BỘ CÔNG THƯƠNG

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THỰC PHẨM TP. HCM**

**KHOA CÔNG NGHỆ THÔNG TIN**

---------------------------



**LUẬN VĂN ĐẠI HỌC**

**TÌM HIỂU MỘT SỐ PHƯƠNG PHÁP TRÍCH CHỌN ĐẶT TRƯNG CHO NHẬN DẠNG CHỮ VIẾT TAY**

**GVHD : Ths TRẦN VĂN THỌ**

**SINH VIÊN THỰC HIỆN:**

**VÒNG VĨNH SHÌN 2001140339 05DHTH2**

**NGUYỄN AN NHÀN**

*TP. HỒ CHÍ MINH, tháng 05 năm 2018*

**LỜI CAM ĐOAN**

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi. Các số liệu, kết quả nêu trong Luận văn là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Tôi xin cam đoan rằng mọi sự giúp đỡ cho việc thực hiện Luận văn này   
đã được cảm ơn và các thông tin trích dẫn trong Luận văn đã được chỉ rõ nguồn gốc.

**Sinh viên thực hiện Luận văn**

Vòng Vĩnh Shìn

Nguyễn An Nhàn

**LỜI CÁM ƠN**

Trong suốt quá trình làm đồ án, em đã nhận được nhiều sự giúp đỡ, đóng góp ý kiến và hướng dẫn nhiệt tình của thầy cô, bạn bè.

Lời đầu tiên em xin gửi lời cám ơn sâu sắc đến thầy Trần Văn Thọ, giảng viên chuyên ngành công nghệ phần mềm – Trường Đại Học Công Nghiệp Thực Phẩm TP.HCM, người đã hướng dẫn nhiệt tình và giúp đỡ chúng em trong suốt quá trình làm khóa luận.

Em xin gửi lời cảm ơn tới các thầy cô trong Ban Giám Hiệu – Trường Đại Học Công Nghiệp Thực Phẩm TP.HCM, các thầy cô giáo trong trường đã truyền đạt kiến thức, giúp chúng em có được cơ sở lý thuyết vững vàng và tạo điều kiện tốt nhất cho chúng em trong suốt quá trình làm khóa luận.

Chúng em cũng xin gửi lời cám ơn chân thành đến gia đình và bạn bè đã luôn là nguồn động viên to lớn, giúp đỡ chúng em vượt qua những khó khăn trong suốt quá trình học tập cũng như thực hiện khóa luận này.

Cám ơn các bạn thuộc lớp 05DHTH2 nói riêng và khóa 05DHTH nói chung đã hỗ trợ, giúp đỡ nhóm nhiệt tình.

Nhóm thực hiện

Vòng Vĩnh Shìn

Nguyễn An Nhàn

**TÓM TẮT**

Nhận dạng và xử lý ảnh là một trong những lĩnh vực có nhiều ứng dụng trong thực tiễn như: hệ thống thông tin địa lý, y học, quân sự… Hiện nay, dù công nghệ thông tin khá phát triển, nhưng trên những mặc khác, việc viết tay vẫn rất cần thiết để lưu trữ thông tin cũng như tính tiện lợi của nó. Cụ thể hơn trong lĩnh vực đào tạo, nhìn thấy được vấn đề trong việc lưu trữ thông tin điểm thi của sinh viên. Để thay thế quá trình nhập liệu điểm, ta cần phải có một phần mềm có thể nhận diện được bảng điểm và tách nguồn thông tin cần thiết để lưu trữ, điều này giúp ích cho việc tăng hiệu suốt công việc và khả năng linh hoạt khi lưu trữ

Khóa luận này tập trung vào tìm hiểu và nghiêm cứu các phương pháp trích chọn đặc trưng cùng trên các kỹ thuật xử lý hình ảnh. Để từ đó áp dụng vào việc tách và nhận dạng nội dung trên bảng điểm

**Chương 1:**

**Chương 2:**

**Chương 3:**

**Chương 4:**

**Chương 5:**

Mặc dù đã rất cố gắng nhưng do kiến thức còn nhiều hạn chế nên Luận văn này khó tránh khỏi những sai sót. Mong quý thầy cô và bạn bè đóng góp ý kiến để Luận văn được hoàn thiện tốt hơn

**NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN**

**NHẬN XÉT CỦA GIẢNG VIÊN PHẢN BIỆN**

**MỤC LỤC** (size 16)

**Danh mục các từ viết tắt**

**Danh mục các bảng**

**Danh mục các biểu đồ, đồ thị, sơ đồ, hình ảnh**

**CHƯƠNG 1:**

**TỔNG QUAN**

* 1. **TỔNG QUAN VỀ XỬ LÝ ẢNH**

**Xử lý ảnh** là một phân ngành trong xử lý số tín hiệu với tín hiệu xử lý là ảnh. Đây là một phân ngành khoa học mới rất phát triển trong những năm gần đây. Xử lý ảnh gồm 4 lĩnh vực chính: xử lý nâng cao chất lựong ảnh, nhận dạng ảnh, nén ảnh và truy vấn ảnh. Ngày nay xử lý ảnh đã được áp dụng rất rộng rãi trong đời sống như: photoshop, nén ảnh, nén video, nhận dạng biển số xe, nhận dạng khuôn mặt, nhận dạng chữ viết, xử lý ảnh thiên văn, ảnh y tế,....

**Việc xử lý phân tích ảnh nhằm mục đích khai thác giá trị nổi dung của ảnh**

* 1. **CÁC QUÁ TRÌNH XỬ LÝ ẢNH**

Hậu xử lý

Hệ quyết định

Lưu trữ

Đối sánh rút ra kết luận

Thu nhận ảnh (Scanner, Camera, Sensor)

Tiền xử lý

Trích chọn đặt trưng

Khối thu thập ảnh: Có nhiệm vụ tiếp nhận ảnh đầu vào.

Khối tiền xử lý: Có nhiệm vụ xử lý nâng cao chất lượng ảnh như giảm nhiễu, phân vùng, tìm biên, v.v…

Khối trích chọn đặc trưng: Có nhiệm vụ trích chọn các đặc trưng quan trọng của các bức ảnh đã được tiền xử lý để sử dụng trong hệ quyết định

Khối hậu xử lý: Có nhiệm vụ xử lý các đặc điểm đã trích chọn, có thể lược bỏ hoặc biến đổi các đặc điểm này để phù hợp với các kỹ thuật cụ thể sử dụng trong hệ quyết định

Khối hẹ quyết định và lưu trữ: có nhiệm vụ đưa ra quyết định (phân loại) dựa trên dự liệu đã học lưu trong khối lưu trữ

Khối kết luận: đưa ra kết luận dựa vào quyết định của khối quyết định

* 1. Một số khái niệm cơ bản:.
     1. Điểm ảnh

Điểm ảnh (Pixel) là một phần tử của ảnh số tại toạ độ (x, y) với độ xám hoặc màu nhất định. Kích thước và khoảng cách giữa các điểm ảnh đó được chọn thích hợp sao cho mắt người cảm nhận sự liên tục về không gian và mức xám (hoặc màu) của ảnh số gần như ảnh thật. Mỗi phần tử trong ma trận được gọi là một phần tử ảnh.

* + 1. Độ phân giải của ảnh

Định nghĩa: Độ phân giải (Resolution) của ảnh là mật độ điểm ảnh được ấn định trên một ảnh số được hiển thị.  
  
Theo định nghĩa, khoảng cách giữa các điểm ảnh phải được chọn sao cho mắt người vẫn thấy được sự liên tục của ảnh. Việc lựa chọn khoảng cách thích hợp tạo nên một mật độ phân bổ, đó chính là độ phân giải và được phân bố theo trục x và y trong không gian hai chiều.  
  
Ví dụ: Độ phân giải của ảnh trên màn hình CGA (Color Graphic Adaptor) là một lưới điểm theo chiều ngang màn hình: 320 điểm chiều dọc \* 200 điểm ảnh (320\*200). Rõ ràng, cùng màn hình CGA 12” ta nhận thấy mịn hơn màn hình CGA 17” độ phân giải 320\*200. Lý do: cùng một mật độ (độ phân giải) nhưng diện tích màn hình rộng hơn thì độ mịn (liên tục của các điểm) kém hơn.

* + 1. Ảnh màu:

Ảnh màu là ảnh mà tại mỗi điểm là một cấu trúc gồm nhiều kênh màu khác nhau (thường là 3 kênh màu)

Một số không gian màu

RGB (red green blue)

thường được sử dụng trong việc hiển thị trên máy tính

RGBA(red green blue alpha)

Alpha là kênh màu trong suốt.

CMYK (cyan magenta yellow black)

Thường dùng trên máy in

HSV(hue saturation vale)

(Không gian màu này còn có tên khác là HSI (intensity), HSL (lightness)

Mô hình này do AR Smith tạo ra vào năm 1978. Nó dựa trên các đặc tính màu trực quan như sắc (tint), bóng (shade) và và tông màu (tone); nói cách khác là họ màu, độ thuần khiết, và độ sáng

* + 1. Ảnh xám (grayscale)

hay còn gọi là ảnh đơn sắc (monochromatic), là ảnh mà tại mỗi điểm ảnh có một giá trị mức xám

Ảnh 8 mức xám sẽ có giá trị mỗi điểm ảnh nằm trong đoạn [0, 7].  
Ảnh 256 mức xám sẽ có giá trị mỗi điểm ảnh nằm trong đoạn [0, 255].

