Postal Envelope Address Block Location by Fractal-Based Approach

Luiz Felipe Eiterer
Jacques Facon
PUCPR-Pontifícia Universidade Católica do Paraná, Rua Imaculada Conceição 1155, Prado Velho
80215-901 Curitiba-PR, Brazil

David Menoti
UFMG Universidade Federal de Mis Gerais, DCC Departamento de Ciência da Computação
Av. Antônio Carlos, 6627 31.270-010, Belo Horizonte-MG, Brazil
{eiterer,facon,menoti}@ppgia.pucpr.br

Abstract

In this paper, we propose an approach based on fractal dimension to automatically locate address blocks in postal envelopes. First, the fractal dimension of each pixel of a postal envelope image is computed using the 2D variation procedure. The K-means clustering technique is then used to label pixels as background, noise and semantic objects like stamps, postmarks, and address blocks. A database composed of 200 *postal envelope images, with no fixed position for the address block, postmark and stamp is used to evaluate the efficiency of the proposed approach. For each envelope image, the ideal result (ground-truth segmentation) regarding each class has been generated. The comparison between the ground-truth segmentation and the results obtained through the proposed methodology is carried out pixel by pixel. Experiments showed significant and promising results. By using the 2D variation procedure for three ranges of neighbor window sizes (*$r = \{3, 5\}$*,* $r = \{3, 5, 7\}$*, and* $r = \{3, 5, 7, 9\}$*), the proposed approach reached a success rate over than* 90% *on average.*

1. Introduction

It is well known that the proper addressing allows mail pieces to be processed quickly and more efficiently. We also know that this task is not a mystery. When the mail piece meets size requirements, address block and zip code are filled in correctly and the proper amount of postage is obeyed, it moves easily through the mechanized sorting process saving labor and time.

Therefore, despite of all rules mentioned above, why is so difficult to increase mail sorting automation? Some reasons like wide variety of postal envelope attributes (layouts, colors and texture), the handwritten address block which appears mixed up with postmarks or stamps are factors that increase the complexity of an efficient mail sorting system.

Several authors have dealt with the problem of locating address blocks in envelope images. The existing methods are based on Artificial Intelligence techniques, on geometrical characteristics or on texture analysis by filtering.

Wang in [13] presents an expert system for automatic sorting of envelope images. The system uses a blackboard to keep the geometric attributes of blocks obtained after processing different types of envelope images. A statistical database of geometric features and a rule-based inference engine are used to assign a score to each address block candidate. Finally, the candidate with the highest score is selected as the address block. A database of over 3000 mail pieces was used to develop the rules associated with the tools. Experiments with a database of 174 mail pieces indicated a success rate of over 81% in locating the address block.

Jain and Bhattacharjee in [7] present a method for locating address blocks based on a multi-channel filtering approach (two-dimensional Gabor Filters). The text in the envelope image defines a textured region. Non-text components, including blank spaces, are considered as regions of different texture types. Gabor filtering decomposes an input image into a number of filtered images, each of which contains intensity variations over a narrow range of frequency and orientation. A three-cluster solution (*k=3*) identifies the regions consisting of printed text, non-text (blank spaces and regions which have slow intensity variations), and the boundaries between the text regions and the non-text regions. The cluster corresponding to the text is isolated and a connected component analysis is used to identify individual blocks of text in the envelope image. The address block is identified from these blocks. Unfortunately, it is impossible to evaluate the accuracy of the proposed method since few images were tested and no success rate in locating the address block was provided.

