Image skew detection and correction in regular images and document images

Sukumar Maji

111CS0582

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India.

Images skew detection and correctioin in regular images and document images

Thesis submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Technology

in

Computer Science and Engineering

by

Sukumar Maji

(Roll: 111CS0582)

under the guidance of

Prof. Pankaj Kumar Sa

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India.
February' 2015

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Orissa, India.

May 9, 2015

Certificate

This is to certify that the research work entitled as ***Image skew detection and correction of unbalanced skew angle of regular and document images*** by ***Sukumar Maji,111CS0582 (2014-2015)*** is a record of project work for the partial fulfillment of the requirements of the degree of Bachelor of Technology in Computer Science and Engineering.

Pankaj Kumar Sa
Asst. Professor

Acknowledgment

I would be grateful to acknowledge my guide, Prof. Pankaj Kumar Sa for his continuous enthusiasm and motivation during the total completion of the work. His reference towards the similar group of work and suggestion for different kind of case study was a constant flow of motivation.

I should be thankful to Prof. S.K.Rath HOD,CSE for granting me this project and belief in me.

I would like to thank all my batch mates for their help in case when i needed them.

And at last i would like to express my gratitude to NIT Rourkela for providing me such a vast environment and surroundings which helped to bring a lot knowledge and experience that i have gained through out this work.

Sukumar Maji

Abstract

During any Document scanning and processing of regular images in our daily life activities image skew is a very important part that should be kept in mind before processing the images. Skew is generally referred to the degree of rotation of an image in comparison with its actual position . So before proceeding to any further activity with the images we need to assure the skew of an image is correct or not. So detection of skew of an image would be the first thing to be applied to regular images some times and specially scanned documents when transforming them to appropriate format.

There are different algorithms for detection of skew of an image that have been implemented in different kind of works. The basic and very commonly used one is Scan line based skew detection. In this technique several lines are passed through the image from left to right, right to left, top to bottom and bottom to top and then the number of black pixels encountered in different projection of line are counted. The projection with maximum black pixels encountered is to be taken to consider the skew of the image.

There is another approaches like Hough transform, Base-point method etc. In hough transform method the pixel value is calculated for each value of θ. The angle producing maximum variance is considered to be the skew angle of the image.

These two algorithms have been implemented and the results have been represented to compare the accuracy.

Contents

List of Figures

Chapter 1

Introduction

1.1 Image Processing

In order to detect skew of an image we need to do operation with the images. But we can not directly operate with the images, we have to pass the image through a sequence of steps where the input is given image we wish to perform our operation. The output of such system is another image or a group of parameters or information generally called as feature set.

As in a system an image is nothing but a 2-dimensional frame points, an image generally is expressed a two dimensional function of x co-ordinate and y co-ordinate

$$I = \phi(x, y)$$

where x and y represents and amplitude ir value of ϕ is the intensity value of the image [1]. For a digital image these spatial co-ordinates and amplitude values are discrete.

Digital image processing is a part of digital signal processing. Generally there are two types of DIP commonly used in different group of works.

The first one is a processing system where both input and output are images. There are techniques such as image acquisition, image enchantment, image restoration, color image processing etc. which uses the above idea.

Then there is the another type where output of an input image is attribute value based on some particular data. There are morphological processing, object recognition etc. which are based on the second idea.

1.2 Document Image Processing

Document image processing is the more specific term of digital image processing where the input image is limited to the scanned documents. In our daily life there are many cases where we need to work with the soft copy of documents. There are cases where we might have got some disturbance with the scanned data. So at first we need to remove those disturbance present in the scanned documents. The output of Document IP is a compatible format which is easy to process and access. [2]

Generally old techniques for copying and manually checking each documents and then correcting is a very rigorous and slow process, very often which my take several months to complete such type of works. The use of computer in such case is really a new era of document image processing.

There is a common technique for document image processing which is called OCR. It is further divided into two steps. First of all we capture the text information out of the scanned documents. Which comprises of the orientation of text, color, font variation, tables etc. After collecting these information we need to process the graphical information which basically works on drawing, separating lines and paragraphs, logos and similar type of image representation.

A word processing system can handle documents with

- a better image quality
- differentiating the text with background images
- conversion of font size
- separation of document type script and hand written script

Current methods generally perform the job by changing the stream of characters into their ASCII values. This technique is quite familiar in bank in case of reading the code on the checks and also in the field of postal address gathering.

