G. Nagy, D. Lopresti, Form similarity via Levenshtein distance between ortho-filtered logarithmic ruling-gap ratios, Procs. SPIE/IST Document Recognition and Retrieval, Feb. 201

Form similarity via Levenshtein distance between ortho-filtered logarithmic ruling-gap ratios

George Nagy[a] and Daniel Lopresti[b*]

[a] RPI, Troy, NY USA;
[b] Lehigh University, Bethlehem, PA USA

ABSTRACT

Geometric invariants are combined with edit distance to compare the ruling configuration of noisy filled-out forms. It is shown that gap-ratios used as features capture most of the ruling information of even low-resolution and poorly scanned form images, and that the edit distance is tolerant of missed and spurious rulings. No preprocessing is required and the potentially time-consuming string operations are performed on a sparse representation of the detected rulings. Based on edit distance, 158 Arabic forms are classified into 15 groups with 89% accuracy. Since the method was developed for an application that precludes public dissemination of the data, it is illustrated on public-domain death certificates.

Keywords: Office forms, form identification, rulings, edit distance

1. INTRODUCTION

We propose a translation-, scale- and rotation-invariant metric for classifying forms. The method is independent of language and script. It is remarkably immune to common types of document degradation, especially those incurred in copying and scanning. Even the Matlab implementation is fast enough for production runs. Our program accepts a set of scanned form images and produces a matrix of pairwise similarity measures based on the configuration of detected rulings. Geometric invariance is achieved by using ratios of the horizontal and vertical distances between rulings. Insensitivity to false negatives and false positives in ruling detection is effected by the application of a string comparison algorithm rather than heuristics to match rulings in different forms.

Layout analysis and interpretation of tabular structures attracted many researchers as soon as technology became available to scan and process page-sized images[1,2,3,4,5]. Form reading achieved commercial viability after a decade of experimentation[6,7,8,9]. Specialized algorithms were crafted to detect parallel rulings in large forms or drawings[10,11]. Research on table recognition was aptly surveyed by Zanibbi, Blostein and Cordy[12]. The benefits of exploiting the colossal amount of information contained in web tables were recently proclaimed by prominent researchers[13].

Forms differ from other types of document because the blank forms are produced by one party, and distinguishing information is added by a different party. Amano called them "interactive documents"[14]. Old forms used widely or over a long period often show variability due to multiple reprinting and even photocopying or carbon paper. Blank forms are far from blank: they typically contain rulings that delimit data fields or act as columns separators or provide writing lines. Rulings may be solid or dashed and vary in weight. Forms normally also have preprinted guide labels or row or column headers, instructions, letterhead-type indications of the source organization, check boxes, logos, etc. Essential information must often be entered in duplicate, or there may be natural redundancy among the data items. Commercial form reading operations can benefit from their large databases that already contain part of the form information.

The form-filler may type, handprint, or write the necessary information and may also add stamps, serial numbers, scribbles, gratuitous checkmarks, blots, and punch holes. The final source of image variability is the result of decentralized scanning: samples of the same type of form – especially widely dispersed historical forms - may be scanned with different auto-feed scanners at various sampling rate, contrast and color settings.

The current project is an offshoot of a US Government endeavor to read forms that were found among the 5.5 million digitized Anfal[†] documents[15,16] deposited for safe-keeping in the Archives of the University of Colorado in 1998 after

[*] Further author information: (Send correspondence to G.N.)
G.N., E-mail: nagy@ecse.rpi.edu, Telephone: 1 518 271 6885
D. L., E-mail: lopresti@cse.lehigh.edu, Telephone: : 1 610 758-5782

[†] *Anfal* means "spoils of war" (from the Quran)

the Kurdish uprising. It has often been noted that old ("historical") forms are more difficult to interpret than contemporary forms[17,18.]. That is also the case for some forms recently generated in less-developed countries. They usually lack a form identification number and the labels are often too poorly printed or scanned to OCR. Storefront letterpress printers with ancient equipment may print rulings as horizontal or vertical lines of tiny type (Figure 1).

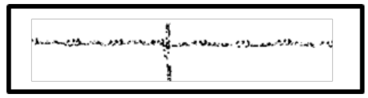

Fig. 1 An intersection of rulings from an Anfal form.

