**CÔNG VIỆC 2: XÂY DỰNG MODULE XÁC ĐỊNH GÓC XOAY TÀI LIỆU**

# **LỜI NÓI ĐẦU**

Trong số các thông tin con người thu nhận từ thế bên ngoài, có đến hơn 79% được ghi nhận bằng mắt tức là ở dạng ảnh. Vì vậy xử lý ảnh là một ngành khoa học đã và đang và sẽ phát triển mạnh cò ứng dụng rộng rãi trong khoa học và đới sống thực tiễn. Các hệ thống xử lý ảnh cho phép con người thu nhận lưu trữ, phân tích và nhận dạng ảnh. Một bộ phận quan trọng của xử lý ảnh là xử lý văn bản. Một trong những nhiệm vụ và là đối tượng chính của xử lý ảnh văn bản là tự động hoá công việc văn phòng.

Một trong những vấn đề đầu tiên và kinh điển trong xử lý ảnh văn bản là bài toán góc xoay tài liệu. Nguyên nhân đẫn đến tài liệu bị xoay một góc xuất phát từ quá trình quét ảnh hoặc copy ảnh. Do đặt ảnh vào bệ máy quét và máy in là một công đoạn được thực hiện bằng tay lên ảnh có thể bị lệch so với bệ máy một góc mà mắt thường không nhận thấy được, đẫn đến ảnh bị lệch đi một góc tương ứng. Tài liệu bị lệch có ảnh hưởng rất lớn đến các quá trình xử lý ảnh tiếp theo, vì vậy việc phát hiện và chỉnh sửa góc xoay tài liệu là nhiệm vụ quan trong đấu tiên trong xử lý ảnh văn bản.

Từ hai thập kỷ trở lại đây, cùng với sự phát triển mạnh mẽ của xử lý ảnh văn bản, đã có nhiều phương pháp và thuật toán được đề xuất để giải quyết bài toán xác định góc xoay của tài liệu. Có rất nhiều hướng tiếp cận cho bài toán góc nghiêng tài liệu từ trước tới nay như: Phương pháp phân tích hình chiếu( Projection Profile), Phương pháp dựa vào trọng tâm( Center of Gravity), Phương pháp dựa trên biến đổi Hough( Hough Transform)…. Các thuật toán phát hiện góc nghiêng thường được xây dựng cho các hệ thống phân tích ảnh tài liệu khác nhau nên giải quyết cho những loại ảnh tài liệu cụ thể. Tài liệu này trình bày cơ sở lý thuyết và nhận xét, đánh giá một số phương pháp để lựa chọn phương pháp tối ưu giải quyết bài toán góc nghiêng văn bản và kết quả thực nghiệm của phương pháp lựa chọn.

# **MỤC LỤC**

[LỜI NÓI ĐẦU 2](#_Toc177082405)

[MỤC LỤC 3](#_Toc177082406)

[DANH MỤC HÌNH ẢNH 4](#_Toc177082407)

[DANH MỤC BẢNG BIỂU 5](#_Toc177082408)

[DANH MỤC TỪ VIẾT TẮT 6](#_Toc177082409)

[CHƯƠNG 1: TỔNG QUAN VỀ BÀI TOÁN OCR VÀ PHÁT HIỆN GÓC NGHIÊNG 7](#_Toc177082410)

[1.1 Giới thiệu về Nhận diện Ký tự Quang học 7](#_Toc177082411)

[1.2 Tổng quan về bài toán phát hiện góc xoay trong OCR 8](#_Toc177082412)

[1.3 Một số cách tiếp cận phát hiện góc xoay tài liệu 10](#_Toc177082413)

[CHƯƠNG 2: CÁC PHƯƠNG PHÁP XÁC ĐỊNH GÓC XOAY TRONG TÀI LIỆU 13](#_Toc177082414)

[2.1 Phương pháp dựa vào thuật toán phân cụm láng giềng gần nhất 13](#_Toc177082415)

[2.2 Phương pháp phân tích hình chiếu( Projection Profile) 17](#_Toc177082416)

[2.3 Phương pháp phân tích dựa vào trọng tâm( Center of Gravity) 19](#_Toc177082417)

[2.4 Phương pháp dựa vào biến đổi Hough(Hough Transform) 23](#_Toc177082418)

[CHƯƠNG 3: THỬ NGHIỆM VÀ ĐÁNH GIÁ 26](#_Toc177082419)

[KẾT LUẬN 30](#_Toc177082420)

[TÀI LIỆU THAM KHẢO 31](#_Toc177082421)

# **DANH MỤC HÌNH ẢNH**

[Hình 1.Minh họa quy trình hoạt động của OCR 8](#_Toc177083504)

[Hình 2.Một ảnh bảng điểm bị nghiêng góc 9](#_Toc177083505)

[Hình 3.Những hình ảnh tài liệu mà ở đó các thành phần kết nối được giới hạn 14](#_Toc177083506)

[Hình 4.Các NNC của hình 3a 15](#_Toc177083507)

[Hình 5.Các NNC của hình 3b 15](#_Toc177083508)

[Hình 6.Hình chiếu dọc và ngang 17](#_Toc177083509)

[Hình 7.Đa giác 6 đỉnh và trọng tâm được xác định 19](#_Toc177083510)