But these idea fails in case of recuperating the texts from an ancient book or script . In this case in stead manually recovering the texts or words form the script (like manual typing) is more fruitful rather than processing their materials through the approached algorithm mentioned above. This idea also can not determine the result where we need to search for a particular word inside the document.

1.3   Problem Definition

There are many uses of documentary image processing and there are different methods to perform those. The term skew is introduced in case of document scanning to detect the degree of deviation of an image from its correct position either in horizontal or vertical manner. So sequence wise detection of skew and then correction of the wrong skew is the main task to be performed.

1.4   Motivation

A literature survey analysis that is already been published indicates tw types of results.

- Result that provides accuracy in terms of skew angle is a bit slow.
- Result that runs fast comparatively has a lower accuracy.

It is user's choice to choose between the two above idea according to the type of result he wishes to get.

Chapter 2

Scan Line Based Skew Detection

2.1 Scan Line

To apply Scan Line to any particular document image we need to cast a horizontal lines across the image from one side to the other say from top to bottom or left to right. In this algorithm for document type of images we need to cast lines from left to bottom and then count the pixels encountered by those rays and then again performing the same operation with the same lines being rotated by a small amount. We track that angle for which the pixels encountered is highest as the skew of the images. [3]

In case of regular image we need to first convert the regular image into the its corresponding edge and then we perform the same operation mentioned above and check the skew.

2.1.1 Algorithm

1. First find out the coordinate of those scan lines that are at slope of θ in the plane. We can do this by bresenham's algorithm.

2. Count the number of pixels encountered by each of those scan line.

3. Then calculate variance value v of the total black pixels that is cut by the scan line for angle θ.

4. pick up that theta for which v is maximum.

Table

2.2 Time Complexity

This algorithm runs in linear time. If total pixel as counted is N then complexity occurs to be $O(N)$

2.3 Hough transformation

Hough transformation is one of the basic boundary detection and edge linking technique to identify location of a special class of shape. Main motive of this technique is to find the number of object or instances of a particular class of such instances or object drawn on a voting process. The process is done through the introduction of parameter space. For a particular parameter space objects are collected as local maxima. [4]The algorithm performs a total check then it computes the global maxima.

We know that any straight line can be expressed through the equation

$$y = mx + c$$

we call m the slope of the straight line and c the intercept og the y-axis. In hough transformation model we represent a straight line as the slope intercept model.

The same equation can be written in polar form

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

Where θ is the vector orthogonal to the line pointing towards the half upper plane. Now we can write it in a different way

$$r = x \cos \theta + y \sin \theta$$

so each edge of an image can be represented with a unique value of r and θ. Form these things it comes to be quite clear that a horizontal line will have θ equals to 0 and a vertical line will have θ equals to 90.

When we proceed for finding the skew first of all we have to interpret boundaries between which θ and r lies. we can notice that θ can lie between -90 to $+90$ degrees and r can lie between $-D$ and D. Where D is the maximum distance between opposite
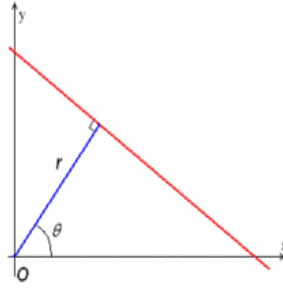
5

Figure 2.1: Representation of line
taken from Wikipedia

corner of the objects.

$-90 <= \theta <= 90$

$-D <= r <= D$, where D is the maximum distance between opposite corners of the object.

We take each pixel background or non-background and analyze. The pixel which is located at (i, j) are set with an accumulator value $A(i, j)$. Each pixel is the square associated with the point with r and θ values as (r_i, θ_j). We initialize each square to zero. Then we encounter each non-background point (x_k, y_k) in the XY plane. Now for each possible value of θ we compute the equation

$$r = x_k \cos \theta + y_k \sin \theta$$

we take all value of r for each possible value of θ and round off each r value to the near most point along r axis.

For an example let us take θ_m and we get from that $r = r_n$, then we update the value $A(m, n) = A(m, n) + 1$. The more subdivision we consider in the r, θ plane the more accurate result we get .

These are the sequential step to find out skew of an image through hough transformation method.

1. take each of the non-background pixel $P('x'_i, 'y'_j)$ in account.