Giá trị điểm ảnh = 0 nghĩa là điểm ảnh đó tối (đen), giá trị điểm ảnh lớn nhất nghĩa là điểm ảnh đó trắng. Nói cách khác, giá trị mỗi điểm ảnh càng lớn thì điểm ảnh đó càng sáng

Cường độ sáng được tính theo công thức (chuyển đổi từ RGB sang grayscale):  
**Độ sáng = 0.2989R + 0.5870G + 01.1140B**

các hệ số có thể được làm tròn thành 0.3, 0.59 và 0.11.

* + 1. Ảnh nhị phân:

Giá trị xám của tất cả các điểm ảnh chỉ nhận giá trị 1 hoặc 0 như vậy mỗi điểm ảnh trong ảnh nhị phân được biểu diễn bởi 1 bit. Ảnh xám: Giá trị xám nằm trong [0, 255] như vậy mỗi điểm ảnh trong ảnh nhị phân được biểu diễn bởi 1 byte.

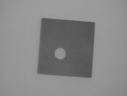
CHƯƠNG 2:

MỘT SỐ KỸ THUẬT TIỀN XỬ LÝ ẢNH

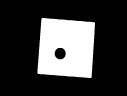
2.1. Tách ngưỡng:

Nguỡng (Threshold) là một khái niệm khá quen thuộc trong xử lý ảnh cũng như rất nhiều giải thuật khác. Nó dùng để chỉ một giá trị mà người ta dựa vào để phân hoạch một tập hợp thành các miền phân biệt.

Giá trị ngưỡng thường được xác định dựa vào những điểm đặc biệt (ví dụ ở trung bình), dựa vào kinh nghiệm khảo sát. Nếu dựa vào số lượng Ngưỡng áp dụng cho cùng một tập dữ liệu người ta sẽ phân ra các phương pháp ứng dụng ngưỡng đơn, ngưỡng kép, hay đa ngưỡng. Nếu dựa vào sự biến thiên của giá trị Ngưỡng, trong cùng phạm vi ứng dụng người ta sẽ phân ra các phương pháp dùng ngưỡng cố định (Constant|Fixed Threshold) và không cố định (Adaptive Threshold). Ngưỡng không cố định nghĩa là giá trị của nó sẽ thay đổi tùy theo sự biến thiên của tập dử liệu theo không gian và thời gian. Thông thường giá trị này được xác định thông qua khảo sát tập dử liệu bằng phương pháp thống kê. Để dễ hình dung hơn về ứng dụng khái niệm Threshold, sau đây chúng ta sẻ xét một ví dụ bộ lọc ngưỡng (Threshold Filter) đơn giản trong xử lý ảnh.

[](http://4.bp.blogspot.com/_b1r1UJs_aVQ/SyEQrj6aT6I/AAAAAAAAAK8/f69bd5rft1U/s1600-h/wdg2.GIF)

Với mỗi pixel trong hình đa mức xám (grayscale) ở trên giá trị sẽ trong khoảng 0 - 255 vậy pixel nào lớn hơn ngưỡng là 120 ta gán giá trị cho nó thành đen (0), ngược lại gán giá trị trắng (255). Kết quả thu được như sau:

[](http://4.bp.blogspot.com/_b1r1UJs_aVQ/SyEQ8iPGV6I/AAAAAAAAALE/wc-vPBI0Zp8/s1600-h/wdg2thr3.GIF)

2.2. Xử lý histogram

Lược đồ mức xám của một ảnh là một hàm cung cấp tần suất xuất hiện của mỗi mức xám (grey level).

Lược đồ xám được biểu diễn trong một hệ toạ độ vuông góc x,y. Trong hệ toạ độ này, trục hoành biểu diễn số mức xám từ 0 đến N, N là số mức xám (256 mức trong trường hợp chúng ta xét). Trục tung biểu di n số điểm ảnh cho một mức xám (số điểm ảnh có cùng mức xám). Cũng có thể biểu diễn khác một chút: trục tung là tỷ lệ số điểm ảnh có cùng mức xám trên tổng số điểm ảnh.

Số điểm ảnh

Mức xám

Số điểm ảnh

Mức xám

Lược đồ xám cung cấp rất nhiều thông tin về phân bố mức xám của ảnh. Theo thuật ngữ của xử lý ảnh gọi là tính động của ảnh. Tính động của ảnh cho phép phân tích trong khoảng nào đó phân bố phần lớn các mức xám của ảnh: ảnh rất sáng hay ảnh rất đậm. Nếu ảnh sáng, lược đồ xám nằm bên phải (mức xám cao), còn ảnh đậm luợc đồ xám nằm bên trái(mức xám thấp).

2.2.1. Cân bằng histogram

Ở những vùng tối giá trị điểm ảnh sẽ không thay đổi nhiều , tuy nhiên với vùng sáng giá trị mức sáng được đẩy lên với bước đẩy tăng dần. Với phép biến đổi này sự khác nhau về giá trị xám của các điểm ảnh tăng lên đủ để có thể phân biệt được Qua quan sát kết quả thu được nhận thấy: Phép biến đổi này thực hiện tốt đối với những ảnh tối hoặc quá sáng.

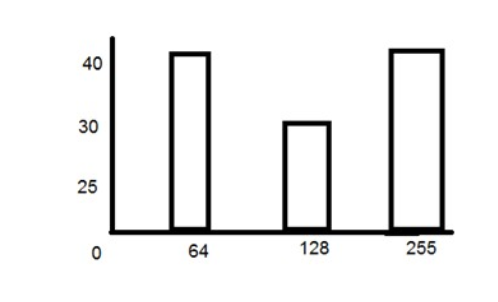
Việc cân bằng histogram sẽ giúp Hiệu chỉnh, làm cân bằng sự phân bố các giá trị bậc sáng trong bức ảnh.

* **Cân bằng histogram toàn cục:**

Để thực hiện cân bằng mức xám ta cần thực hiện qua 4 bước sau.

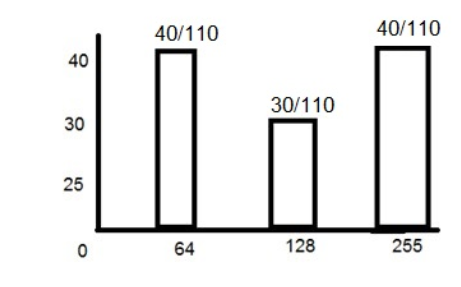
**Bước 1: Probability Mass Function (PMF) (Tính xác xuất)**

Ta xét ví dụ với Histogram sau:

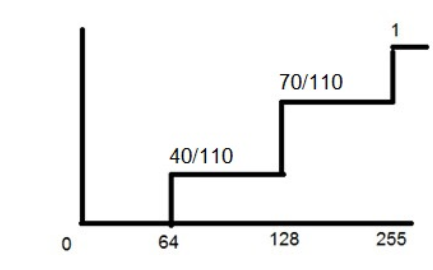


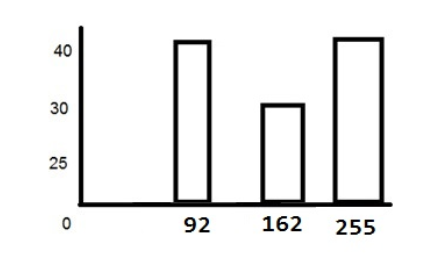
Ta thấy h(64) = 40, h(128) = 30, h(255) = 40. Và có tất cả 110 điểm ảnh.

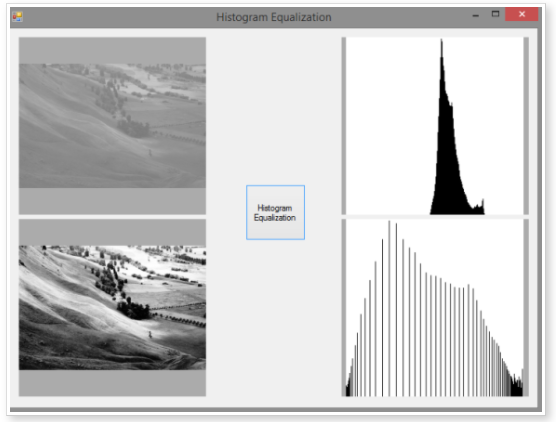
Trước tiên ta cần tính được xác xuất xuất hiện của các mức xám trong ảnh.



Sau bước 1 ta có tần xuất xuất hiện hay xác xuất của các mức xám tương ứng.

p(64) = 40/110  
p(128) = 30/110  
p(255) = 40/110  
**Bước 2: Cumulative Distributive Function** Bước này chỉ đơn giản là cộng dồn tần xuất của mức xám hiện tại với tần xuất của các mức xám trước nó. Ta có CDF tương ứng với PMF phía trên như sau:  
Ta thấy:  
c(64) = 40/110  
c(128) = 30/110 + 40/110 = 70/110  
c(255) = 40/110 + 70/110 = 1  
**Bước 3: Multiplize**Trong bước này chúng ta sẽ nhân CDF của mức xám với giá trị max của mức xám ở đây là 255.  
Tương ứng với ví dụ trên ta có kết quả như sau:

m(64) = 40/110 \* 255 = 92  
m(128) = 70/110 \* 255 = 162  
m(255) = 1 \* 255 = 255  
Kết quả sau khi nhận được bước này là độ xám mới của các mức xám.  
**Bước 4: Mapping**  
Bước này chúng ta chuyển đổi tất cả các pixel có độ xám mới.  
Điểm có mức xám 64 chuyển hết thành 92,  
Điểm có mức xám 128 chuyển hết thành 162,  
Điểm có mức xám 255 chuyển hết thành 255.  