We report here the results of a small pilot experiment formulated with the help of the National Institute of Standards and Technology. The form images cannot be published because many of the Anfal forms contain confidential information. We therefore reluctantly illustrate our algorithms only with images of death certificates harvested from the web. Although death certificates may be of interest *per se*, particularly for genealogical research, we selected them because their general configuration, condition, and degradation exhibit some similarity to the Anfal forms. We have not conducted any large scale experiments and present our ideas here in case some members of the DRR community find the combination of geometric invariance with a metric for variable-sized feature vectors of intrinsic interest. We have not found any experiments or proposals directly comparable to ours.

In the next section we describe line detection, where the only interesting wrinkle is an ortho-normal ruling filter. In Section 3, we justify the use of "minimal" gap-ratio features and logarithmic gap ratios. We present the string-matching aspects of our approach in Section 4. In Section 5 we describe briefly our experiment on the Anfal data. Section 6 is a short summary.

2. LINE DETECTION

The ruling lines are detected with standard image processing algorithms in the following sequence:

1. Extract edgelets
2. Perform Hough Transform that yields the radius-angle (*rho-theta*) coordinates of ruling candidates
3. Extract dominant near-horizontal and near-vertical lines
4. Apply ortho-filter to retain only orthogonal rulings
5. Sort the horizontal and vertical rulings separately according to their rho (radial) coordinate
6. Obtain the ruling intervals (*gaps*) from the difference of the rho coordinates of adjacent rulings

The Prewitt filter is used for edge extraction. It is of course linear in the number of pixels and has a low constant of proportionality[19,]. Other well-known edge filters that we tried (Sobel, Roberts[20]) yielded comparable results. There is no need for anything more complex like a Canny filter. We retained the Prewitt filter partly because of our admiration for its inventor.

The Hough Transform[21] was applied with two-pixel horizontal and four-pixel vertical rho bins, and one- degree theta bins. The angular bins spanned a range within ±30 degrees of the vertical and ±30 degrees of the horizontal under the assumption that the maximum skew introduced by scanning or copying would be less than 30 degrees. Alternative methods of line detection could, of course, have been used instead, but the Hough transform is convenient because it yields the radial and angle coordinates of the lines that we use further on.

The major peaks in transform space (those larger than one tenth of the magnitude of the largest peak) are retained. A suppression region proportional to the image size is used to discard "shadow" line segments caused by degraded bitmaps and imperfect Prewitt edgelets. The usual gap-suppression and gap-filling parameters are also set proportional to the image size. None of these Hough Line parameters are sensitive: they can be readily doubled or halved without perceptible change in the final outcome. Typically ten times more peaks were generated than there were rulings on the form, with each peak representing a potential line in rho-theta (radius, angle) coordinates. We retain only the top 60 horizontal lines and the top 20 vertical lines because a larger number of rulings would hardly leave enough space to enter information.

The rulings were extracted with an ortho-filter, i.e., the dominant set of orthogonal rulings ware found simultaneously as follows. The near-vertical line segments were rotated 90 degrees by adding 90 degrees to their theta coordinates. Then all the angles were histogrammed, and the lines within one degree of the highest peak were retained as rulings. Although extracted at the same time, the horizontal and rotated vertical rulings were kept separate. Each of the two sets had nearly equal near-zero theta coordinates. The information used in subsequent steps consists of their rho coordinates which, after sorting, provide the spacing (gaps) between successive rulings. Examples of the lines extracted from a death certificate and of the rulings retained by ortho-filtering are shown in Figure 2.

