International Journal of Electrical and Computer Engineering (IJECE)
Vol. 6, No. 5, October 2016, pp. 2140~2149
ISSN: 2088-8708, DOI: 10.11591/ijece.v6i5.10785

An Unsupervised Classification Technique for Detection of Flipped Orientations in Document Images

Vijayashree C. S[1], N. Shobha Rani[2], Vasudev T[3]
[1]PES Research Center, PES Institute of Technology, Mandya, Karnataka, India
[2]Department of Computer Science, Amrita Vishwa Vidyapeetham, Mysore, Amrita University, India
[3]Maharaja Research Foundation, University of Mysore, Maharaja Institute of Technology, Mysore, India

Article Info

Article history:

Received Apr 10, 2016
Revised Jul 4, 2016
Accepted Jul 20, 2016

Keyword:

Directional Gradient
Document Image
Flipped Text Orientation
Unsupervised Learning

ABSTRACT

Detection of text orientation in document images is of preliminary concern prior to processing of documents by Optical Character Reader. The text direction in document images should exist generally in a specific orientation, i.e., text direction for any automated document reading system. The flipped text orientation leads to an unambiguous result in such fully automated systems. In this paper, we focus on development of text orientation direction detection module which can be incorporated as the perquisite process in automatic reading system. Orientation direction detection of text is performed through employing directional gradient features of document image and adapts an unsupervised learning approach for detection of flipped text orientation at which the document has been originally fed into scanning device. The unsupervised learning is built on the directional gradient features of text of document based on four possible different orientations. The algorithm is experimented on document samples of printed plain English text as well as filled in pre-printed forms of Telugu script. The outcome attained by algorithm proves to be consistent and adequate with an average accuracy around 94%.

Corresponding Author:

N. Shobha Rani,
Departement of Computer Science,
Amrita Vishwa Vidyapeetham, Mysore,
Amrita University, India.
Email: n.shoba1985@gmail.com

1. INTRODUCTION

Has come out clearly from the study on automatic reading through Optical Character Recognition (OCR) systems that the OCRs have limitations in reading and recognizing text documents which are not in proper textual orientation i.e. other than 00 text orientation. It is quite possible to have document images with text oriented at differing orientations due to the improper direction in feeding the documents to scanning device by the operators. The document images with text oriented in other than orientation are flipped documents and are not suitable to extract the information by OCRs. Such flipped document images need to be processed to fix the image into proper/normal orientation.

Normally document image suffers from either skew or improper orientation. A clear distinction lies between skewed documents and improper oriented documents. Skew is the small angular rotation/tilt to the document in the normal direction, whereas improper oriented document/flipped document corresponds to a total rotation to a document in a different direction other than the normal direction. The flipped documents conflicts the basic operating procedure for various document reading systems like character recognition systems, printers, photo copying systems and other imaging systems. With respect to OCRs, the improper text orientation introduces ambiguity in reading the text and results in erroneous recognition outcomes whereas in photocopying process introduces ambiguity in document feeding process to the devices and thus leads to non-uniform text orientations in the same photo copied document. The manual procedure of correcting document orientation intervenes and slows the basic operation of the document reading systems and other imaging devices. Especially the document reading systems like OCR, the most crucial stages of document image processing like segmentation, feature extraction and document classification are sensitive to the flipped orientation of the document images [1]. Flipped orientation in a document image is due to the error in placing/feeding the document into the scanning device [2],[3]. During the process of scanning, there is a possibility to feed the document to scanner by placing the document in wrong directions leading to generation of document images in corresponding flipped orientations. Improper oriented text is text which is not at 0o orientation to the page. Improper oriented document can be oriented at 90o or 180o or 270o to the page. The different types of text oriented samples are shown in Figure 1.

