Accuracy of Gradient based Skew Estimation

Florian Kleber, Markus Diem and Robert Sablatnig
Computer Vision Lab
Vienna University of Technology
Vienna, Austria
{kleber, diem, sab}@caa.tuwien.ac.at

***Abstract*—Document preprocessing is the first step of document image analysis systems and comprise a skew estimation of documents. The gradient information can be used to estimate the skew due to the characteristic of latin characters and specific fonts which are composed of horizontal and vertical strokes. This paper shows the accuracy of a gradient orientation measure on synthetic data and for printed text documents. Also the difference of the accuracy on a binarized and grayvalue dataset is presented.**

I. Introduction

Document skew detection estimates the global orientation of a document. A combination of gradient based and focused nearest neighbour clustering of interest points skew estimation is introduced in [1]. An extension of [1], tailored for machine printed documents, which incorporates line and paragraph information is presented in [2]. The following section shows the introduced errors of the gradient orientation measure presented in [1], [2].

II. Accuracy Evaluation

The gradient orientation estimation is a pixel based method. Its key concept is that script comprises mostly vertical or horizontal strokes. This assumption can be verified if solely printed text is considered. However, for handwritten text with slant, the modal angle corresponds to the slant and not to the text line angle. Nevertheless, this methodology has the advantages of considering additional information like ruling lines or underlines, provides an accurate angle estimation (median error $< 0.08\,^\circ$; PRIMA 2009 dataset) and can also deal with sparsely inscribed documents. The only preprocessing of the gray value image is a smoothing with a Gaussian kernel ($\sigma = 12$, see Table I) to suppress the gradient information of noise and background clutter. In contrast to a binarization a Gaussian smoothing cannot introduce any new structures. To make a directional gradient analysis, all gradients are accumulated into an orientation histogram $H_n(\phi)$ that consists of n bins representing the global skew domain $\phi\ [-\frac{\pi}{2}\ \frac{\pi}{2})$. The number of bins n define the angle resolution $r = n/180$ of the orientation histogram. For a detailed description see [2].

To show the accuracy of the gradient orientation measure, the method is evaluated on a synthetic dataset comprising a line, which is rotated from $-90\,^\circ$ to $+90\,^\circ$ with steps of $0.1\,^\circ$ and smoothed with different Gaussian kernels, referred as line dataset. Table I shows the mean, the median error and the CE (Correct Estimations) of the line dataset. It is shown that the mean and median error decreases with a higher σ from $0.12\,^\circ$ ($\sigma = 2$) to $0.007\,^\circ$ ($\sigma = 14$). Thus, the gradient orientation measure achieves a mean error below $0.008\,^\circ$ for a σ higher than 12. Figure 1 shows the angular error of the line dataset over the entire angle range for $\sigma = 12$. Two errors are accumulated: the first one has a periodic pattern with a frequency of $1\,^\circ$. It is introduced since the final angle is chosen as the median value of all angles of the highest bin in the orientation histogram. Thus, the smallest error occurs at the exact angle related to the corresponding bin. Angles in the range between two bins have a higher error because of the uneven distribution of the angles between two bins (see also Figure 3). The second error with a frequency of $45\,^\circ$ is introduced due to aliasing effects.

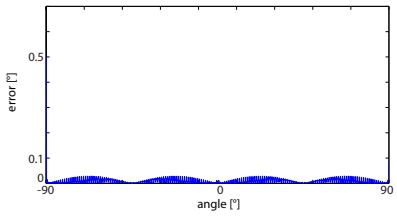

Fig. 1. Angular error of the line dataset ($\sigma = 12$).

To show the accuracy of the gradient orientation measure for printed text documents, and that the same errors apply for printed text, a synthetic test set comprising a text page of the PRIMA 2009 dataset is rotated from $-90\,^\circ$ to $+90\,^\circ$ with steps of $0.1\,^\circ$ and smoothed with different Gaussian kernels, referred as single page dataset. Table I summarizes the gradient error of the single page dataset for different σ. The mean and the median error decreases with a higher σ. The same effect applies to the line dataset. However, the error is higher for the same σ due to the distribution of the angles. For all further tests a σ value of 12 is chosen as a tradeoff between accuracy and the computational effort.