Nguồn : https://namlunoy.wordpress.com/2015/01/25/can-bang-muc-xam-histogram-equalization/

**Cân bằng histogram cục bộ**

* Thay vì cân bằng trên toàn bộ bức ảnh, thuật toán này lại quan tâm đến từng phần của bức ảnh.
* Giúp những phần quá tối, quá sáng của bức ảnh được cân bằng tốt hơn.
* Hạn chế của phương pháp là phải lặp đi lặp lại nên việc tính toán trên những phân vùng nhỏ nên thời gian tính toán rất lâu

2.3. Bộ lọc trong miền không gian

Lọc không gian thông thường được thực hiện để khử nhiễu hoặc thực hiện một số kiểu nâng cao ảnh. Các thao tác này được gọi là lọc không gian để phân biệt chúng với lọc tần số

Có 2 kiểu lọc thông dụng được trình bày ở đây là:

- Lọc trung bình.

- Lọc trung vị.

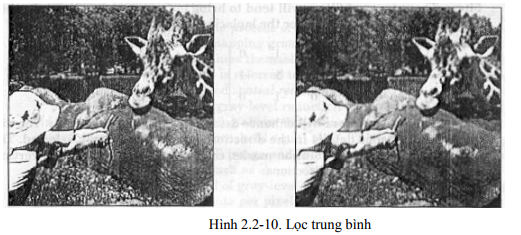
Ngoài ra còn có các bộ lọc khác như :

- Lọc nâng cao.

- Lọc Max

- Lọc Min

Hai kiểu lọc đầu tiên chủ yếu được sử dụng để che đậy hoặc khử nhi u, mặc dù chúng cũng có thể được sử dụng cho các ứng dụng cụ thể khác. Chẳng hạn, một bộ lọc trung bình làm cho ảnh "mịn hơn" như hình 2.2-10. Các bộ lọc nâng cao thường để làm sáng các cạnh và các chi tiết nằm trong ảnh



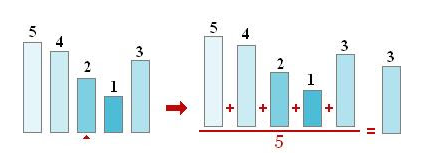
Nhiều bộ lọc không gian được thực hiện thông qua phép cuộn. Bởi vì thao tác mặt nạ cuộn cung cấp kết quả là tổng trọng số của các giá trị của một pixel và các láng giềng của nó, nên được gọi là là một bộ lọc tuyến tính.

Một khía cạnh thú vị của các mặt nạ cuộn là hiệu ứng tổng thể có thể dự đoán được dựa trên mẫu của chúng. Chẳng hạn, nếu các hệ số của mặt nạ có tổng bằng 1, thì độ sáng trung bình của ảnh sẽ được giữ nguyên. Nếu tổng các hệ số bằng 0, thì độ sáng trung bình sẽ mất đi và trả về một ảnh tối. Hơn nữa, nếu các hệ số có cả âm và dương, thì mặt nạ là một bộ lọc trả về chỉ thông tin cạnh, còn nếu các hệ số đều dương, thì nó là một bộ lọc làm mờ (blur) ảnh the image.

**Lọc trung bình: (tìm hiểu thêm về các mặt nạ cuộn cung như áp dụng)**

Lọc trung bình là kĩ thuật lọc tuyến tính, hoạt động như một bộ lọc thông thấp. Ý tưởng chính của thuật toán lọc Trung vị như sau: ta sử dụng một cửa sổ lọc (ma trận 3x3) quét qua lần lượt từng điểm ảnh của ảnh đầu vào input. Tại vị trí mỗi điểm ảnh lấy giá trị của các điểm ảnh tương ứng trong vùng 3x3 của ảnh gốc "lấp" vào ma trận lọc. Giá trị điểm ảnh của ảnh đầu ra là giá trị trung bình của tất cả các điểm ảnh trong cửa sổ lọc. Việc tính toán này khá đơn giản với hai bước gồm tính tổng các thành phần trong cửa sổ lọc và sau đó chia tổng này cho số các phần tử của cửa sổ lọc.

Sẽ dễ hình dung hơn bằng mô tả trong hình dưới đây:



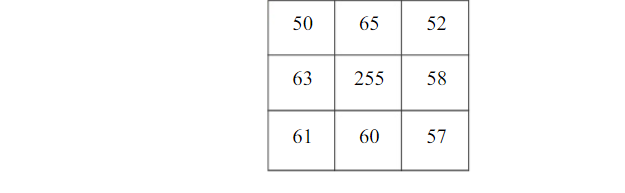
Sơ lược một cách ngắn gọn các bước của giải thuật:

1. Quét cửa sổ lọc lần lượt lên các thành phần của ảnh đầu vào; điền các giá trị được quét vào cửa sổ lọc.
2. Xử lý bằng cách thao tác trên các thành phần của cửa sổ lọc.
3. Tính giá trị trung bình các thành phần trong cửa sổ lọc.
4. Gán giá trị trung bình này cho ảnh đầu ra.

**Lọc trung vị**

Lọc Trung vị là một kĩ thuật lọc phi tuyến (non-linear), nó khá hiệu quả đối với hai loại nhiễu: nhiễu đốm (speckle noise) và nhiễu muối tiêu (salt-pepper noise). Kĩ thuật này là một bước rất phổ biến trong xử lý ảnh.

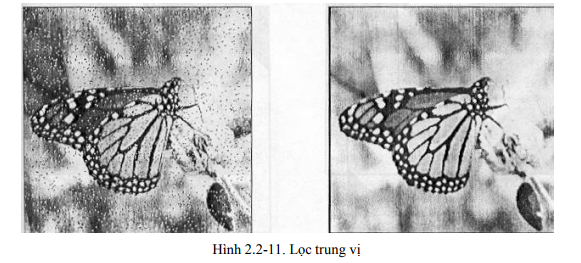
Ý tưởng chính của thuật toán lọc Trung vị như sau: ta sử dụng một cửa sổ lọc (ma trận 3x3) quét qua lần lượt từng điểm ảnh của ảnh đầu vào input. Tại vị trí mỗi điểm ảnh lấy giá trị của các điểm ảnh tương ứng trong vùng 3x3 của ảnh gốc "lấp" vào ma trận lọc. Sau đó sắp xếp các điểm ảnh trong cửa sổ này theo thứ tự (tăng dần hoặc giảm dần tùy ý). Cuối cùng, gán điểm ảnh nằm chính giữa (Trung vị) của dãy giá trị điểm ảnh đã được sắp xếp ở trên cho giá trị điểm ảnh đang xét của ảnh đầu ra output.



Các giá trị sau khi sắp xếp



Kết quả là giá trị 255 sẽ được thay thế bằng giá trị 60



2.4. Xử lý hình thái ảnh

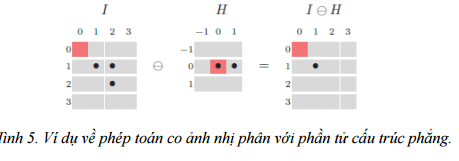
2.4.1. Phép co ảnh và giãn ảnh đối với ảnh nhị phân

Phép co ảnh

Xét tập hợp A và tập hợp B (Phần tử cấu trúc), phép co ảnh nhị phân của tập hợp A bởi phần tử cấu trúc B được kí hiệu A B và viết dưới dạng công thức như sau:

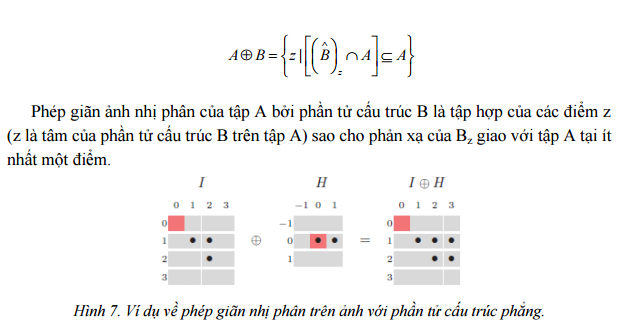
A B ={z| A}

Phép co ảnh nhị phân của tập hợp A bởi phần tử cấu trúc B là tập hợp các điểm z (z nằm ở tâm điểm của phần tử cấu trúc B) sao cho là tập con của A.



Phép giản ảnh:

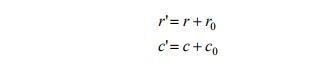
Gọi A là ảnh gốc, B là một phần tử cấu trúc. Phép giãn nhị phân của ảnh A với phần tử cấu trúc B được kí hiệu A⊕ B và chúng ta có thể biểu diễn phép toán co ảnh dưới dạng phép toán tổ hợp như sau:



2.4.2. **Tịnh tiến và quay ảnh:**

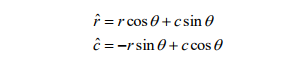
Hai thao tác đáng chú ý khác trong hình học ảnh là tịnh tiến và quay.

Quá trình tịnh tiến có thể được thực hiện với các phương trình sau:



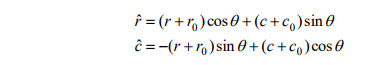
trong đó r’ và c’ là các toạ độ mới, r và c là các toạ độ ban đầu, và r0 và c0 là các khoảng cách cần dịch chuyển (tịnh tiến) ảnh.

Quá trình quay cần sử dụng các phương trình sau:



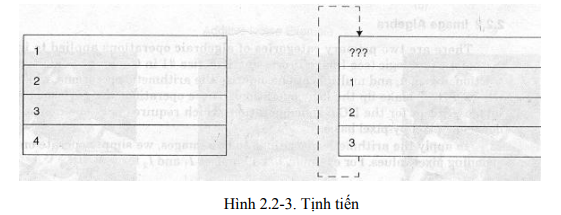
trong đó r ˆ ,c ˆ là các toạ độ mới, r và c là các toạ độ ban đầu, θ là góc quay. θ được định nghĩa theo chiều quay kim đồng hồ tính từ trục hoành trong ảnh mà gốc toạ độ nằm ở góc trái trên.