2. Calculate the value of r for each value of θ i.e. $-90 <= \theta_i <= 90$ an round off the values obtained as result of r to the near most valid square cell value through the r-axis.

3. Increment the related matrix cell by one.

6

By computing all possible values of θ we get a matrix that is known as Hough matrix each cell of which (i, j) gives the number of points that lies on the line with the parameters set as (r_i, θ_j). The column cells of the matrix gives all og those points which lie in parallel line for whatever the values of r be. Now we have to find the variance of all the values in each column giving the result as variance of number background pixels that lie on a set of parallel line. Now th angle of skew is the angle at which the variance is maximum.

2.3.1 Implementation

If we take an image and perform Hough transformation then we will notice that the value of θ increases in the rows from left to right and the value of r increases in column from top to bottom. variance for each column is calculated and taken in comparison. Take the column which has maximum variance and choose the θ value for the corresponding column.

the following are the results.

2.4 Time Complexity

Time complexity depends upon the degree of precision we consider. It depends upon the number of θ we take.

Let the number of θ be N_θ and the number of non-background pixels be N_{nb}. Then the complexity is calculated to be $O(N_\theta N_{nb})$.

Chapter 3

Skew Detection for Regular Images

Scan line detection or Hough transformation is a good choice for skew detection for scanned documents or text images where we detect the skew of the image by casting rays along the image horizontally and vertically through different angles. generally in text images or scanned document type of images all texts are oriented in a particular direction so only object is text.

But in case of regular images there are many objects inside an image so we can not detect any particular object and then calculate its orientation.In that case we have to use edge detection technique. But then again in an complex image there are thousands of edges. Which edge do we chose to find the orientation ? That is a great question over this topic.

To solve this we have to divide the input image in to three color range i.e. Red, Green and Blue. for each image we have to observe all the color variation and then decide that taking the edge of which color part can make the task easier. For an example if we take a picture of sea side view and want to find its orientation then we have to take the blue part of that image to detect the edges. Where the horizon between sea and sky will be a good choice of edge to detect its skew. [5]

3.1   Canny Edge Detection Technique

The Canny edge indicator is an edge recognition process that uses a calculation in multi-step to distinguish variety of edges in pictures. It was created by John F. Watchful in 1986.

Edge discovery, particularly step edge location has been broadly connected in different diverse PC vision frameworks, which is a vital strategy to concentrate helpful auxiliary data from distinctive vision items and significantly diminish the measure of information to be transformed. Vigilant has discovered that, the necessities for the utilization of edge location on differing vision frameworks are moderately the same. Consequently, an improvement of an edge discovery answer for location these necessities can be actualized in an extensive variety of circumstances. The general criteria for edge identification incorporates. [6]

The first criteria of Canny's edge detection is low error rate, it should detect as many edges as possible from the input image. [7]

The edge point identified from the administrator ought to precisely limit on the focal point of the edge.

A given edge in the picture ought to just be checked once, and where conceivable, picture commotion ought not make false edges.

To fulfill the prerequisites Canny utilized the math of varieties a method which discovers the capacity which advances a given useful. The ideal capacity in Canny's indicator is portrayed by the total of four exponential terms, yet it can be approximated by the first subsidiary of a Gaussian.

Among the edge recognition techniques grew in this way, careful edge location calculation is a standout amongst the most entirely characterized strategies that gives great and dependable identification. Attributable to its optimality to meet with the three criteria for edge recognition and the straightforwardness of methodology for usage, it turns into a standout amongst the most prevalent calculations for edge discover.

3.1.1   Canny's Algorithm

The Process of Canny edge identification calculation can be separated to 5 unique steps:

1.Apply Gaussian channel for smoothening the picture keeping in mind the end goal to remove all noises.

2.Discover the power slopes or the intensity gradients of the picture.

3.Apply non-greatest concealment to dispose of spurious reaction to edge location

4.Apply twofold edge to focus potential edges.

5.Tracking the edge through hysteresis: Finalize the edge location by stifling the various edges that are feeble and not joined with solid edges.

