

Postal Envelope Segmentation using Learning-Based Approach

Horacio A. Legal-Ayala

Universidad Nacional de Asunción, Facultad Politécnica,
San Lorenzo, Paraguay
hlegal@pol.una.py

Jacques Facon

Pontifícia Universidade Católica do Paraná,
Programa de Pós-Graduação em Informática Aplicada,
Curitiba, Brazil, 80215-901
facon@ppgia.pucpr.br

and

Benjamín Barán

Universidad Nacional de Asunción, Facultad Politécnica,
San Lorenzo, Paraguay
bbaran@pol.una.py

Abstract

This paper presents a learning-based approach to segment postal address blocks where the learning step uses only one pair of images (a sample image and its ideal segmented solution). First, this approach learns the available knowledge among pixels (each gray level) in an input image and its corresponding output in the ideal segmented solution. A classification array is generated which is re-utilized during the segmentation of new images. Features are extracted and updated by means of an adaptive square neighborhood. At the moment of new image segmentation, the submitted images are segmented by means of a *k-Nearest Neighbor* (*k-NN*) algorithm that seeks, for each pixel, the best solution in the classification array. Tests on a database of 200 complex envelope images were performed and a pixel to pixel accuracy measure validates the new approach. Results compared to other approaches for the same database show the efficiency and performance of the proposed learning-based approach. Success rates achieved for address block, stamps, rubber stamps and noise suggest that the features used in the proposed approach improves the segmentation results.

Keywords: Artificial Intelligence, Computational Learning, Postal envelope, Image Segmentation, k-NN.

1. INTRODUCTION

Postal automation tries to get the mail from the sender to the recipient quickly, in a reliable and economical process. Even if this area was the object of an increasing research in the last years, localization of address blocks in postal envelopes still remains a challenging problem in any image analysis system developed for postal automation.

To correctly send and receive a postal envelope, the Brazilian Post Office Agency suggests some rules for the correct filling of a postal envelope:

- Stamps and rubber stamps must occupy the upper right side of the envelope.
- The address block must be filled in such a way that it does not superimpose to the stamps and rubber stamps.

In a practical way, the above rules are not very often followed, impairing any segmentation technique based on the relative position of the composing classes of a postal envelope.

Other problems can increase the challenge in developing efficient postal image segmentation systems, such as:

- Imperfections of the envelopes themselves, due to handling.
- Presence of drawings, making the background of the envelopes more complex.
- Problems in the acquisition step, which can add borders that do not belong to the postal envelopes.

Several authors have dealt with the problem of locating address blocks in envelope images. One of the earliest works in address location [10] first applies a digital Laplacian operator on an image to separate light and dark regions. In [3], to identify regions in envelope image candidates for being the destination address, the authors apply a texture segmentation based on Gabor filters. In [8] a dedicated hardware for postal address block location is presented. The system is designed as a blackboard architecture that invokes image processing and block analysis tools in a rule-based order. They report tests on 174 mail pieces with success rates of 81% for the Destination Address Block. In [9], another address block location method for working with both machine and hand printed addresses is proposed. The method is based on dividing the input image into blocks where the homogeneity of each block gradient magnitude is measured. In their tests, 1600 machine printed address and 400 hand printed ones were used, reporting over 91% successful location.

Recently, Yonekura and Facon [11] segment postal envelopes combining a 2-D histogram and morphological clustering. The proposed clustering is based on the watershed transform and morphological grayscale reconstruction filtering. Experiments on a database composed of complex postal envelope images, with creased background and no fixed position for the handwritten address blocks, postmarks and stamps show that the method is successful in locating the correct address block in 75%.

Eiterer *et al* [2] propose an address block segmentation approach based on fractal dimension. After computing the local fractal dimension using overlapping square windows of length r , a clustering technique based on K -means is used to label pixels into semantic objects. The evaluation of the efficiency is carried out from a total of 200 postal envelope images with no fixed position for the address block, postmark and stamp. A ground-truth strategy is used to achieve an objective comparison. Experiments reach a success rate over than 90% on average.

Another approach to segment address block location based on feature selection in wavelet space is presented by Menoti *et al* [7]. They run an experimental setup by separating and classifying blocks in the envelopes, and validating results by a pixel to pixel accuracy measure using a ground truth database of 440 original images with different layouts and backgrounds. Success rate for address block location reached is over 85%.