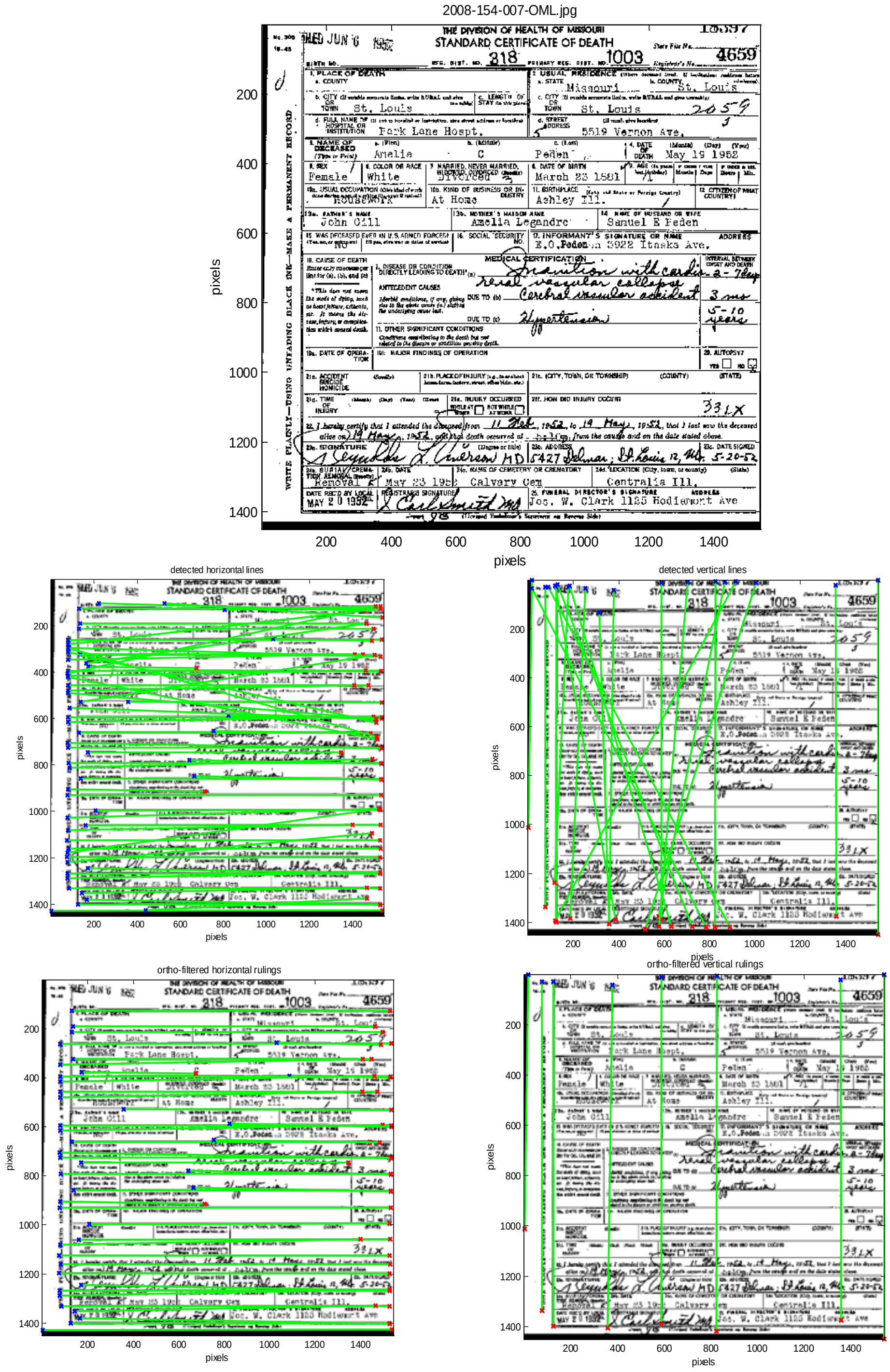

Fig. 2 Horizontal and vertical ruling candidates and detected rulings of the digitized form shown at the top. The rules are considered infinite lines because only the distance between them is used in form classification.

3. GAP REPRESENTATION

Processing is performed separately on horizontal and vertical rulings. Because the *position* of the rulings, i.e., their rho coordinates, can be detected far more reliably than their end points, we consider all rulings as infinite lines. Therefore each ruling is represented by a single scalar value (the rho coordinate of the corresponding Hough peak). In the following discussion, we consider only the scalar position-coordinates (x-values) of either set of rulings. Fig. 3 illustrates a few x-values to help visualize the gap-ratio calculation.

It is well-known that the ratios of intervals $(x_s - x_t)/(x_u - x_v)$ between pairs of points on a line are invariant to linear transformations $x' = ax + b$. This means that the ruling gap ratios are invariant to translation and scale. Because the differences of rho-values represent perpendicular distances between rulings, the ratios are also rotation invariant. But what is a minimal set of gap ratios that contain all of the information that can be derived from the given set of rulings?