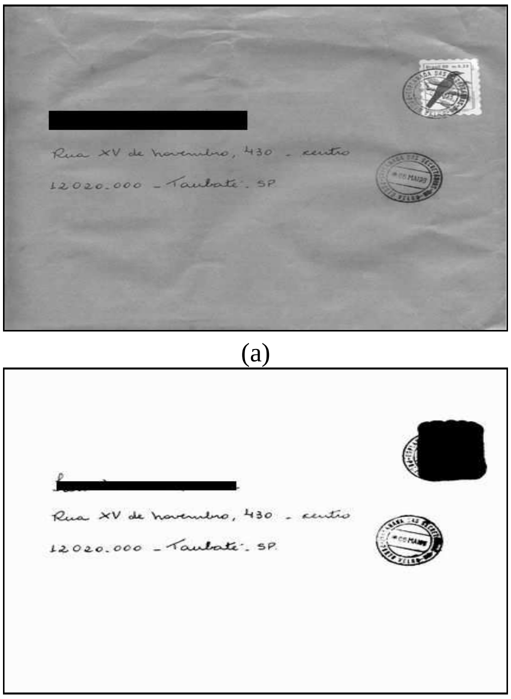
Yu et al in [15] present a method for locating address blocks from magazine envelopes based on the following assumptions: The address is written on a light-colored label, generally skewed. Moreover, the label may be stuck on the magazine, under the plastic envelope or displayed in a plastic window provided in the envelope. The address block follows the left alignment rule. The contrast between the ink and the spaces between characters varies according to the equipment used (laser or matrix printer, or even a typewriter). And the magazine and envelope ensemble contains other text messages in addition to the address block. The authors segment magazine envelopes using an Otsu's modified method which reduces the influence of the grayscale distributions. Then a connected component analysis based on BAG strategy is used to identify individual blocks of text. Some heuristics are used to eliminate false address block candidates. Finally, a recognizer is used to locate the address block. Experiments on 53 IBM magazine envelopes and 52 other ones showed that the method was successful in locating the correct address block in 71.70% and 92.86%, respectively.

Yonekura et al [14] present a method for postal envelope segmentation combining the 2-D histogram and morphological clustering. A new filter based on the morphological grayscale dual reconstruction was proposed to the 2-D histogram calculation and the proposed clustering was based on the watershed transform criterion. Experiments on a database composed of 300 complex postal envelope images, with creased background and no fixed position for the handwritten address blocks, postmarks and stamps showed that the method was successful in locating the correct address block in 75%.

Some methods presented above are highly dependent on some *a priori* knowledge of the envelopes such as orientation and position of the address block. Consequently, they are not really suitable for complex envelopes. Due to the lack of quantitative evaluation criteria in some publications, an objective comparison is a difficult task.

The goal of this paper is to propose and to evaluate a new postal envelope segmentation method based on Fractal Dimension clustering. The aim is to locate and segment handwritten address blocks, postmarks and stamps, as illustrated in figure 1, with no *a priori* knowledge of the envelope images. The Fractal Dimension clustering is carried out by means of the Kmeans. A ground-truth segmentation is employed to evaluate the accuracy of this approach.

This paper is organized as follows: section 2 describes some methods of grayscale image fractal dimension, section 3 shortly reviews fractal-based techniques to solve the segmentation challenge. The proposed method based on the K-means is detailed in section 4. Section 5 describes our approach and details the used fractal dimension method and the input parameters used in the K-means. Section 6 presents some experimental results and discussions. In addition, the proposed segmentation method is evaluated by means of ground-truth images in sections 6.1 and 6.2. Finally, some conclusions are drawn in section 7.



(a)

(b)

Figure 1. Ideal envelope image segmentation: (a) Envelope image, (b) Ground-truth handwritten address block, postmark and stamps.

2. Fractal Dimension for grayscale images

Fractals are very useful and have become popular in modeling roughness and self-similarity properties of surfaces in image processing. Mandelbrot [9] proposed fractal geometry and is the first one to notice its existence in the natural world. The fractal dimension gives a measure of the roughness of a surface. Intuitively, the larger the fractal dimension, the rougher the texture is. A deterministic fractal is defined using this concept of self-similarity as follows. Given a bounded set A in a Euclidean n-space, the set A is said to be self-similar when A is the union of N distinct (non-overlapping) copies of itself, each of which has been scaled down by a ratio of r. The fractal dimension FD is related to the number N and the ratio r as follows:

$$FD = \frac{logN}{log(1/r)} \qquad (1)$$

In fact, most natural surfaces and in particular textured surfaces are not deterministic as described above but have a statistical variation. So this makes the computation of grayscale image fractal dimension FD more difficult. In literature, on can meet some methods for estimating the FD for grayscale images. Probably the most popular one is the differential box-counting method introduced by Chaudhuri et al [2]. Consider an $M \times M$ pixel image as a surface in (x, y, z) space where $P(x, y)$ represents the pixel position and z is the pixel intensity. We now partition the (x, y) space into a grid of $s \times s$ size pixels. An estimate of the relative scale is $r = s/M$. At each grid position, we stack cubes of size s, numbering each box sequentially from 1 up to the box containing the highest intensity in the image over the $s \times s$ area. Denoting the minimum and maximum grey levels for the image at position (i, j) by k and l respectively, one can define $n_r(i, j) = l - k + 1$. The parameter N in equation 1 is then estimated from $N(r)$ by summing over the entire grid as $N(r) = \sum_{(i,j)} n_r(i, j)$. The above procedure is repeated for a number of values of $r(s)$. The slope estimation of the graph $log(N(r))$ versus $log(1/r)$ by the least-squares method provides the fractal dimension FD.

The 2D variation procedure, proposed by [5] [3], similar to the differential box-counting method, analyses the pixel environment at different distances r. The squares with different sizes r are running pixel by pixel across the image from left to right and from top to bottom. The algorithm determines the minimum min and maximum max grey values within the square of size r. Since the intensity value v of a pixel p is related to the height of the objects in the image, one can meaningfully define at p a volume $r \times r \times v$. By defining the difference volume of p as $r \times r \times (max - min)$, one can denote $V(r)$ as the sum of differences between the maximum and minimum for scale r for the entire image, resulting in a power law: $V(r) = const\ r^s$. In the Richardson-Mandelbrot plot ($log(V(r))$ versus $log(r)$), the dependence of this volume $V(r)$ should be linear with the square size r. By using the least-squares method to this regression line, one can compute the slope s of this line and then deduce FD:

$$FD = 3 - s/2 \qquad (2)$$

Maragos and Sun [10] have introduced the measurement of FD by morphological covers for a discrete-time finite length signal $f[m]$, $m = 0, 1, ...M$ by using covers at discrete scales $t = 1, 2, 3, ..., t_{max}$. By using dilations and erosions with structuring element g, they computed the Cover Area $A_g[t]$:

$$A_g[t] = \sum_{m=0}^{M} (\delta_g^t(f) - \varepsilon_g^t(f))[m] \qquad (3)$$

where $t = 1, 2, 3, ..., t_{max} \leq M/2$. By fitting a straight line using least-squares to the graph of $log(A_g[t]/(t')^2)$ versus $log(1/t')$, where $t' = 2t/M$, one can compute the slope s which is an approximate estimate of FD for the signal $f[m]$. In the case of images, one can deduce FD by equation 2.

3. Fractal-based techniques for segmentation challenge

To clearly show the interest in fractal-based methods for image processing, a number of recent practical applications in segmentation challenges are shortly reviewed. For instance, Diayana et al [4] have compared three microcalcification detection methods in digital mammogram. They showed that the fractal analysis based on morphological dilation and erosion in 8×8 image blocks is a promising approach.

Andrew Ringler [11] has used a local FD approach to segment oceanic images into *water* and *no-water* classes. By using the 2D variation procedure for each image pixel, the author has successfully segmented ocean waves to detect dolphin and swimmers and to perform human search and rescue.

Samarbandu et al [12] have proposed an initial study of bone X-Rays images through a morphological fractal analysis. In this approach, where only the dilation is used, the FD computation by morphological covers is easier than in the Maragos and Sun's one. They have demonstrated the feasibility of detecting bone structures independent of scale.