The method we present in this paper is an approach through a computational learning strategy for postal image segmentation; it is a general and robust segmentation method not restricted to a particular envelope layout. The main idea is to use and take advantage of the user's knowledge and his expertise to segment a gray level image into several classes. The method consists of two stages, learning and segmentation. In the first stage, a sample image and its ideal segmented image is submitted in order to extract the relevant features around each pixel neighborhood. The term "ideal image" means the expected image generated by an expert. From this step, one feature vector is generated. The association of this feature vector set with its respective ideal output is named the classification array. The second stage consists in segmenting new images by seeking the best solution for each pixel in the classification array.

In this paper we have extended an early brain resonance magnetic image (RMI), signature and fingerprint image segmentation approach [4, 5] by modifying and optimizing the best set of feature selection and optimizing the best choice of k parameter into Nearest Neighbor strategy. Also, the success rate of this learning-based approach is compared to other known methods over the same database [2, 7, 11].

The rest of this paper is organized as follows. Section 2 details our learning-based approach. Section 3 shows results from an experimental setup we organized using original envelopes with different backgrounds and comparing the success rates achieved by different approaches over the same database. Section 4 points to the conclusions of this work.

2. THE LEARNING-BASED APPROACH

2.1 Mathematical Definitions

A digital image I_x is composed of pixels arranged in a rectangular array with a certain height and width. Each pixel may consist of one or more bits of information, representing the brightness of the image at that point and possibly including color information encoded.

In the context of this work, the digital image I_x is a gray level image where each pixel x_i is codified by 8 bits [0, .., 255] (1). A binary segmented image is formed by an array of pixels codified by only two values: {0} or {255} (2). Thus, we can define:

$$\text{Gray level image: } x_i \in [0, \dots, 255] \quad (1)$$

$$\text{Binary segmented image: } x_i = \{0\} \text{ or } \{255\} \quad (2)$$

being: $x_i=0$, black level; $x_i=255$, white level; and $0 < x_i < 255$ gray level pixels with N the total number of region pixels and $Region$ being either an image x_i or its neighbor.

2.2 View of the Learning-Based Approach

A general scheme of the learning-based segmentation approach is depicted in Figure 1. First, a gray level sample input image (a) along with its expected ideal solution (b) is submitted to the learning algorithm. At this stage, feature vectors are extracted for each pixel compounding a decision array (c). In a next stage, new gray level images (d) are submitted to the decision array. After extracting feature vectors for each pixel from the new images, the segmentation algorithm seeks through the *k-NN algorithm* the best result for every one of these pixels generating segmented images (e).

The nearest neighbor (NN) rule is a learning model based on instances. An early formalization of this rule can be found in [1]. The 1-NN approach can be extended to the k nearest neighbors, or *k-NN*.

The nearest neighbor rule requires the definition of a distance between two elements. Distance used in this work is the Euclidian Distance (3) (4). The value of this distance between two vectors \underline{x} and \underline{y} is given by:

$$D(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\| = \sqrt{(\underline{x} - \underline{y})^T (\underline{x} - \underline{y})} \quad (3)$$

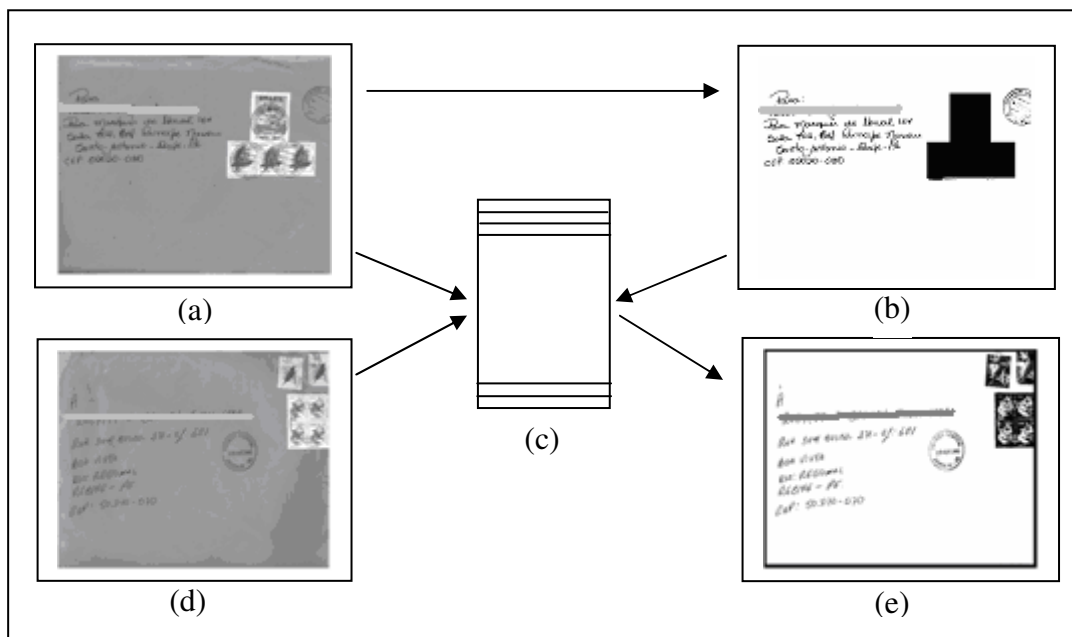
$$= + \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4)$$