We intuitively chose to use only the ratios between successive ruling lines. *N* ruling lines have *N-1* gaps, and *N-2* ratios $R_i = (x_{i+2} - x_{i+1})/(x_{i+1} - x_i)$ of pairs of consecutive gaps. We now show that any other gap ratio $(x_s - x_t)/(x_u - x_v)$ can be obtained from these, therefore the N-2 R_i ratios are sufficient.

The complete derivation is tedious, but the underlying idea is that any desired ratio can be expressed in terms of only adjacent intervals. First note that any ratio between two pairs of adjacent rulings can be computed from the quotient of the products of two sets of consecutive intervals follows:

$$\frac{x_s - x_{s-1}}{x_u - x_{u-1}} = \frac{x_s - x_{s-1}}{x_2 - x_1} \times \frac{1}{\frac{x_u - x_{u-1}}{x_2 - x_1}} = \frac{\prod_{i=1}^{s-2} R_i}{\prod_{i=1}^{u-2} R_i}$$

Now we express the ratio of two pairs of arbitrary ruling gaps as the quotient of the sum of adjacent gaps;

$$\frac{x_s - x_t}{x_u - x_v} = \frac{(x_s - x_{s-1}) + (x_{s-1} - x_{s-2}) + \cdots + (x_{t+2} - x_{t+1}) + (x_{t+1} - x_t)}{(x_u - x_{u-1}) + (x_{u-1} - x_{u-2}) + \cdots + (x_{v+2} - x_{v+1}) + (x_{v+1} - x_v)}, \quad \text{where } s > t \text{ and } u > v$$

Here all the parenthesized terms represent gaps between adjacent rulings. Further manipulation of this expression yields:

$$\frac{x_s - x_t}{x_u - x_v} = \frac{(R_1 \times R_2 \times \ldots \times R_{t-1})(1 + R_t[1 + R_{t+1}[\ldots.[1 + R_{s-1}]\ldots.])}{(R_1 \times R_2 \times \ldots \times R_{v-1})(1 + R_v[1 + R_{v+1}[\ldots.[1 + R_{u-1}]\ldots.])}$$

This equation proves that any ratio of the distance between two arbitrary pairs of rulings can be expressed in terms of only pairwise gap ratios R_i of adjacent intervals; therefore no other information is needed to characterize the relative positions of a set of parallel rulings. (Several terms in the formula can be cancelled when the actual values of the gap parameters are known.) Consider, for example, $(x_7 - x_4)/(x_6 - x_3)$ with the values given in Figure 3, where $R_1 = 2/1$, $R_2 = 3/2$, $R_3 = 3/3$, $R_4 = 4/3$, $R_5 = 5/4$. The formula, with $s = 7$, $t = 4$, $u = 6$ and $v = 3$, yields the expected value of *6/5*, which can be verified by inspection. We emphasize that we never use this formula: it is presented only by way of proof that the pairwise gap ratios R_i contain all the information needed to characterize form configuration (except for the start and end points of the rulings).

Any ratio of ruling gaps can be larger or smaller than unity. We want to give equal weight to a ratio regardless of the sense in which it is measured (changing sense yields the reciprocal of the original ratio). This is accomplished by taking the logarithm of the R_i. The base 10 logarithmic ratios are then uniformly quantized into 20 bins ranging from -1 to +1. Any ratio larger than 10 or smaller than 0.1 is assigned to the border bins. Depending on its bin, each ratio is now assigned to a symbol of some fixed alphabet in preparation for the string matching phase described in the next section.

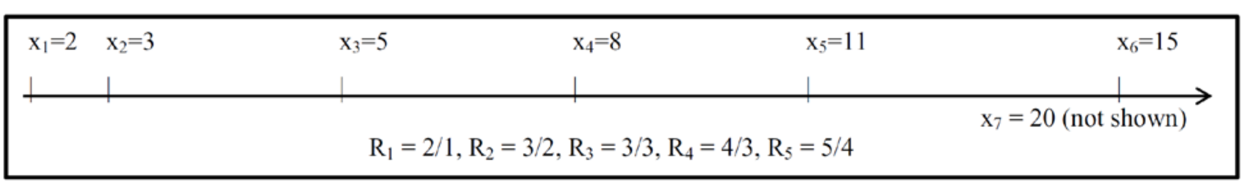

Figure 3. Pairwise gap ratios. The x_i's are the rho coordinates of the parallel rulings culled from the Hough Transform.