4. Clustering

We are aiming at automatic clustering, where the handwritten address block, postmark and stamp classes should be extracted from the image through the fractal dimension. For this purpose, we use the K-means clustering which is a robust unsupervised algorithm often applied to the problem of image texture segmentation. The K-means clustering is an algorithm for partitioning N data points into k disjoint clusters C_j containing N_j data points so as to minimize the sum-of-squares error. Each cluster in the partition is defined by its member data points and by its centroid. The centroid for each cluster is the point to which the sum of distances from all data points in that cluster is minimized. The

K-means clustering uses an iterative algorithm that minimizes the sum of distances $S_d = \sum[C_j(P) - P]^2$ from each data point P to its cluster centroid C_j, over all clusters. This algorithm moves data points between clusters until the sum cannot be decreased further. The result is a set of clusters that are as compact and well-separated as possible. It is comprised of a simple re-estimation procedure as follows [1] [8]:

1. First, the data points are assigned at random to the k sets;
2. Second, the centroid is computed for each set.

These two steps are alternated until a stopping criterion is reached (i.e. when there is no further change in the assignment of the data points). Details of the minimization can be controled using several optional input parameters to K-means, including ones for the initial values of the cluster centroids, and for the maximum number of iterations.

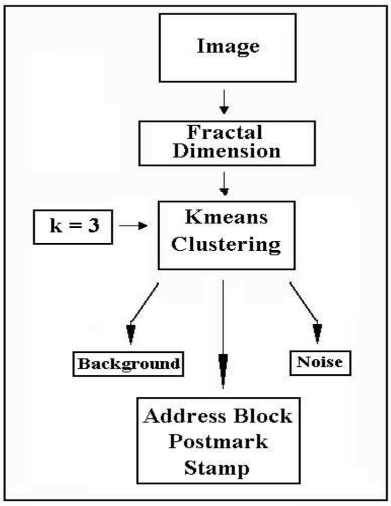


Figure 2. Fractal Segmentation Approach Scheme

5. The Fractal-based approach

As already explained before, our address block segmentation approach is based on fractal dimension computation and class labeling through K-means clustering technique (Figure 2).

5.1. Fractal dimension computation

Among the available FD computation techniques and Fractal-based approaches, one could observe that the 2D variation procedure is less complex than other ones (for instance morphological covers) and has shown to be promising in complex images (see [11]). It is the reason why it has been chosen. We applied the 2D variation procedure onto postal envelope images for three ranges of neighbor window sizes ($r = \{3, 5\}$, $r = \{3, 5, 7, \}$, and $r = \{3, 5, 7, 9\}$). It results in a I_{FD} image where the FD of each pixel stays in the range $[2.0, 3.0]$ (see equation 2).

5.2. K-means parameters

The input parameters used in K-means clustering are:

- Number of features: only one, the FD itself;
- Number of clusters : 3 clusters which are the handwritten address block, postmark and stamp one, the background one and the noise one;
- Number of iterations: 10;
- Convergence criterion: 0.0001;

It is well-known that the K-Means algorithm suffers from initial starting conditions effects (initial clustering effects). On can meet in the literature some researches about the initialization optimisation of iterative partitional clustering [6]. In our approach, min, $(max + min)/2$ and max are used to initialize the values of cluster centroids, where max and min respectively are the FD buffer maximum and minimum values.

The convergence of the K-Means algorithm onto the image I_{FD} results in the labeled image I_{Kmeans} where each pixel is labeled into the 3 clusters.

5.3. Selection of the handwritten address block, postmark and stamp cluster

The segmentation of the handwritten address block, postmark and stamp cluster requires its identification among the 3 clusters in I_{Kmeans}. First, the spatial distribution measurement of each class in the labeled image is carried out by means of a co-ocurrence matrix. In our approach, the co-ocurrence matrix of I_{Kmeans} is a 3×3 dimensional matrix $C = [c_{ij}]$ which contains statistical information regarding the number of cluster occurrences of two pixels separated by a displacement (d_x, d_y). The co-occurrence matrix $C = [c_{ij}(d_x, d_y)]$ from any pair of pixels *(x,y)* and *$(x + d_x, y + d_y)$* representing respectively the clusters *(i,j)* can be formalized as follows:

$$c_{ij}(d_x, d_y) = \#\{(x, y)/ I_{Kmeans}(x, y) = i, I_{Kmeans}(x + d_x, y + d_y) = j\} \quad (4)$$

where $\#$ denotes cardinality and where $i, j \in \{1, 2, 3\}$.