2.3 The Learning Stage

The learning stage depicted in Figure 3-(a) consists in manipulating one pair of images, the grayscale learning sample and its ideal segmented version. Both images are analyzed simultaneity where each pair of pixels proceeding from both learning sample image and the ideal one is learned. A set of features mapping each pair of pixels and its respective neighborhood is computed. A classification array for each gray level is then generated that will be forward re-utilized during the image segmentation. The learning stage can be decomposed as follows:

- The feature set computation for each learning sample image pixel.
- The assignment of the corresponding input grayscale level to the feature set.

- The assignment of the ideal binary output to the feature set.
- The classification verification which consists in detecting two pixels with equal feature set and different ideal binary outputs that characterizes a learning error. This situation is solved by automatically enlarging the neighborhood window (initially 3x3 pixels) and by re-applying the learning step until the learning error elimination.
- The storage of the converging learning window size.
- Finally, the storage of the classification array.



(a) Learning input image, (b) Ideal output image, (c) Decision array,
(d) New submitted image, (e) Segmented image

Figure 1: General scheme of the learning-based segmentation approach

2.4 Choice of Features

It is well known that the choice of features is primordial in any learning process. In our case, it is necessary to define the feature set that better defines the correspondence between the triple (pixel, gray-level, neighbor) in the learning sample image and the ideal segmented one. Efforts were taken to select features available in literature that:

- (a) better define the relationship between the learning triple and the ideal output;
- (b) be able to reproduce the learning results in images of high complexity; and
- (c) reduces the computational cost and learning errors.

The feature set is: gray pixel level (x_i), mean (μ), variance (σ^2 or ν), skewness (γ or s) and kurtosis (c). They represent the first (5), second (6), third (7) and fourth moments (8), respectively. Corresponding formulas are denoted below.

$$\mu = E(X) = \frac{1}{N} \sum_{i=1}^N x_i \quad (5)$$

$$\sigma^2 = E(X^2) = \frac{1}{N} \sum_{i=1}^N [x_i - E(X)]^2 \quad (6)$$

$$\gamma = \frac{E[(X - E(X))^3]}{N\sigma^3} \quad (7)$$

$$c = \frac{E[(X - E(X))^4]}{N\sigma^4} \quad (8)$$

2.5 Adaptive learning window sizing

The performance of any learning process can be measured by the efficiency in solving any ambiguity. In our case, in order to overcome this challenge, an adaptive square neighborhood was chosen to adjust the learning to the context.

Beginning with a 3x3 window, the learning process computes all features for each pixel. In order to detect learning conflicts, a verification step detects if there exist two or more input pixels with the same feature set but for different ideal outputs. If this is the case, the window size is automatically increased (Figure 3). This process is automatically repeated until no learning conflict remains. After this step is concluded, the converging learning window size is stored (Figure 2).

The classification array stores, accordingly to the gray levels, the normalized feature vectors obtained for all the pair of pixels of the ideal input and output images, excluding redundancies.

