

# MỘT SỐ PHƯƠNG PHÁP HIỆU CHỈNH GÓC NGHIÊNG CỦA ẢNH VÀ ỨNG DỤNG

Nguyễn Trọng Vinh<sup>1</sup>, Ngô Quốc Tạo<sup>2</sup>, Trần Hành<sup>1</sup>, Huỳnh Cao Tuấn<sup>1</sup>, Trần Đức Toàn<sup>1</sup>, Nguyễn Bình Trọng<sup>1</sup>

<sup>1</sup> Đại học Lạc Hồng

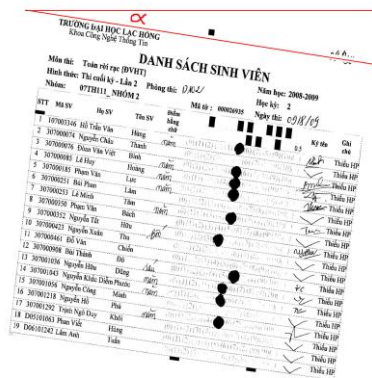
<sup>2</sup> Viện Công nghệ Thông tin, Viện Khoa học và Công nghệ Việt Nam

**Tóm tắt.** Bài báo này miêu tả chi tiết về cách tiếp cận áp dụng biến đổi Hough để giải quyết bài toán góc nghiêng ảnh văn bản. Dựa trên cơ sở trích chọn các đối tượng đặc trưng trên ảnh, gồm ba pha: 1) Trích chọn các đối tượng đặc trưng (là ký tự) trên ảnh văn bản; 2) Áp dụng biến đổi Hough lên điểm đại diện cho mỗi đối tượng đặc trưng; 3) Xác định góc nghiêng của ảnh dựa vào các đường thẳng Hough tìm được và xoay ảnh. Cách tiếp cận này sẽ giải quyết được bài toán góc nghiêng của ảnh văn bản khi ảnh có ít ký tự, hình vẽ và ký tự xen lẫn nhau, ảnh màu, cũng như vẫn tồn tại nhiều góc nghiêng trên ảnh.

**Từ khóa:** góc nghiêng, biến đổi Hough, điểm đặc trưng.

## 1 Đặt vấn đề

Hiện nay, thị giác máy tính đang là vấn đề quan tâm của nhiều nghiên cứu. Việc sử dụng máy tính để xử lý và nhận dạng ảnh giúp cải thiện tốc độ trong quá trình nhập dữ liệu ví dụ: nhận dạng chữ in, nhận dạng mẫu phiếu đánh giá hay bảng điểm sinh viên.... Vấn đề là khi quét vào máy tính, chúng ta không thể thu nhận được tài liệu như mong muốn bởi nhiều lý do khách quan khiến cho trang tài liệu bị nghiêng ngả, mờ nhoè, .... làm giảm hiệu quả của khâu nhận dạng sau này. Một giải pháp được nghĩ đến ngay đó là xây dựng các hệ thống hiệu chỉnh góc nghiêng văn bản đối với cả ảnh màu và ảnh trắng đen thuần túy.



Hình 1. Một ảnh bảng điểm bị nghiêng góc

Hiệu chỉnh góc nghiêng văn bản là một bài toán kinh điển trong xử lý ảnh văn bản. Giải quyết bài toán phát hiện góc nghiêng là nhiệm vụ tiên quyết của bất kỳ một hệ thống xử lý ảnh văn bản nào. Vì lẽ đó, cùng với sự phát triển của xử lý ảnh nói chung và xử lý ảnh văn bản nói riêng, bài toán góc nghiêng văn bản cũng được quan tâm ngày càng nhiều và dưới nhiều góc độ khác nhau. Có rất nhiều hướng tiếp cận cho bài toán góc nghiêng văn bản 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 vào biến đổi Hough (Hough Transform).... Các thuật toán phát hiện góc nghiêng thường

được xây dựng cho các hệ thống phân tích ảnh văn bản khác nhau nên chỉ giải quyết cho những loại ảnh văn bản cụ thể.