4. STRING METRIC

Wagner and Fischer published an optimal algorithm for computing a class of edit distances in 1974[22]. It was used widely in signal coding, genome sequencing, and text comparisons. String matching is now a mature field with many alternative algorithms and programs available[23,24,25,26], but we believe that this is its first application to form rulings.

Given two symbols strings, for example A = (b,d,e,e,e,b,g,k) and B = (b,f,e,e,b,g,c,k), the Levenshtein metric (a specific edit distance) is a count of the minimum number of insertions, deletions and substitutions necessary to transform the first string into the second or vice-versa[27]. Here one deletion (‘e’), one substitution (‘d’ → ’f’) and one insertion (‘c’) are needed to transform A into B, for a total edit distance of 3.

To determine the similarity of the ruling configuration of each pair of forms, we computed the Levenshtein metric independently for horizontal and vertical rulings. We used the public-domain EDIT DISTANCE WEIGHTED program posted in 2010 by B. Schauerte[28]. We kept the default parameters for equal weighting. An example of the distance matrix for 16 death certificates is shown in Figure 4 (the matrix for the 158 Anfal forms is much too large for page display). Because the matrix is symmetric, only (16× 15)/2= 120 edit distances were actually calculated. Fig. 5, shows examples of the pairs of death certificates (from a different set) with the smallest and largest edit distance between them.

It may be advantageous to weight substitutions according to how much the ratios involved differ. In given applications more weight could be given to deletions than to insertions. If, for example, a clean master form is available for comparison, then its scanned image is more likely to miss rulings than to contain spurious rulings. On the other hand, densely typed lines are likely to give rise to spurious rulings. It is also possible to take into account the constraints on the changes of gap ratios caused by the appearance and disappearance of rulings.

distforms_H_2 =

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		16	12	31	27	22	25	29	13	17	16	15	15	10	13	19
2			15	32	27	25	27	32	12	17	14	15	13	15	13	16
3				29	27	27	26	32	15	16	14	11	12	12	14	17
4					32	29	32	28	33	30	35	29	30	32	31	40
5						27	29	33	28	31	29	27	27	27	28	37
6							26	27	28	23	33	24	26	25	25	35
7								27	27	27	31	25	22	27	25	37
8									32	31	35	29	29	31	29	41
9										17	14	15	14	14	15	16
10											17	14	16	13	15	19
11												14	17	16	13	16
12													15	15	10	18
13														14	14	18
14															11	17
15																18

distforms_V_2=

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		4	3	2	4	3	3	3	4	4	4	4	4	4	3	4
2			2	4	6	3	3	3	1	1	3	1	1	1	1	1
3				3	5	3	4	1	2	1	2	2	2	2	2	2
4					4	3	4	3	4	4	3	4	4	4	4	4
5						5	5	5	6	5	4	6	6	6	6	6
6							2	3	3	3	3	3	3	3	2	3
7								4	4	4	4	4	4	4	3	4
8									3	2	3	3	3	3	3	3
9										1	3	1	0	0	1	0
10											2	1	1	1	1	1
11												3	3	3	2	3
12													1	1	1	1
13														0	1	0
14															1	0
15																1

Fig. 4.. Distance matrix of gap ratios for 16 forms. Left: between horizontal rulings; Right: between vertical rulings

5. “EXPERIMENT”

The procedures described in the last three sections were applied in the MADCAT project to assign one of 15 model types to 158 ground-truthed forms provided by a U.S. government agency. The ground truth consisted of the model type for each form. These forms had many more horizontal rulings than vertical, so we used only the distances between horizontal rulings. Including the vertical distances made little difference, but using the vertical distance alone nearly doubled the error rate. We classified the forms with a leave-one-out nearest-neighbor classifier using the Levenshtein metric as distance measure. The resulting error rate was 11% (18 errors).