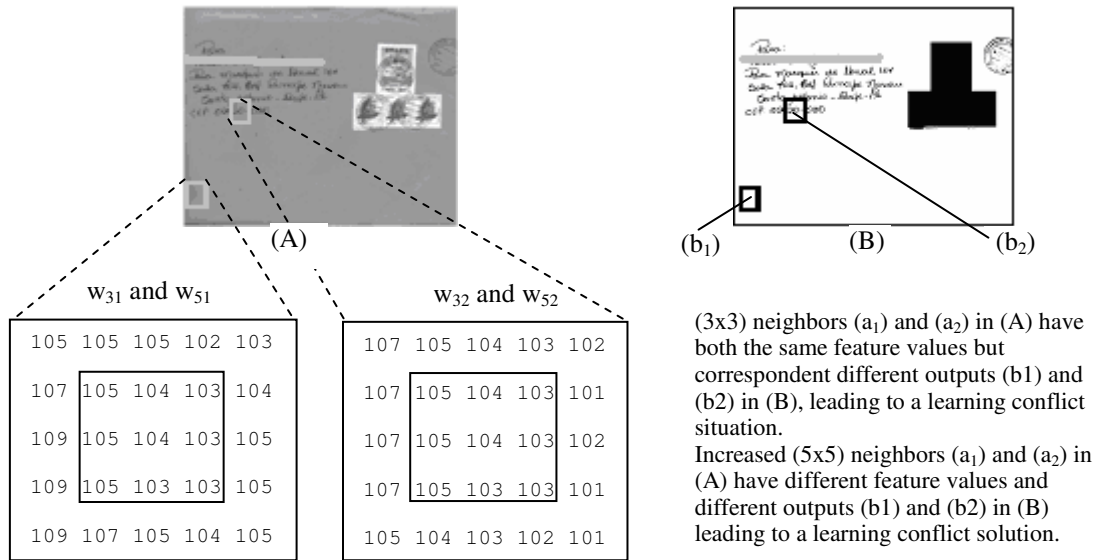


Figure 2: Illustration of the adaptive window sizing in learning stage

2.6. Segmentation by learning step

The steps of the segmentation stage by learning depicted in Figure 3-(b) are:

- Assign the stored learning window size to segmentation one.
- Compute the feature set for each pixel in the new image to be segmented.
- Assign the feature set to the corresponding input grayscale level.
- Define the “most likely” pair (feature vector, output) in the classification array. This step is performed by using the *k-Nearest Neighbor* strategy.
- Assign to each pixel the output value stored in the classification array.