[Hình 8.Hình chữ nhật ngoại tiếp ảnh văn bản 20](#_Toc177083511)

[Hình 9.Minh hoạt thuật toán Center of Gravity 22](#_Toc177083512)

[Hình 10.Đường thẳng Hough và trực tọa độ Oxy 22](#_Toc177083513)

[Hình 11.Biểu diễn đường thẳng Hough qua 3 điểm 23](#_Toc177083514)

# **DANH MỤC BẢNG BIỂU**

# **DANH MỤC TỪ VIẾT TẮT**

|  |  |  |
| --- | --- | --- |
| **STT** | **Từ viết tắt** | **Từ viết tắt đầy đủ** |
| 1 | NCC | Nearest neighbour chain |
| 2 | OCR |  |
| 3 |  |  |
| 4 |  |  |
| 5 |  |  |
| 6 |  |  |
| 7 |  |  |
| 8 |  |  |

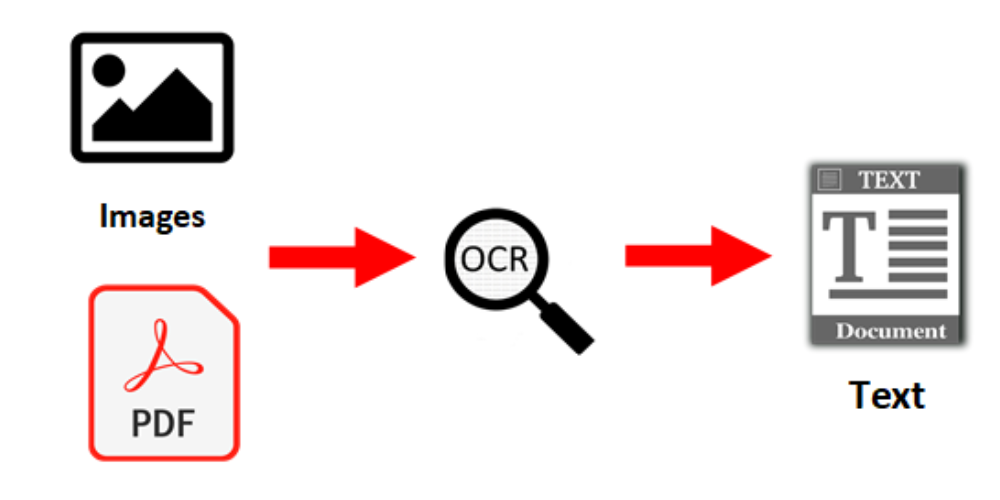
# **CHƯƠNG 1: TỔNG QUAN VỀ BÀI TOÁN OCR VÀ PHÁT HIỆN GÓC NGHIÊNG**

## **1.1 Giới thiệu về Nhận diện Ký tự Quang học**

Nhận diện Ký tự Quang học (OCR) là một công nghệ quan trọng trong việc chuyển đổi các tài liệu văn bản từ định dạng hình ảnh thành văn bản số để dễ dàng chỉnh sửa, lưu trữ và tìm kiếm. OCR có ứng dụng rộng rãi trong các lĩnh vực như số hóa tài liệu, quản lý hồ sơ, nhận diện văn bản trong ảnh, và nhiều ứng dụng khác trong ngành công nghệ thông tin và quản lý tài liệu.

Nguyên lý hoạt động của OCR:

* Nhận diện hình ảnh: Hệ thống OCR đầu tiên chuyển đổi hình ảnh văn bản thành dữ liệu số, thường thông qua việc phát hiện các ký tự và từ ngữ.
* Tiền xử lý: Giai đoạn này bao gồm việc làm sạch và chuẩn hóa hình ảnh để cải thiện độ chính xác của việc nhận diện ký tự.
* Nhận diện ký tự: Sử dụng các thuật toán học máy hoặc mạng nơ-ron để nhận diện và phân loại các ký tự từ hình ảnh.
* Hậu xử lý: Cải thiện kết quả nhận diện bằng cách sử dụng các kỹ thuật như kiểm tra chính tả và phân tích ngữ nghĩa.

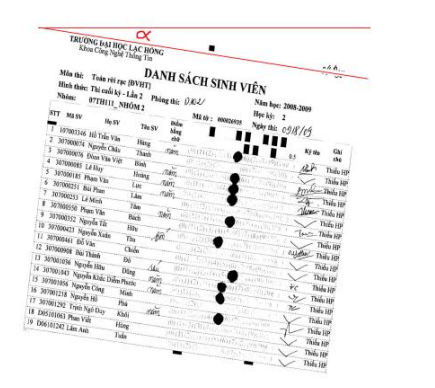


Hình 1.Minh họa quy trình hoạt động của OCR

## **1.2 Tổng quan về bài toán phát hiện góc xoay trong OCR**

Trong quá trình tiền xử lý OCR, việc phát hiện và hiệu chỉnh góc xoay của tài liệu là bước rất quan trọng. Các tài liệu có thể được quét hoặc chụp ở các góc khác nhau, dẫn đến sự nghiêng hoặc xoay của văn bản. Điều này có thể gây khó khăn cho hệ thống OCR trong việc nhận diện chính xác văn bản. Việc phát hiện góc xoay và hiệu chỉnh một cách chính xác tài liệu trước khi đưa vào hệ thống OCR có thể đem lại một số lợi ích như:

* Cải thiện chất lượng nhận diện: Khi văn bản được căn chỉnh chính xác, hệ thống OCR có thể nhận diện ký tự và từ ngữ một cách rõ ràng hơn. Điều này làm giảm tỷ lệ lỗi nhận diện và cải thiện chất lượng của kết quả nhận diện.
* Tăng cường hiệu quả xử lý: Việc chuẩn hóa văn bản giúp hệ thống OCR hoạt động hiệu quả hơn, giảm thiểu số lượng lỗi cần phải được xử lý trong các bước tiếp theo. Điều này không chỉ cải thiện kết quả nhận diện mà còn giúp tiết kiệm thời gian và công sức trong quy trình xử lý tài liệu.
* Khả năng mở rộng và tính chính xác cao hơn: Đối với các hệ thống OCR cần xử lý nhiều loại tài liệu và ngôn ngữ khác nhau, việc hiệu chỉnh góc xoay giúp đảm bảo rằng tất cả các tài liệu đều được xử lý theo cách đồng nhất. Điều này giúp cải thiện khả năng mở rộng của hệ thống OCR và nâng cao tính chính xác của nhận diện trên nhiều loại tài liệu và ngôn ngữ khác nhau.
* Ứng dụng rộng rãi trong thực tế: Việc phát hiện và hiệu chỉnh góc xoay có ứng dụng quan trọng trong nhiều lĩnh vực, bao gồm số hóa tài liệu, quản lý hồ sơ, và xử lý hình ảnh trong các ứng dụng thương mại. Đặc biệt trong các hệ thống cần xử lý số lượng lớn tài liệu, việc hiệu chỉnh góc xoay giúp tăng cường hiệu quả và độ tin cậy của quy trình xử lý tài liệu.



Hình 2.Một ảnh bảng điểm bị nghiêng góc

Như vậy, việc phát hiện và hiệu chỉnh góc xoay không chỉ là một bước cần thiết trong quá trình tiền xử lý OCR mà còn là yếu tố quyết định đến chất lượng và hiệu quả của toàn bộ hệ thống nhận diện ký tự. Các phương pháp và công cụ để thực hiện việc này ngày càng được cải thiện, giúp nâng cao độ chính xác và khả năng xử lý của hệ thống OCR trong các ứng dụng thực tế.

## **1.3 Một số cách tiếp cận phát hiện góc xoay tài liệu**

Trong việc phát hiện và hiệu chỉnh góc xoay tài liệu, có nhiều phương pháp tiếp cận khác nhau, mỗi phương pháp đều có những ưu điểm và hạn chế riêng. Dưới đây là tóm lược các phương pháp phổ biến:

* Phương Pháp Dựa Vào Thuật Toán Phân Cụm Láng Giềng Gần Nhất. Phương pháp này sử dụng kỹ thuật phân cụm để nhóm các đối tượng văn bản dựa trên sự gần gũi của chúng. Sau khi xác định các đối tượng và láng giềng gần nhất, phương pháp này tính toán vector định hướng giữa các cặp láng giềng. Các góc xoay được xác định bằng cách tích lũy các vector định hướng vào một histogram và chọn góc tương ứng với đỉnh cao nhất trong histogram. Phương pháp này có thể gặp khó khăn khi các đối tượng văn bản không đồng nhất hoặc có sự chồng chéo giữa các đối tượng.
* Phương Pháp Phân Tích Hình Chiếu (Projection Profile).Phương pháp phân tích hình chiếu dựa trên việc tính toán histogram cho các góc lệch khác nhau bằng cách phân tích hình chiếu của văn bản. Histogram cho mỗi góc là số lượng điểm đen nằm trên các đường thẳng có hướng tương ứng với góc đó. Góc nghiêng của tài liệu được xác định bằng cách áp dụng hàm chi phí cho các giá trị histogram, với góc có giá trị hàm chi phí cực đại được coi là góc xoay của tài liệu. Phương pháp này đơn giản và dễ thực hiện nhưng có thể kém hiệu quả đối với tài liệu có độ nghiêng nhỏ hoặc có nền phức tạp.
* Phương Pháp Phân Tích Dựa Vào Trọng Tâm (Center of Gravity). Phương pháp dựa vào trọng tâm tính toán góc nghiêng của tài liệu bằng cách xác định trung tâm trọng lực của các đối tượng văn bản. Các trọng tâm được xác định và góc nghiêng của tài liệu được tính toán dựa trên sự phân phối của các trọng tâm này. Phương pháp này hữu ích trong việc phát hiện góc nghiêng của các tài liệu có cấu trúc đồng nhất nhưng có thể gặp khó khăn với các tài liệu có hình dạng không đồng đều hoặc phức tạp.
* Phương Pháp Dựa Vào Biến Đổi Hough (Hough Transform).Biến đổi Hough là một kỹ thuật mạnh mẽ để phát hiện các đường thẳng trong hình ảnh. Trong bối cảnh phát hiện góc xoay, biến đổi Hough ánh xạ các đường thẳng trong mặt phẳng ảnh thành các cặp (r, θ) trong không gian Hough, với r là khoảng cách từ gốc tọa độ và θ là góc nghiêng của đường thẳng. Góc nghiêng của tài liệu được xác định bằng cách tìm góc có tổng số điểm lớn nhất nằm trên các đường thẳng có cùng lệch góc. Phương pháp này rất hiệu quả trong việc xử lý các tài liệu với đường thẳng rõ ràng nhưng có thể tốn kém về tính toán đối với các tài liệu phức tạp.

