Correlation Filter: An Accurate Approach to Detect and Locate Low Contrast Character Strings in Complex Table Environment

Yi Li, Zhiyan Wang, *Member*, *IEEE*, and Haizan Zeng

Abstract—Correlation has been used extensively in object detection field. In this paper, two kinds of correlation filters, Minimum Average Correlation Energy (MACE) and Extended Maximum Average Correlation Height (EMACH), are applied as adaptive shift locators to detect and locate smudgy character strings in complex tabular color flight coupon images. These strings in irregular tabular coupon are computer-printed characters but of low contrast and could be shifted out of the table so that we cannot detect and locate them using traditional algorithms. In our experiment, strings are extracted in the preprocessing phase by removing background and then based on geometric information, two correlation filters are applied to locate expected fields. We compare results from two correlation filters and demonstrate that this algorithm is a high accurate approach.

Index Terms—Document analysis, graphics recognition, pattern analysis, correlation theory.

1 Introduction

A major problem in many document processing systems is to locate character string in irregular tabular forms. Even if tabular layout is specified, it could not be solved easily because of complexity of irregular tabular columns and rows. In our case, we must tackle a more complicated situation—characters are pervasively smudgy. Printings in coupon can be classified into two categories: 1) preprinted blank coupon table with background and 2) filled-in computer printed strings (Fig. 1).

Airline coupons have the following features [1]: 1) Coupons are carbon copies, low in contrast, noisy, and smudgy. 2) For computer-printed characters, the entire string could be shifted out of the table and skewed due to improper paper feeding in printing phase. Therefore, filled-in data often appear outside the table fields. We cannot solve the registration problem by just retrieving, locating, and discarding the preprinted table because this problem is not caused by our scanning phase. 3) Pitches between characters vary from one coupon to another, but they are the same in one coupon. Therefore, there are no distinct or apparent pitches to separate adjacent fields so that even if the first category of printings is completely removed, it is still difficult to locate filled-in characters and to group them into fields.

In this paper, we present a high-accurate technique to detect and locate character strings. Without losing generality, we sample 3,740 pieces red-green Billing and Settlement Plan (BSP) coupons of different image qualities in our coupon image database. Our input is color images scanned by high-speed scanner and deskewed, and the output has two fields—Airport Code field with 3-characters and TAX field (Fig. 1) for demonstrate and comparison.

The rest of this paper is organized as follows: In Section 2, some related materials are reviewed, including coupon reading system, table processing classification, and other relevant topics. Section 3 presents a brief review of the correlation filter theory and Section 4 presents our preprocessing in color space. Sections 5 and 6 describe our algorithm, steps, and experimental result. In these two sections, discussion and comparison between two correlation filters are presented. Conclusions are presented in the Section 7.

2 Related Research in Tabular Document Understanding

In the early age of document layout analysis, researchers simply assume that tabular information fields are in a fixed position. However, this assumption does not hold in most cases such as checks, stubs, etc. [5]. Researchers recently began to develop algorithms for form classification and analysis [6], [7], [8]. These works can be divided into model-driven [9] and data-driven approaches [10]. The key to form analysis is to extract and locate expected filled-in data from the known [11] or unknown forms [12].

Without form's layout information, data fields are useless in most cases. Therefore, data extraction and location have become the focus of automatic form analysis. There are two fundamental approaches to retrieve the data: 1) first drop forms then register (map between character strings and their corresponding table elements) the fields [5], [10], [13] or 2) retrieve the fields directly and recover [1]. Because form borders interfere with the analysis of contents, most researchers choose the former approach. In our previous researches [14], [15], we extend traditional gray scale to color space, and achieve form dropout by applying Principal Components Analysis (PCA) [16] and Learning Vector Quantization (LVQ) [17] neural network in complex background. Because it is very difficult to develop a general system for fields' registration, we build a specific purpose approach for tabular images with computer-printed characters instead of a complex one for all sorts of documents.

Most practical form processing systems are for specific purposes and only few are for coupon image processing research. Mao et al.'s work [1] is among the first papers to be published worldwide and focus on coupon image, and we can employ coupon's geometric information in layout analysis thanks to their observation on coupons. However, they only demonstrated their system's performance using about 110 tickets with 80 percent accuracy rate only when context redundancy is high, and did not present further discussion on different image qualities.

Moreover, we address three major defects in their paper: 1) They pass all the neighborhood of every expected string to Optical Character Recognition (OCR) module, and match the expected string by an expert-system-like module after recognizing every character in this neighboring area. We consider that it is time-consuming and prone to error so that we try to detect and locate strings in their neighborhood using automatic target recognition (ATR) approach. 2) They do not drop the form before OCR. Considering it with previous defect, OCR requirement is rather critical and human-computer interaction is inevitable. We try to improve the OCR accuracy rate and process coupons automatically by better location results. 3) They employ redundant information that exists only in coupon images so that it is not easy to generalize their method to other document images with computer-printed characters, i.e., immigration pass.