Six of the errors were confusions between two configurations of essentially the same form that differed only slightly in the width of two columns. Only three labeled samples were available from one of these subclasses. The other misclassifications were fairly widely distributed, with several due to faint or cut-off rulings (near an edge of the form). The average between-groups distance was about twice the average within-group distance. In hindsight, it may have been better to normalize the raw distance by the number of detected lines.

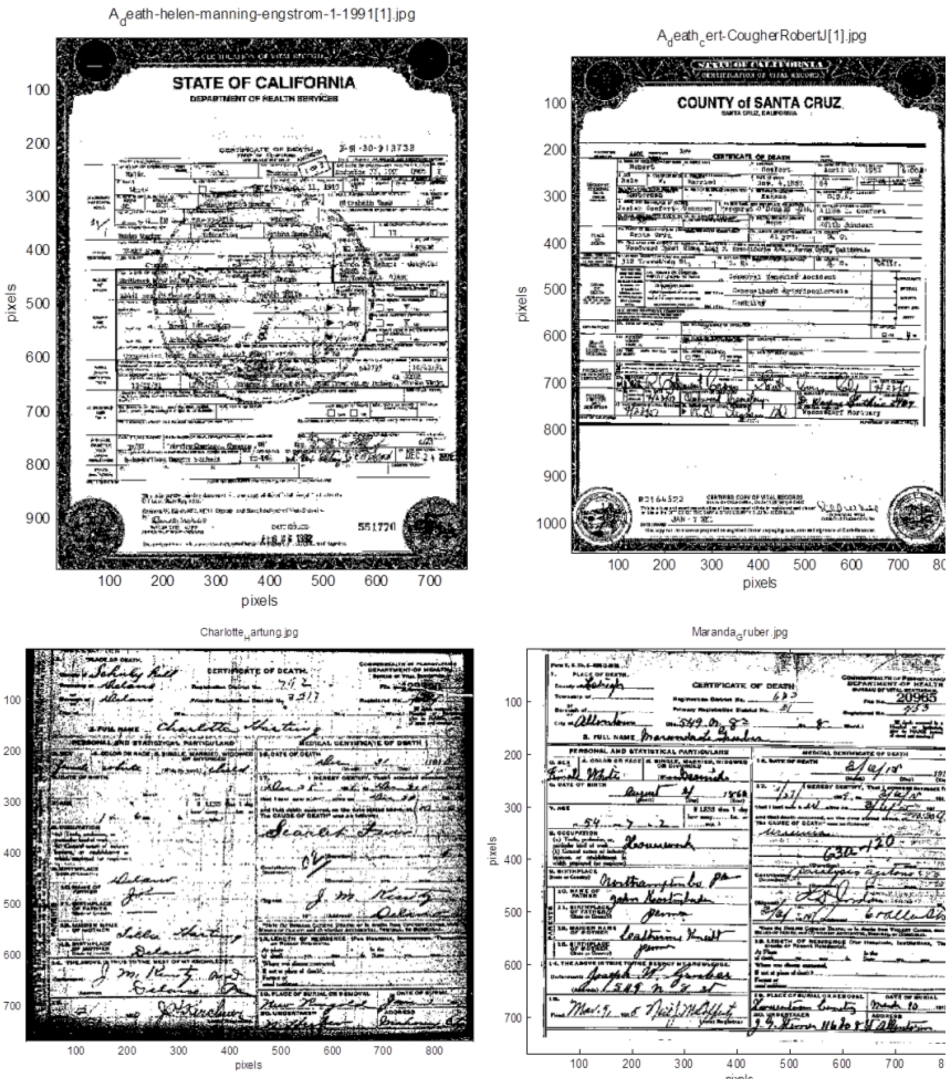

Fig. 5. According to the distance matrix, the ruling structure of the top two forms are more similar to each other, and the bottom forms are more similar to each other, than the top forms are to the bottom forms. The gap ratios (13 and 15), between the horizontal rulings of similar forms are smaller than the gap ratios (18, 19, 19, 21) between dissimilar forms. However, the two death certificates from California (top) are far from identical, and the two certificates below from Michigan are from different counties.

As mentioned, the Anfal data is not currently available for public distribution, but scanned death certificates from several states are posted on the web. No attempt was made to tune the program for the transition from the Anfal forms to the death certificates although some improvement could be expected from changing the line detection thresholds to fit the gross statistics of the line distributions and of the observable scanning parameters. We do not, however, consider these low-resolution death certificates suitable for meaningful systematic experimentation.