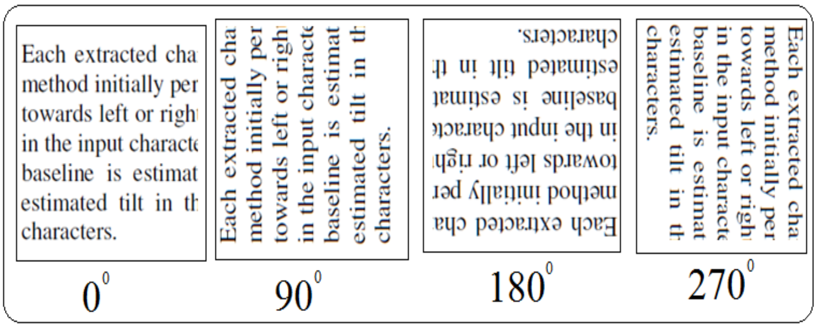


Figure 1. Different text orientations

The improperly oriented document images need to be preprocessed or converted to 00 orientation for correct reading by OCR. The methods reported in literature with respect to Document Image Analysis (DIA) is variant to rotation and demonstrate lower efficiency if the document image suffers from rotation. This demands the need for the input document images to be rotation free. In addition, OCRs also demonstrate a marked reduction in recognition rate if the input text images are oriented in any direction other than 00. Hence there is a need for a preprocessing stage which detects and corrects the flipped orientation of document image before it is subjected for reading by OCR.

We have focused in this research to detect the flipped orientation of the document image and transform it to achieve orientation at to the document page to prepare the transformed text more suitable for processing by OCR.

2. LITERATURE SURVEY

Considerable number of research works are noticed during literature survey on document orientation detection and correction. With the availability of apriori knowledge about the layout of the document, the page orientation detection would be made simplified to certain extent. But this apriori knowledge would not be available for generic cases and requires unsupervised classification approaches. The generic approaches for detection of page orientation direction calls for extraction of certain global features from the document. In this direction researchers have made attempts in detecting the page orientation direction and are reported in literature. Few related works referred are discussed in this section.

Caprari [4] has proposed a method for page up/down orientation detection model using Bayesian classifier. The algorithm operates on a bit-mapped text pattern array to determine the up/down orientation of the page, i.e. whether the page is upright or inverted by exploiting an up/down asymmetry of passages of text composed of roman letters and Arabic numerals. This approach is limited to classify only orientations in 00 and 1800. Nestares [5] has filed a patent for the detection of dominant orientation estimation by using seven steerable filters. The orientation detection is performed using steerable filters which provide an energy versus orientation curve of the image data. A maximum of the energy curve may indicate the amount of angular rotation that may be corrected by the orientation corrector. A methodology was proposed by Aditya et.al [6] to detect orientation in non-textual images by adopting Bayesian classifier for estimating the orientation and method is not extendable to text images.

D.X. Le et. al [7] proposed an algorithm to detect page orientation and skew in document using projection profile. The limitation of the methodology is that it can detect the orientation of the document in portrait and landscape direction but do not distinguish between normal and flipped with in portrait or landscape direction. Avila and Lins [8] suggested a fast method to detect skew and orientation in complex monochromatic document images. The method is limited only to detect skew between to to normal position but does not address the detection of flipped orientations. Vasudev et al [9] proposed a skew correction followed by orientation detection and correction approach. A non-rotational transformation model is applied in two stages. This approach works in detection of flipped orientation in monochrome text images with orientations. Murali et. al [10] have proposed the skew correction in the first stage which is based on line transformation model. In the second stage a simple x-cut and y-cut technique to determine the orientation of the document. A finer decision on orientation is made based on the domain knowledge of the pixel distribution in the document image. T Asano et al [11] presented an algorithm for rotating a sub image in place without using any extra working array. They overwrite pixel values with interpolated values. Only linear interpolation is considered and the correctness for large window sizes is not guaranteed. You Guang Chen et al [12] proposed a method for document orientation detection and classification by using Support Vector Machine (SVM) and then the orientation of unknown document images is classified. Veena et al [13] proposed a skew correction followed by orientation detection and correction of vehicle number plate. The proposed hybrid model work in two stages. It first detects the skew using Radon transformation and then the document orientation is detected using auto correlation. This approach works on skew and orientation detection in vehicle number plate images with orientations, the method is ideal for the images with very small number of characters and cannot be extended for images with large amount of text.

It is observed that few of the approaches reported in the literature are devised partly by employing knowledge base and remaining are micro level feature based approaches. The use of domain knowledge base or extraction of micro level features using point processing or block processing may slow this very basic operation i.e., detection of text orientation and correcting it. This motivates us to devise a macro level feature approach without employing any pre-built knowledge base for detection of flipped documents. Section 3 describes the methodology proposed for the detection and correction of flipped orientation in document images.