Bài báo 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 ảnh văn bản và kết quả thực nghiệm của phương pháp lựa chọn.

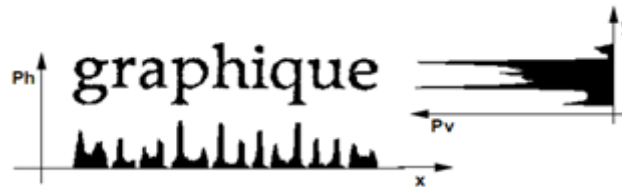
## 2 Các phương pháp phát hiện góc nghiêng ảnh văn bản

### 2.1 Phương pháp phân tích hình chiếu (Projection Profile)

Ý tưởng chính 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 số thuật toán - fiducial point) trong ảnh sao cho các điểm 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 của 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 Saiwai, 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 2. Hình chiếu dọc và ngang

Các thuật toán pháp 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  $\Phi$  có một mảng  $A_{\Phi}[]$  dùng lưu số điểm đại diện. Mảng  $A_{\Phi}[]$  là mảng một chiều, phần tử  $A_{\Phi}[r]$  sẽ cho biết số điểm đại diện nằm trên đường thẳng tạo với trục  $Ox$  góc  $\Phi$  và khoảng cách từ gốc tọa độ tới đường thẳng đó là  $r$ .

- + Khi tính được mảng  $A_{\Phi}[]$ , ta áp dụng hàm tối ưu hóa  $\Omega$  cho các giá trị của mảng này theo một tiêu chuẩn nào đó (có thể là tổng bình phương các giá trị trong mảng  $A_{\Phi}[]$  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  $\Omega$ .

\* 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  $\pm 15^\circ$ . 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.

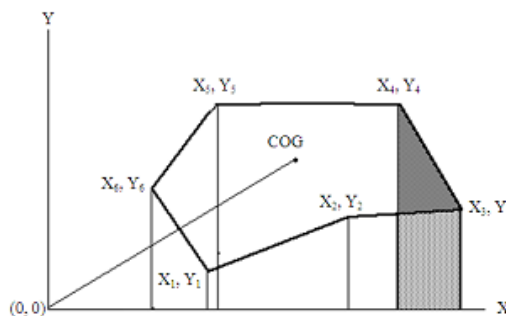
Gần đây, người ta kết hợp phương pháp hình chiếu và phương pháp các đối tượng nhằm giải quyết vấn đề giới hạn góc lệch, nhưng phụ thuộc nhiều vào khoảng cách giữa các dòng văn bản và chỉ xử lý được với những ảnh có chứa nhiều dòng văn bản và kích thước bé cỡ  $512 * 512$  pixels.

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 trọng tâm khối văn bản.

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

Ý tưởng chính của phương pháp là xây dựng một 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 mới trong thuật toán là áp dụng lên tất cả các từ nội tiếp trong đa giác. Trọng tâm của đa giác với góc 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.



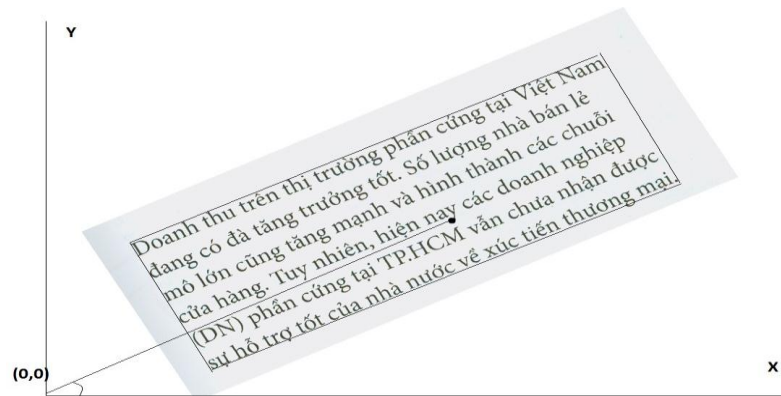
**Hình 3.** Đa giác 6 đỉnh và trọng tâm được xác định

Hình 3, một đa giác có 6 đỉnh được tìm thấy và trọng tâm của đa giác được xác định bằng công thức [7]:

$$C_x = \frac{1}{6A} \sum_{i=0}^{n-1} (x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i)$$

$$C_y = \frac{1}{6A} \sum_{i=0}^{n-1} (y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \quad (1)$$

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 như trong hình 4 cũng được mô tả như là một cách để xây dựng đường cơ sở giúp xác định góc nghiêng văn bản.