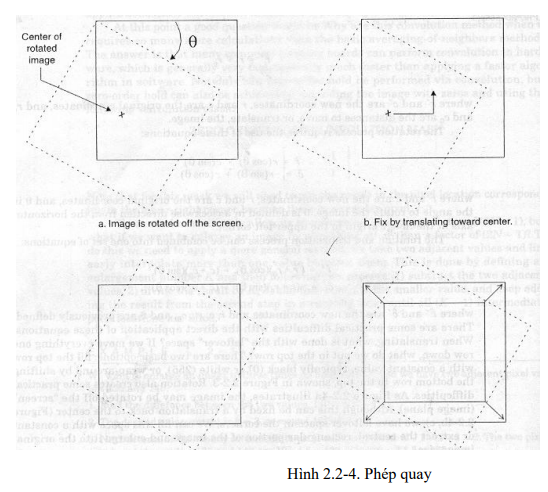
Cả hai quá trình quay và tịnh tiến có thể được tổ hợp vào thành cùng một phương trình như sau:



với r ˆ ,c ˆ là các toạ độ mới, r, c, r0, c0 và θ như đã định nghĩa ở trên.

Có một số khó khăn thực tế khi áp dụng trực tiếp các phương trình trên. Khi tịnh tiến, sẽ làm gì với không gian “thừa”. Nếu ta dịch chuyển mọi thứ trên hàng đi xuống, thì ta sẽ đặt cái gì vào hàng trên cùng? Có hai tuỳ chọn cơ bản: tô đầy hàng trên cùng với một giá trị không đổi, thường là đen (0) hoặc trắng (255), hoặc quấn lại bằng cách dịch hàng dưới đáy lên trên cùng, như hình Hình 2.2-3. Phép quay cũng tương tự, như hình 2.2-4a minh hoạ ảnh có thể được quay rotated off "screen" (mặt phẳng ảnh). Mặc dù điều này có thể cố định lại bằng cách tịnh tiến trở lại tâm (Hình 2.2-4b, c), nhưng ta vẫn có những không gian thừa ở các góc. Ta có thể tô đầy không gian này bằng hằng số, hoặc cắt ra phần trung tâm là phần hình chữ nhật của ảnh rồi mở rộng ra kích thước ảnh ban đầu



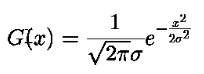


2.5 Phương pháp phát hiện biên

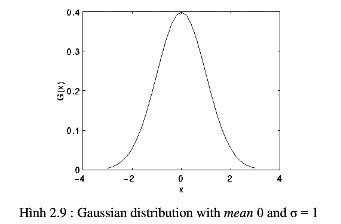
2.5.1. Phương pháp làm nổi biên ảnh

Gaussian blurring là cách làm mờ một ảnh bằng hàm Gaussian. Phương pháp này được ứng dụng rộng rãi và hiệu quả trong các phần mềm xử lý đồ họa. Nó cũng là công cụ phổ biến để thực hiện quá trình tiền xử lý hình ảnh dùng để làm dữ liệu đầu vào tốt cho các phân tích cao cấp hơn trong Computer Vision, hoặc cho các giải thuật được thực hiện trong một tỉ lệ khác của hình đã được cho.

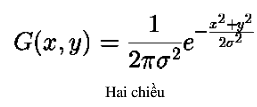
Như vậy, ta có thể nói Gaussian blurr là một loại bộ lọc làm mờ ảnh, sử dụng lý thuyết hàm Gaussian(cũng được biết đến như là dạng phân tán chuẩn (Normal Distribution) trong thống kê) để tính toán việc chuyển đổi (Transformation) mỗi điểm ảnh của hình, giúp làm giảm nhiễu và mức độ chi tiết (không mong muốn) của hình ảnh. Đây là phương trình hàm Gaussian (Gaussion Distribution) trong không gian một chiều

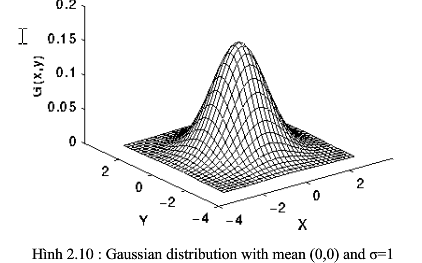


trong đó σ phương sai chuẩn của phân tán Gaussian

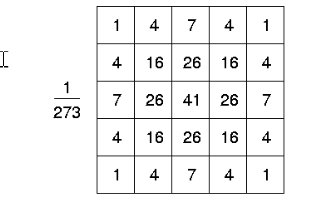


Và trong không gian hai chiều





Với x, y là tọa độ theo hai trục đứng và ngang. Trong không gian hai chiều, công thức này sản sinh ra những đường viền là những đường tròn đồng tâm, tuân theo logic phân tán Gausian từ điểm trung tâm. Giá trị từ hệ thống phân tán này sẽ được sử dụng để xây dựng một ma trận tích chập dùng tính toán phép tích chập với ảnh gốc.



Hình 2.11: Discrete approximation to Gaussian function with σ =1.0

Giá trị mới sau khi tính tích nhân chập với cửa sổ (kernel) đại diện cho hàm Gaussian có thể coi là trung bình lượng giá của các điểm ảnh xung quanh nó. Ta thấy rằng giá trị lượng của phần tử trung tâm kernel tương ứng với điểm ảnh đang xét là lớn nhất, giá trị này sẽ nhỏ hơn đối với các phần tử tương ứng với các điểm ảnh kế cận một cách đối xứng và tỉ lệ thuận với khoảng cách của phần tử này với trung tâm. Tính chất này giúp giữ lại đường viền và đường biên cũng như làm mờ một cách đồng bộ hơn so với các phương pháp khác.

Trong lý thuyết, hàm Gaussian tại mỗi điểm trên hình là khác 0. Điều này có nghĩa là Gaussian kernel nên có kích thước bằng với hình ảnh và giá trị tại mỗi phần luôn khác 0. Tuy nhiên trong thực hành, do việc tính toán dựa trên xấp xỉ rời rạc (Discrete Appoximation) cho nên giá trị của các phần tử trên bề mặt Gaussian ở khoảng cách lớn hơn 3σ so với trung tâm gần như không đáng kể (tiệm cận 0). Do vậy các Gaussian distribution ngoài bán kính này sẽ bị bỏ qua, đó cũng là lí do mà thông thường Gaussian kernel có kích thước giới hạn 3, 5, 7..(Cái này tùy thuộc vào giá trị phương sai chuẩn mà bạn chọn). Khoảng cách giữa hai điểm gần nhau trong Gausssian kernel là σ

Do đặc tính cấu trúc như những hình tròn đối xứng, với hình hai chiều, Gaussian blur có thể được áp dụng như là hai phép tính toán một chiều độc lập (Độc lập tuyến tính- Linear Separable). Có nghĩa là hiệu quả thu được từ tính toán trên ma trận hai chiều có thể tương đương với cách ứng dụng tính toán một loạt các Gaussian một chiều theo hướng ngang và đứng. Điều này rất hữu ích trong việc giảm chi phí tính toán và thể hiện khá rõ trong hai công thức tương ứng sau:

**O(ωkernelhkernelωimagehimage)** và **O(ωkernelωimagehimage)+O(hkernelωimagehimage)**

Một vấn đề nữa cần quan tâm của Gaussian blur là nếu chúng ta sử dụng lần lượt nhiều Gaussian cho một ảnh thì kết quả cũng tương đương với khi bạn dùng một Gaussian lớn hơn có bán kính bằng căn bậc hai tổng bình phương các bán kính của các Gaussian đã dùng, ví dụ: . Cũng vì mối quan hệ này mà thời gian tính toán sẽ không được tiết kiệm khi bạn chia nhỏ các Gaussian. Do đó khuyến cáo nên sử dụng Gaussian đơn trong tính toán nếu có thể.

Ngoài ra Gaussian blur cũng được sử dụng để giảm kích thước hình ảnh. Khi tiến hành xử lý giảm tỉ lệ lấy mẫu tín hiệu cho ảnh (Downsampling), người ta thường áp dụng bộ lọc low pass filter trước khi tái lấy mẫu. Điều này là để chắc chắn những thông tin không mong muốn, tần số cao sẽ không xuất hiện trong hình đã được Downspamling.

2.5.2. Kỹ thuật tách biên Canny

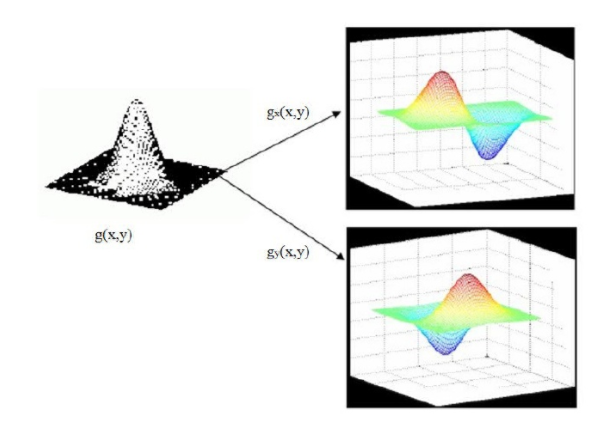
Kỹ thuật phát hiện biên bằng phương pháp canny là phương pháp dò biên trực tiếp rất hiệu quả áp dụng cho loại ảnh nhiễu Phương pháp này do John Canny. ở phòng thí nghiệm MIT khởi xướng năm 1986. Bộ tách biên Canny là bộ tách biên mạnh nhất cung cấp bởi hàm edge.