To better measure the interaction between the 3 clusters, the co-ocurrence matrix is built by union of neighbouring pixel occurrences in horizontal, vertical, and diagonal directions, respectively named C_{hor}, C_{ver} and C_{diag}. The final co-occurrence matrix $C = [c_{ij}]$ used in our approach is

$$C = C_{hor} + C_{ver} + C_{diag} = \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{bmatrix} \quad (5)$$

For each cluster *k*, an accumulator Acc_k is defined as follows:

$$Acc_k = \sum_{i=1}^{3} c_{i,k} + \sum_{j=1}^{3} c_{k,j}, \text{ where } 1 \leq k \leq 3$$

The background cluster is identified, from Acc_k, by selecting the cluster *k* for which Acc_k is maximum:

$$k_{background} = Arg_{\ k} \max{(Acc_k)}$$

After the background cluster identification, one can identify the address block cluster as the cluster with the smallest interaction with the background cluster and with more than 1.5% of pixels related to entire image I:

$$k_{addressblock} = Arg_{\ k} \min{[c_{k,k_{background}} + c_{k_{background},k}]}$$

where $1 \leq k \leq 3$ and $k \neq k_{background}$ and $N_l > 1.5\% \times N_I$. N_l represents the number of pixels that belong to cluster l, and N_I is the total number of pixel in the original image I. The factor 1.5% was obtained after an empirical study of envelope images. It means that if the cluster is very small, it probably is a noisely cluster.

6. Experiments

A database composed of 200 complex postal envelope images, with no fixed position for the handwritten address blocks, postmarks and stamps was used to evaluate the efficiency of the proposed approach. Each image has about 1500×2200 pixels and was digitized at 300 dpi. We could verify that the address blocks, stamps and postmarks represent only 1.5%, 4.0% and 1.0% of the envelope area, respectively and that the great majority of pixels of these images belong to the envelope background (approximatively 93.5%).

6.1. Evaluation strategy of the proposed method

A ground-truth strategy was employed to evaluate the accuracy of the proposed approach. The ideal result (ground-truth segmentation) regarding each class (handwritten address block, postmarks and stamps) has been generated for each envelope image (figure 1).

The comparison between the ground-truth segmentation and the results obtained by the proposed approach was carried out pixel by pixel. We have computed each obtained result with the ideal segmentation in terms of identical pixels at the same location. We have also computed the average of all obtained results for each class.

6.2. Numerical results and discussion

Figures 3, 4 and 5 show interesting address block, stamp, postmark segmentations in case of envelope images with creased background and wrong layout (Figure 3), with many stamps and postmarks (Figure 4), and without stamp (Figure 5). In spite of the background complexity, and independently of the semantic classes, the address blocks, stamps and postmarks are very well segmented. Little noise appears in the results. It is possible to conclude that, independently of the layout and background in the input images, the segmentation recovered the address blocks, stamps, postmarks with great success. As explained before, the evaluation of the approach was carried out by comparing pixel by pixel the ground-truth segmentation and the results obtained. The average segmentation rates for each class are showed in Table 1. This table shows how the accuracy of the algorithm changes if the box size range r is changed. Best values are obtained for range $r = \{3, 9\}$, where the segmentation recovered address blocks ($97.24\% \pm 13.64\%$) postmarks ($91.89\% \pm 17.22\%$) with great success. The most important figures are the ones related to the address block, since for practical purposes the classes of stamp, and postmark would be discarded in the end and it would not harm the result. It can be seen that the segmentation did not recovered the stamps as expected. It is caused by the fact that the stamps contain some complex drawings, gray level differences (dark objects and bright background) decreasing a good rate.