In addition, we compare backgrounds between our work and references [5], [10], [11], [12], [18] because of their similarities. Line dropout is focused on in [5], [10], [12]. Therefore, although an improved method of [10] further demonstrated its effect on online removal in an airline ticket [18], they do not address the registration step clearly as an important part; we conclude that it may result from 1) gaps between filled-in strings are significant and 2) useful strings are handwritten. Therefore, expected strings may be grouped by connected-component method or other simple approaches. However, there are no distinct or apparent pitches to separate adjacent fields in coupon images so that these researches do not provide applicable solutions for registration. Cesarini et al. [11] create attributed relational graphs by relation between instruction and information fields. However, although there are instruction fields in

- *The authors are with the School of Computer Science and Engineering, South China University of Technology, Wushan, Guangzhou, China, 510640.*
E-mail: liyi@umd.edu, wzhyan@ieee.org, hzzeng@szse.cn.

Manuscript received 20 Apr. 2003; revised 8 Apr. 2004; accepted 14 May 2004.
Recommended for acceptance by J. Weng.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0044-0403.

Fig. 1. A Typical BSP Flight Coupon, its (a) Airport Code and (b) Tex field.

coupon images, they are too small to be recognized easily. Moreover, instruction and information fields in coupon images belong to two categories of printings; therefore, geometric relation of information and instruction fields frequently changes in regard to different coupon images due to the shift of filled-in information strings.

Other related algorithms that are widely used in drawings and maps include graph-text [2] or character-map [3] separation. None of them can be used in our case because they cannot process accurately on strings that are crossed by bold line or low contrast. Most researches on color space and complex background mainly focus on color segmentation [4], which can be used in cases that texts and background have obvious differences so that we cannot use these algorithms directly.

In our case, we can retrieve an imprecise neighboring rectangle, which contains expected and other character strings by geometric information after preprocessing. We then adopt two correlation filters [19], [20], [21], [22], [23], [24] to avoid registration error caused by interference of neighboring fields and noises. We treat different character strings in the same field as different profiles of the same "virtual object," therefore, we convert our problem to multiclass pattern recognition cases [19]. In our research, we extend gray scale to color space in the preprocessing phase to retrieve strings, and we integrate the object recognition approach in frequency domain with geometric information in temporal domain together to detect and locate target object.

3 MACE and EMACH Correlation Filter

Correlation operation serves as an effective match-based segmentation technique for a long time and many works [20], [21] discuss correlation and its application in template-match problems. Kumar et al. [22], [23], [24] explore the properties of correlation by using numerical method theory in frequency domain and deduce many applicable correlation filters. Correlation filters have been used extensively for distortion invariant recognition, ATR and other problems.

The minimum average correlation energy (MACE) filter [22] is developed to minimize the average correlation energy of the correlation outputs by training images while simultaneously satisfying the properties of correlation peak constraints at the origin. The effect of minimizing the average correlation energy is that the resulting correlation planes would yield values close to zero everywhere except at the location of a trained object, where it would produce a strong peak.

When we adopt correlation theory to 2D image processing, we have:

$$H = D^{-1}X(X^*D^{-1}X)^{-1}c. \tag{1}$$

Here, we continue using the symbols in [24] and only present the deduction result of MACE filter. Details of deduction and the meaning of each symbol can be found in [22].

By this definition, we can draw conclusion that MACE satisfies least-mean-square property of training images in frequency domain. However, MACE is later observed that it is over reliance on the mean training image and could cause difficulties with clutter rejection, which finally leads to the introduction of the Extended Maximum Average Correlation Height (EMACH) filter [23]. In order to reduce the overemphasis of the mean training image in the filter, EMACH tries to solve a more complicated equation. Basic concepts of EMACH are shown below.

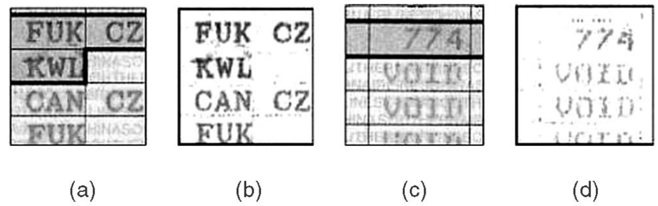

Fig. 2. Fragment result of text extraction. (a) Before preprocess. (b) After preprocess. (c) Before preprocess. (d) After preprocess.

Let H denote an EMACH correlation filter, therefore, we have:

$$J(h) = \frac{H^*C_x^{\beta}H}{H^*(I+S_x^{\beta})H}. \tag{2}$$

H should be the eigenvectors corresponding to the most dominant eigenvalues of matrix $(I+S_x^{\beta})^{-1}C_x^{\beta}$ theoretically. Here, we continue using the symbols in [23] and only present the deduction result of EMACH filter. Details of deduction and the meaning of each symbol can be found in [23].

