RAPID ANALYTICAL VERIFICATION OF HANDWRITTEN ALPHANUMERIC ADDRESS FIELDS

C. K. LEE   G. LEEDHAM

School of Computer Engineering, Nanyang Technological University, Singapore 639798
Email: travck@yahoo.com, asgleedham@ntu.edu.sg

This paper presents a combination of fuzzy system and dynamic analytical model to deal with imprecise data derived from feature extraction in handwritten address images which are compared against postulated addresses for address verification. A dynamic building-number locator is able to locate and recognise the building-number, without knowing exactly where the building-number starts in the candidate address line. The overall system achieved a correct sorting rate of 72.9%, 27.1% rejection rate and 0.0% error rate on a blind test set of 450 cursive handwritten addresses.

1 Introduction

Automatic handwritten address reading has a number of commercial applications, including automation of the handwritten letter-mail stream. To keep the error rate low while retaining a high read rate is a critical aspect for these applications. The address reading task is made particularly difficult because of the large variety of writing styles and abbreviations, and the presence of uncertainty in handwritten addresses. The alphanumeric fields found in the address, such as building-number, can play a crucial role in compensating for the estimation errors in holistic features.

Fuzzy set theory is a generalisation of abstract set theory. Hence, fuzzy set theory has a wider scope of applicability than abstract set theory in solving complex or ill-defined problems that involve, to some degree, subjective evaluation. A fuzzy logic system was described by Gader et al. (1995) for handwritten street numbers location. Madhvanath et al. (1997) used a coarse-grained holistic model to improve the error rates of a lexicon-driven analytical classifier for rapid verification of handwritten phrases.

In this paper, we have applied our ideas and strategies to the specific task of locating and recognising the alphanumeric address fields in Singapore handwritten addresses to reduce mail being wrongly recognised by the OCR of the 6-digit postcode. This paper presents a system which employs a method which dynamically locates and recognises the building-number in Singapore handwritten addresses. The proposed method uses a fine recognition model to maximise correct sorting and minimise errors of a fuzzy-based coarse holistic model (Lee and Leedham 1999).

2 Script Address Pre-Processing and Interpretation

The automated reading of an address can be subdivided into the following principal tasks: preprocessing, recognition and interpretation. The preprocessing step starts from the scanned letter mail to address block location, to line location and finally to segmentation of lines into words and if possible into characters.

Handwritten addresses typically contain a mixture of isolated hand-printed characters, and connected cursively written or printed words. Figure 1 shows two typical address samples of the Singapore addresses we are attempting to recognise. A

In: L.R.B. Schomaker and L.G. Vuurpijl (Eds.), Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition, September 11-13 2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 571-576

priori knowledge of possible address layout is used to locate a postcode candidate and segment it into discrete numerals. These numerals are passed to a back-propagation neural network numeral recogniser. The results consist of numeric alternatives ranked by a confidence measure. Postcode candidates are reconstructed from the individual character recognition results and multiple postcode hypotheses are generated. The confidence value of the reconstructed postcode is derived from the confidence of the individual recognised numerals. By use of a complete country-specific directory for postcode-related addresses, a lexicon of address words is generated. The task of the holistic verifier is to rank those postulated addresses by matching it with a set of features extracted from the input image. Once the postcode is verified a similar strategy can be applied to building numbers. To ensure the error rate is low, this implies the capability to reject a wrong address. This may occur either if the image is not what it was supposed to be, caused by incorrect word segmentation, or incorrect estimation of word features particularly in the number of characters and closed/near loops.

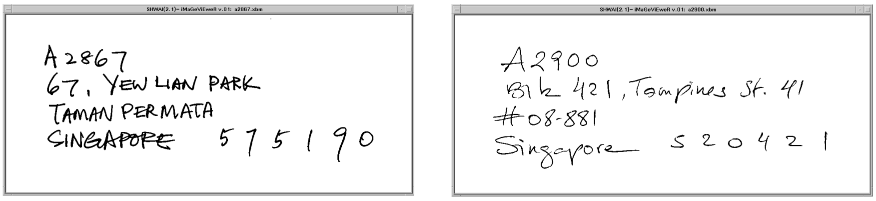

Figure 1: Typical images of handwritten addresses

3 Estimation of Grades of Membership for Handwritten Address Words

Let $X = \{x\}$ denote a space of objects. Then a fuzzy set of type 2, A in a universe of discourse X is characterised by a *fuzzy membership function* μ_A, as a set of ordered pairs