Các phương pháp phát hiện góc xoay tài liệu có nhiều điểm mạnh và hạn chế khác nhau. Trong phần tiếp theo của tài liệu, các phương pháp sẽ được làm rõ cụ thể, đánh giá ưu nhược điểm của chúng trong việc phát hiện và hiện chính góc xoay của các tài liệu cụ thể.

# **CHƯƠNG 2: CÁC PHƯƠNG PHÁP XÁC ĐỊNH GÓC XOAY TRONG TÀI LIỆU**

## **2.1 Phương pháp dựa vào thuật toán phân cụm láng giềng gần nhất**

Thuật toán dựa vào phương pháp phát hiện góc nghiêng tài liệu không yêu cầu sự hiện diện của một vùng văn bản nổi bật và không tùy thuộc vào hạn chế góc nghiêng. Tuy nhiên, nói chung sự chính xác của những phương pháp này không hoàn hảo. Yue Lu, Chew Lim Tan[1] giới thiệu phương pháp láng giềng gần nhất cải tiến dựa theo cách tiếp cận thực hiện ước lượng chính xác góc nghiêng tài liệu. Kích thước hạn chế được giới thiệu với các phát hiện của những cặp láng giềng gần nhất. Sau đó lựa chọn những chuỗi với số lượng lớn nhất của các cặp láng giềng gần nhất và tính toán những độ dốc của chúng để đưa ra góc nghiêng của tài liệu ảnh. Kết quả thí nghiệm trên các loại tài liệu có chứa chữ viết khác nhau và bố trí đa dạng cho thấy hướng tiếp cận đưa ra đã đạt được tính chính xác cải thiện cho việc ước lượng góc nghiêng tài liệu ảnh và có lợi thế tồn tại ngôn ngữ độc lập.

Định nghĩa 1: Khoảng cách trọng tâm giữa hai thành phần được định nghĩa là:

Trong đó △x=|| △y=||

Định nghĩa 2: Khoảng cách giữa hai thành phần được định nghĩa là:

Định nghĩa 3: Thành phần là láng giềng gần nhất của thành phần , nếu:

1. với , hoặc với
2. với , hoặc với

Trong đó, β không đổi, và được đặt là 1.2

Sau đó các cặp láng giếng gần nhất liền kề sẽ tạo thành một chuỗi láng giềng gần nhất nếu chúng có cùng chiều rộng hoặc chiều cao.

Định nghĩa 4: Chuỗi K láng giềng gần nhất(K-NCC) (NNC: Nearest neighbour chain) được định nghĩa là 1 chuỗi có chứa K thành phần [, trong đó là láng giềng gần nhất của với i=1,2,3,…K-1.

Theo định nghĩa, một ảnh tài liệu có thể được chia thành một số mức độ khác nhau gồm các NNC với một hằng số K. Hình 3 cho hai ảnh tài liệu( một là tài liệu tiếng Anh và một là tài liệu tiếng Trung), ở đó những thành phần được nối đã được bao trong những hình chữ nhật ngoại tiếp. Hình 4 (a-c) và hình 5 (a-c) minh họa các K-NNC với K=2, K=3 và K>=4 tương ứng. Nói ngắn gọn quá trình các K-NNC với tất cả K>=4 được đưa vào một hình ở đây. Hình 4 (d-f) và hình 5 (d-f) cho thấy đường kết nối NNC của hình 4 (a-c) và hình 5 (a-c) tương ứng. Có thể thấy được rằng góc của những đường dốc này phản ánh độ góc xoay tài liệu nói chung, đặc biệt là với những tài liệu có K lớn hơn.

Định nghĩa 5: Giả thiết là K-NNC thứ n (n=1,2,3,…N), độ dốc của nó được định nghĩa như sau:

Đối với một hằng số K, có thể thu được giá trị trung bình hoặc trung tuyến các đường dốc của các NNC. Giá trị có thể được dùng để đại diện cho góc xoay của tài liệu. Sử dụng giá trị đối với một K lớn hơn giá trị góc xoay tài liệu, tùy theo điều kiện số lượng các K-NNC được rút ra lớn hơn một ngưỡng đã định nghĩa từ trước. Ngưỡng sử dụng ở đây để đảm bảo có đầy đủ các NNC cho K đặc biệt, với mực đích tránh ảnh hưởng của nhiễu.

|  |  |
| --- | --- |
| *(a)Tài liệu tiếng anh* | *(b)Tài liệu tiếng Trung* |

Hình 3.Những hình ảnh tài liệu mà ở đó các thành phần kết nối được giới hạn

|  |  |  |
| --- | --- | --- |
| *(a)K=2* | *(b)K=3* | *(c)K>=4* |
| *(d)Đường kết nối với K=2* | *(e)Đường kết nối với K=3* | *(f)Đường kết nối với K>=4* |