3.1.2 Gaussian Filter

As a matter of fact that location of all edge results are effortlessly influenced by picture commotion, it is crucial to channel out the clamor to anticipate false recognition brought on by commotion. For smoothening the picture, a Gaussian channel is connected for convolution with the picture. This step will somewhat smooth the picture to lessen the impacts of evident clamor on the edge locator. The mathematical statement for a Gaussian channel part with the span of $2l + 1 * 2l + 1$ is indicated as takes after. [8]

$$H_{ij} = \frac{1}{2\pi\sigma^2} * \exp\left(-\frac{(i - l - 1)^2 + (j - l - 1)^2}{2\sigma^2}\right)$$

3.1.3 Image Intensity gradient

An edge in a picture generally points in a mixture of directions, for this purpose the Canny calculation utilizes four channels to identify diagonal edges with horizontal and vertical in the obscured picture. The edge identification administrator gives back a worth for the first subsidiary in the even course (Hx) and the vertical bearing (Hy). From this the edge angle and course can be dead set

$$H = \sqrt{H_x^2 + H_y^2}$$

$$\theta = atan(H_y, H_x)$$

where H is defined using the hypot function and atan is the inverse tangent function with H_x and H_y. The angle of direction for the edge is adjusted to any four direction say the two diagonal and the vertical and horizontal (0, 45, 90 and 135 degrees for instance). An edge heading in every shading area would be given a particular point values, for instance alpha lying in yellow district will be set to 0 degree.

3.1.4 Edge Thinning

It is actually a technique for edge thinning.

This technique is implemented for thinning the edges. In the wake of applying inclination estimation, the edge separated from the angle quality is still very obscured. As for criteria 3, there ought to just be one precise reaction to the edge. Accordingly non-greatest concealment can help to smother all the slope qualities to 0 aside from the neighborhood maximal, which demonstrates area with the most honed change of power quality. [9]The calculation for every pixel in the slope picture is:

Analyze the edge quality of the present pixel with the edge quality of the pixel in the positive and negative angle headings.

In the event that the edge quality of the present pixel is the biggest contrasted with alternate pixels in the veil with the same direction(i.e, the pixel that is indicating in the y bearing, it will be analyzed the pixel above and underneath it in the vertical hub), the worth will be safeguarded. Something else, the worth will be stifled.

In a few executions, the calculation classifies the nonstop inclination bearings into a little arrangement of discrete headings, and afterward moves a 3x3 channel over the yield of the past step (that is, the edge quality and slope headings). At each pixel, it smothers the edge quality of the middle pixel (by setting its esteem to 0) in the event that its greatness is not more noteworthy than the extent of the two neighbors in the inclination course.

3.1.5 Impure Edge Removal

After utilization of non-greatest concealment, the edge pixels are truly exact to present the genuine edge. In any case, there are still some edge pixels as of right now created by clamor and shading variety. So as to dispose of the spurious reactions from these troubling elements, it is vital to channel out the edge pixel with the powerless slope esteem and save the edge with the high angle esteem. In this manner two edge qualities are situated to elucidate the diverse sorts of edge pixels, one is called high limit esteem and the other is known as the low edge esteem. In the event that the edge pixel's slope worth is higher than the high limit esteem, they are stamped as solid edge pixels. On

the off chance that the edge pixel's slope worth is littler than the high limit esteem and bigger than the low edge esteem, they are checked as powerless edge pixels. On the off chance that the pixel quality is littler than the low edge esteem, they will be stifled. The two limit qualities are observationally decided qualities, which will need to be characterized when applying to distinctive pictures.

3.1.6   Edge Tracking By Hysteresis

As such, the solid edge pixels ought to absolutely be included in the last edge picture, as they are extricated from the genuine edges in the picture. Notwithstanding, there will be some open deliberation on the powerless picture pixels, as these pixels can either be separated from the genuine edge, or the clamor/shading varieties. To accomplish an exact result, the frail edges brought on from the last reasons ought to be uprooted. The criteria to figure out which case does the powerless edge fits in with is that, as a rule the feeble edge pixel brought about from genuine edges will be associated with the solid edge pixel. [10] To track the edge association, Binary Large Object-examination is connected by taking a gander at a frail edge pixel and its 8-associated neighborhood pixels. The length of there is one solid edge pixel is included in the BLOB, that frail edge point can be distinguished as one that ought to be protected.