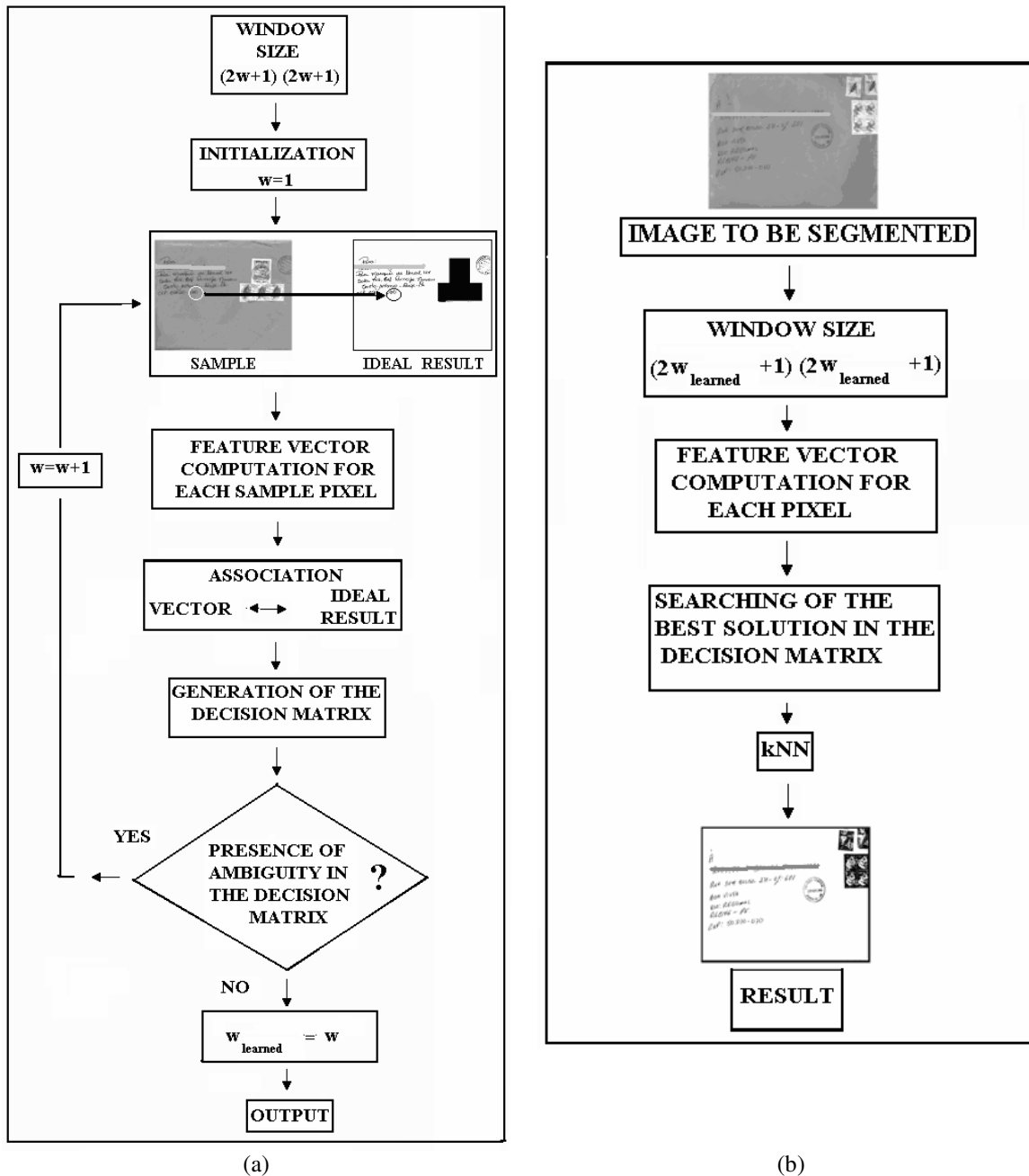


Figure 3: (a) Scheme of the learning stage (b) Scheme of the segmentation stage

3. EXPERIMENTAL RESULTS

3.1. Postal Envelope database

The testes presented here were carried out using the database used by [2] and [7]. This database is composed of 200 complex postal envelope images, with no fixed position for the handwritten address blocks, postmarks and stamps. Each grayscale image of approximately 1500×2200 pixels, was digitized at 200 dpi. It was verified that the address blocks, stamps and postmarks represent only 1.5%, 4.0% and

1.0% on average of the envelope area, respectively. The great majority of pixels of these images belong to the envelope background (approximately 93.5%).

3.2. Selection of best feature set and best k nearest neighbor parameter

Our novel approach includes an automatically selection of the best feature set and of k parameter in Nearest Neighbor strategy. To adaptive select the features and the k parameter, a random selection of ten postal envelope images with their corresponding ground-truth solution was carried out and the ten images were submitted to the learning stage. The procedure consists of the following tasks:

- a) Define the classification arrays using the combination of the four features and odd values of k parameter (from 1 to 15).
- b) Perform the cross validation by segmenting the remaining 9 selected images using each classification array.
- c) Compare each image segmented and its correspondent ideal one (Ground truth one).
- d) Compute the segmented rates.
- e) Automatically select the feature set and k parameter that result in the best segmentation rate.
- f) Update the selected classification array with the relevant information computed during the learning and adaptive selection stages: window size, feature set and k -nearest neighbor parameter.

Table 1 and Figure 4 show the partial numerical results of the adaptive feature set selection algorithm. Table 2 and Figure 5 show the partial numerical results of the adaptive k -NN parameter selection algorithm. The best feature set and k parameter computed by the automatic selection algorithm for this specific tested database were:

- Feature set: mean and variance;
- k -NN parameter: 1.

3.3. Approach evaluation

Once all the parameters were computed, and the classification array was built, the segmentation stage itself was performed over the 200 images.

A ground-truth strategy was employed to evaluate the accuracy of the proposed approach. The ideal result (ground-truth segmentation) regarding each class (address block, postmarks and stamps) has been generated for each envelope image. A score of segmentation was computed by comparing identical pixels at the same location in the ground-truth images and segmented ones (true positive).

Table 3 gives the average success accuracy considering pixel by pixel for the address block, stamps, rubber stamps and the noise left in the background. Also, a comparison between this results and prior approaches over the same database is shown (2-D histogram approach [11], best solution of the fractal dimension approach [2] and wavelet approach [7]). Success rates achieved for address block, stamps, rubber stamps and noise suggest that the features used in the proposed learning-based approach improves the segmentation results.

Figure 6 shows the learning sample/ideal pair and one envelope segmentation result.

4. CONCLUSION

This paper presented a modified learning-based approach for segmentation of postal envelopes and address block location. Success rates achieved for address block, stamps, rubber stamps and noise suggest that the features used in the proposed approach improves the segmentation results. The improvement of the segmentation rates achieved in this approach relies on the adaptive feature set and k nearest neighbor parameter selection. It must be remarked that the learning and parameter selection stage is performed once for the database and correspondent values are stored in the classification array built. The main advantages of this approach are the easiness to faithfully reproduce the objectives of the user without requiring neither the use of heuristic parameters nor the interaction of a specialist user after the learning process.