$$A = \{ x,\ \mu_A(x) \}, \quad x \in X, \tag{1}$$

where $\mu_A(x)$ is the grade of membership of x in A on the unit interval [0,1]. The grades 1 and 0 representing, respectively, *full membership* and *non-membership* in a fuzzy set reflect an "ordering" of the objects in the universe where the grade-of-membership value $\mu_A(x)$ can also be interpreted as the degree of compatibility of the predicate associated with A and the object x.

In general, we distinguish three kinds of inexactness found in handwritten addresses, namely, *generality*, that a concept applies to a variety of consequences; *ambiguity*, that it describes more than one distinguishable sub-concept; and *vagueness*, that precise boundaries are not defined. All three types of inexactness are represented by a fuzzy set:

- Generality occurs when the universe is not just one point,
- Ambiguity occurs when there is more than one local maximum of a membership function, and
- Vagueness occurs when the function takes values other than just 0 and 1.

A *Euclidean space* is a vector space over $\mathbf{R}$ together with a specified inner product in the space. Consider a word from the address image as a vector, $\alpha = (x_1, x_2, \dots x_n)$ and a word from the postulated address, $\beta = (y_1, y_2, \dots y_n)$, where x_i measures the i^{th} feature in the word α and y_i measures the i^{th} feature in the word β, are elements of $\mathbf{R}^n$. Let V be a finite-dimensional Euclidean space over $\mathbf{R}$ and let B = $(\alpha_1, \ldots, \alpha_n)$ be an ordered basis of V. The numerical-valued function $(\ |\ )$ on $V \times V$ into $\mathbf{R}$ defined by

$$(\alpha \mid \beta) = \sum_{i=1}^{n} x_i y_i , \tag{2}$$

where $\alpha = \sum_1^n x_i \alpha_i$ and $\beta = \sum_1^n y_i \alpha_i$ is an *inner product* in V. Notice that in this Euclidean space

$$(\alpha_i \mid \alpha_j) = \begin{cases} 1 & if\ i = j, \\ 0 & if\ i \neq j, \end{cases} \quad 1 \leq i, j \leq n.$$

From the symmetry, $(\alpha \mid \beta) = (\beta \mid \alpha)$ and the linearity, $(a\alpha + b\beta \mid \gamma) = a(\alpha \mid \gamma) + b(\beta \mid \gamma)$, for all $\alpha, \beta, \gamma \in V$ and all $a, b \in \mathbf{R}$, of an inner product, the bilinearity of an inner product is the inferred linearity in the second argument. Hence, Eq. (2) becomes

$$(\alpha \mid \beta) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i (\alpha_i \mid \alpha_j) y_j \tag{3}$$

where $\alpha = \sum_1^n x_i \alpha_i$ and $\beta = \sum_1^n y_j \alpha_j$ are elements of V. It follows that the inner product is determined by the n^2 scalars

$$c_{ij} = (\alpha_i \mid \alpha_j), \qquad 1 \leq i, j \leq n. \tag{4}$$

These scalars define the matrix $C = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}$,

which is called the *matrix of the inner product* $(\ |\ )$ relative to the (ordered) B-basis. By virtue of the properties symmetry and $(\alpha \mid \alpha) > 0$, if $\alpha \neq 0$ of an inner product $c_{ij} = c_{ji}$ and $c_{ii} > 0$ for $1 \leq i, j \leq n$. On writing the vector elements of α relative to the B-basis as the $1 \times n$ matrix and the vector elements of β as the $n \times 1$ matrix, Eq. (3) can be expressed as the matrix product

$$(\alpha \mid \beta) = (x_1, \ldots, x_n)(c_{ij})\,(y_1, \ldots, y_n)^t. \tag{5}$$

If $(\alpha_1, \ldots, \alpha_n)$ is an orthonormal basis of a Euclidean space V, then for all α, β in V

$$\alpha = \sum_{i=1}^{n} (\alpha \mid \alpha_i)\alpha_i \quad \text{and} \quad (\alpha \mid \beta) = \sum_{i=1}^{n} (\alpha \mid \alpha_i)(\beta \mid \alpha_i). \tag{6}$$