Hình 4.Các NNC của hình 3a

|  |  |  |
| --- | --- | --- |
| *(a)K=2* | *(b)K=3* | *(c)K>=4* |
| *(d)Đường kết nối với K=2* | *(e)Đường kết nối với K=3* | *(f)Đường kết nối với K>=4* |

Hình 5.Các NNC của hình 3b

Thuật toán ước lượng góc nghiêng được tổng kết như sau:

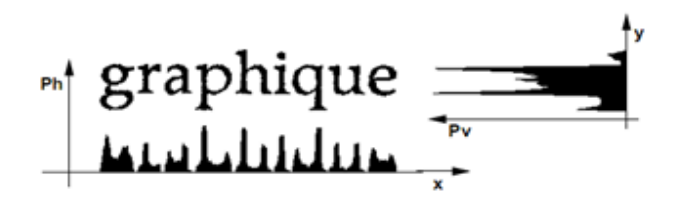
1. Phát hiện tất cả các thành phần trong ảnh, và hợp nhất hai thành phần kết nối nếu một thành phần được bao quanh bởi một thành phần khác.
2. Phát hiện láng giềng gần nhất của mỗi thành phần, theo định nghĩa 3. Chú ý rằng một số thành phần không thể tìm được các láng giềng gần nhất như đã đề cập.
3. Xác định chuỗi láng giềng gần nhất theo định nghĩa 4.
4. Khởi tạo K là số lượng các thành phần lớn nhất trong tất cả các NNC tạo ra từ bước 3.
5. Tính số (N) của các K-NNC.
6. Nếu N lớn hơn ngưỡng đã định trước, đến bước 7, nếu không thì K=K-1, đến bước 5.
7. Tính mỗi đường dốc K-NNC (n=1,2,…N) theo định nghĩa 5.
8. Thu được đường dốc tài liệu bằng cách sử dụng giá trị trung bình hoặc trung tuyến của đường dốc từ bước 7.
9. Tính góc nghiêng θ=arctan(ℼ

## **2.2 Phương pháp phân tích hình chiếu( Projection Profile)**

Ý tưởng của phương pháp phân tích hình chiếu là tính histogram cho tất cả các góc lệch. Histogram của một góc là số điểm ảnh đen( hoặc số điểm đại diện trong một thuật toán-fiducial point) trong ảnh sao cho tất cả các điểm đen nằm trên những đường thẳng có cùng một hướng tương ứng với góc đó.

Xây dựng hàm chi phí cho các giá trị histogram. Góc nghiêng văn bản tương ứng với góc có giá trị hàm chi phí cực đại.

Các tác giả: Akiyama và Hagita, Bard, Bloomberg, Nakano, Kanai và Bagdanov, Komukai và Saiwait, Lam và Zandy, Shutao Li, Jun Sun[3], Messelodi và Modena, Pavidis và Zhou, Potsl[5] và Spitz đề xuất các thuật toán phát hiện góc nghiêng.



Hình 6.Hình chiếu dọc và ngang

Các thuật toán phát hiện góc nghiêng dựa vào hình chiếu thường bao gồm các bước chính sau:

+Dùng hàm rút gọn F để chuyển ảnh đầu vào thành tập các bộ ba(x,y,w) trong đó (x,y) là tọa độ của một điểm ảnh đại diện cho một đối tượng , w là trọng số của điểm. Ở đây, điểm đại diện được hiểu theo nghĩa là điểm biểu diễn các ký tự trong đối tượng của ảnh. Trọng số w thường phụ thuộc vào từng thuật toán.

+Một hàm P dùng chiếu các điểm tìm được vào một mảng đếm A[] theo các góc chiếu khác nhau. Ứng với mỗi góc Φ có một mảng [] dùng lưu số điểm đại diện. Mảng [] là mảng một chiều, phần tử [r] sẽ cho biết số điểm đại diện nằm trên đường thẳng tạo với trực OX góc Φ và khoảng cách từ gốc tọa độ tới đường thẳng đó là r.

+Khi tính được mảng [], ta áp dụng hàm tối ưu Ω cho các giá trị của mảng này theo một tiêu chuẩn nào đó( có thể là tổng bình phương các giá trị trong mảng [] hoặc là số các giá trị bằng 0 trong mảng,…). Góc lệch của văn bản là góc tương ứng có giá trị hàm tối ưu hóa cực đại.

Sự khác nhau giữa các thuật toán theo phương pháp này chính là việc xây dựng các hàm rút gọn F và hàm tối ưu hóa Ω.

\*Phát hiện góc lệch văn bản bằng cách chiếu các góc là một phương pháp đơn giản và dễ hiểu, tuy nhiên những thuật toán này còn hạn chế về độ chính xác với các góc lệch lớn. Baird cho rằng để thuật toán có độ chính xác cao thì góc lệch văn bản phải giới hạn trong khoảng ±. Nếu văn bản có nhiều nhiễu và các đối tượng phi văn bản như bảng biểu, hình ảnh thì độ chính xác của thuật toán giảm đáng kể.