7. Conclusions

A new postal envelope segmentation method based on fractal dimension was proposed. After computing the fractal dimension for each image pixel, the resulting fractal image is then clustered by K-means procedure. By observing the obtained results in Table 1, it is possible to conclude that the proposed approach is robust if we consider the *handwritten address block and postmark* class without *a priori* knowledge about the position of each semantic class. The 2D variation procedure was shown to be appropriate in our purpose. In spite of its simplicity, the centroid initialization in K-means was shown to be efficient in the segmentation challenge. The use of of co-ocurrence matrix in the automatic selection of the handwritten address block, postmark and stamp cluster was positive. In addition, since

Region Class Box sizes r= $\{3,5\}$	Accuracy pixel by pixel ($\mu \pm \sigma$)
Address Block	92.61% $\pm$ 13.61%
Stamp	48.86% $\pm$ 13.84%
Postmark	82.02% $\pm$ 19.66%
Region Class Box sizes r= $\{3,5,7\}$	**Accuracy pixel by pixel ($\mu \pm \sigma$)**
Address Block	96.05% $\pm$ 10.89%
Stamp	59.76% $\pm$ 15.68%
Postmark	88.97% $\pm$ 17.58%
Region Class Box sizes r= $\{3,5,7,9\}$	**Accuracy pixel by pixel ($\mu \pm \sigma$)**
Address Block	97.24% $\pm$ 13.64%
Stamp	66.34% $\pm$ 16.90%
Postmark	91.89% $\pm$ 17.22%

Table 1. Average results with identification of regions (pixel by pixel accuracy).

this method does not use a priori knowledge, it can be employed to segment other types of document images. Future works will focus on the use of other features (like lacunarity, skewness, kurtosis) to improve the segmentation task by K-means and the evaluation of segmented classes in handwritten recognition systems for full automatic classification.

Acknowledgments

We would like to acknowledge support for this research from PUCPR, CNPq/MCT and the Brazilian Post Office Agency (Correios Brasileiros).

References

[1] C. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.

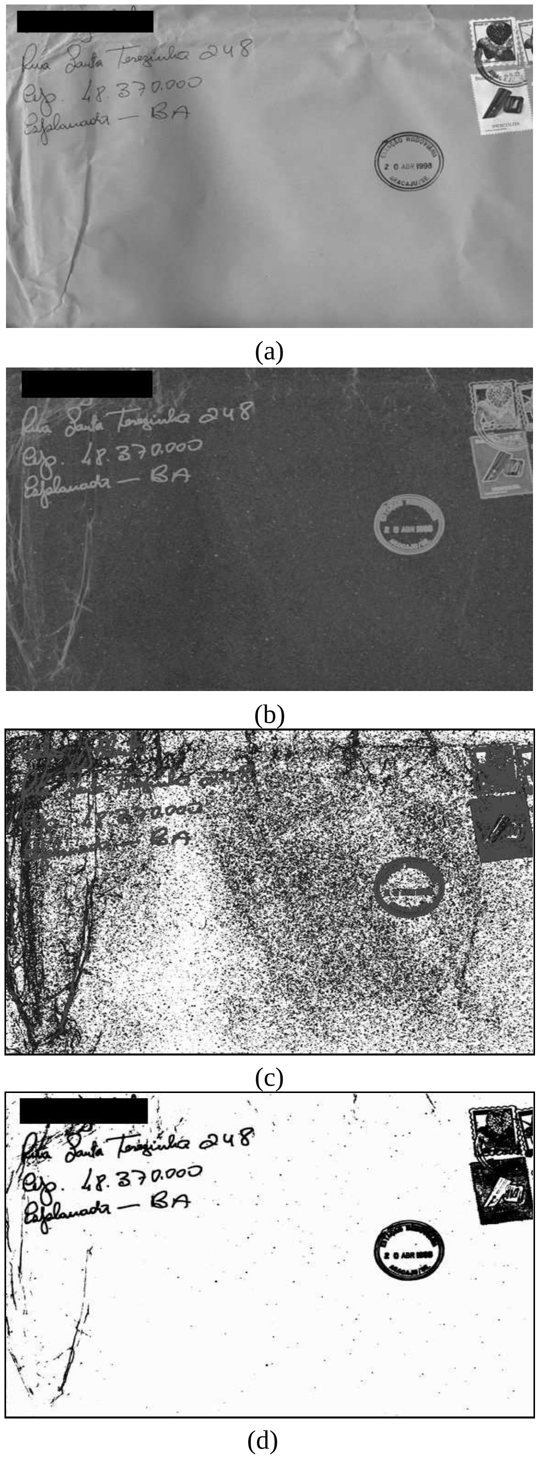
[2] B. Chaudhuri, N. Sarkar, and P. Kundu. Improved fractal geometry based texture segmentation technique. *IEE Proceedings*, 140:233–241, 1993.