Chapter 4

Result

14849090
18401391
18275881
17897514
18373699
13760834
18129604
16942013
18378043
14962057

(a)

14849090
18401391
18275881
17897514
18373699
13760834
18129604
16942013
18378043
14962057

(b)
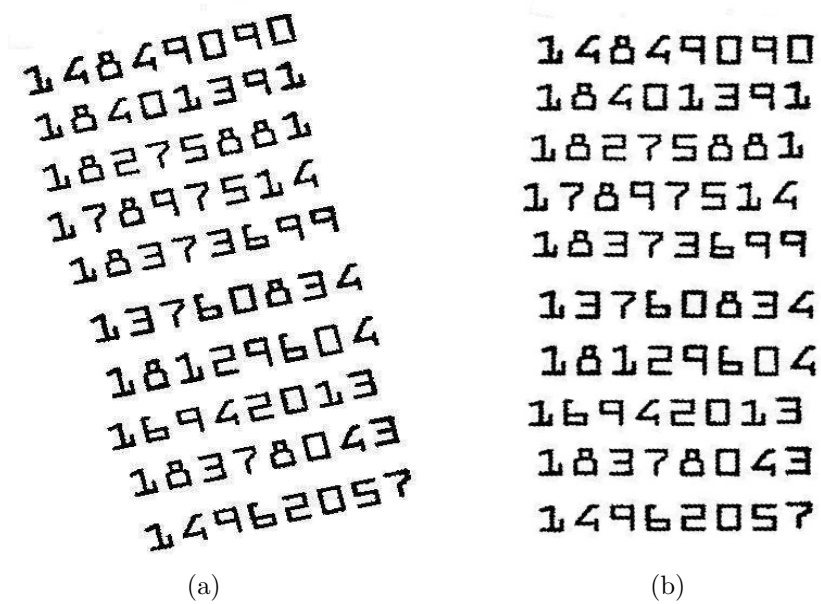
Figure 4.1: (a) Before experiment, (b) After experiment.
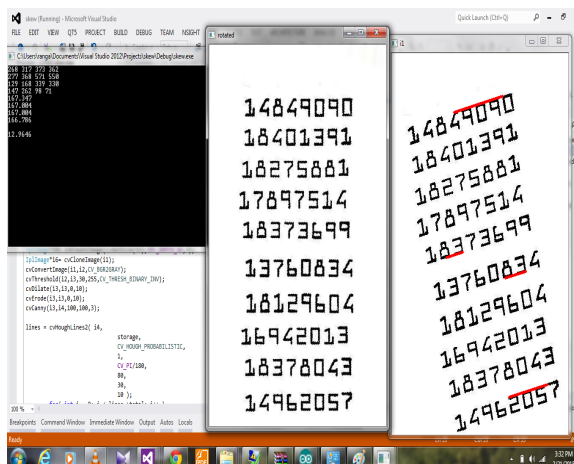
Figure 4.2: output in opencv

(a)   (b)

Figure 4.3: (a) Before experiment, (b) After experiment.
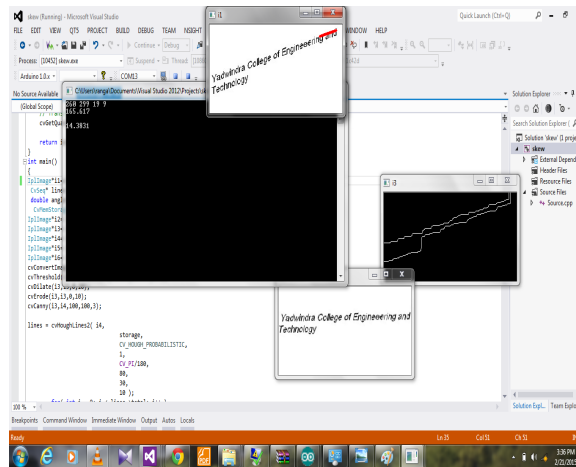
Figure 4.4: output in opencv

(a)   (b)

Figure 4.5:   (a) Before experiment, (b) After experiment.
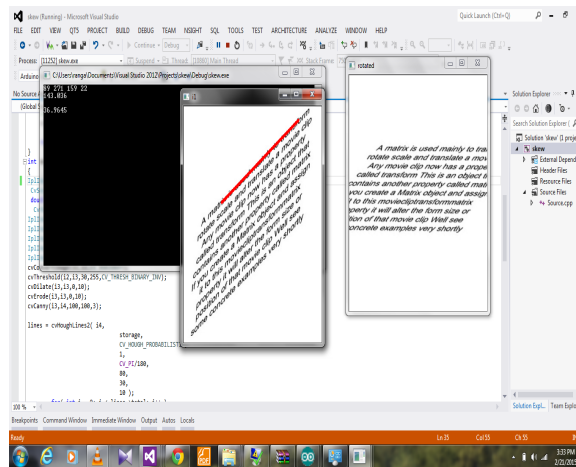
Figure 4.6: output in opencv