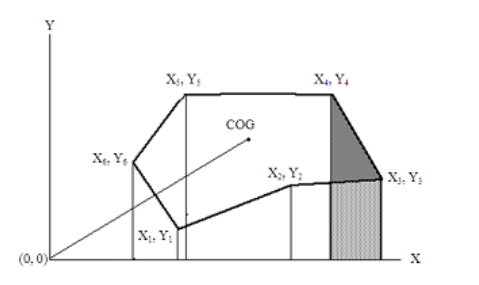
Gần đây, người ta kết hợp phương pháp hình chiếu và phương pháp các đối tượng nhằm giải quyết vấn đề giới hạn góc lệch, nhưng phụ thuộc vào khoảng cách giữa các dòng văn bản và chỉ xử lý được với những hình ảnh chứa nhiều dòng văn bản và kích thước bé cỡ 512\*512 pixels.

Với những trang ảnh tài liệu có nhiều đối tượng ký tự và phi ký tự xen lẫn nhau thì phương pháp trình bày sẽ gặp khó khăn. Sau đây trình bày một cách tiếp cận khác để giải quyết bài toán góc nghiêng dựa vào việc xác định trong tâm khối văn bản.

## **2.3 Phương pháp phân tích dựa vào trọng tâm( Center of Gravity)**

Ý tưởng chính của phương pháp là xây dựng một khối đa giác từ các điểm cực biên của văn bản.. Một đường thẳng được xây dựng từ tọa độ trọng tâm của đa giác đến gốc tọa độ và góc lệch của đường thẳng này so với trục hoành chính là góc nghiêng của văn bản.

Theo hướng tiếp cận này, việc xác định đường cơ sở là bước quan trọng nhất của quá trình và giải pháp trong thuật toán là áp dụng lên tất cả các từ nội tiếp trong đa giác. Trọng tâm của đa giác với gốc tọa độ sẽ tạo thành một đường thẳng lệch một góc nào đó với trục ngang. Góc được xác định cũng chính là góc nghiêng của từ, đoạn văn và cả ảnh văn bản.

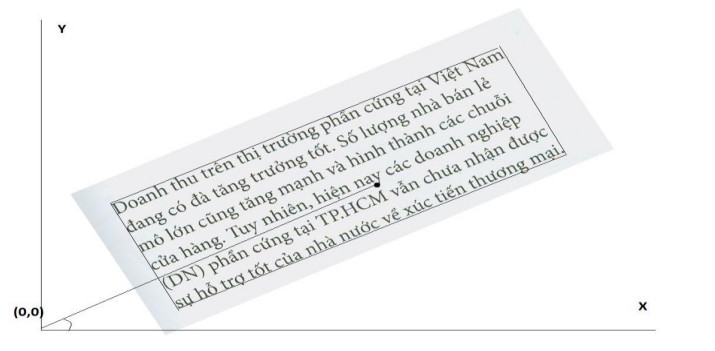


Hình 7.Đa giác 6 đỉnh và trọng tâm được xác định

Hình 7, một đa giác có 6 đỉnh được tìm thấy và trọng tâm của đa giác được xác định bằng công thức:

(1)

Như vậy tùy theo đa giác xác định bởi điểm xa nhất theo các hướng mà ta áp dụng thuật toán. Hình chữ nhật được thay thế cho đa giác trong hình 7 cũng được mô tả như là một cách để xây dựng đường cơ sở giúp xác định góc nghiêng văn bản.



Hình 8.Hình chữ nhật ngoại tiếp ảnh văn bản

Thuật toán gồm các bước:

+Đầu vào: ảnh văn bản bị nghiêng giống hình 8

+Đầu ra: ảnh đã được hiệu chỉnh góc nghiêng

+Bước 1: Xác định những điểm xa nhất trong cả bốn hướng. Hình 9 cho thấy hình ảnh quét điểm xa nhất.

+Bước 2: Tìm trọng tâm bằng cách sử dụng bốn điểm vừa xác định được ở bước 1, bốn điểm trước đại diện các góc đa giác và trung tâm đa giác(COG) có thể được tính bằng cách sử dụng công thức (1)

+Bước 3: Để có được đường cơ sở, tiến hành kẻ đường thẳng nối trọng tâm đến gốc tọa độ. Hình 7 cho thấy đường cơ bản được tìm thấy.

+Bước 4: Tìm góc của đường cơ sở so với trục ngang để phát hiện góc nghiêng. Hình 8 cho thấy việc phát hiện góc nghiêng trên ảnh văn bản.

+Bước 5: Xoay ảnh với góc nghiêng tìm được theo chiều ngược chiều kim đồng hồ để được ảnh văn bản ngay ngắn, dễ nhìn.