Bước 1: Làm trơn ảnh bằng cách nhân chập với bộ lọc Gauss. Sau khi tiến hành nhân chập chúng ta có đầu ra là dữ liệu ảnh S đã được nhân chập.

Bước 2: Lấy đạo hàm bậc nhất của kết quả ở bước 1.

Ta có: ∇s = ∇(g\*1)= (∇g)\*1

Như vậy, kết quả ảnh bước hai chính là sự tổng hợp của đạo hàm của Gauss theo hướng x nhân với ảnh 1 và đạo hàm của Gauss theo hướng y nhân với ảnh I. Nghĩa là ta có thể đạo hàm hàm Gauss theo hai hướng rồi mới tiến hành nhân xoắn với ảnh thay vì nhân xoắn ảnh với hàm Gauss rời mới đạo hàm. Có thể minh họa như sau đạo hàm Gauss theo hai hướng x và y như sau:



Như vậy cách thức thực hiện bước thứ hai như sau:

Sau khi tiến hành nhân xoắn ảnh I với bộ lọc Gauss ở bước 1 ta có một ảnh mới S được làm trơn. Tiến hành thực hiện bước hai bằng cách lấy đạo hàm ảnh mới đó theo hai hướng x và y rồi tổng hợp kết quả lại như đã biết, phương pháp Gradient là phương pháp dò biên cục bộ dựa vào cực đại của đạo hàm, đó chính là phương pháp đạo hàm bậc nhất. Chính vì vậy ta có thể thực hiện việc đạo hàm ở bước 2 bằng cách nhân ảnh kết quả S ở bước 1 với các mặt nạ trong phưuong pháp Gradient dựa theo các toán tử như Sobel, Pixel, Difference. Ở đây ta tiến hành nhan xoắn ảnh S với hai mặt nạ của phương pháp Sobel theo hai hướng x và y như sau:



Sau khi tiến hành nhân xoắn ảnh theo hai hướng x và y ta được hai ảnh theo hai hướng là Sx và Sy, ta tiến hành tổng hợp hai kết quả đó để cho ra kết quả cuối cùng S’:



Với  là hướng của đường biên

Bước 3: Loại bỏ một số điểm dư thừa:

Đối với mỗi điểm ảnh trên ảnh S’ ta tiến hành so sánh giá trị của điểm đó với giá trị của hai điểm lân cận điểm đó. Hai điểm lân cận này là hai điểm nằm trên đường thằng chứa hướng của đường biên . Công thức tính hướng của đường biên  nằm ở bước

Bước 3:

**CHƯƠNG 2:**

**PHƯƠNG PHÁP TRÍCH CHỌN ĐẶT TRƯNG**

**2.1. Trích chọn đặc trưng**

**1.1. Đặc trưng bất biến**

Để nhận dạng các dạng biến đổi của cùng 1 ký tự, đặc trưng bất biến được sử dụng như là tính chất để đảm tính chắc chắn việc biến đổi vẫn có giá trị sử dụng. Bất biến là một đặc trưng mà các vật mẫu sẽ có cùng giá trị tương đương

Ví dụ: dịch chuyển, co dãn, xoay, thay đổi kích thước, nghiên hoặc phản chiếu, Tuy nhiên không phải tất cả thay đổi của ký tự đều được nhận diện nhờ vào đặc trưng bất biến.

Đặc trưng bất biến kích thước và dịch chuyển dề dàng được phát hiện, Việc phân vùng các ký tự riêng biệt cũng có thể cung cấp giá trị ước lượng của kích thước và vị trí. Nhưng phương pháp trích chọn đặc trưng vẫn thường xuyên cung cấp tính chính xác cao về giá trị ước lượng

Bất biến định hướng rất quan trọng trong việc nhận biết ký tự từ bất kỳ hướng nào. Tuy nhiên nếu tất cả ký tự điều cùng định hướng, thì việc sử dụng đặc trưng định hướng có thể nhận diện được sự khác biệt giữa '6' và '9'. Một khả năng khác để sử dụng đặc trung định hướng là nhận diện góc xoay của ký tự.

Nếu góc xoay bị hạn chế, ta có thể ước lượng sai giữa -45o và 45o. Khi đó việc xoay 180o với mỗi ký tự sẽ khác nhau.

Đối với đặc trưng bất biến kích thước được sử dụng để nhận diện sử khác biệt giữa ',' và '9', '.' , 'o' và '0'

Đặc trưng bất biến – nghiêng có thể hữu dụng đối với chữ viết tay khi ký tự đó có độ nghiên, và đối với chữ đánh máy có sử dụng font nghiên.

Bất biến – đối chiếu hình ảnh thì không được mong đợi trong việc nhận diện ký tự, đối chiếu hình ảnh thường tạo ra các ký tự không hợp lý hoặc 1 ký tự khác

Đối với trích chọn đặc trưng từ hình ảnh đa cấp xám, bất biến sẽ làm rõ sự tương phản đối với chữ in và nền, từ đó ta thấy được sử cần thiết của hình ảnh đa cấp xám,...

Nếu như đặc trưng bất biến không thể tìm thấy, một lựu chọn khác là tiêu chuẩn hóa đầu vào của hình ảnh để ta có 1 kích thước chuẩn, góc xoay, phản chiếu. Tuy nhiên, 1 điều cần lưu ý là điều này sẽ nãy sinh ra lỗi rời rạc

**1.2. Khả năng khôi phục**

Đối với phương pháp trích chọn đặc trưng, ký tự có thể được tái cấu trúc lại. Cách thức này sẽ đảm bảo tính toàn vẹn thông tin của ký tự

Mặc dù đối với nhiều phương pháp, tái cáu trúc thực sử cần nhiều đặc trưng, nhưng đối với kết quả xấp xĩ gần đúng của ký tự gốc, ta chỉ cần 1 số lượng đặc trưng có tính tinh cậy cao, nhưng đặc trưng này cần được quan tâm hơn.

Bằng cách tái cấu trúc của các ký tự hình ảnh bằng phương pháp trích gọn đặc trưng, ta cần 1 cuộc kiểm tra bằng 1 số lượng đặc trưng để nhận diện được cấu trúc cần thiết của ký tự, tái cấu trúc được sử dụng để đảm bảo việc thực hiện nhận dạng là đúng

**2.2. Trích chọn đặc trưng từ ảnh đa cấp xám**

Vấn đề đặc ra của các phương thức trong ảnh đa cấp xám là có thể xác định được vị trí của ký tự. Có thể dùng phương thức nhị phân hóa thích nghi cục bộ để tạo ra ảnh nhị phân raster và sử dụng các thành phần liên kết với kích thước ký tự dự kiến để xác định các ký tự. Các phương thức trong ảnh đa cấp xám sẽ được sử dụng khi mà việc nhận diện từ ảnh nhị phân không khả quan, Chúng ta cần sử dụng các phương pháp tiếp cận tất cả vị trí có thể trong ảnh. Tuy nhiên trước đó ta cần giả định kích cỡ tiêu chuẩn cho một ký tự

Kết quả mong muốn của việc khoang vùng hoăc phân vùng là có thể xác định được một ký tự bên trong và không bao gồm các đối tượng khác trong ảnh. Tuy nhiên khi các ký tự in có vị trí gần nhau trong hình ảnh phân tích, mục tiêu khoang vùng (hoặc phân vùng) không phải lúc nào cũng đạt được. Thông thường các ký tự khác có thể vô tình xuất hiện bên trong ảnh con (Hình 3), việc này có thể bóp méo việc trích chọn đặc trưng, đây là một trong những lý do tại sao mỗi hệ thống nhận dạng ký tự đều có tùy chọn từ chối.

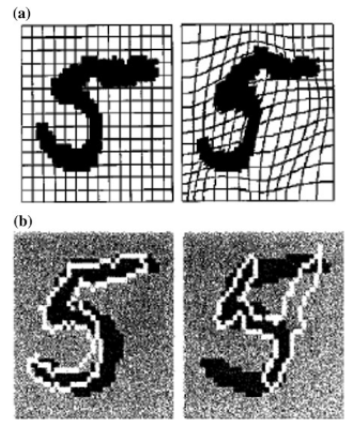
**2.2.1. Đối sánh mẫu:**

Chúng ta không thể nhận biết được ký tự quang học bằng phương pháp đối sánh mẫu trên hình ảnh đa cấp xám.

**2.2.2. Mẫu biến dạng:**

Ý tưởng của phương pháp là sẽ biến đổi hình dạng của tập mẫu và so sánh với hình ảnh đang xét. Việc biến đổi sẽ theo quy tắt biến đổi biên của mẫu đè lên biên của hình ảnh xét.

Dữ liệu sẽ được giảm thiểu bằng cách lặp lại các thông số biến đổi để phù hợp nhất với đối tượng đang xét



2.2.3. Biến đổi Unitar

2.3.3.1. Khái niệm

2.3.3.2. Cách biến đổi

2.3.3.2.1. Ma trận trực giao và ma trận unitary

\* Cho A là một ma trận vuông

\* A trực giao khi : A-1 = AThay AAT = I

Trong đó A-1 là ma trận đảo của A.

AT là ma trận chuyển vị của A.