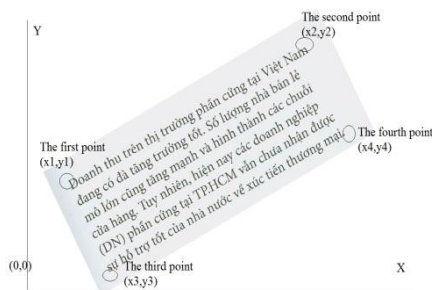
**Hình 4.** Hình chữ nhật ngoại tiếp ảnh văn bản

Thuật toán gồm các bước:

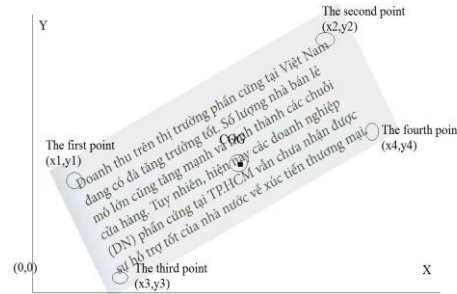
- + Đầu vào: ảnh văn bản bị nghiêng giống như hình 4
- + Đầu ra: ảnh đã được hiệu chỉnh góc nghiêng.
- + Bước 1: Xác định những điểm xa nhất trong tất cả bốn hướng. Hình 5 cho thấy hình ảnh quét điểm xa nhất
- + Bước 2: Tìm trọng tâm bằng cách sử dụng bốn điểm vừa xác định được ở bước 1, bốn điểm trước 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 sở 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 ảnh văn bản được scan vào từ tạp chí, sách giáo khoa, báo chí và tài liệu viết tay, với độ phân giải khác nhau, phong chữ khác nhau với tỷ lệ chính xác khá cao. Phương pháp khá 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 phong

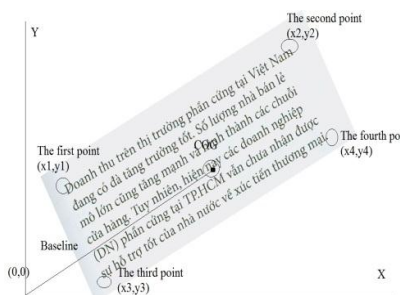
chữ khác nhau và cả các văn bản có độ phân giải khác nhau.



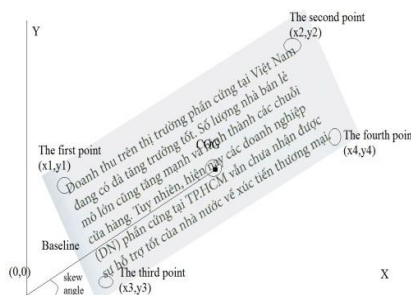
**Hình 5.** Tìm các điểm xa nhất theo các hướng trên ảnh



**Hình 6.** Trọng tâm được xác định dựa vào các điểm xa nhất



**Hình 7.** Đường cơ sở được nối từ trọng tâm đến gốc tọa độ



**Hình 8.** Xác định góc nghiêng ảnh văn bản

### 2.3 Phương pháp dựa vào biến đổi Hough (Hough Transform)

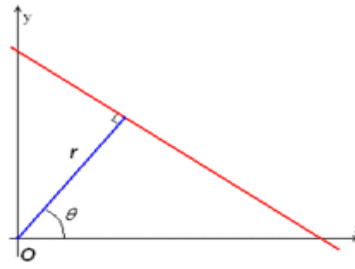
Một hướng tiếp cận khác cho bài toán phát hiện góc nghiêng văn bản là phương pháp dùng biến đổi Hough là xác định một số điểm đen và dùng biến đổi Hough tác động lên các điểm đó.

