình đường thẳng.

Bước 2: Đặt ra một ngưỡng nào đó. Trong bài toán này, ngưỡng được xác định là khoảng cách từ điểm bất kì đến mô hình đường thẳng đã xây dựng.

Bước 3: Từ tập N-2 điểm còn lại, tiến hành tính khoảng cách của từng điểm đến đường thẳng. Điểm nào có khoảng cách vượt quá ngưỡng sẽ được coi lại nhiễu.

Thực hiện lặp đi lặp lại các bước trên sau k lần, chú ý là ngưỡng của tất cả các lần lặp là như nhau. Khi đó, mô hình nào có nhiều “inlier” nhất sẽ là mô hình tốt nhất phù hợp với dữ liệu đầu vào.



Thuật toán:

input:

data: một tập dữ liệu ban đầu.

model: một mô hình phù hợp với data.

n: số lượng dữ liệu nhỏ nhất để thiết lập nên mô hình.

k: số lần lặp lại của thuật toán.

t: một giá trị ngưỡng để xác định mốc tính toán cho mô hình.

d: giá trị các điểm còn lại trong tập data đùng để kiểm chứng mô hình phù hợp với data( d = data - các điểm ngẫu nhiên )

output:

best\_model: các tham số mô hình tạo nên mô hình phù hợp nhất với data

best\_consensus\_set: Tập các điểm khớp với best\_model

best\_error: Tập các điểm outlier

iterations := 0

best\_model := nil

best\_consensus\_set := nil

best\_error := infinity

while(iterations < k){

maybe\_inliers := n chọn ngẫu nhiên n điểm từ data ban đầu.

maybe\_model là mô hình được tính toán từ tập maybe\_inliers đã chọn.

consensus\_set := maybe\_inliers

for(d){

nếu điểm đó có giá trị nhỏ hơn ngưỡng t thì thêm điểm đó vào tập consensus\_set

}

Nếu số lượng phần tử của consensus\_set > d (có nghĩa là đấy là mô hình tốt)

Thực hiện việc kiểm chứng mô hình.

This\_model := là mô hình phù hợp với tất cả các điểm trong consensus\_set.

This\_error :=

If(This\_error < best\_error){

Ta đã tìm được mô hình tốt nhất kể từ khi lặp. Giữ cho đến khi tìm được mô hình tốt hơn.

best\_model := this\_model

best\_consensus\_set := consensus\_set

best\_error := this\_error

}

Interations++;

}

return best\_model, best\_consensus\_set, best\_error

1. Tính ma trận Homography bằng RANSAC

CHƯƠNG IV

THỰC NGHIỆM