* Ma trận A được gọi là ma trận Unitar nếu:

A-1 = A\*T hay AA\*T = I

A\* là ma trận liên hợp của A

Các phần tử của A\* được xác định như sau với aik = x + jy thì a\*ik = x – jy

(dạng số phức tổng quát)

Nhận xét:

Nếu các phần tử của ma trận A có giá trị là số thực thì

A trực giao ⇔ A Unitar

Ví dụ về ma trận Unitar



Ta có





A trực giao ⇒ A Unitar

Ví dụ về ma trận không phải ma trận Unitar



Nhận xét





Tuy nhiên





2.3.3.2.2. Phép biến đổi Unitar một chiều

Cho vector  và ANxN là ma trận Unitar. Ta có ảnh

 của  qua phép biến đổi Unitar thuận.

 hay 

Ví dụ:



, ma trận unitary

Ta có:



 Suy ra: 

Hay ta có công thức:



Trong đó:



Kết luận: với hình ảnh cơ sở a k là cột k của ma trận A\*T, ta tách  ảnh cơ sở thông qua các hệ số của 

Hệ số phân tích



Các hình ảnh cơ sở

2.3.3.2.3. Phép biến đổi Unitar 2 chiều

Cho ma trận Unitar ANxN, với ảnh s(m,n) ta có công thức biến đổi unitary của ảnh S như sau:

Cặp biến đổi Unitar 2 chiều:

V = ASAT (Xác định hệ số phân tích)

S = A\*TVA\* (Xác định ảnh cơ sở)

Hay S = , với  là hình ảnh cơ sở

 = ak\* al\*T

Trong đó: ak\* và al\* là các cột thứ k và l của A\*T

Ví dụ: xác định các ảnh cơ sở của biến đổi

 và 

Ta có: V = ASAT = 



\* Xác định các 

Ta có:  và 

, 

, 

\* Như vậy S có thể biểu diễn qua các hình ảnh cơ sở như sau:



Hình ảnh cơ sở

Ví dụ 2:

Cho ma trận Unitar A và ảnh S, xác định V và A\*k,l

 và 

Giải:





* Tính với  và 



2.2.4. bất biến hình học

2.2.5. bất biến Zenite.

2.3. Trích chọn đặc trưng từ ảnh nhị phân:

Hình ảnh nhị phân thu được bằng cách nhị phân hóa thích nghi toàn cục hoặc nhị phân hóa thích nghi cục bộ của ảnh đa cấp xám. Trong nhiều trường hợp, phân đoạn các ký tự được thực hiện đơn giản bằng cách cô lặp các thành phần kết nối. Tuy nhiên đối với hình ảnh phức tạp, các ký tự có thể sẽ chạm hoặc chồng chéo lên nhau. Một vấn đề khác là khi các ký tự được phân mảnh thành 2 hoặc nhiều thành phần kết nối. Vấn đề này có thể giảm bớt được phần nào bằng cách sử dụng phương pháp nhị phân hóa thích nghi cục bộ.

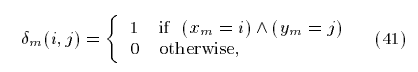
Biểu diễn nhị phân của ký tự là một dạng đơn giản hóa của biễu diện đa cấp xám. Phương thức hình ảnh Z(x,y) với 2 giá trị 0 và 1, thay vì hình ảnh đa cấp xám cần tới 256 giá trị. Điều này có nghĩa là tất cả các phương thức được phát triển trên ảnh đa cấp xám đều có thể áp dụng cho ảnh quét nhị phân.

2.3.1. Đối sánh mẫu

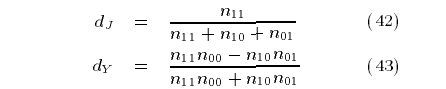
trong đối sánh mẫu nhị phân, Để phát hiện tính tương đồng, đặt ni j là số vị trí pixel, với pixel x của mẫu là i và pixel y của hình ảnh là j (i ,j thuộc {0,1});



Trong đó



i,j thuộc {0,1} và ym và xm là pixel thứ m của ảnh nhị phân Y và X, chúng sẽ được dùng để so sánh. Tubbs đã đánh giá 8 khoảng cách và tìm thấy khoảng cách Jaccard d j và khoảng cách Yule d y là hai khoảng cách tốt nhất.



Tuy nhiên, biến thể hình dạng thiếu tính bên vững được đề cập ở trường hợp đa cấp xám vẫn được áp dụng. Tubbs cố gắng để cải thiện sử thiếu sót bằng cách giới thiệu về trọng lượng của các vị trí pixel khác nhau. Nó được thay thế bằng:



Tại đó pm (*k* | *i*) là xác xuất mà hình ảnh Y tương thích với mẫu Xk, pixel thứ m trong mẫu Xk là i,.

pm(*k | i*) sấp xĩ với số lượng mẫu (bao gồm mẫu Xk) có cùng giá trị pixel tại vị trí m giốg với mẫu Xk,  chia cho tổng số mẫu

Tuy nhiên, chúng ta nghi ngờ rằng tính linh hoạt thu được = p(*k|i*)là chưa đủ để đối phó với sự biến đổi trong hình dạng ký tự từ ký tự viết tay hoặc ký tự in được chỉnh sửa font. Một hứa hẹn khác được Gader et al để cập, người đã sử dụng mẫu của mỗi lớp ký tự và 1 phương thức để lựa chọn mẫu cơ bản trên tập huấn luyện

2.3.2. Mẫu biến dạng

2.3.3. Biến đổi ảnh Unitar

2.3.3.1. Khái niệm

2.3.3.2. Cách biến đổi

2.3.3.2.1. Ma trận trực giao và ma trận unitary

\* Cho A là một ma trận vuông

\* A trực giao khi : A-1 = AThay AAT = I

Trong đó A-1 là ma trận đảo của A.

AT là ma trận chuyển vị của A.

* Ma trận A được gọi là ma trận Unitar nếu:

A-1 = A\*T hay AA\*T = I

A\* là ma trận liên hợp của A

Các phần tử của A\* được xác định như sau với aik = x + jy thì a\*ik = x – jy

(dạng số phức tổng quát)

Nhận xét:

Nếu các phần tử của ma trận A có giá trị là số thực thì

A trực giao ⇔ A Unitar

Ví dụ về ma trận Unitar



Ta có





A trực giao ⇒ A Unitar

Ví dụ về ma trận không phải ma trận Unitar



Nhận xét





Tuy nhiên





2.3.3.2.2. Phép biến đổi Unitar một chiều

Cho vector  và ANxN là ma trận Unitar. Ta có ảnh

 của  qua phép biến đổi Unitar thuận.

 hay 

Ví dụ:



, ma trận unitary

Ta có:



 Suy ra: 

Hay ta có công thức:



Trong đó:



Kết luận: với hình ảnh cơ sở a k là cột k của ma trận A\*T, ta tách  ảnh cơ sở thông qua các hệ số của 

Hệ số phân tích



Các hình ảnh cơ sở

2.3.3.2.3. Phép biến đổi Unitar 2 chiều

Cho ma trận Unitar ANxN, với ảnh s(m,n) ta có công thức biến đổi unitary của ảnh S như sau:

Cặp biến đổi Unitar 2 chiều:

V = ASAT (Xác định hệ số phân tích)

S = A\*TVA\* (Xác định ảnh cơ sở)

Hay S = , với  là hình ảnh cơ sở

 = ak\* al\*T

Trong đó: ak\* và al\* là các cột thứ k và l của A\*T

Ví dụ: xác định các ảnh cơ sở của biến đổi

 và 

Ta có: V = ASAT = 



\* Xác định các 

Ta có:  và 

, 

, 

\* Như vậy S có thể biểu diễn qua các hình ảnh cơ sở như sau:



Hình ảnh cơ sở

Ví dụ 2:

Cho ma trận Unitar A và ảnh S, xác định V và A\*k,l

 và 

Giải:





* Tính với  và 



2.3.4 Bất biến hình học.

2.4. Trích chọn đặc trưng từ biên ảnh:

2.4.1. Chiếu nghiêng

Sự thúc đẩy sử dụng chiếu nghiêng là mỗi 1/2 đường cong có thể xấp xỉ bằng 1 hàm rời rạc của một không gian các biến x hoặc y. Sau đó, các đặc trưng có thể được trích từ các hàm rời rạc . Chúng ta có thể dùng nghiêng dọc hoặc nghiêng ngang, và chúng có thể nghiêng ngoài hoặc nghiêng trong.

Để nghiêng theo chiều dọc, đầu tiên xác định vị trí các điểm ảnh cao nhât và thấp nhất trên đường cong. Đường cong là cắt 2 điểm này. Để có được nghiêng ngoài, mỗi giá trị y, chọn giá trị x ngoài cùng trên mỗi nữa đường cong. Để đặt được nghiêng trong, mỗi giá trị y, giá trị x trong cùng sẽ được lựa chọn. Chiếu nghiêng theo dạng ngang được trích chọn kiểu tương tự, chia đường cong ở trên và giảm nữa.

Chiếu nghiêng phụ thuộc vào gốc xoay. Vì vậy các đặc trưng được tìm thấy từ chiếu nghiêng cũng phụ thuộc gốc xoay..

2.4.2. tách vùng

Kimura và Shridhar đã dùng tách vùng trên đường cong biên. Trong mỗi vùng, các phần đường cong giữa các điểm ảnh lân cận là nhóm theo dạng ngang (0) dọc (90) và 2 hướng của đường chéo(45,135).

Takahashi cũng đã dùng biểu đồ hướng từ tách vùng, nhưng dùng các đường dọc , ngang chéo như các các vùng. Các hướng được trích chọn từ đường cong bên ngoài đến đường cong bên trong khi làm biểu đồ.