After the filter is trained, we can compute the correlation of filter and target an image whose size is the same as the filter. We can further determine the input image by its response. To determine it, we first retrieve the peak position and its value $Peak$, locate a 20×20 neighboring field, and then compute the mean value $Mean$ and standard deviation Std of the neighboring field except the center 5×5 area. Therefore, we calculate Peak-to-Sidelobe-Ratio (PSR):

$$PSR = {(Peak - Mean)}\Big/{Std}. \tag{3}$$

If the PSR value is determine, we can detect and classify target images to source training image classes or vice versa. For example, PSR of true class is more than 20 in human face validation, while PSR of false class is lower than 10, when correlation filter is used [24].

Generally, the correlation filter serves as a tool to detect and recognize known signals in target signal. In our case, it serves as an adaptive shift locator to locate expected strings accurately in an imprecise neighboring area retrieved by geometric information.

4 Preprocessing—Text Extraction

Before we adopt correlation filters, we perform preprocessing to extract texts out of background accurately based on the information provided by color space, using the approach presented in [14]. Here, we present two fragments of preprocessing result in Fig. 2.

Figs. 2a and 2b are of high image quality and the other two are of poor [14]. The result shows that most text pixels exist after preprocessing while the background pixels are removed in our case; therefore, this preprocessing algorithm provides an excellent input of key step in location—correlation filter.

5 Detect and Locate Tabular Fields in Coupon Using Correlation Filter and Geometric Information

As stated in Section 1, character strings can be shifted to various positions in a neighboring area. Therefore, we can roughly locate a neighboring area of expected field by priori geometric information and then apply correlation filter to locate strings by determining the value of (3). In our case, correlation filter is applied as an effective adaptive shift locator. All images discussed in this section are processed using preprocessing algorithm in [14] (see Section 4).

Fig. 3. Samples of various qualities and contents for training.

5.1 Fundamental Algorithm to Apply Correlation Filter

In the following description, vector [r, c] denotes a point where its origin is upper left corner of the whole image (r denotes the row number and c denotes the column number, respectively), and vector $[RStart, CStart, REnd, CEnd]$ denotes a rectangle image whose coordinate of upper left corner is $[RStart, CStart]$ and lower right corner is $[REnd, CEnd]$ in the original coupon image, respectively. Without losing generality, we only demonstrate how to detect and retrieve one field in the coupon in this part. Details of experiment steps are presented in Section 5.2.

5.1.1 Train Correlation Filters

By using priori geometric information, we sample some typical images of expected field based on result of preprocessing, as shown in Fig. 3.

These character strings are of different image qualities and contents in Airport Code field. Each sample size is the same as:

$$size = (width, height). \tag{4}$$

For each image in the training set, we adopt Fourier Transformation and scan the result from left to right and from top to bottom and form x_i. Input x_i to MACE and EMACH, solve (1) and (2) to train two filters, h_{MACE} and h_{EMACH}, respectively.

5.1.2 Detect and Locate Expected String Object Using Correlation Filters

After the preprocessing phase, we can retrieve an initial neighboring area of expected field by geometric information (See Section 5.2 for detail):

$$neighborRegion = [RowS, ColS, RowE, ColE]. \tag{5}$$

These numeric values, $RowS$, $RowE$, $ColS$, and $ColE$, are predefined and imprecise, as shown in first column in Fig. 4. This neighboring field imprecisely contains (but not limits to) expected field and may contain noises and part of other fields.

Note that the size of correlation filter is exactly equal to training images, therefore, it is not the same as $neighborRegion$ in our case, and we cannot use the correlation filter directly. To solve this problem, we can apply property of correlation operation and slide window technique. We first apply correlation operation to whole $neighborRegion$ using correlation filter, and then locate some reasonable positions based on correlation result. Considering edge effect, for each possible position we retrieve its neighboring area that center at this possible position and has the same size with the filter, compute PSR value of correlation result and determine similarity between training set and this neighboring area. These steps can be expressed as:

1. Apply correlation operation of h_{MACE} and h_{EMACH} to (5), respectively, that is:

$$result = neighborRegion \circ h_*, \tag{6}$$

   where $\circ$ denote correlation operation, and $* = MACE$ or $* = EMACH$.

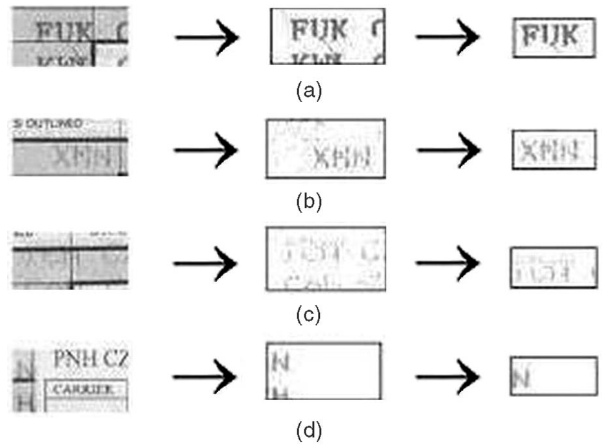

Fig. 4. Example—Process of different kinds of coupons. First Column: original color coupon (only fragments corresponding to the second column are shown in this figure). Second Column: neighboring area of expected field after preprocessing. Third Column: located strings by applying correlation filters. (a), (b), (c), and (d) are examples of high, average, poor quality, and false alarm coupon, respectively.