The gradient orientation histogram of the single page dataset skewed with $-34\,^\circ$ is shown in Figure 2. The highest peak is at the correct bin, corresponding to $-34\,^\circ$. By determining the main orientation by choosing the median value of all angles of the highest bin, the same angular errors occur as shown within the line dataset. The angular error of the single page dataset is visualized in Figure 3, where also a detail of the angular error at $-34\,^\circ$ is shown. Figure 3 shows also the ratio of the highest bin and its neighbor bins of the correct angle $-34\,^\circ$ and the angular neighborhood. In order to improve the

TABLE I
GRADIENT ORIENTATION MEASURE RESULTS ON THE LINE AND SINGLE PAGE DATASET, VARYING SIGMA.

line dataset									
sigma	2	4	6	8	10	12	14		
mean error [°]	0.12	0.05	0.01	0.01	0.01	0.008	0.007		
median error [°]	0.11	0.05	0.01	0.007	0.006	0.006	0.005		
CE [%]	42.47	99.88	99.88	99.88	99.88	99.88	99.88		
single page dataset									
sigma	2	4	6	8	10	12	14	16	18
mean error [°]	0.46	0.25	0.25	0.20	0.17	0.14	0.13	0.11	0.10
median error [°]	0.40	0.21	0.20	0.17	0.15	0.13	0.12	0.10	0.09
CE [%]	10.82	21.37	25.37	29.42	31.81	38.42	41.69	48.75	51.69

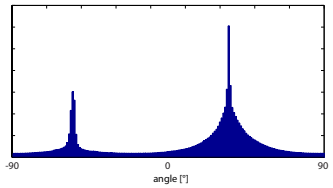

Fig. 2. Gradient orientation histogram of a document page with an orientation of $-34\,^\circ$ ($\sigma = 12$).

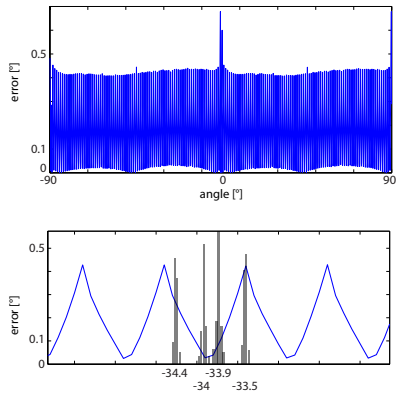

Fig. 3. The angular error of the single page dataset and a detail of the angular error at $-34\,^\circ$. Gray bins show the distribution at different angles. Note that the distribution is asymmetric for angles close to $.5\,^\circ$ while it is symmetric for integers.

result, the distribution of the angles of the highest bin and its neighbors is taken into account. The ratio of the left neighbor bin and the highest bin is denoted by r_l and the ratio of the right neighbor bin and the highest bin is denoted by r_r. Finally the angles of the three highest bins are sorted according to their values. All angular values from the highest bin are taken into account. The number of angles from the neighboring bins taken into account depends on the ratio r_r and r_l (i.e. $r_l\%$ of the angles of the left bin and $r_r\%$ of the angles of the right bin). The final orientation is determined by the median of the determined angle set, which is equally distributed. This leads to a mean error of $0.05\,^\circ$, a median error of $0.039\,^\circ$ and a CE of 94.72%. Thus, the mean error is improved by $0.09\,^\circ$ and the median error by $0.091\,^\circ$. The CE improves from 38.42% to 94.72%. To show the influence of a binarization, the line dataset and PRIMA 2009 dataset have been binarized. The line dataset has a mean error of $0.008\,^\circ$ compared to a mean error of $0.037\,^\circ$ if the dataset is binarized. For the PRIMA 2009 dataset, the mean error is $0.087\,^\circ$ compared to $0.158\,^\circ$ (binarized).

TABLE II
GRADIENT ORIENTATION ERRORS ($\sigma = 12$) ON SINGLE CHARACTERS OF THE ALPHABET.

	mean error [°]	median error [°]	CE [%]
alphabet, lower case	3.39	0.02	57.69
alphabet, upper case	1.41	$0.4 \cdot 10^{-5}$	65.38

To demonstrate the performance of the gradient orientation measure, the method has been tested on all single characters (Times New roman, 12 pts, 300 dpi, upper and lower case) of the alphabet. Table II shows the results for the alphabet printed in lower and upper case. The maximal error of $43\,^\circ$ and $39.9\,^\circ$ occur for the characters *o* and *x*. Without regarding these 2 characters the mean error for the lower case alphabet is $0.2\,^\circ$. For the upper case alphabet, the maximal error of $10\,^\circ$ and $21\,^\circ$ occur for the characters *C* and *L*. If the 2 characters are not considered, the mean error of the upper case alphabet is $0.16\,^\circ$. The results of the gradient orientation measure on images of single characters show the ability to determine the correct orientation for sparsely inscribed documents.

III. CONCLUSION

This paper evaluates the accuracy of a skew detection based on gradient information and an orientation histogram. The analysis of the gradient orientation measure accuracy depending on the Gaussian smoothing and the type of content as well as the influence of a binarization step is shown.[1]

REFERENCES

[1] M. Diem, F. Kleber, and R. Sablatnig, "Skew Estimation of Sparsely Inscribed Document Fragments," in *Proceedings of the International Workshop on Document Analysis Systems (DAS)*, Goldcoast, Australia, 2012.

[2] F. Kleber, M. Diem, and R. Sablatnig, "Robust skew estimation of handwritten and printed documents based on grayvalue images," in *In Proc. of the 22nd International Conference on Pattern Recognition (ICPR)*, 2014, pp. 1–6.

[1]This project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 674943