Phương pháp có hiệu quả khi phát hiện và hiệu chỉnh góc nghiêng của văn bản được scan vào từ tạp chí, sách giáo khoa, báo chí và tài liệu viết tay, với độ phân giải khác nhau, phông chữ khác nhau với tỷ lệ chính xác khá cao. Phương pháp này đơn giản và độ phức tạp thấp dẫn đến thời gian thực hiện quá trình xử lý nhanh, không bị ảnh hưởng bởi nhiễu và phù hợp với văn bản có phông chữ khác nhau và cả các văn bản có độ phân giải khác nhau.

|  |  |
| --- | --- |
| *(a)Tìm các điểm xa nhất theo các hướng trên ảnh* | *(b)Trọng tâm được xác định dựa vào các điểm xa nhất* |
| *(c) Đường cơ sở được nối từ trọng tâm đến gốc tọa độ* | *(d)Xác định góc nghiêng ảnh văn bản* |

Hình 9.Minh hoạt thuật toán Center of Gravity

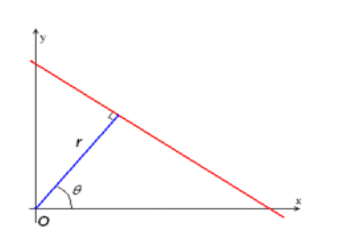
## **2.4 Phương pháp dựa vào biến đổi Hough(Hough Transform)**

Một phương pháp tiếp cận khác cho bài toán phát hiện góc nghiêng văn bản là phương pháp biến đổi Hough là xác định một số điểm đen và dùng biến đổi Hough để tác động lên các điểm đó.

Biến đổi Hough[6] ánh xạ một đường thẳng trong mặt phẳng thành các cặp (r,θ) trong không gian Hough với r là khoảng cách từ gốc tọa độ tới đường thẳng và θ là góc nghiêng của đường thẳng đó so với trục tung. Sử dụng tham số này thì phương trình đường thẳng có thể được viết là:

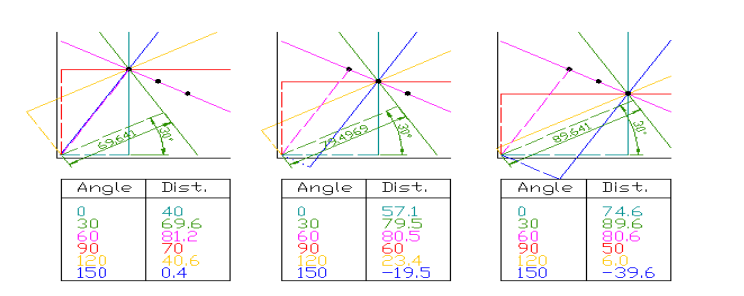
(2)

Do đó với mỗi đường thẳng được xác định trong không gian Hough sẽ có duy nhất một cặp (r, θ). Như vậy mỗi điểm bất kỳ trên mặt phẳng với trực tọa độ ( giả sử là (x0,y0)) thì các đường đi qua nó có dạng r(θ)=x0\*cos θ+y0\*sin θ với r được xác định bởi θ.



Hình 10.Đường thẳng Hough và trực tọa độ Oxy

Góc nghiêng của văn bản là góc có tổng số điểm nằm trên những đường thẳng cùng lệch góc lớn nhất. Số các điểm đen được áp dụng biến đổi Hough là tất cả các điểm đen hoặc chỉ những điểm thỏa mãn một số ràng buộc nào đó hoặc là đáy của các đối tượng ảnh.

**

Hình 11.Biểu diễn đường thẳng Hough qua 3 điểm

Những thuật toán của các tác giả: Hinds, Jiang, Dianel Le, Sugwara, Nakano, nhóm Srihari và Govindaraju, nhóm Yu và Jain, nhóm Admin, Fischer [4], Parkison và Rischy có liên quan đến cách tiếp cận nêu trên.

Nhưng chỉ có phương pháp của Srihari và Govindaraju là áp dụng biến đổi Hough cho tất cả các điểm đen của ảnh.. Việc áp dụng không có loại trừ một điểm nào dẫn đến chi phí tính toán rất lớn và ảnh hưởng đến độ chính xác của thuật toán. Để giảm thời gian và tăng chính xác, Hinds chỉ áp dụng biến đổi Hough cho một số ít điểm hơn bằng phân tích chạy dài theo chiều dọc. Mục đích của nén chạy dài theo chiều dọc trong thuật toán là để lấy ra các điểm đáy của các dòng văn bản, loại bỏ đi những điểm đen khác kể cả chùng thuộc vào một ký tự và áp dụng biết đổi Hough lên điểm đen đó. Tuy nhiên, chi phí của thuật toán này vẫn còn lớn và việc áp dụng biến đổi Hough cho tất cả các điểm đen ở đáy có thể dẫn đến kết quả sai trong trường hợp ảnh đầu vào có nhiều đối tượng phi ký tự như nhiễu, bảng biểu hay hình ảnh..

\*Trên cơ sở các thuật toán đã nghiên cứu để giải quyết bài toán góc nghiêng bằng cách áp dụng biến đổi Hough không xuất phát từ tốc độ xử lý mà chủ yếu căn cứ vào góc lệch tìm được tương đối chính xác hơn so với các phương pháp khi cùng xử lý trên 1 trang ảnh văn bản.