Biến đổi Hough [6] ánh xạ một đường thẳng trong mặt phẳng thành các cặp  $(r, \theta)$  trong không gian

Hough với  $r$  là khoảng cách từ gốc tọa độ tới đường thẳng và  $\theta$  là góc nghiêng của đường thẳng đó so với trục tung. Sử dụng các tham số này thì phương trình đường thẳng có thể được viết là:

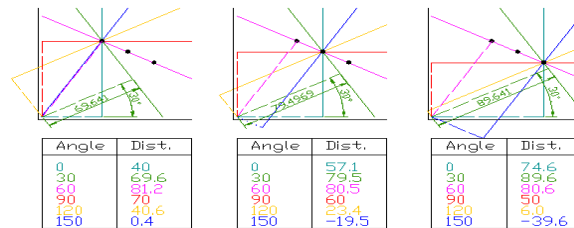
$$y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \left(\frac{r}{\sin\theta}\right) \tag{2}$$

Do đó với mỗi đường thẳng được xác định trong không gian Hough sẽ có duy nhất một cặp  $(r, \theta)$ . Như vậy mỗi điểm bất kỳ trên mặt phẳng ảnh với trục tọa độ (giả sử là  $(x_0, y_0)$ ) thì các đường đi qua nó có dạng:  $r(\theta) = x_0 \cdot \cos\theta + y_0 \cdot \sin\theta$  với  $r$  được xác định bởi  $\theta$ .



**Hình 9.** Đường thẳng Hough và trục tọa độ

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 10.** 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 Amin, Fischer [4], Parkison và Risky 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 chúng thuộc vào một ký tự và áp dụng biến đổi Hough lên điểm đen đó. Tuy nhiên, chi phí tính toán 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 những 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 khác khi cùng xử lý trên 1 trang ảnh văn bản.

Một thử nghiệm được tiến hành trên trang ảnh văn bản nhiều màu bị nghiêng một góc 10<sup>0</sup> 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 và thời gian xử lý giữa các phương pháp tiêu biểu giải quyết bài toán góc nghiêng ảnh văn bản.

**Bảng 1.** So sánh hiệu quả giữa các phương pháp phát hiện góc nghiêng

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 [2]	10	10.6	1.173
Morphology [1]	10	8.5	1.12
Hough Transform	<b>10</b>	<b>10.12</b>	<b>1.098</b>

Sau đây, chúng tôi trình bày phương pháp áp dụng biến đổi Hough phát hiện góc nghiêng dựa trên những nghiên cứu này.

### 3 Áp dụng biến đổi Hough phát hiện góc nghiêng ảnh văn bản

Về cơ bản thuật toán phát hiện góc nghiêng văn bản dựa vào biến đổi Hough gồm các bước chủ yếu sau:

Bước 1: Duyệt ảnh theo thứ tự từ trên xuống và từ trái qua phải, với mỗi điểm ảnh.

Với mỗi điểm vùng  $(x,y)$ , ta tiến hành kiểm tra:

Nếu đã xét qua thì cập nhật lại mảng tích lũy và lựa chọn điểm vùng tiếp theo.