   Therefore, based on the correlation result, we can locate the point whose value is maximal, and denote it $maxP = [R_{\max}, C_{\max}]$. We may draw a conclusion that $maxP$ is the maximal likelihood point. It is true in some cases in our experiment, but note that 1) the determinant in correlation theory is definition (3) and 2) our proposal is not only to locate target but also to detect its existence. Therefore, it leads to next step.

2. Using slide window technique to determine the maximal-likelihood position.

   We can retrieve a neighboring area around the point of the maximal value:

$$SlideWindow = [R_{\max} - \Delta_1, C_{\max} - \Delta_2, R_{\max} + \Delta_1, C_{\max} + \Delta_2],$$

   where Δ_1, Δ_2, (and Δ_3 in context) are experimental constants.

   For each pixel in the $SlideWindow$, that is, $Pix_i = [r, c]$ $i \in SlideWindow$, we retrieve $[r - height/2, c - width/2, r + height/2, c + width/2]$, where the values of $width$ and $height$ are the same as (4). Calculate its PSR based on the definition of (3). Therefore, finally, we get a $2\Delta_1$ by $2\Delta_2$ real matrix as the result. Select the point of maximal value in this real matrix, and compare it with a prespecific threshold called $preSpecifiedThresh$.

   If the value of maximal likelihood point is smaller than $preSpecifiedThresh$, we conclude that no known string objects in target image, output its position vice versa. By this judgment, we can detect whether a target image contain a known character object, and locate character string if it exists. In our case, we set $preSpecifiedThresh = 12$ based on experience and ROC curves (Fig. 7).

5.2 Experiment Steps

In our experiment, we sample 3,740 coupon images in our database, which can be classified as 17 types by its content of character strings in Airport Code field (Character Content Class). Each type contains 20 coupons whose string in "Airport Code" is missing due to the reason that it is covered by other luggage tickets, edited by airport clerk or else (i.e., Fig. 4d, first column). We also classify these coupon images into another 10 classes based on their images quality (Image Quality Class, class 1 is the highest quality and 10 is the poorest). Statistical data of Image Quality Class are presented in Table 1, column "Coupon Number."

TABLE 1
Detection and Location Result, Image Quality Class
($preSpecifiedThresh = 12$)

Quality	Coupon Number	EMACH			MACE		
		Hit	Miss	False Alarm	Hit	Miss	False Alarm
1	516	501	9	5	498	11	5
2	361	346	10	4	344	12	6
3	312	293	11	6	281	20	8
4	388	361	14	9	362	17	11
5	279	258	13	7	250	19	9
6	204	190	14	6	182	16	8
7	463	420	28	16	411	29	17
8	354	314	25	13	310	27	16
9	278	240	26	15	237	26	18
10	245	206	24	16	201	28	21
Total	3400	3129	174	97	3076	205	119

To train correlation filters, we select 20 images per character content class, whose character string exists in field and image quality varies from poor to high. Another 200 images per character class are used as target images to be tested. In our experiment steps, we process two fields—TAX field and Airport Code field (Figs. 1 and 5).

In a local area of gray-scale coupon image, we first project pixel values to y-axis to locate an imprecise y-axis coordinate. In practical steps, string in the lower right corner of the coupon always exists by observation and its feature is very dominant. We can observe that its relative distance to other character strings in y-axis is almost constant although its characters in x-axis vary from one airline to another. Therefore, we can project it to y-axis and locate the y coordinate $Y_{\max}$ (Fig. 5b).

By priori geometric information, we have the value of relative distance, d, between expected field and $Y_{\max}$ in row, and set a pair of initial values, Col_1 and Col_2, in column (Figs. 5b and 5c). Therefore, we can retrieve the neighboring area of expected field and (5) becomes:

$$neighborRegion = [Y_{\max} + d - \Delta_3, Col_1, Y_{\max} + d + \Delta_3, Col_2].$$

We can use algorithm in previous part to detect and locate character string in $neighborRegion$ image and result images and data are presented in next section. These steps show that we only need to specify table by imprecise geometric information of layout, and we can apply correlation filter to locate accurate position.

To simplify our discussion on an experimental result without losing generality, we just discuss the result of one of these two fields, "Airport Code" field, in next section. In fact, by registering one field precisely, we can locate other fields solely by priori layout information.