[3] K. Clarke. Computation of the fractal dimension of topographic surfaces using the triangular prism surface area method. *Computers and Geosciences*, 12(5):713–722, 1986.

[4] W. M. Diayana, J. Larcher, and R. Besar. A comparison of clustered microcalcifi cations automated detection methods in digital mammogram. *ICASSP*, pages 385–388, 2003.

[5] B. Dubuc, J. Quiniuo, C. Roques-Carmes, C. Tricot, and S. Zucker. Evaluating the fractal dimensions of profi le. *Phys. Rev.*, 39:1500–1512, 1989.

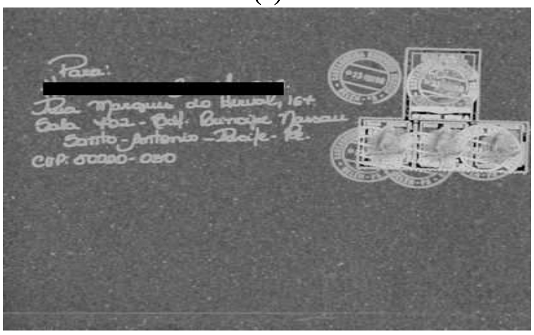
[6] U. Fayad, C. Reina, and P. Bradley. Initialization of iterative refi nement clustering algorithms. *4th KDD98- Interna-*



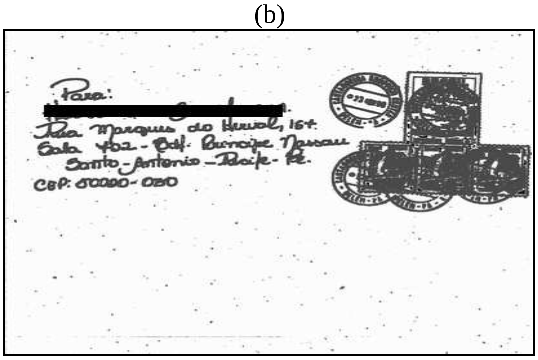
(a) (b) (c) (d)

Figure 3. Results for an envelope image with creased background and wrong layout: (a) Envelope image, (b) Fractal Image,(c) K-means clustering, (d) Handwritten address block, postmark and stamps.

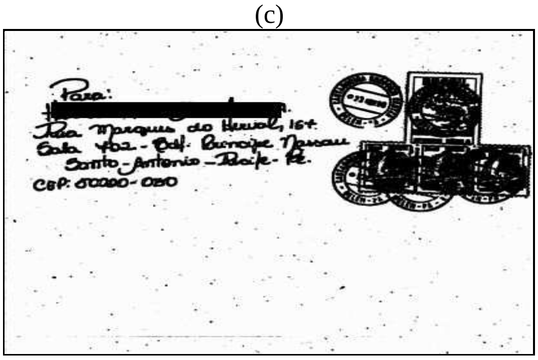
(a)



(b)

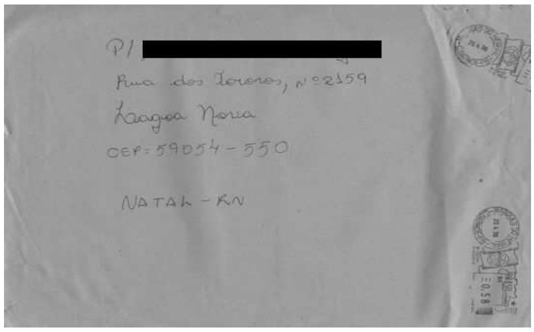


(c)