3. PROPOSED METHOD

The detection of text image orientation in the proposed work assumes the input as a plain English text document. Initially the input image is subject to pre-processing procedures followed by the directional gradient feature computation. Finally the features computed are subjected to a multi-level unsupervised classifier that predicts the orientation of text in the image. The block diagram of the text image orientation is shown in the Figure 2.

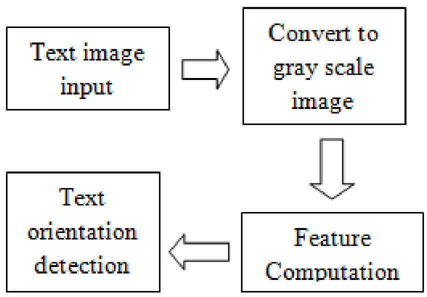


Figure 2. Block diagram for text orientation detection

The invocation of text orientation detection algorithm begins by acquiring text image as an input. The input image can be read in either 00, 900, 1800 or 2700 orientation of text. The computation of features and multi-level classification of text orientation is elucidated in the sub sections A and B.

3.1. Feature Analysis

The gray scale image is processed to obtain the gradient of the image [14]. The gradient information of an image with respect to a text orientation points to largest possible intensity increase in that text direction. Each pixel of a gradient image measures the change in intensity of the same point in the original image with regard to the orientation of text. The magnitude of the gradient represents how rapidly the intensity changes from one point to another point in the corresponding direction. The gradient of the image is given by (1).

$$\nabla f = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \partial f / \partial x \\ \partial f / \partial y \end{bmatrix} \tag{1}$$

where $\frac{\partial f}{\partial x}$ is the gradient in the x-direction and $\frac{\partial f}{\partial y}$ is the gradient in the y-direction. The gradient direction is computed by (2)

$$\theta = \tan^{-1}\left[\frac{g_y}{g_x}\right] \tag{2}$$

The gradient information of a typical text image with respect to 0^0, 90^0, 180^0 and 270^0 orientations is represented in Figure 3.

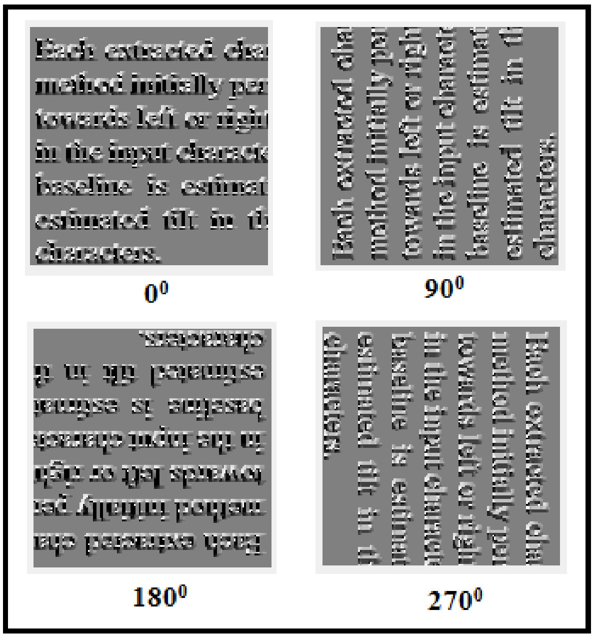


Figure 3. Text gradient at 0^0, 180^0, 90^0 and 270^0

The gradient features of input text images are computed with respect to four possible orientations. The gradient features are interpreted further for the selection of aggregate features. The feature selection is accomplished via summarizing the gradient features to statistical quantities like kurtosis and sum. The statistical quantities represent the aggregated gradient features of the image. The features of input image are computed for all possible four orientations. Further the selected features are forwarded for classification to detect the orientation of text in the input image.

Initially, the algorithm assumes the input document orientation as R_1 irrespective of its text orientation and obtain the rotated versions of input document in other three orientations R_2, R_3 and R_4 respectively. Figure 4 shows the input image and its various rotated versions for comprehension

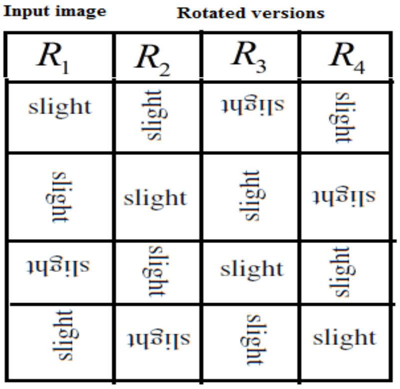


Figure 4. Input image and its rotated versions

From Figure 4, it is evident that processing of an input image considers the image at all its rotated orientations as indicated. The features of the rotated versions are employed for detection of text orientation direction in the input image R_1.