Fig. 6. Examples of successfully located strings.

6 Experimental Result and Discussion

In this section, we present a table and some diagrams of result data to illustrate our experiment and discuss correlation filters based on our observation. In our experiment, two categories of report are output: String Detected (SD) with corresponding string coordinate and String Missing (SM). Fig. 6 presents some examples of "Airport Code" field that are successfully located.

Fig. 7 is ROC curves of two correlation filters, "Hit Rate" denotes probability of SD report when the field really contains

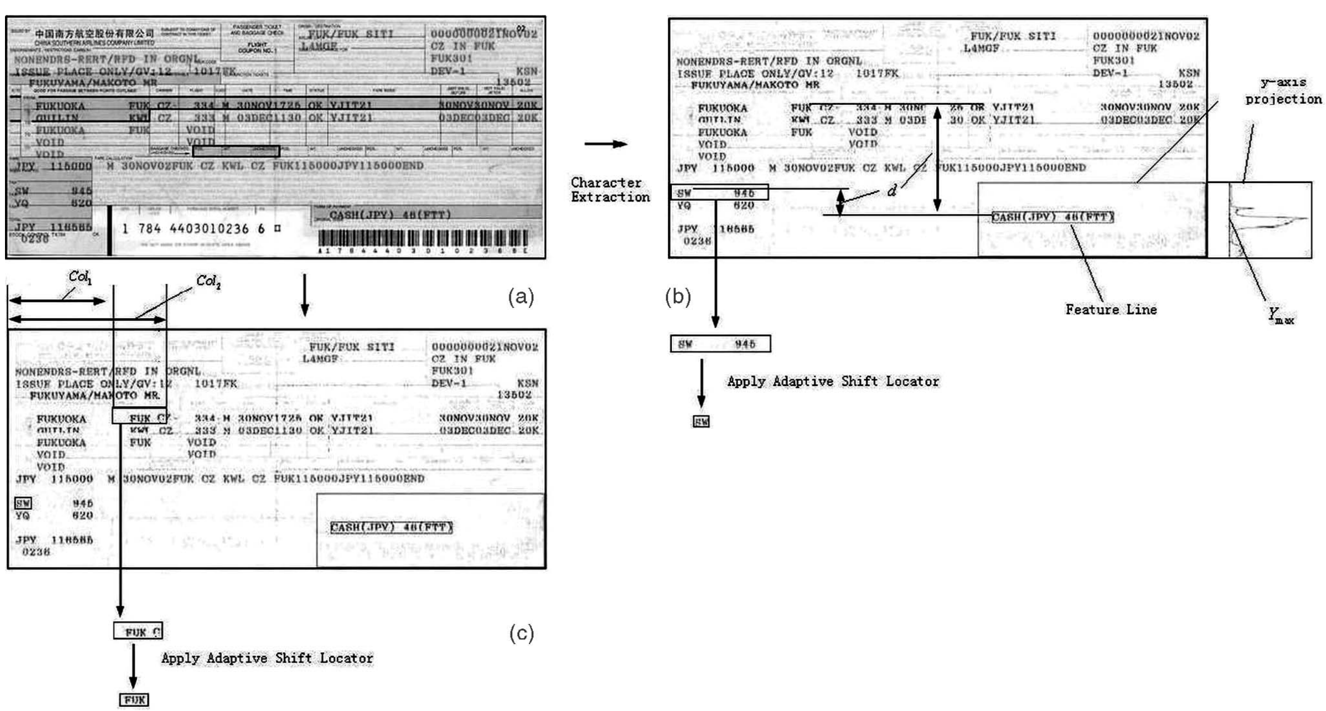

Fig. 5. Experimental steps in character strings detection and location. Explanation: (a) Original coupon images. (b) Image after preprocessing and initial position in row. In the lower right corner, the curve of y-axis projection in local image and corresponding maximal value $Y_{\max}$ are presented. (c) Image after preprocessing and initial position in column.

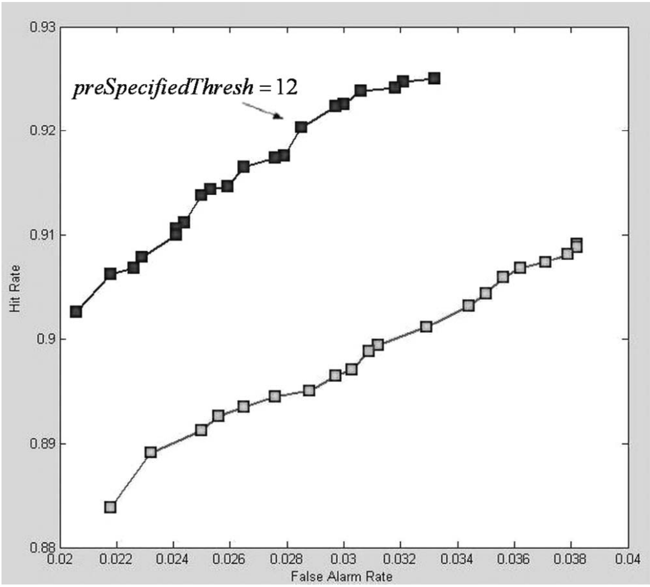

Fig. 7. ROC Curve ($preSpecifiedThresh = [6, 7, \ldots, 24, 25]$). EMACH: upper dotted curve with dark gray rectangles. MACE: lower dotted curve with light gray rectangles.