Run time of the 2012a Matlab m-code under Windows 7 on a 2.5 GHz laptop computer is less than one second per form, including reading the sizeable Anfal form-images from disk and generating the figures and printouts. As expected from the successful use of dynamic programming algorithms on long strings of natural language text and bio-informatics, the run time of the edit distance routine on a few dozen ruling gap ratios is insignificant. For the 158 Anfal forms, the roughly 25,000 horizontal and vertical string distance calculations take 0.06 s. Although computing the distance matrix is necessarily quadratic in the number of forms, in most applications there would be no reason to compute the Levenshtein metric for every pair of inputs.

6. CONCLUSION

The modest contributions of this communication are twofold: (1) efficient use of an angle-filter to find the largest orthogonal set of detected line segments, and (2) introduction of the edit distance between logarithmic gap ratios as a measure of translation- and scale-invariant similarity between parallel sets of rulings.

The Hough Transform on edgelets, combined with ortho-normal filtering, appears to be a viable candidate for detecting horizontal and vertical rulings in tables or forms. The radial components of the resulting line coordinates are invariant to rotation. The ratios of distances between parallel rulings are invariant to translation and scaling, and can be calculated efficiently from the radius-angle representation provided by the Hough Transform. It was shown that the N-2 ratios of N-1 adjacent intervals offer a complete representation of the relative positions of N parallel lines. Logarithmic scaling of the ratios allows uniform quantization for the conversion of the numeric values of the ratios to strings of symbols. String comparison, one of the few effective methods for classifying objects with a variable number of features, was chosen to provide a similarity measure for pairs of forms because standard statistical classifiers require fixed-length feature vectors.

Tuning the weights would certainly improve classification. We did not, however, have enough samples to weight each type of insertion, deletion and substitution according to discriminative statistics learned from a training set. Another possible improvement, suggested by one of the referees, is to apply this method recursively to detect local similarities and differences. This would seem particularly appropriate for boxy forms with many short rulings.

In many current applications (medical claims forms, insurance applications, election ballots, etc.) forms can be readily identified by a form identification number prominently printed near a corner or an edge. In the absence of form numbers, the preprinted labels that specify what information should be entered are usually sufficient to identify the form. Relying on the configuration of rulings would be a poor third choice. In any case processing *scanned* forms is becoming obsolete (except for archival applications) because most organizations find it much cheaper and faster to collect information through web forms. US income tax returns, for instance, are filed on paper by less than 20% of the population. Nevertheless, in many endeavors to conserve noisy historical documents, rulings provide the most dependable means of form identification. Pixel-based ruling detection may also occasionally be appropriate for PDF or other representations of computer generated forms.

Like many other ideas in Document Image Analysis, this method is currently seeking a home in a suitable application.

ACKNOWLEDGMENTS

George Nagy is grateful for support from the Computer Science Department of Lehigh University during the MADCAT phase of this work. The work of Daniel Lopresti is supported by a DARPA IPTO grant administered by Raytheon BBN Technologies.

REFERENCES

1 O. Hori, , D.S. Doermann, "Robust table-form structure analysis based on box-driven reasoning," In: Proceedings of the Third International Conference on Document Analysis and Recognition (ICDAR'95),. Montréal, Canada, 218–221, 1995.

2 K. Itonori, "A table structure recognition based on textblock arrangement and ruled line position," in Proceedings of the Second International Conference on Document Analysis and Recognition (ICDAR'93), 765–768. Tsukuba Science City, Japan (1993)

3 T.A. Bayer, "Understanding structured text documents by a model based document analysis system," in Proceedings of the Second International Conference on Document Analysis and Recognition (ICDAR'93), pp. 765–768. Tsukuba Science City, Japan 448-453 (1993).

4 K-C Fan, Y-K Wang, and M-L Chang, "Form document identification using line structure based features," in 6th International Conference on Document Analysis and Recognition, ICDAR'2001, 704-708, 2001.

5 B. Krüpl, M. Herzog, and W. Gatterbauer. "Using Visual Cues for Extraction of Tabular Data from Arbitrary HTML Documents," WWW, ACM, 1000-1001, 2005.