Let I is a gray scale image and R_1, R_2, R_3 and R_4 represents the rotated images of I in four different orientations respectively. The G_1, G_2, G_3 and G_4 represents the gradient features computed from R_1, R_2, R_3 and R_4 and $S_i * K_i$ is the features selected from the computed gradient features for the i^{th} orientation, where $i = 1, 2, 3, 4$. The proposed model for feature analysis and selection process of an image is shown in Figure 5.

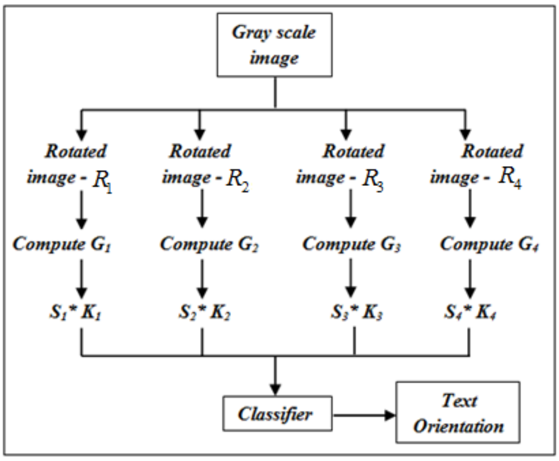


Figure 5. Proposed model for feature analysis and selection

The proposed algorithm demonstrates the process of feature computation and selection.

Algorithm_Feature_Analysis

1. Read an image I
2. Obtain the rotated images R_1, R_2, R_3 and R_4 of I.
3. Compute gradient feature vectors G_i where $i = 1, 2, 3, 4$ of rotated images.
4. Compute S_i, the sum of gradients and K_i the kurtosis of gradients where $i = 1, 2, 3, 4$.

   where $S_i = \sum_{j=1}^{n} \left[G_j \right]_i$ and $K_i = \frac{1}{n} \sum_{j=1}^{n} \left(\frac{[G_j]_i - \mu}{\sigma(G_i)} \right)^4$

'n' is the length of gradient and $[G_j]_i$ is a feature value at index 'j' of gradient feature vector G_i where, $i = 1,2,3,4$.

'μ' and $\sigma(G_i)$ is the mean and standard deviation of the gradient G_i with $i = 1,2,3,4$ respectively.

5. Compute the Orientation coefficient $\wp_i = S_i * K_i \;\; \forall \;\; G_i$ where $i = 1,2,3,4$.
6. Proceed $\wp_i \; \forall \; G_i$ for decision analysis, where $i = 1,2,3,4$.

The orientation coefficient $\wp_i$ is the product of sum and kurtosis features $(S_i * K_i)$ for each G_i and is further directed for detection of text orientation. The decision analysis for detection of text orientation is as discussed in the subsequent section.

3.2. Decision analysis for detection of text orientation

Decision analysis is one of the crucial procedures in any image processing system and also considered as the final stage that decides the efficiency of the system. In the proposed system the orientation detection of text in an image is done through constructing a decision tree, each level fulfills the criteria for one of the text orientation. The decision rules are devised by identifying a set of inferences from the features selected in the feature analysis stage. In regard with the orientation coefficient $\wp_i$ determined at orientations $i = 1,2,3,4$ the inferences 1, 2, 3 and 4 are derived for detection of text orientation.

Inference1: $\wp_1 < \wp_3 \; \&\& \; \wp_2 < \wp_4$

Inference 2: $\wp_1 < \wp_3 \; \&\& \wp_2 > \wp_4$

Inference 3: $\wp_1 > \wp_3 \; \&\& \wp_2 > \wp_4$

Inference 4: $\wp_1 > \wp_3 \; \&\& \wp_2 < \wp_4$

The Figure 6 depicts the decision tree for text orientation detection by employing the inferences obtained from orientation coefficient $\wp_i$ at $i = 1,2,3,4$.