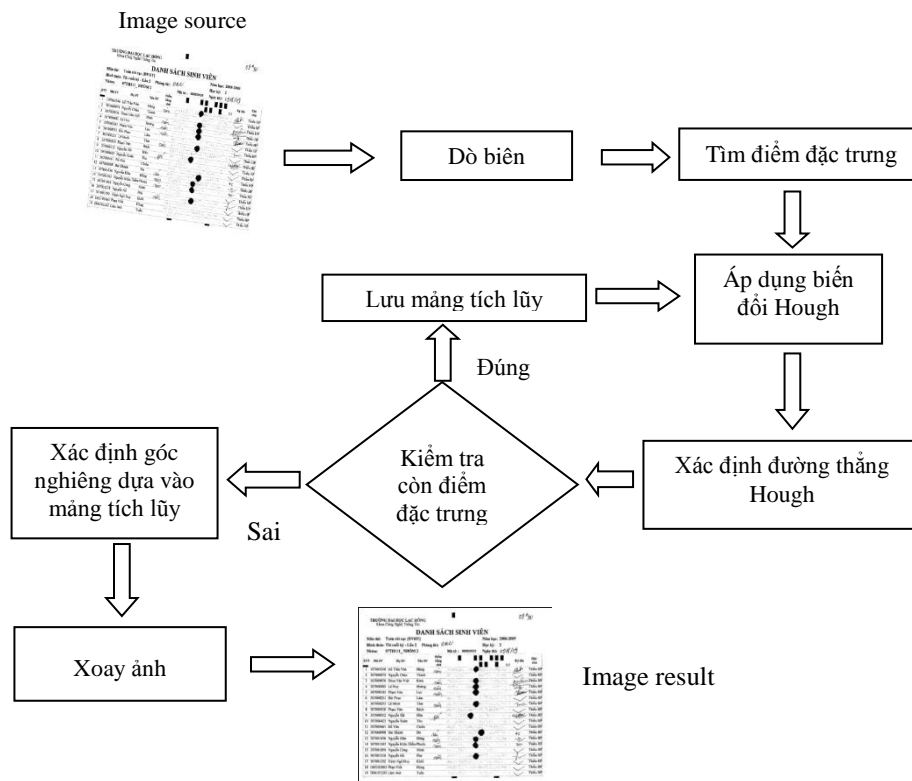
Nếu chưa xét thì xây dựng hình chữ nhật ngoại tiếp dựa vào thuật toán xác định chu tuyến. Áp dụng biến đổi Hough lên trung điểm cạnh đáy của hình chữ nhật ngoại tiếp. Cập nhật lại mảng tích lũy.

Bước 2: Dựa vào kết quả của phép biến đổi Hough ước lượng góc nghiêng cho ảnh văn bản.

Bước 3: Áp dụng thuật toán xoay ảnh lên trang văn bản với góc nghiêng tìm được ở bước 2.

Nếu áp dụng biến đổi Hough cho tất cả các đối tượng của ảnh thì thuật toán không chính xác hoặc là tốn nhiều thời gian thực hiện hoặc cả hai. Vì vậy chúng ta không áp dụng biến đổi Hough cho tất cả các đối tượng sau khi tìm được chu tuyến của chúng mà loại ra những đối tượng có kích thước quá lớn, hoặc là rất bé so với ký tự thường.

Sơ đồ tổng quát:



Hình 11. Sơ đồ tổng quát

#### 4 Kết quả thực nghiệm

Chúng tôi cài đặt thử nghiệm thuật toán trên một số ảnh văn bản (cả ảnh màu và ảnh đen trắng thuần túy) bị nghiêng và bước đầu cho kết quả khá tốt với góc nghiêng có thể xác định được là trong khoảng  $\pm 20^0$ . Như vậy với phương pháp này thì việc xác định góc nghiêng có thể áp dụng trên ảnh văn bản có cả ký tự và phi ký tự ứng dụng vào hỗ trợ nhận dạng các mẫu bảng điểm trắc nghiệm và tăng tính thẩm mỹ cho các trang tài liệu điện tử.

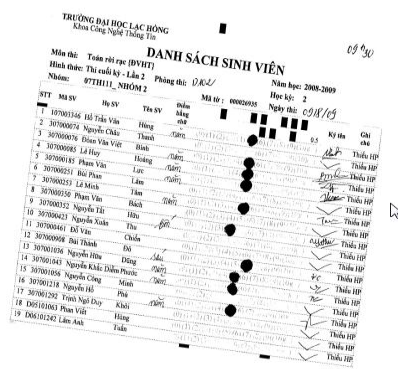
Chương trình được thực nghiệm trên nhiều ảnh màu với các đối tượng ký tự và phi ký tự xen lẫn nhau. Kết quả đạt được rất khả quan.