known character object. "False Alarm Rate" denotes probability of SD report when the field does not contain any known character object. By analyzing ROC curves, we can find that the performance of EMACH is slightly better than MACE, in our case, and $preSpecifiedThresh = 12$ is the best point of trade-off.

For simplification, we only draw figures under the condition that $preSpecifiedThresh = 12$ in Fig. 8. Our experiment shows that curves also have similar tendency when $preSpecifiedThresh$'s value ranges from 6 to 25.

We calculate overall accuracy of target images and get 90.44 percent (MACE) and 92.03 percent (EMACH) accuracy when $preSpecifiedThresh$ is 12, respectively. It is reasonable and acceptable in most practical cases.

In Fig. 8, we find that one of the remarkable results is the "Total Hit" number of "Airport Code" field of every character content class achieves high accuracy, which infers a high accuracy both in detection and location with different character strings. This result demonstrates the effectivity and high performance of adaptive shift locator using MACE and EMACH. Therefore, we can conclude that our algorithm can be used for variant coupon types, and the same field in different types can be regarded as different profiles (or classifications) of the same object.

We can draw some further conclusions from Fig. 8. By comparing Figs. 8a and 8c, we can find that although two correlation filters provide reasonable result regardless character classes, they are sensitive to quality of images. In Figs. 8b and 8d, we find that two rates both rise with quality going down.