Ngoài ra, Takahashi đã xác đỉnh các điểm cong cao dựa vào các đường cong ngoài và trong. Mỗi điểm này, giá trị cong, đường cog tiếp tuyến và vị trí vùng của điểm được trích chọn. Một lưới đều được dùng làm các vùng. Cao et al đã quan sát thấy rằng đường cong gần biên , các biên thể nhỏ trong đường cong có thể dẫn đến các biến thể lớn trong trích chọn đặc trưng. Họ đã thử mô tả điều sử dụng các biên mờ. Các điểm gần vùng biên là  gửi các giá trị thành viên mờ đến các vùng 2 hoặc 4 và các điểm này có tổng là 1.

2.4.3. xấp xỉ đường cong,

Sekita et al xác định các điểm cong cao , được gọi là các điểm gãy, trên đường cong ký tự ngoài và đường cong xấp xỉ giữa các điểm gãy với một hàm mỏng. Sau đó, cả các điểm gãy và các tham số đường cong mỏng được sử dụng là các đặc trưng.

Taxt et al đường cng ngoài với một đường cong mỏng là xấp xỉ, đó là làm mịn. Đường cong mỏng được làm mịn được chia thành M phần của độ dài đường cong bằng nhau. Trong mỗi phần, độ cong trung bình là được tính toán. Them vào đó, khoảng cách từ số học trung bình của các điểm đường trên biên đến N các điểm đều nhau trên đường cong là được đo lường. Kích thước đường cong mỏng của ký tự xấp xỉ 1 size chuẩn trước các đặc trung được đo lường, các đặc trưng trở thành size bất biến.

2.4.4. Mô tả Fourier.

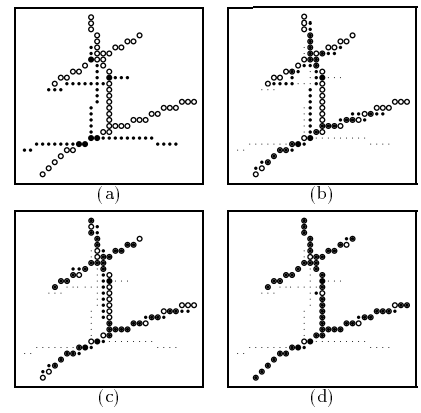
2.5. Trích chọn đặc trưng từ biểu diễn véc tơ

2.5.1. Đối sánh mẫu

Đối sánh mẫu không phù hợp đối với ký tự xương. Trường hợp ngẫu nhiên khá nhỏ để pixels của các nhánh của xương ký tự trừng khớp với pixel của xương mẫu.

2.5.2. Mẫu biến dạng

Mẫu biến dạng được sử dụng bởi Burr và Wakahara cho việc nhận dạng ký tự xương. Số lượng và kiểu biến dạng trước khi đối xứng được thu thập bằng cách sử dụng một cách đo khác giữa mỗi mẫu và input pattern



Hình 20: Cách tiếp cận phù hợp với mô hình biến dạng

của Wakahara .

Với:

` . '= Các điểm mẫu ban đầu không có trong mẫu biến đổi

`•'= mẫu biến đổi,

`○' = mẫu đầu vào;

`●'= các điểm ảnh phổ biến của mẫu chuyển đổi và input pattern

(a) Mẫu và input pattern của ký tự trung quốc  
(b) - (d) sau 1, 5, và 10 lần lặp lại, tương ứng của biến đổi cục bổ trên một

bản sao mẫu.

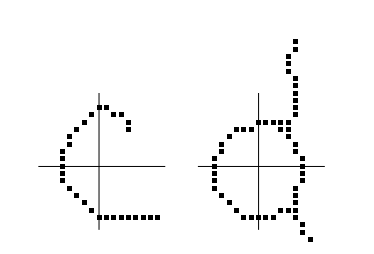
2.5.3. Đặc trưng rời rạc

Từ các ký tự mỏng, các tính năng sau đây có thể được trích xuất: số vòng lặp; số lượng kết - T; số lượng kết - X; số điểm uốn; tỷ lệ chiều rộng / chiều cao của khung hình chữ nhật kèm theo; sự hiện diện của dấu chấm cô lập; toàn bộ số điểm kết thúc, và số điểm cuối trong mỗi bốn hướng N, S, W, và E; số vòng bán nguyệt ở mỗi hướng; và số điểm giao của trục dọc và ngang, tương ứng, các trục đặt vào trung tâm. Hai tính năng cuối cùng được giải thích trong Hình 21.

Có thể sử dụng điểm giao giữa các đường chồng lên như các đặc trưng, và trên thực tế, điều này đã được thực hiện trong các hệ thống OCR trước đó. Tuy nhiên, những tính năng này không dẫn đến các hệ thống nhận dạng mạnh mẽ; khi số lượng các đường chồng lên được tăng lên, các kết quả đặc trưng ít tin cậy hơn so đối với các phông chữ khác nhau (đối với các ký tự được in bằng máy) và sự biến đổi trong các hình dạng ký tự và cách viết (cho các ký tự viết tay).

2.5.4. Biểu diễn đồ thị

Đối với nhận diện chữ tiếng hoa, một vài tác giả đã sử dụng biểu diện đồ thị hoặc là biểu diện xương như là các đặc trưng



Hình 21: Các chữ cái 'c' và `d 'thưa thớt. Các trục đứng và trục ngang được đặt ở giữa trọng tâm. Cả `c' và ` d' đều có hình bán nguyệt theo hướng Tây. `c 'có một đường thẳng đứng và hai đường ngang.

Rút ra các đồ thị được phân cấp theo thứ bậc để giải quyết các biến thể trong độ dài nét và sự liên kết do kiểu viết khác nhau của người viết khác nhau. Cheng và cộng sự đã sử dụng biến đổi Hough trên các hình ảnh ký tự đơn lẽ để trích xuất các đường nét từ khung xương của các ký tự tiếng hoa.

2.5.5. Biểu diễn Fourier.

**CHƯƠNG 3**

**PHÂN TÍCH BẢNG ĐIỂM**

**3.1. GIỚI THIỆU BÀI TOÁN**

**Bảng điểm của trường ĐH Công Nghiệp Thực Phẩm gồm có các trường thông tin là stt, họ và tên, ngày sinh, tên lớp, điểm số, điểm chữ, ký tên, ghi chú**



3.2. PHÂN TÍCH VÀ TÁCH CÁC TRƯỜNG THÔNG TIN CÓ TRONG BẢNG ĐIỂM

Đối với yêu cầu đặt ra là cần phân tích là trích chọn được nội dung của các cột mssv, điểm số và điểm chữ

Khi tính toán vị trí của 3 trường thông tin trên ta được các số liệu:

* Mã SV nằm tại vị trí chiều rộng \* 0.111
* Điểm số nằm tại vị trí chiều rộng \* 0.689
* Điểm chữ nằm tại vị trí chiều rông \* 0.778

3.2.1. Quy trình xử lý ảnh đầu vào

Đối với ảnh đầu vào sẽ có các trường hợp xảy ra:

* ảnh mờ do camera có độ phân giải thấp hoặc mật in bị mờ khi scane
* ảnh có tối hoặc quá sáng do địa điểm chụp ở nơi tối hoặc ở nơi quá sáng
* ảnh không trọn vẹn do nhân tố con người

Cần phải can thiệp và giải quyết các trường hợp trên để có thể trích xuất được thông tin ảnh một cách đầy đủ

Đề xuất đặt ra để quản lý ảnh đầu vào phù hợp

Đối với ảnh chụp

* Khi chụp ảnh, cần sử dụng camera có độ phân giải cao (khuyến nghị trên 10px)
* Khi chụp ảnh nên lựa chọn vị trí thích hợp, không chụp ở nơi quá tối hoặc quá sáng
* Bảng điểm được chụp với góc thằng đứng so với góc chụp

Đối với ảnh scane

* Kiểm tra độ đẩm của mực in khi scane, tránh tình trạng in quá đậm hoặc quá mờ

Nhân tố con người

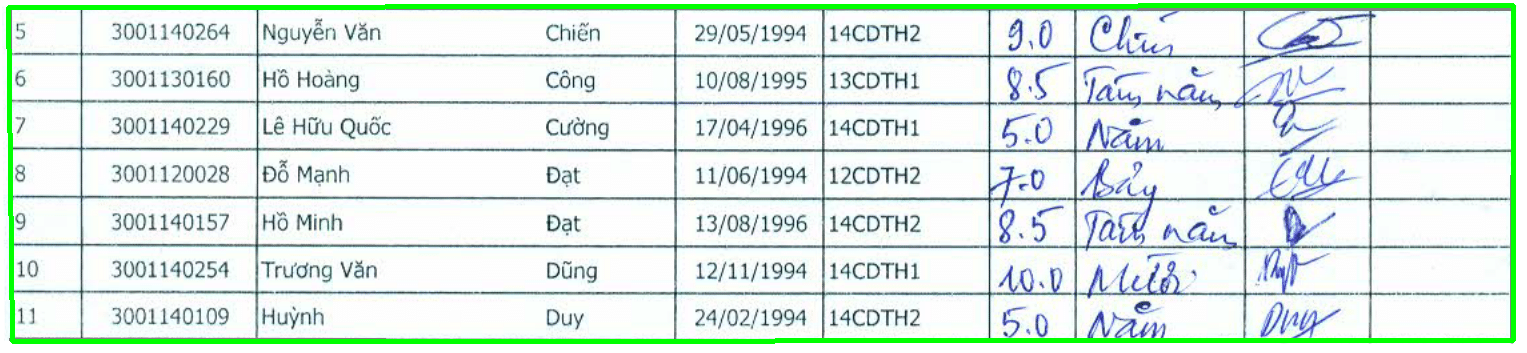
* Không gạch xóa bảng điểm, như vậy sẽ làm nhiễu thông tin
* Giữ bảng điểm sạch sẽ