If $\alpha, \beta \in V$, then from Eq. (2), $\alpha = \sum_1^n x_i \alpha_i$ and $\beta = \sum_1^n y_i \alpha_i$, where $x_i = (\alpha \mid \alpha_i)$ and $y_i = (\beta \mid \alpha_i)$, $1 \leq i \leq n$ (see Eq. (6)). It follows that $\alpha - \beta = \sum_1^n (x_i - y_i)\alpha_i$ and hence

$$(\alpha - \beta \mid \alpha - \beta) = \sum_{i=1}^{n} (x_i - y_i)^2 \tag{7}$$

Following through the steps from Eq. (2) to Eq. (5), and from Eq. (7) we get

$$(\alpha - \beta \mid \alpha - \beta) = \sum_{i=1}^{n} (x_i - y_i)^2\, c_i \tag{8}$$

The *distance* $d(\alpha, \beta)$ between vectors α and β with inner product $(\ |\ )$ is

$$d(\alpha, \beta) = \sqrt{(\alpha - \beta \mid \alpha - \beta)} \tag{9}$$

Finally, the *fuzzy membership function*, μ_A is defined as

$$\mu_A(x) = [1 + \{ d(\alpha, \beta) / F_1 \}^{F_2}]^{-1} \tag{10}$$
$$= \left[1 + \left\{ \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2 c_i} \; / F_1 \right\}^{F_2} \right]^{-1}$$

where the fuzzifiers F_1 and F_2 have their significance for creating fuzziness of a set (Pal and Majumder 1977). The optimal matches between words in the input address and the postulated address is achieved by solving the assignment problem in linear programming. The overall score for each address is computed as the normalised sum of the optimal word scores.

4 Dynamic Classifier based on Semantic Information

The building-number found in Singapore postal addresses usually appears in the second line of the handwritten address. The semantics of the alphanumeric building-number can be summarised as follows:-

- It can be a suffix of a fixed pattern, namely the word "Block" or "Blk".
- It takes the pattern of *n*+*a** where *n* represents a numeric character and *a* represents an alphabetic character. It must consists of at least one or more numeric character (*n*+) followed by an optional alphabet character (*a**). In certain case, the pattern *a** is also possible. That is, the building-number consists of only one alphabet character.

The dynamic classifier approach is that the first and last letters of the image were separately recognised by two classifiers, one numeric and the other an alphanumeric classifier. The rest of the letters were recognised by a numeric classifier. The recognition result is the *best* path (maximising confidence) through the first letter, middle letter(s), and last letter, in which the analytical scores represent the local constraints at the first and last stages. The confidence value of the word is the product of the top confidence values of each letter recognised. It is dynamic because for each building-number image the analytical classifier generates a completely new content. Thus, the location of the building-number can be determined easily no matter where it is along the observed line.

Segmentation of touching, overlapping and connected characters is also a problem as shown in Figure 2 which requires a numeral string segmentation algorithm (Shi et al. 1997).

(a) (b) (c)

Figure 2: Examples of (a) touching, (b) overlapping, and (c) connected characters.

5 Analytical Rapid Verification of Alphanumeric Address Fields

The confident analytical score of the building-number further verifies the holistic approach finalised postcode at the final stage of the address verification (Lee and Leedham, 1999). This is implemented through a set of heuristic rules derived from the analyses of the error images in the training set.

The building-number is normally written on the line directly below the addressee's line. So, the detection of this alphanumeric field is concerned with finding the correct position within that line. Dynamic location of the building-number:-

- "BLK" pattern detection. This pattern is detected by recognising the first letter of the first word as a letter "B" or,
- Existence of the hash symbol, "#". Note that in Singapore Housing and Development Block (HDB) addresses, apartment units are written as "#03-22" for a 3rd floor and number 22 unit.
- Merging of "broken" building-number by exploiting the next word to the recognised building number (e.g. "6" and "B" become "6B").

In the presence of the word "Blk" or a hash symbol "#" in the address image, the finding of the building-number starts from the second word. The following word is recognised as well to determine whether there is a "broken" building number. This frequently occurs in cases when an alphabetic character is a suffix to the numerals, e.g. "Blk 232 A". Hence, a merge of the letter "A" to the recognised "232" is essential for the analytical paradigm to produce correct results.