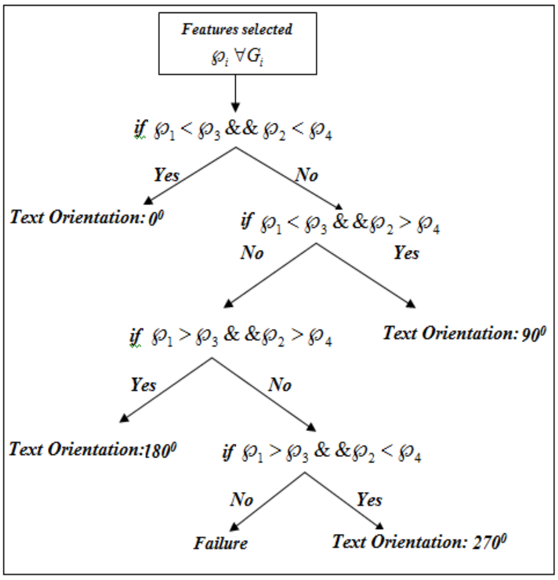


Figure 6. Proposed model for text orientation detection

At each level of decision, the analysis is performed with the observed inferences to determine the orientation of the text in the image. The observations drawn from the selected are employed as decision makers for the detection of text orientation. The experimental inferences drawn with respect to product features $\wp_i$ at $i = 1,2,3,4$ and inference 1 is plotted as Figure 7 and Figure 8.

Figure 7. Orientation coefficients- $\wp_{0^0} < \wp_{180^0}$

Figure 8. Orientation coefficients- $\wp_{90^0} < \wp_{270^0}$

The inferences of experimentation for inference 2 are plotted as Figure 9 and Figure 10.

Figure 9. Orientation coefficients- $\wp_1 < \wp_3$

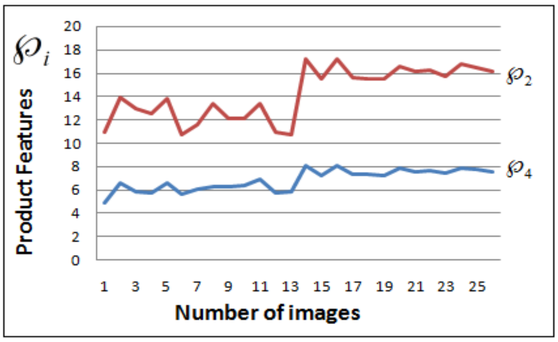


Figure 10. Orientation coefficients- $\wp_2 > \wp_4$

The inferences of experimentation for inference 3 are plotted as Figure 11 and Figure 12.

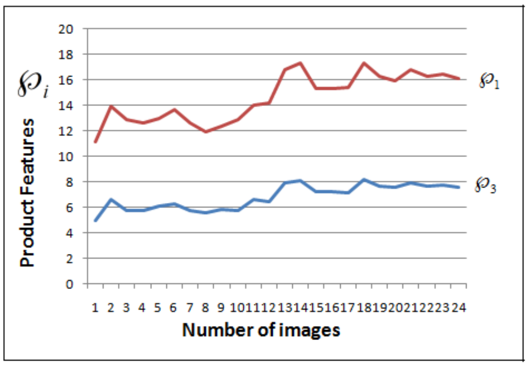


Figure 11. Orientation coefficients- $\wp_1 > \wp_3$

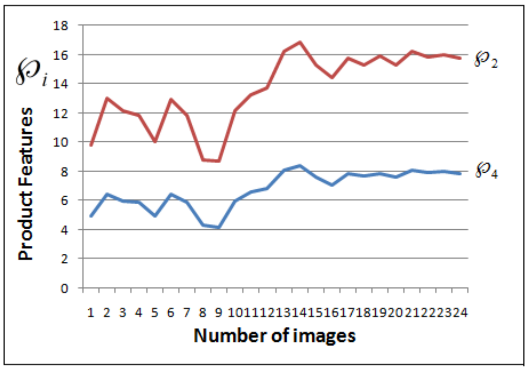


Figure 12. Orientation coefficients- $\wp_2 > \wp_4$

The inferences of experimentation for inference 4 are plotted as Figure 13 and Figure 14. From the interpretation of experimental inferences drawn from Figure 6 through Figure 13, it is evident that text orientation is discriminant from one class to another class with respect to the decision rules identified.