(a)   (b)

Figure 4.7: (a) Before experiment, (b) After experiment.
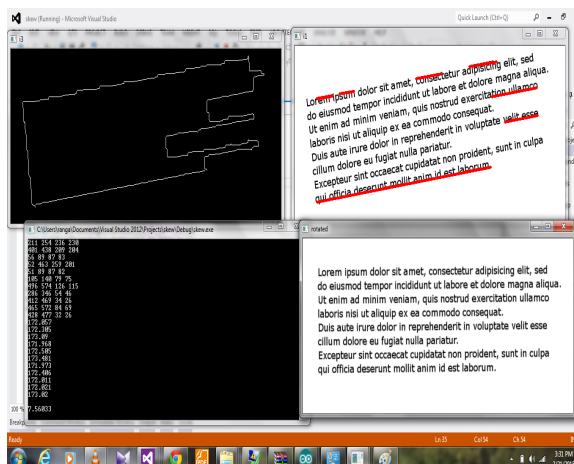
Figure 4.8: output in opencv

(a)

(b)

Figure 4.9: (a) Before experiment, (b) After experiment.

Figure 4.10: output in opencv

(a)

(b)

(c)

Figure 4.11: (a) Before experiment, (b) After experiment. (c) Detecting edge in blue region

18
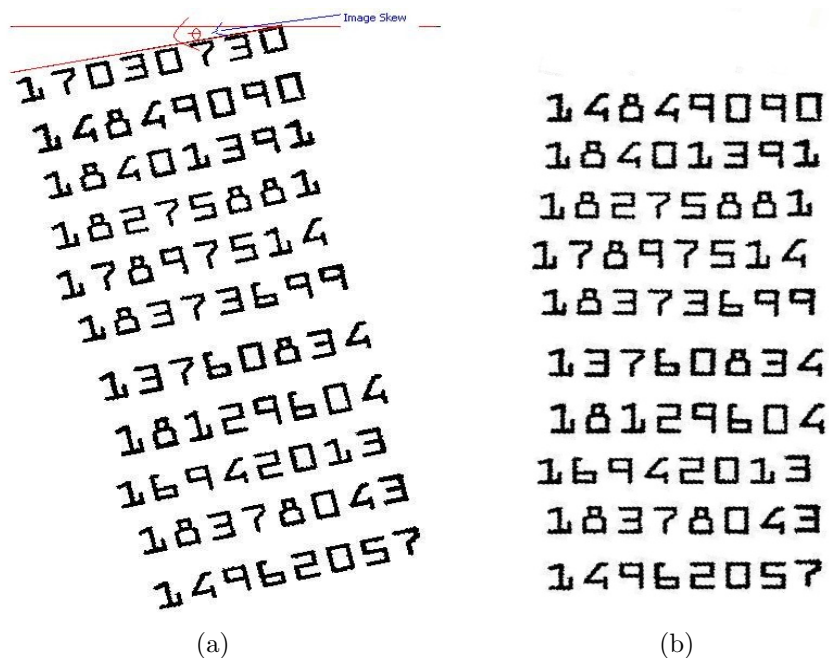
(a)   (b)

(c)

Figure 4.12: (a) Before experiment, (b) After experiment. (c) Detecting edge in blue region

(a)   (b)

(c)

Figure 4.13: (a) Before experiment, (b) After experiment. (c) Detecting edge in red region

(a) (b) (c)

Figure 4.14: (a) Before experiment, (b) After experiment. (c) Detecting edge in blue region

(a) (b)

(c)