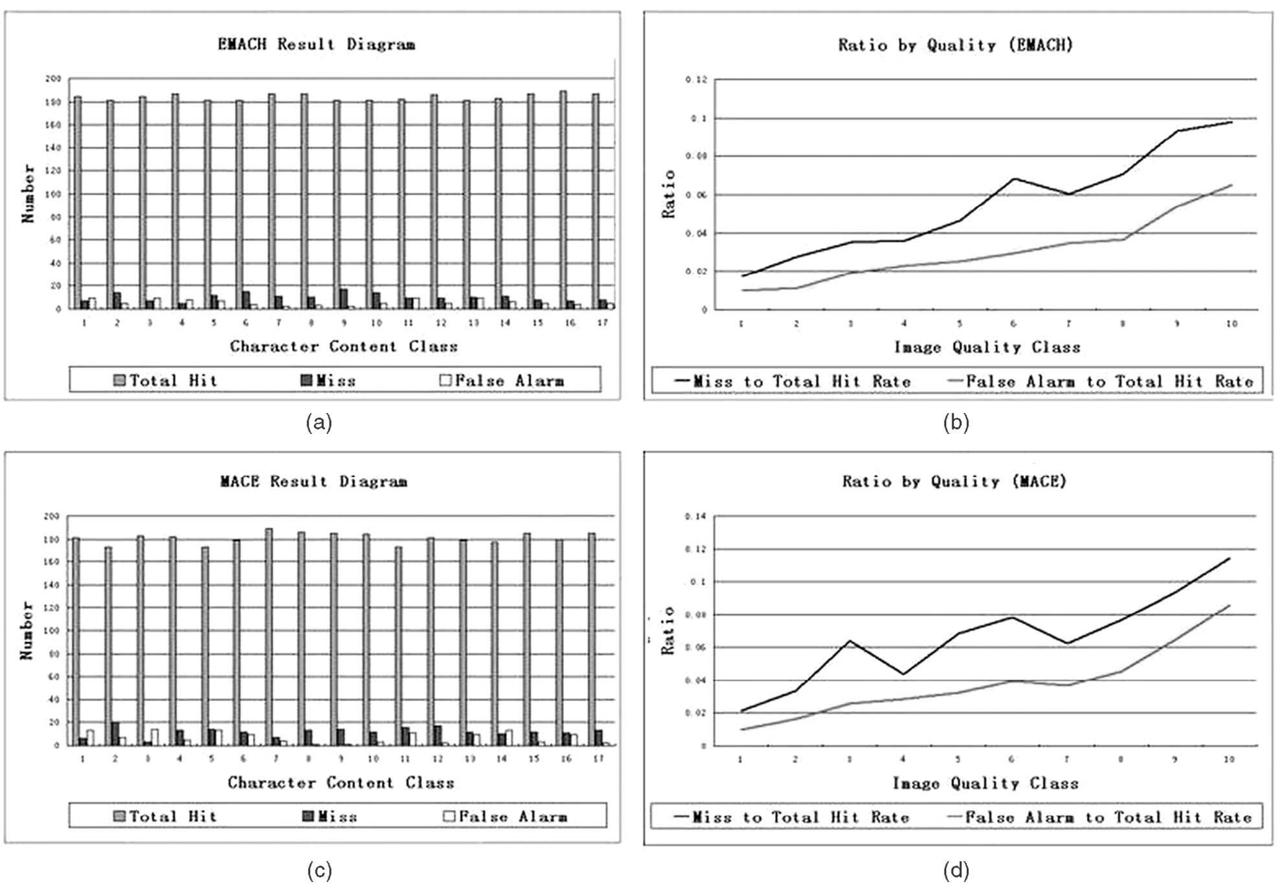

Fig. 8. Diagrams of detection and location results ($preSpecifiedThresh = 12$). Explanation: (a) and (b) are statistical results of EMACH correlation filter. (c) and (d) are statistical results of MACE correlation filter. (a) and (c) are charts of different character content classes. "Total Hit" denotes the total output number of each class whose report is SD if string exists and SM if field's string is missing, "False Alarm" denotes the total output number of each class whose report is SD but string is missing (See Fig. 4d), and "Miss" denotes total output number of each class whose report is SM while string exists. (b) and (d) are curves of different image quality classes. "Miss to Total Hit Rate" denotes the ratio of "Miss" to "Total Hit" for each image quality class. "False Alarm to Total Hit Rate" denotes the ratio of "False Alarm" to "Total Hit" for each image quality class. Definition of "Miss," "Total Hit," and "False Alarm" are the same as those definitions in (a) and (c) except that the statistical classes are image quality classes instead of character content classes. The statistical data of different image qualities are presented in Table 1.

By comparing Figs. 8b and 8d, we find that if coupon's image quality is high, there are not any differences between two correlation filters. However, with poor image quality, EMACH is better than MACE in some sense. Note that poor quality and skewed images can be treated as a kind of variation of distortion so that this phenomenon shows that EMACH performs better than MACE in the situation of distortion, smudgy, and noisy environment. This result also shows that accuracy can be improved if we integrate a restoration phase to enhance input images.

Comparison can be also made between filters' speeds. As seen in (1) and (2), we can find that two filters both need to compute an invertible matrix in training phase, which is costly. Besides, the invertible matrix size of EMACH is larger than MACE, and EMACH needs to compute the eigenvalue and eigenvector of a large size matrix, which are both a time-consuming operation. However, they are in a close temporal and spatial complexity in application phase when two correlation filters are applied in target classification, which is because correlation operation in frequency domain is relatively fast.

7 Conclusion

This paper presents an algorithm to detect and locate expected field in multifield irregular table using correlation filter by specifying imprecise geometric information. We extend the conventional processing from gray scale to color space, and apply a method with LVQ and PCA effectively as a text extractor to achieve the best effect of text extraction in preprocessing. Then, we convert our problem to ATR by treating the same field in different coupon types as different profiles of the same "object."

We apply correlation filter as an adaptive shift locator to locate specific fields based on geometric information in our experiment. It is because: 1) The features of low contrast texts cannot be obtained easily with some techniques such as edge detection. 2) Correlation filter avoids registration error of connected components. 3) Correlation filter provides a high accurate result of detection and location of known string object.

We also compare two correlation filters and conclude that: 1) MACE can satisfy our requirements in most situations, but the performance of EMACH is better than MACE in poor quality images at the price of timely computation during training correlation filter. 2) Both filters provide a reasonable temporal complexity in application phase because correlation operation is fast in frequency domain.

The main contributions of our work include: 1) We solve one category of form registration problems in complex tabular background environment by detecting known objects in tables. 2) We treat the same field in different coupon types as different profiles of the same "virtual object" so that we can detect and locate them using ATR, which is used traditionally in military. 3) Although the concept and idea of correlation have been proposed for many years, quite a lot of new applications are still being developed. This paper is one among these new interdisciplinary applications. To best of our knowledge, this is a first trial on flight coupon table registration using pattern recognition approach.

Generally speaking, we can conclude that correlation filter serves as a statistical tool in frequency domain to extract, detect, and locate images that have the same or close features with training images. In sum, our algorithm improves the accuracy location in the coupon table at a reasonable temporal-spatial cost.

Acknowledgments

The authors would like to thank Dr. Liu Wenyin, assistant professor in the Department of Computer Science at the City University of Hong Kong, and Dr. David Doermann, associate researcher in the Institute of Advanced Computer Studies at the University of Maryland in College Park, for their long-term support of this work. The authors would also like to thank the anonymous reviewers for their constructive comments. This research is funded by the China Innovation Fund for Small Technology-based Firms (02C26214400224), and Guangdong Government (2002A1020104).

References

[1] J. Mao, R. Lorie, and K. Mohiuddin, "A System for Automatically Reading IATA Flight Coupons," *Proc. Fourth Int'l Conf. Document Analysis and Recognition,* pp. 153-157, 1997.