References

- [1] Cover, T.M. and Hart P.E. Nearest Neighbor Pattern Classification. *IEEE T. Information Theory*, Vol. IT-13, no. 1, pp. 21-27, 1967.
- [2] Eiterer L.F., Facon, J. and Menoti, D. Segmentation of Envelope Address Blocks through Fractal-Based Approach, *5th Iberoamerican Meeting on Optics (RIO) and 8th Latin American Meeting on Optics, Lasers and Their Applications (OPTILAS), Proceedings SPIE*, Vol. 5622, 2004, pp. 1265-1270.
- [3] Jain, A. and Bhattacharjee, S. Address block location on envelopes using gabor filters. *Pattern Recognition*, Vol. 25, No. 12, 1992, pp. 1459-1477.
- [4] Legal-Ayala, H.A. and Facon, J. Automatic Segmentation of Brain MRI Through Learning by Example, *Proceedings of 11th International Conference on Image Processing (ICIP)*, IEEE Signal Processing Society, Singapore, vol 2, 2004, pp. 917-920.
- [5] Legal-Ayala, H.A. and Facon, J. Comparative Results of Image Segmentation By Learning in Different Types of Images. *Proceedings of ICISP'2003 - IAPR International Conference on Image and Signal Processing*, Agadir, Morocco, v2, 2003, pp 381-388.
- [6] Legal-Ayala, H.A. Uma Abordagem de Segmentação de Imagens por Aprendizagem, *PhD Thesis, PUCPR*, Curitiba, 2004, 151p. In <http://www.ppgia.pucpr.br/teses/TesePPGIa-HoracioLegal-2004.pdf> [Apr, 23, 2007].
- [7] Menoti, D.; Facon, J.; Borges, D.L.; Brito Jr., A. Segmentação de Envelopes Postais para Localização do Bloco Endereço: uma abordagem baseada em seleção de características no espaço Wavelet. *In: Workshop of Thesis and Dissertation in SIBGRAPI 2004*, Anais Eletrônicos do WTD-SIBGRAPI2004, Curitiba, Brasil., 2004. pp. 1-8.
- [8] Palumbo, P., Srihari, S. and Soh J. Postal address block location in real time. *IEEE Computer*, July 1992, pp. 34-42.
- [9] Wolf, M., Niemann, H., and Schmidt. W. Fast address block location on handwritten and machine printed mail-piece images, *ICDAR'1997 - Fourth IEEE-IAPR International Conference on Document Analysis and Recognition*, Ulm, Germany, Aug. 1997, pp.753-757.
- [10] Yeh, P., Antoy, S., Litcher, A. and Rosenfeld, A. Address Location on envelopes. *Pattern Recognition*, Vol. 20, No. 2, 1987, pp. 213-227.
- [11] Yonekura, E. and Facon, J. Postal envelope segmentation by 2-d histogram clustering through watershed transform, *ICDAR 2003 - 7th International Conference on Document Analysis and Recognition*, 2003, pp. 338-342.

Feature-set [mvsc]	Segmentation (% $\mu \pm \sigma$)	Noise (% $\mu \pm \sigma$)
0001	70.01 \pm 15.08	3.52 \pm 2.61
0010	66.84 \pm 12.44	3.11 \pm 2.28
0011	71.25 \pm 14.66	3.09 \pm 1.67
0100	86.10 \pm 12.00	2.23 \pm 1.24
0101	85.04 \pm 12.23	1.94 \pm 1.08
0110	85.35 \pm 12.29	1.74 \pm 1.01
0111	84.37 \pm 12.55	1.76 \pm 1.08
1000	79.99 \pm 11.75	3.91 \pm 2.86
1001	77.86 \pm 12.59	2.17 \pm 1.41
1010	77.19 \pm 13.51	2.09 \pm 1.30
1011	76.09 \pm 14.09	2.03 \pm 1.35
1100	86.50 \pm 12.23	1.74 \pm 1.15
1101	85.37 \pm 12.28	1.47 \pm 0.86
1110	85.28 \pm 12.14	1.48 \pm 0.89
1111	84.03 \pm 12.39	1.55 \pm 1.04

Table 1: Numerical results of the adaptive feature set selection algorithm. A codification [mvsc] was use for mean (m), variance (v), skewness (s) and kurtosis (c), where 1 means “feature on” and 0 means “feature off”.