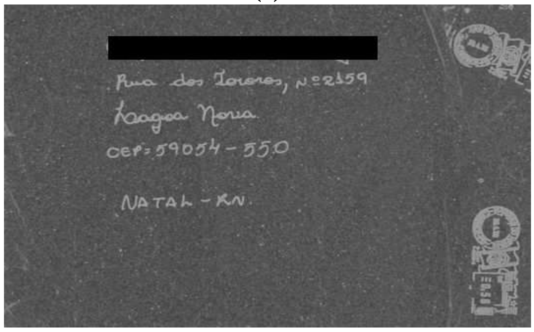


(d)

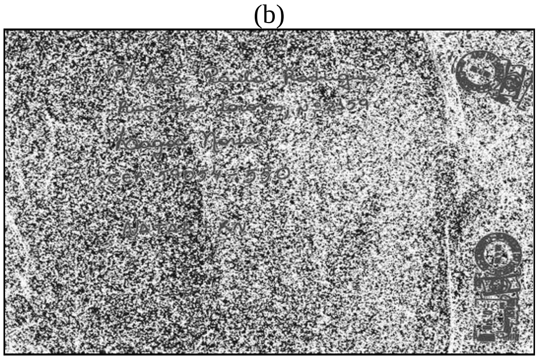
Figure 4. Results for an envelope image with many stamps and postmarks: (a) Envelope image, (b) Fractal Image,(c) K-means clustering, (d) Handwritten address block, postmarks and stamps.



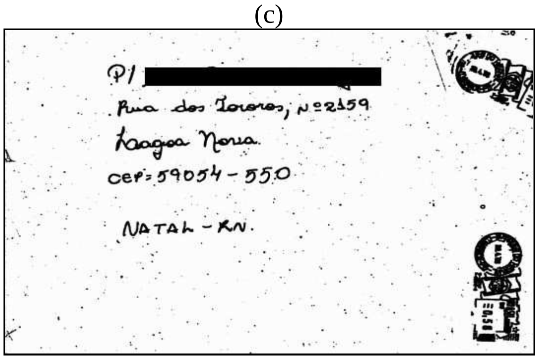
(a)



(b)



(c)



(d)

Figure 5. Results for an envelope image with creased background and no stamp: (a) Envelope image, (b) Fractal Image,(c) K-means clustering, (d) Handwritten address block and postmarks.

tional Conference on Knowledge Discovery and Data Mining, 1998.

[7] A. Jain and S. Bhattacharjee. Address block location on envelopes using gabor filters. *Pattern Recognition*, 25(12):1459–1477, 1992.

[8] A. Jain and R. Dubes. *Algorithms for Clustering Data*. Prentice Hall: New Jersey, 1988.

[9] B. Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman And Company, New york, 1983.

[10] P. Maragos and F. K. Sun. Measuring the fractal dimension of signals: Morphological cover and iterative optimization. *IEEE Trans. on Signal Processing*, 41(1):108–121, 1993.

[11] A. Ringler. Texture segmentation by local fractal dimension as applied to oceanic search and rescue. *IEEE ICICS, International Conference on Information Communications Signal Processing*, 2:975–979, 2003.

[12] J. Samarbandu, R. Acharya, E. Hausmann, and K. Allen. Analysis of bone x-rays using morphological fractals. *IEEE Transactions on Medical Imaging*, 12(3):466–474, 1993.

[13] C. Wang, P. Palumbo, and S. Srihari. Object recognition in visually complex environments: An architecture for locating address blocks on mail pieces. *Proc. Ninth Intl. Conf. on Pattern Recognition*, pages 365–367, 1988.

[14] E. Yonekura and J. Facon. Postal envelope segmentation by 2-d histogram clustering through watershed transform. *ICDAR 2003, 7th International Conference on Document Analysis and Recognition*, 1:338–342, 2003.

[15] B. Yu, A. Jain, and M. Mohiuddin. Address block location on complex mail pieces. *Technical Report MSUCPS:TR97-12, Dept. of Computer Science, Michigan State University*, 1997.