# **CHƯƠNG 3: THỬ NGHIỆM VÀ ĐÁNH GIÁ**

Để xác định hiệu quả của phương pháp tiếp cận, các thí nghiệm đã được tiến hành trên nhiều loại tài liệu có độ rộng khác nhau với bố trí đa dạng và mức độ góc xoay khác nhau. Những tài liệu này bao gồm không chỉ có văn bản, mà còn có đồ họa, bảng biểu, sơ đồ, công thức toán hoặc. 280 hình ảnh tài liệu được kiểm tra được sử dụng trong các thí nghiệm. Trong số này, 32 văn bản được lựa chọn từ cơ sở dữ liệu hình ảnh tài liệu Tiếng Anh UW, và 78 tài liệu được thu thập từ các tài liệu quét của học sinh( cơ sở dữ liệu NUSST) được thu thập từ dữ liệu thư viện của các trường đại học, 4 văn bản số là hình ảnh của máy fax. Góc nghiêng của các tài liệu này thường nhỏ, ví dụ khoảng [. Ngoài ra có thêm các dữ liệu được quét từ 6 tài liệu tiếng trung với kết quả 100 DPI, cũng có chứa một số bảng biểu hoặc đồ họa. Bên cạnh văn bản Trung Quốc, một số tài liệu cũng chứa văn bản tiếng Anh. Các dòng văn bản ngang và dọc có thể xuất hiện trong phạm vi một tài liệu, và có thể là ký tự tiếng Trung Quốc được đơn giản hóa hoặc ký tự tiếng Trung Quốc truyền thống. Ngoài ra, 3 tài liệu tiếng Tamil được quét để kiểm tra thêm khả năng xử lý các chữ viết khác nhau. Các hình ảnh tài liệu được quét này, cũng như một số lựa chọn từ các cơ sở dữ liệu UW và cơ sở dữ liệu NUSST, sau đó đã được xoay ở các góc khác nhau được chọn trước theo cả hai hướng cùng chiều kim đồng hồ và ngược chiều kim đồng hồ, từ đến , bằng cách sử dụng Adobe Photoshop. 166 hình ảnh tài liệu thu được thông qua cách này.

|  |  |  |  |
| --- | --- | --- | --- |
| **Actual Angle** | **Phương pháp Hashizume** | **Phương pháp Jiang** | **Phương pháp được đề xuất sử dụng giá trị trung bình** |
| 40 | 45.0000 | 38.9783 | 39.4325 |
| 30 | 26.4532 | 29.0956 | 30.4532 |
| 20 | 21.8767 | 20.6756 | 20.3124 |
| 10 | 10.7843 | 9.8485 | 9.8444 |
| 5 | 5.4403 | 4.9617 | 4.9732 |
| 2 | 0.0000 | 2.0934 | 3.0011 |
| -2 | 0.0000 | -1.9412 | -1.4391 |
| -5 | -5.7106 | -4.9063 | -10.9321 |
| -10 | -9.4623 | -10.5371 | -10.9321 |
| -20 | -18.4349 | -19.2619 | -19.8427 |
| -30 | -26.5651 | -29.4423 | -30.4917 |
| -40 | -39.5597 | -39.5675 | -39.8317 |

*Một số kết quả tiêu biểu về ước lượng góc xoay*

Một thử nghiệm được tiến hành trên trang văn bản nhiều màu bị nghiêng một góc với các đối tượng ký tự và phi ký tự xen lẫn nhau. Kết quả cho thấy, việc áp dụng biến đổi Hough cho kết quả xác định góc lệch tốt hơn các phương pháp phát hiện góc nghiêng khác tuy thời gian xử lý không phải là tối ưu nhất nhưng vẫn chấp nhận được. Bảng sau đây so sánh góc lệch tìm được vào thời gian xử lý giữa các phương pháp tiêu biểu giải quyết bài toán góc nghiêng văn bản.

|  |  |  |  |
| --- | --- | --- | --- |
| **Phương pháp** | **Góc lệch thực tế** | **Góc lệch tìm được** | **Thời gian(s)** |
| Profile Projection | 10 | 9.19 | 0.889 |
| Center of Gravity | 10 | 9.5 | 1.146 |
| Nearest Neighbour Clustering | 10 | 10.6 | 1.173 |
| Morphology[2] | 10 | 8.5 | 1.12 |
| **Hough Transform** | **10** | **10.12** | **1.098** |

***Bảng1.*** *So sánh hiệu quả giữa các phương pháp phát hiện góc nghiêng*

Phần tiếp theo của tài liệu trình bày phương pháp biến đổi Hough phát hiện góc nghiêng dựa trên những nghiên cứu này.

# **KẾT LUẬN**

# **TÀI LIỆU THAM KHẢO**

**Tài liệu tham khảo**

1. A.K. Das, B.Chada. A fast algorithm for skew detection of document images using morphological. Proc of International Journal on Document Analysis and Recognition, vol.4, 2001.

2. X. Jaing, H. Bunke, D. Widmer-Kljajo. Skew detection of document image by focused nearest-neighbourclustering. Proc. Of the 5th International Conference on Document Analysis and Recognition, Bangalore. 1999.

3. Shutao Li, Qinghua Shen and Jun Sun. Recognition Letters, Volume 28, Issue 5, 1 April 2007.

4. A Amin and S. Fischer, A Document Skew Detection Method Using the Hough Transform, Pattern Analysis & Applications, 2000.

5. W. P stl, “Detecti n f line blique st uctu es nd skew sc n in digitized d cu ents”. D cu ent An l sis nd Recognition, 1986.

6. S. Srihari and V. Gonvindaraju. Analysis of texual images using hough transform, 1989.

7. Atallah Mahmoud Al-Shatnawi and Khairuddin Omar. Skew Detection and Correction Technique for Arabic Document Images Based on Centre of Gravity. Journal of Computer Science 5, 2009