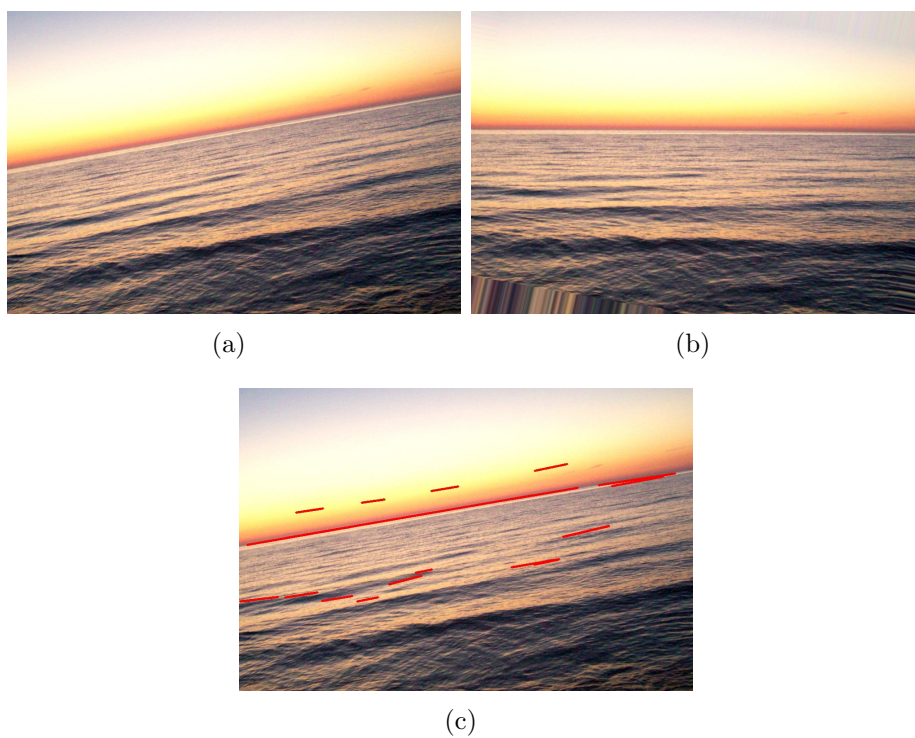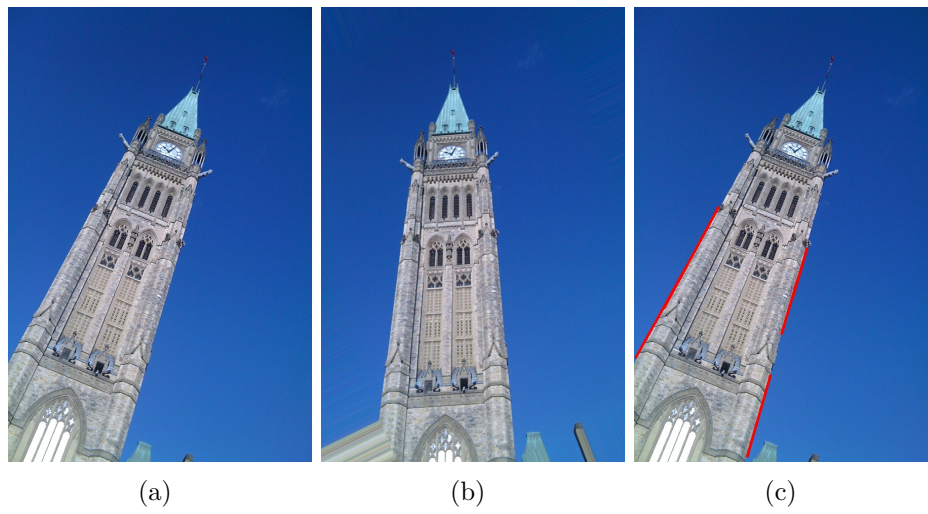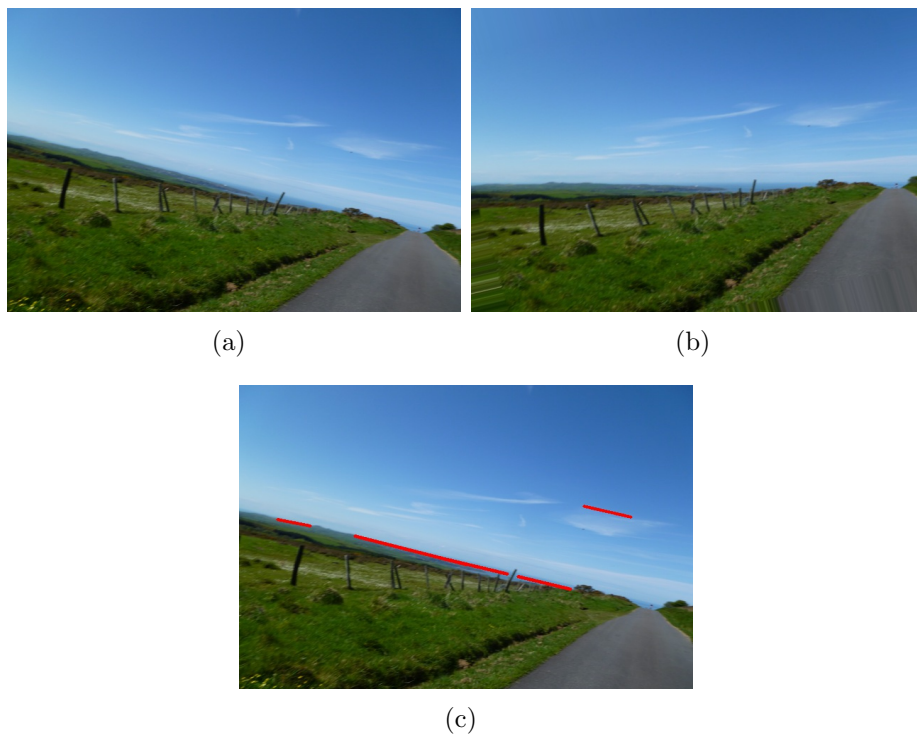Figure 4.15: (a) Before experiment, (b) After experiment. (c) Detecting edge in green region