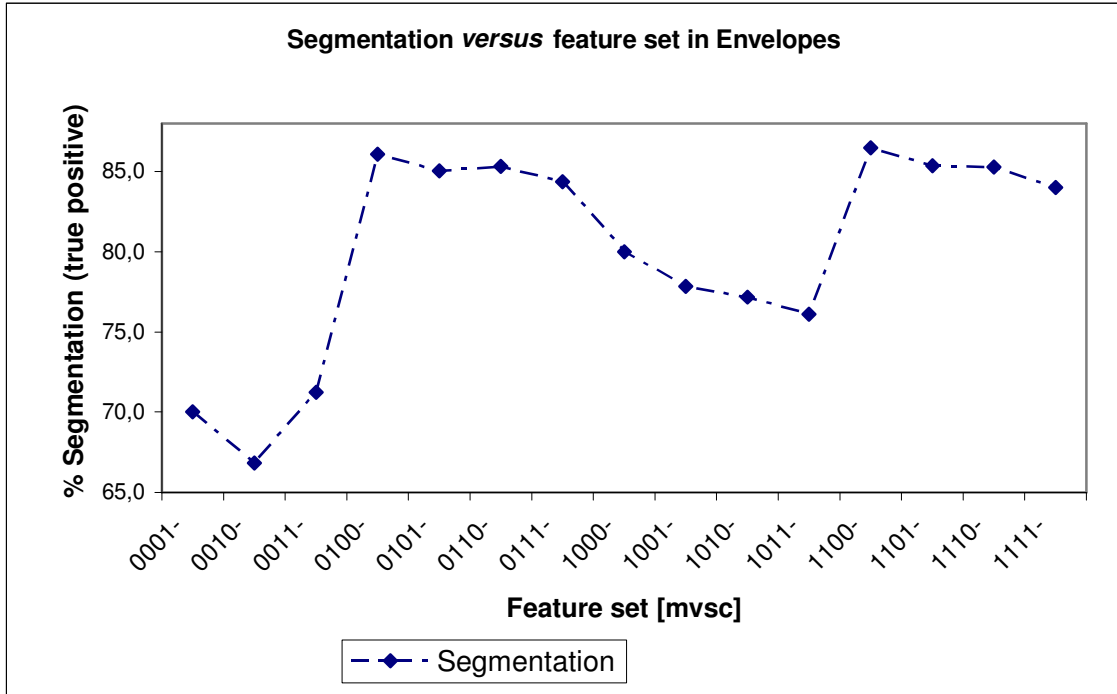


Figure 4: Numerical results of the adaptive feature set selection. A codification [mvsc] was use for mean (m), variance (v), skewness (s) and kurtosis (c), where 1 means “feature on” and 0 means “feature off”.

Feature set: Mean - Variance		
k-NN	Segmentation (% $\mu \pm \sigma$)	Noise (% $\mu \pm \sigma$)
1	80.31 \pm 14.3	1.86 \pm 1.68
3	78.26 \pm 11.6	1.44 \pm 1.61
5	74.18 \pm 11.9	1.21 \pm 1.58
7	71.04 \pm 12.2	1.17 \pm 1.73
9	69.39 \pm 12.2	1.25 \pm 1.95
11	68.70 \pm 12.1	1.41 \pm 2.33
13	68.49 \pm 12.1	1.60 \pm 2.86
15	68.40 \pm 12.0	1.78 \pm 3.46

Table 2: Numerical results of the adaptive k -NN parameter selection algorithm

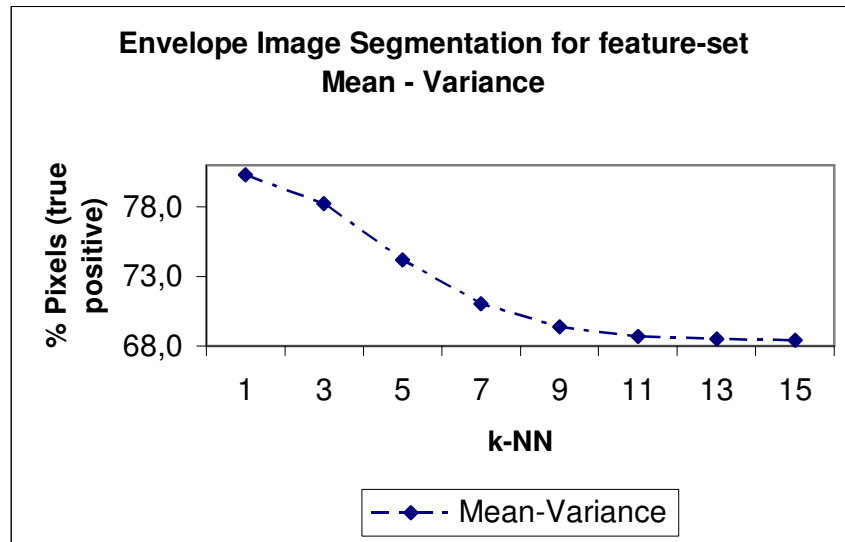
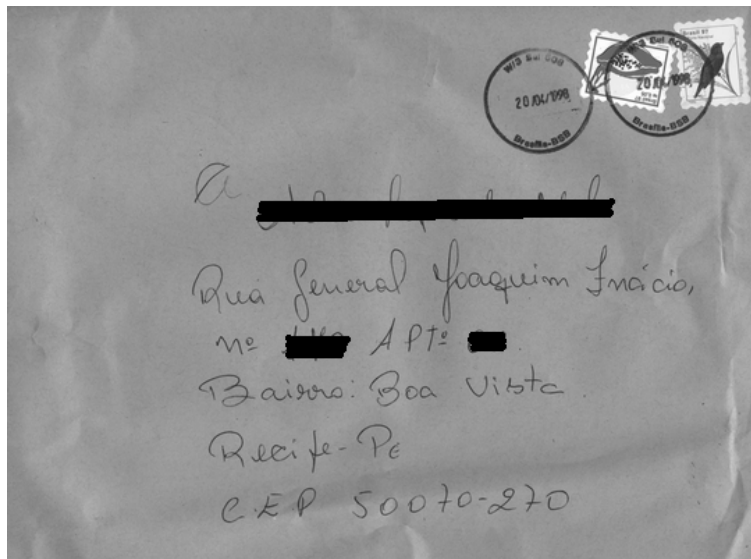