6 A. Denel and G. Barth, “ANASTASIL: hybrid knowledge-based system for document layout analysis,” in Proceedings of the 11th International Joint Conference on Artificial intelligence - Volume 2, 1249-1254, 1989.

7 T. Kieninger and A, Dengel, “Applying the T-Recs table recognition system to the business letter domain,” in Proceedings of the Sixth International Conference on Document Analysis and Recognition. Seattle, WA (2001)

8 B. Klein, S. Agne, A. Dengel, “Results of a study on invoice reading systems in Germany,” in S.Marinai and A. Dengel (eds.) Document Analysis Systems VI. Lecture Notes in Computer Science, vol. 3163, 451–462, Springer-Verlag, Berlin, Germany (2004)

9 A. Dengel, “Towards Understandable Explanations for Document Analysis Systems,” 10th IAPR International Workshop on Document Analysis Systems (DAS), Gold Coast, Queensland, Australia, (2012)

10 J.G. Arias, S. Balasubramanian, A. Prasa, R. Kasturi, A. Chhabra, “Information extraction from telephone company drawings,” in Proceedings of the Conference on Computer Vision and Pattern Recognition, Seattle, Washington, 729–732, (1994)

11 J.F. Arias, A. Chhabra,V. “Efficient interpretation of tabular documents,” In: Proceedings of the International Conference on Pattern Recognition (ICPR’96), vol. III, pp. 681–685. Vienna, Austria (1996)

12 R. Zanibbi, D. Blostein, and R Cordy, “A survey of table recognition: Models, observations, transformations, and inferences,” International Journal on Document Analysis and Recognition, vol. 7, (2004)

13 A.Halevy, P. Norvig, and F. Pereira. "The Unreasonable Effectiveness of Data," IEEE Intelligent Systems 24, 2, 8-12, (2009).

14 A. Amano and N. Asada,”Graph grammar based analysis system of complex table form document,” in Proceedings of the Seventh International Conference on Document Analysis and Recognition, (2003)

15 B.P. Montgomery, “The Iraqi Secret Police Files: A Documentary Record of the Anfal Genocide,” Archivaria 52, 81–82, (2001)

16 B.P. Montgomery, “Returning Evidence to the Scene of the Crime: Why the Anfal Files Should be Repatriated to Iraqi Kurdistan,” Archivaria 69, 143–171, (2010)

17 B. Coüasnon and L. Pasquer, “A real-world evaluation of a generic document recognition method applied to a military form of the 19th century,” Proceedings of the Sixth International Conference on Document Analysis and Recognition, Seattle, WA, 779-783, (2006)

18 B. Coüasnon, “Recognition of Tables and Forms,” in Handbook for Document Image Analysis, D. Doermann and K. Tombre, editors, Springer, (2013)

19 J.M.S. Prewitt, “Object Enhancement and Extraction,” in Picture Processing and Psychopictorics, edited by B.S. Lipkin and A. Rosenfeld, Academic Press, (1970)

20 K. R. Castleman, Digital Image Processing, Prentice Hall, (1996)

21 R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, Wiley, (1973)

22 R.A. Wagner, M.J. Fischer, “The String-to-String Correction Problem”, Journal of the ACM, 21(1), 168–173, (1974)

23 V.I. Levenshtein, “Universal bounds for codes and designs”, in Handbook of Coding Theory, V.S.Pless and W.C.Huffman, Eds., Elsevier, Amsterdam, vol. 1, 499-648, (1978)

24 P.A.V. Hall and G. R. Dowling, “Approximate String Matching,” ACM Computing Surveys, Volume 12, Issue 4, December, 381-402 (1980)

25 D. Sankoff,, J.B.; Kruskal, Time warps, string edits, and macromolecules: The theory and practice of sequence comparison, Addison Wesley (1983)

26 G. Navarro. "A guided tour to approximate string matching", ACM Computing Surveys 33 (1), 31–88, 2001

27 V.I. Levenshtein, “Binary Codes Capable of Correcting Deletions, Insertions and Reversals,” Soviet Physics Doklady, Vol. 10, p.707, (1966)

28 B. Schauerte and G. A. Fink, "Focusing Computational Visual Attention in Multi-Modal Human-Robot Interaction," in Proc. ICMI, (2010)