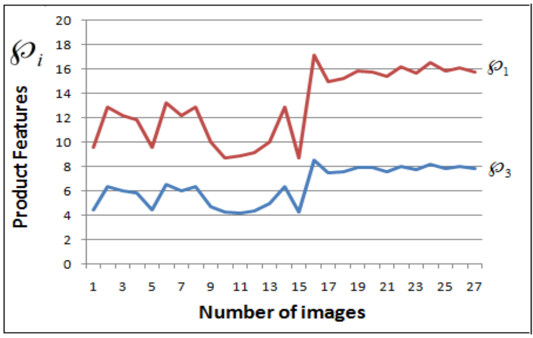


Figure 13. Orientation coefficients- $\wp_1 > \wp_3$

Figure 14. Orientation coefficients- $\wp_2 < \wp_4$

4. EXPERIMENTAL ANALYSIS

The proposed methodology for text orientation detection, the experimentation is performed using the datasets that includes 80 printed English documents and 50 Telugu *application form* documents. Each document is tested with all the orientations $0^0, 90^0, 180^0$ and 270^0. The experimental analysis of the proposed algorithm is discussed as follows.

If D_{ea} represents the total number of documents employed for experimentation in which 'e' indicates printed English documents and 'a'indicates the application form documents in D_{ea}. The number of recognized printed English documents is given by N_e and recognized *application form* documents represents N_a, then the text orientation detection rate is given by equation (1).

$$\text{Text orientation rate, } T_{orient} = \frac{N_e + N_a}{D_{ea}} \tag{1}$$

If the false positive rate of document recognized is given by $F(N_e + N_a)+$, then the true positive rate of orientation detection is given by $T(N_e + N_a)+$ and is depicted in equation (2). True positive rate,

$$T(N_e + N_a)+ = \frac{[(N_e + N_a) - F(N_e + N_a)+]}{D_{ea}} \tag{2}$$

The experimental statistics of the proposed methodology is tabulated in Table 1.

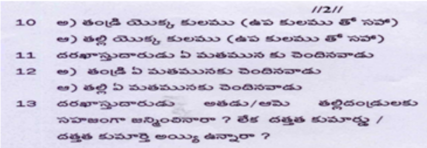
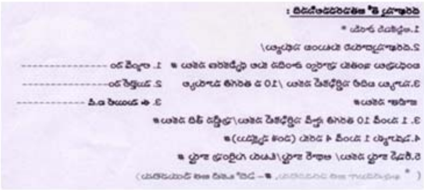
Table 1. Experimental analysis of flipped orientation detection

Orientation	Number of Samples English Document	Number of Samples Telugu Document	Orientation Detection Rate T orient	True Positive Rate $T(N_e+N_a)$
0^0	80x4	50x4	96.15%	93.84%
90^0	80x4	50x4	90.76%	87.69%
180^0	80x4	50x4	96.92%	94.61%
270^0	80x4	50x4	92.30%	88.46%

Some of the input document images considered and the orientation direction detected by the proposed methodology is listed in the Table 2.

The methodology fails to work for some cases like document images with low resolution. It further fails where the document images have more number of pictures/emblems and for documents with very little text. If the document image has varying fonts, this would also result in wrong orientation direction detection. Figure 15 illustrates some of instances of input images where the detection of orientation fails.

Table 2. Samples of document images with detected orientation

Input Sample	Orientation Detected
	0 degree
	180 degree
	270 degree
	90 degree

OZONE OVER SCOTLAND

We've made a unique move here in Scotland that I'm sure is a first in the industry. In 1990, all of the paintball sites in our country will operate without the use of ozone-destroying CO_2 gas for powering guns.

Ever since February of 1987 my company, Alternative Leisure, operating the survival game throughout Scotland, have used air conversion kits on all guns. The system has major cost advantages both to players and site owners and as a result all sites, franchises and independents will be running on air by the fall of 1990.

Now guns provide the capacity of constant air combined with a propellant that produces consistent and accurate shots over a wide range of temperatures. Air bottles are refilled from conventional scuba tanks that have negligible refill costs. These savings are passed on to players.

Not only does this keep us within firearm laws, it's cost effective and the industry collectively is doing its bit to preserve the environment on which we depend for our living.

This action will insure that paintball sites throughout Europe will remain environmentally friendly in the 1990s. If there is a hole in the ozone layer it sure isn't going to be over Scotland!