3.2.2. Quy trình tách nội dung trên bảng điểm

3.2.2.1. điều chỉnh lại góc ảnh, và hình thái ảnh

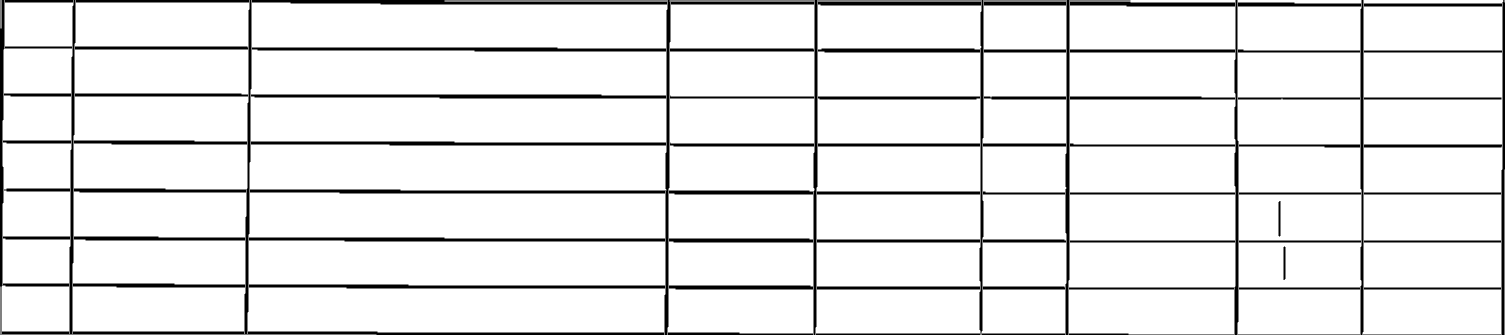
Hình ảnh đầu vào không thể nào có góc chính xác như mong muốn vì trong quá trình chụp, hoặc scane trong không gian ngoài. Việc chụp lại luôn bị lệch một góc nào đó. Yêu cầu đầu tiên là cần phải chính lại góc của ảnh cho đúng để lấy chính xác các trường dữ liệu.

Sử dụng thuật toán tìm viền (contours) đã được trình bày ở trên để tìm kiếm viền của ảnh. Kết quả của thuật toán sẽ trả về nhiều viền, ta chỉ cần chọn viền có diện tích lớn nhất

****

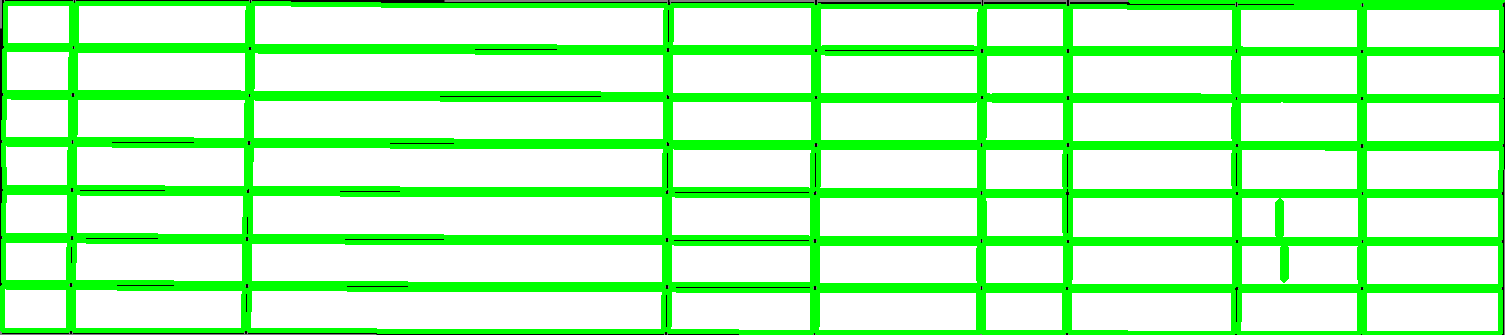
**3.2.2.2. Tách vùng**

**Trước khi tách vùng ta cấn tách khung của ảnh để thuật tiện trong việc trích các khung dữ liệu. Sử dụng thuật toán giãn nở ảnh với 2 kernel [1 1 1 1 1 1 1] và [1 1 1 1 1 1 1] T được trình bày ở trên ta sẽ có được kết quả**

****

**3.2.2.2. Xác định ô dữ liệu và tách dữ liệu**

**Sau khi có được khung điểm dự trên bảng điểm, ta tiến hành khoang viền dựa thuật toán xác định viền ngay trên khung tìm được.**

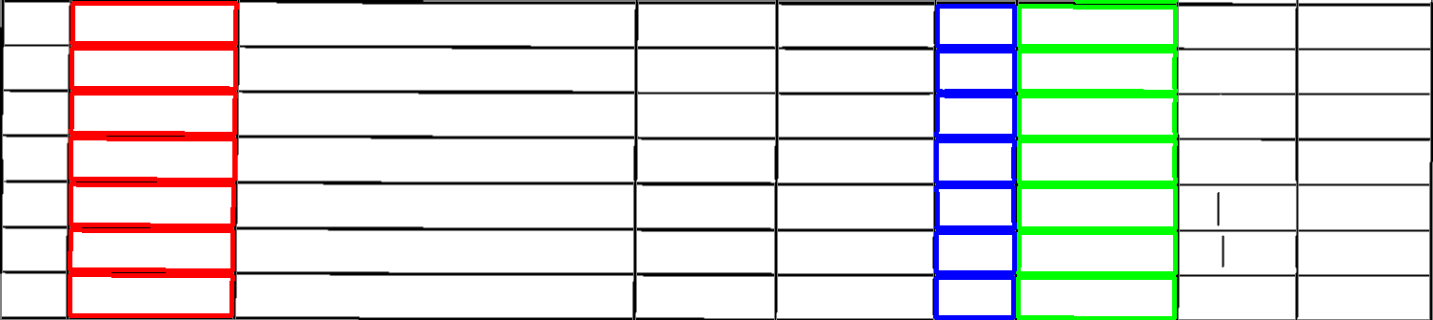
****

**Sau khi có được các ô dữ liệu, việc tiếp theo là xác định vị trí của mssv, điểm số và điểm chữ**

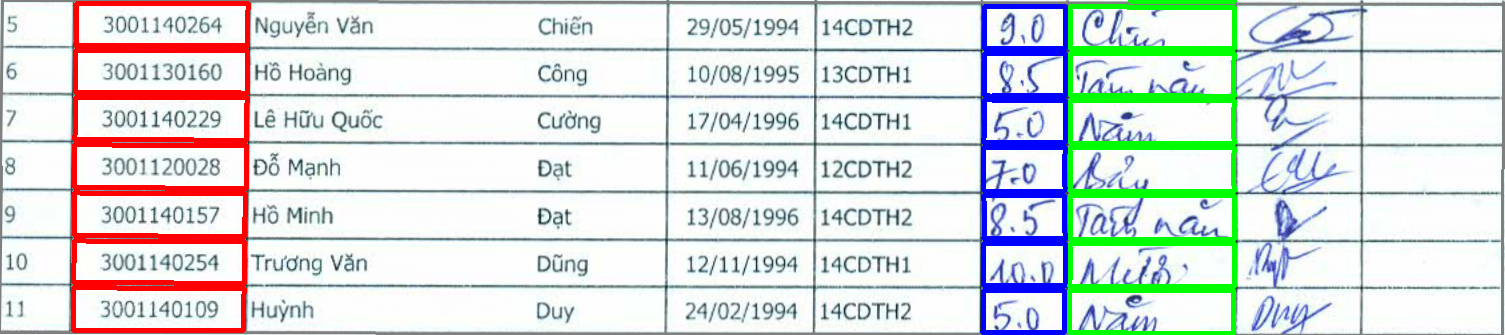
**Dựa trên số liệu phân tích, ta có:**

* Mã SV nằm tại vị trí chiều rộng \* 0.111
* Điểm số nằm tại vị trí chiều rộng \* 0.689
* Điểm chữ nằm tại vị trí chiều rông \* 0.778

Từ đó ta khoang được vị trí của mssv, điểm số và điểm chữ

****

**Sau đó ta ánh xạ các vùng đã được xác định lên ảnh gốc bảng điểm**

****

CHƯƠNG 4:

THỰC NGHIỆM

4.1. Công nghệ, nền tảng sử dụng

Ứng dụng được xây dựng dựa trên các ngôn ngữ Python, C#, Batch Script cùng với thư viện OpenCV. Giao diện chương trình được xây dựng trên winform trên nền tảng .NET 4.5, các thuật toán xử lý được viết trên Python 2.7. Chương trình được làm ra với mục đích có thể nhận diện được bảng điểm của sinh viên, tách bảng điểm ra 3 trường thông tin là MSSV, điểm số và điểm chữ, đồng thời nhận diện và lưu vào file excel

4.2. Các bước xử lý của chương trình

**CHƯƠNG 4:**

**KÊT LUẬN**

**TÀI LIỆU THAM KHẢO:**

[1]. Giáo trình xử lý ảnh – PGS.TS. ĐỖ NĂNG TOÀN – Học viện bưu chính viễn thông

[2]. Cơ bản xử lý ảnh trong matlab

**Tìm hiểu thuạt toán contours**[**http://www.jackyle.xyz/2017/11/tim-hieu-ve-contour-moments-trong-xu-ly.html**](http://www.jackyle.xyz/2017/11/tim-hieu-ve-contour-moments-trong-xu-ly.html)

**Phụ lục** *(không đánh số trang)*