Hơn 300 ảnh được quét từ các tạp chí khoa học, tài liệu, giáo trình, hồ sơ thuộc nhiều thể loại và ngôn ngữ khác nhau. Bên cạnh đó, có khả năng xử lý trên nhiều định dạng ảnh khác nhau.

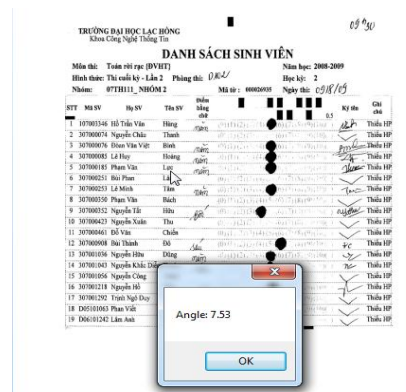


**Bảng 2.** Kết quả thực nghiệm trên một số ảnh văn bản

LOẠI TÀI LIỆU	ĐỊNH DẠNG	TỔNG SỐ ẢNH	XỬ LÝ ĐƯỢC	Tỉ lệ
Bảng điểm	BMP, TIF	150	150	100%
Sách, tạp chí, Giáo trình	JPG, BMP, TIF	100	96	96%
Hồ sơ, văn bản	JPG	40	39	97.5%
Sách ngoại văn	JPG, BMP	50	48	96%



**Hình 12.** Ảnh văn bản bị nghiêng góc



**Hình 13.** Ảnh văn bản sau khi được điều chỉnh góc nghiêng

Ngoài ra chương trình thực nghiệm có khả năng phát hiện được nhiều góc nghiêng trên cùng 1 trang ảnh văn bản. Giải quyết cho trường hợp văn bản bị nghiêng nhiều góc do thao tác quét của người sử dụng cũng có thể là do chính bản thân trang văn bản tồn tại nhiều góc nghiêng cần được hiệu chỉnh lại.



**Hình 14.** Trang ảnh văn bản tồn tại nhiều góc nghiêng được phát hiện

## 5 Kết luận

Một trang ảnh văn bản có thể được kết luận là bị nghiêng chủ yếu dựa vào các đối tượng chủ đạo là ký tự trên ảnh và đường thẳng nối các điểm giữa cạnh đáy của chúng tạo thành một góc lệch nào đó so với 2 trục toạ độ. Xuất phát từ nhận định đó và dựa trên cơ sở nghiên cứu các thuật toán Projection Profile, Center of Gravity, Hough, chúng tôi đã chọn và ứng dụng biến đổi Hough vào việc phát hiện và hiệu chỉnh góc nghiêng văn bản để việc xử lý ảnh văn bản đạt hiệu quả tốt nhất.

Giải quyết bài toán góc nghiêng ảnh văn bản dựa trên biến đổi Hough bước đầu cho kết quả rất khả quan và hỗ trợ tốt hơn cho khâu tiền xử lý ảnh trước khi nhận dạng. Đây là điều kiện tiên quyết của mọi hệ thống xử lý ảnh. Phương pháp áp dụng biến đổi Hough cho phép phát hiện ra nhiều góc nghiêng trên cả ảnh xám và ảnh màu nhưng chỉ dừng lại ở việc phát hiện mà chưa xử lý xoay ảnh với nhiều góc khác nhau trên trang ảnh văn bản. Nghiên cứu tiếp theo, chúng tôi sẽ tập trung vào việc phân vùng trang ảnh văn bản tương ứng với từng góc nghiêng tìm được để xử lý riêng biệt cho từng vùng cũng như nghiên cứu cải tiến một số thuật toán khử nhiễu và nâng cao chất lượng ảnh.

### 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-neighbour-clustering. 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. Postl, "Detection of linear oblique structures and skew scan in digitized documents". Document Analysis and Recognition, 1986.
6. S. Srihari and V. Govindaraju. Analysis of textual 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