6 Results

The database of 900 handwritten addresses was collected from a wide and random population. Three hundred people were asked to write three addresses on special white printed cards, which had dropout colour boxes and lines. The printed boxes were used to write the postcode field while the rest of the address fields were written in a totally unconstrained manner with horizontal guidelines only. In this way, the numeral string segmentation problem can be handled separately. The envelope-like cards were scanned using Hewlett Packard Scanjet flat bed scanner operated at 200 *dpi* (dots per inch) with a scanning zone of 6×14cm for consistency with international postal sorting systems. The digital images were randomly divided into a training set of 450 images and a testing set of 450 images. The training set was used to fine-tune the pre-classification modules, and the testing set was used for evaluating the performance of the final system. Table 1 summarises and compares results of the various techniques used at different stages of the handwritten address interpretation system.

Table 1: Performance with and without analytical rapid verification of alphanumeric fields

Techniques	Correctly Sorted %			Rejected %			Wrongly Sorted %		
	TR	**TE**	**Total**	**TR**	**TE**	**Total**	**TR**	**TE**	**Total**
1. Postcode recognition (OCR alone)	82.1	80.5	81.3	4.5	7.9	6.2	13.4	11.6	12.5
2. OCR + postcode dictionary check	89.0	84.9	87.2	2.7	7.2	4.7	8.3	7.9	8.1
3. Holistic Verification	72.9	71.8	72.3	26.9	27.8	27.4	0.2	0.4	0.3
4. Dynamic Analytical Verification	74.0	72.9	73.4	26.0	27.1	26.6	0.0	0.0	0.0

(TR = training set; TE = testing set)

From Table 1, the hard classifier result shows a totally unacceptable sorting performance with a high error rate of 11.6%. When added with semantic information by checking the validity of the reconstructed postcode against a postal directory, although improvement was observed, the error rates remain high at 7.9%. In address verification, the use of holistic features has significantly reduced the wrongly sorted images to a percentage of below one on the TE set at the expense of more items being rejected. An important observation was found in the last tested Technique 4. An increment of 1.1% in the correct sorting rate (72.9%) and a virtually zero error rate were obtained. This is possible because the analytical model rejects the wrongly verified address caused by the holistic verification by detecting a mismatch in the postulated building number against the recognised building number.

7 Conclusions

The analytical dynamic classifier approach described in this paper exploits semantic information or the likely syntax constructions found in the addresses to improve the location of the building-number and the overall analytical scores. In the first approach fuzzy sets serve as a compressed description of imprecise, generally contradictory pieces of information mainly due to inexact holistic features extraction. We employ a refined analytical model to our previous developed holistic model, based on identifiable alphanumeric address fields to achieve a more precise interpretation of the handwritten addresses. The rapid verification of building-number image supplements the coarse word shape features to *correct* some wrongly rejected images and eradicates errors caused by the holistic verification.

References

1. Gader, Paul, Keller, James M. and Cai, Juliet (1995), "A fuzzy logic system for the detection and recognition of handwritten street numbers," *IEEE Trans. on Fuzzy Systems*, **3**(1), 83-95.
2. Lee, C. K. and Leedham, G. (1999), "Empirical design of a holistic verifier for automatic sorting of handwritten Singapore postal addresses," *Proc. of the Fifth Int. Conf. on Document Analysis and Recognition*, Bangalore, India, 733-736.
3. Madhvanath, S., Kleinberg, E., Govindaraju, V. and Srihari, S. N. (1997), "The HOVER systems for rapid holistic verification of off-line handwritten phrases," *Proc. of the Fourth Int. Conf. on Document Analysis and Recognition*, Ulm, Germany, **2**, 855 –859.
4. Pal, S. K. and Majumder, D. D. (1977), "Fuzzy sets and decision-making approaches in vowel and speaker recognition," *IEEE Trans. on Systems, Man and Cybernetics*, **7**(8), 625-629.
5. Shi, Z., Srihari, S. N., Shin, Y-C. and Ramanaprasad, V. (1997), "A system for segmentation and recognition of totally unconstrained numeral strings," *Proc. of the Fourth Int. Conf. on Document Analysis and Recognition*, Ulm, Germany, **2**, 455 – 458.