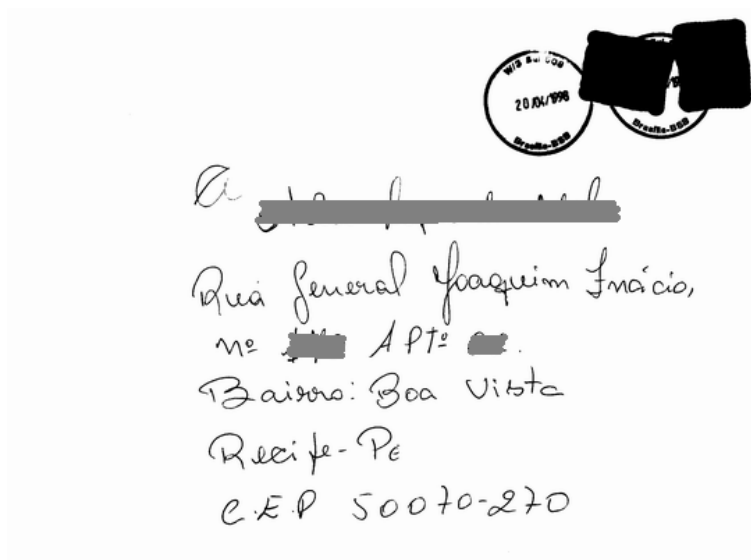
Figure 5: Numerical results of the adaptive k -NN parameter selection algorithm

Region class	Accuracy pixel by pixel (% $\mu \pm \sigma$)			
	Learning-based approach	2-D histogram approach	Fractal dimension approach	Wavelet approach
Address block	98.62 \pm 06.91	75	97.24 \pm 13.64	97.13 \pm 05.86
Rubber stamp	95.36 \pm 13.64		91.89 \pm 17.22	88.12 \pm 19.36
Stamp	87.04 \pm 19.67	25	66.34 \pm 16.90	25.44 \pm 19.29
Noise	02.06 \pm 01.32	N/A	N/A	00.58 \pm 00.88

Table 3: Average results with identification of regions (pixel by pixel accuracy) for the 200 envelope images tested and comparison with prior approaches.



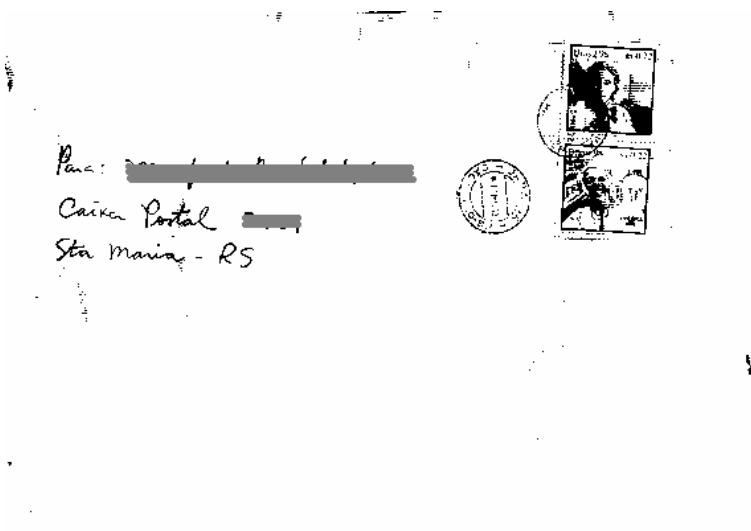
(a) Gray level image sample



(b) Ideal image



(c) New gray level image



(d) Segmented image by learning approach

Figure 6: Learning sample/ideal pair and one envelope segmentation result
(Addresses and names were hidden on purpose)