<center>(a)   (b)   (c)</center>

Figure 4.16: (a) Before experiment, (b) After experiment. (c) Detecting edge in green region

Chapter 5

Conclusion and Future Work

For regular images the technique mentioned above is a good one but erroneous up to a very little extent. In the experiment done by me the the task is very smoothly performed up to the skew detection part. But after detecting the skew when we try to correct the image to its right orientation some pixel value of the image are lost. So some part of the image is mis matched with the original image . But again that is a very little part to observe.

More over this experiment runs on the set of images where there is a sharp edge in the images. Without any sharp edge with a minimum distinction of color we can not find the right edge to do the experiment. So this experiment is limited to landscape images. for more complex image of an unknown object (i.e. human organs) it may not give a correct output.

5.1 Limitations

The scan line algorithm which is based on the platform of calculating the coordinate of the line in the image and then finding the non background pixels in those lines and afterwards calculating the variance to detect the skew of an image is no doubt a good technique and efficient too. But this algorithm works for a range of images and with a range of skew related with them. The main disadvantage of this algorithm is that it is very slow in comparison to other algorithms.

If we take hough transform in account we will observe that the result of hough transform is as efficient as scan line algorithm but faster than scan line algorithm with

a lighter time complexity.

These above mentioned two algorithms work fine for any range of image even if the image consists of a set of parallel line with the text pixels. These two algorithms is best solutions for the images where major part of the image is set of lines and script.

This project can be proved to be a good effort to record script of old literature copy or old book which has very high maintenance cost.

5.2   Further Development

The proposed method works fine for titled skew but not very good for sheared skew as shear skew can not be detected by this method. Further this work can be updated to detect the skew of any regular images with a much higher range and correct the skew.

Bibliography

[1] J. S. Lim, "Two-dimensional signal and image processing," *Englewood Cliffs, NJ, Prentice Hall, 1990, 710 p.*, vol. 1, 1990.

[2] T. Pavlidis, *Algorithms for graphics and image processing.* Computer science press, 1982.

[3] X. Jiang and H. Bunke, "Edge detection in range images based on scan line approximation," *Computer Vision and Image Understanding*, vol. 73, no. 2, pp. 183–199, 1999.

[4] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[5] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6, pp. 679–698, 1986.

[6] J. F. Canny, "Finding edges and lines in images," *Massachusetts Inst. of Tech. Report*, vol. 1, 1983.

[7] R. Deriche, "Using canny's criteria to derive a recursively implemented optimal edge detector," *International journal of computer vision*, vol. 1, no. 2, pp. 167–187, 1987.

[8] G. Deng and L. Cahill, "An adaptive gaussian filter for noise reduction and edge detection," in *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.*, pp. 1615–1619, IEEE, 1993.

[9] A. Neubeck and L. Van Gool, "Efficient non-maximum suppression," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 3, pp. 850–855, IEEE, 2006.

[10] B. Green, "Canny edge detection tutorial," *from web resource. www. pages. drexel. edu/weg22/cantut. html*, 2002.