Andy Park
Alternative Leisure
Glasgow, SCOTLAND

HINDI / ENGLISH
TYPING SERVICE
PROVIDING HINDI & ENGLISH
TYPING SERVICE IN MS-WORD / PDF / JPEG
(IMAGE) FORMAT

Figure 15. Samples of document which resulted in wrong orientation detection

5. CONCLUSIONS

The work presented detects the flipped orientation direction of document image. The method is focused for detecting document orientation for text documents. The methodology has been tested for English text documents and Telugu text documents. Based on the samples tested, this method is successful on most of the text document images including those which contain small emblems. Document images in the form of application forms have also been successfully tested. This method of orientation detection performs effectively with an overall efficiency of 94%.The method does not require any special hardware setup to acquire the document image, the image obtained by flat-bed scanner is sufficient. This method can very well be extended for other language text document images.

REFERENCES

[1] Srihari S. N., "Handwritten address interpretation: A task of many pattern recognition problem," *Journal of Pattern Recognition and Artificial Intelligence*, vol/issue: 14(5), pp. 663-674, 2000.
[2] R. C. Gonzales and R. E. Woods, "Digital Image Processing," 2nd Edition, Pearson Education Publication, 2002.
[3] L. O'Gorman, R. Kasturi, "Executive Briefing: Document Image Analysis," *IEEE Computer Society Press*, 1998.
[4] Caprari R. S., "Algorithm for Text Page Up/Down Orientation Determination," *Pattern Recognition Letters*, vol/issue: 21(4), pp. 311-317, 2000.
[5] O. Nestares, "Automatic Dominant Orientation Estimation in Text Images Based on Steerable Filters," IPC8 Class: AG06K932FI, USPC Class: 38229, Agents: COOL PATENT, P.C.;c/o CPA Global, 2009.
[6] A. Vailaya, *et al.*, "Automatic Image Orientation Detection," *IEEE transactions on Image Processing*, vol/issue: 11(7), 2002.
[7] D. X. Le, *et al.*, "Automated Page Orientation and Skew Angle Detection for Binary Document Images," *Pattern Recognition*, vol/issue: 27(10), pp. 1325-1344, 1994.
[8] B. T. Avila and R. D. Lins, "A fast orientation and skew detection algorithm for monochromatic document images," *Proceedings of the ACM symposium on Document engineering*, pp. 118-126, 2005.
[9] Vasudev T. and P. Nagabhushan, "Detection and Correction of Page Orientation in Monochrome Textual Document Image," *Journal of Computer Science*, vol/issue: 2(1), pp. 465-473.
[10] Murali S., *et al.*, "Language Independent Skew Detection and Correction of Printed Text Document Images: A Non-rotational Approach," *International journal of Artificial Intelligence*, vol/issue: 16(2), pp. 08-15, 2006.
[11] T. Asano, *et al.*, "In-place algorithm for image rotation," *Algorithms and Computation*, pp. 704-715, 2007.
[12] Y. G. Chen, *et al.*, "A method for detecting document orientation by using SVM classifier," *International conference on Multimedia Technology (ICMT)*, pp. 47-50, 2011.
[13] Veena M. N., *et al.*, "Detection of oriented and skewed number plate in vehicle using hybrid approach," *International Journal of multimedia and its applications (IJMA)*, vol/issue: 7(4/5), 2015.
[14] C. L. Liu, "Normalization-cooperated gradient feature extraction for handwritten character recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol/issue: 29(8), 2007.

BIOGRAPHIES OF AUTHORS

Vijayashree C. S, obtained her B.E. Degree in Computer Science from B.I.T, Bangalore and M.E. Degree in Computer Science from U.V.C.E, Bangalore. She is pursuing research towards her Ph. D. Degree in Computer Science of University of Mysore, Mysore, at P.E.S. College of Engineering, Mandya.

N. Shobha Rani is currently working as Assistant professor in Amrita Vishwa Vidyapeetham University, Mysore an also pursuing her Ph.D degree in Maharaja Research Foundation, University of Mysore, Mysore. Her areas of interest include Document image processing, Optical character recognition and Computer vision.

Dr. Vasudev T is Professor, in the Department of Computer Applications, Maharaja Institute of Technology, Mysore. He obtained his Bachelor of Science and post graduate diploma in computer programming with two Masters Degrees one in Computer Applications and other one is Computer Science and Technology. He was awarded Ph.D. in Computer Science from University of Mysore. He is having 30 years of experience in academics and his area of research is Digital Image Processing specifically document image processing and authored about 50 research papers.