[2] Z. Li, "Detection of Text Regions from Digital Engineering Drawings," *IEEE Tran. Pattern Analysis and Machine Intelligence,* vol. 20, no. 4, pp. 431-439, Apr. 1998.

[3] P. Tofani and R. Kasturi, "Segmentation of Text from Color Map Images," *Proc. 14th Int'l Conf. Pattern Recognition,* pp. 945-947, 1998.

[4] C. Zhang, "A New Method of Color Image Segmentation Based on Intensity and Hue Clustering," *Proc. 15th Int'l Conf. Pattern Recognition,* pp. 613-616, 1999.

[5] L.Y. Tseng and R.C. Chen, "Recognition and Data Extraction of Form Documents Based on 3 Types of Line Segments," *Pattern Recognition,* vol. 31, no. 10, pp. 1525-1540, 1998.

[6] S.L. Taylor, R. Fritzson, and J.A. Pastor, "Extraction of Data from Preprinted Forms," *Machine Vision and Application,* vol. 5, pp. 211-222, 1992.

[7] T. Watanabe, Q. Luo, and N. Sugie, "Layout Recognition of Multi-Kinds of Table-Form Documents," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 17, no. 4, pp. 432-445, Apr. 1995.

[8] K.C. Fan, J.M. Lu, and G.D. Chen, "A Feature Point Clustering Approach to the Recognition of Form Documents," *Pattern Recognition,* vol. 31, no. 9, pp. 1205-1220, 1998.

[9] R. Casey, D. Ferguson, K. Mohiuddin, and E. Walach, "Intelligent Forms Processing System," *Machine Vision and Applications,* vol. 5, no. 5, pp. 143-155, 1991.

[10] Y.Y. Tang, H. Ma, J. Liu, B.F. Li, and D. Xi, "Multiresolution Analysis in Extraction of Reference Lines from Documents with Gray-Level Background," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, no. 8, pp. 921-926, Aug. 1997.

[11] F. Cesarini, E. Francesconi, M. Gori, and G. Soda, "Analysis and Understanding of Multi-Class Invoices," *Int'l J. Document Analysis Recognition,* vol. 6, no. 2, pp. 102-114, 2003.

[12] J.L. Chen and H.J. Lee, "Field Data Extraction for Form Document Processing Using a Gravitation-Based Algorithm," *Pattern Recognition,* vol. 34, no. 9, pp. 1741-1750, 2001.

[13] B. Yu and A.K. Jain, "A Generic System for Form Dropout," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 18, no. 11, pp. 1127-1134, Nov. 1996.

[14] Y. Li et al., "String Extraction in Complex Coupon Environment Using Statistical Approach," *Proc. Seventh Int'l Conf. Document Analysis and Recognition,* pp. 289-294, 2003.

[15] Y. Li et al., "High Accurate Character Strings Extraction from Coupon Images Using Hybrid Neural Network," *Computer Eng. and Application,* vol. 40, no. 8, pp. 209-212, 2004. (in Chinese)

[16] S. Haykin, *Neural Networks-A Comprehensive Foundation,* pp. 392-440, second ed. Prentice-Hall Inc., 1998.

[17] T. Kohonen, "The Self-Organizing Map," *Proc. IEEE,* vol. 78, no. 9, pp. 1464-1480, Sept. 1990.

[18] D. Xi and S.-W. Lee, "Reference Line Extraction from Form Documents with Complicated Backgrounds," *Proc. Seventh Int'l Conf. Document Analysis and Recognition,* pp. 1080-1084, Aug. 2003.

[19] C.F. Hester and D. Casasent, "Multivariant Technique for Multiclass Pattern Recognition," *Applied Optics,* vol. 19, pp. 1758-1761, 1980.

[20] M. Sonka et al., *Image Processing, Analysis and Machine Vision,* pp. 191-194, second ed. Thomson Learning, 2000.

[21] R.C. Gonzalez et al., *Digital Image Processing,* pp. 703-704, second ed. Prentice Hall, 2002.

[22] A. Mahalanobis, B.V.K. Vijaya Kumar, and D. Casasent, "Minimum Average Correlation Energy Filters," *Applied Optics,* vol. 26, pp. 3633-3630, 1987.

[23] R. Singh and B.V.K. Vijaya Kumar, "Performance of the Extended Maximum Average Correlation Height (EMACH) Filter and the Polynomial Distance Classifier Correlation Filter (PDCCF) for Multiclass SAR Detection and Classification," *Proc. SPIE,* vol. 4727, pp. 265-276, 2002.

[24] M. Savvides, B.V.K. Vijaya Kumar, and P. Khosla, "Face Verification Using Correlation Filters," *Proc. Third IEEE Automatic Identification Advanced Technologies,* pp. 56-61, 2002.