Deformable Template Matching within a Bayesian Framework for Hand-Written Graphic Symbol Recognition

Ernest Valveny and Enric Martí

Computer Vision Center – Computer Science Department,
Universitat Autònoma de Barcelona
Edifici O - Campus UAB. 08193 Bellaterra. Spain
{ernest,enric}@cvc.uab.es

Abstract. We describe a method for hand-drawn symbol recognition based on deformable template matching able to handle uncertainty and imprecision inherent to hand-drawing. Symbols are represented as a set of straight lines and their deformations as geometric transformations of these lines. Matching, however, is done over the original binary image to avoid loss of information during line detection. It is defined as an energy minimization problem, using a Bayesian framework which allows to combine fidelity to ideal shape of the symbol and flexibility to modify the symbol in order to get the best fit to the binary input image. Prior to matching, we find the best global transformation of the symbol to start the recognition process, based on the distance between symbol lines and image lines. We have applied this method to the recognition of dimensions and symbols in architectural floor plans and we show its flexibility to recognize distorted symbols.

1 Introduction

Bayesian inference and deformable templates have been widely used in many fields of computer vision to reason with uncertainty when prior information about possible values of parameters to be estimated is available. Their application ranges a wide number of computer vision tasks such as object recognition, segmentation, tracking, restoration, etc. [4]. In document analysis, however, and to our knowledge, their use has been restricted to a few applications to hand-written numeral and character recognition [1,5,12].

We argue that a Bayesian framework is also a well-suited method to recognize hand-drawn graphic symbols, such as those found in many kinds of diagrams, maps and line drawings. Hand-drawn symbols are imprecise, with very distorted shapes from their ideal patterns, as it is shown in Fig. 1. Therefore, their recognition must face a high degree of uncertainty. Traditional methods for symbol recognition are generally based on vectorization, feature extraction and structural matching [3,6,8,9,11]. They decrease their efficiency and robustness as long as noise and distortion of hand-drawn symbols increase [2,13] because structural matching cannot recover from feature misdetections and errors introduced

Atul K. Chhabra and D. Dori (Eds.): GREC'99, LNCS 1941, pp. 193–208, 2000.
© Springer-Verlag Berlin Heidelberg 2000

in feature extraction. Bayesian inference can help to overcome the drawbacks of these methods modelling uncertainty though the combination of prior information and likelihood. Prior information can be easily encoded in symbol recognition through the representation of the symbol with a pattern of its ideal shape and the generation of all its possible deformations from this pattern. Then, prior information provides the degree of fidelity of each deformation to the ideal shape of the symbol. On the other hand, likelihood can be seen as a measure of similarity between a given deformation of the symbol and the image. Combining both concepts we can look for the less deformed shape of the symbol that yields the best fit to the image. Therefore, deformable template matching and Bayesian inference arise as an alternative approach to symbol recognition in front of traditional methods. In a previous work [14] we have proposed to use them to recognize hand-drawn symbols in graphic documents. In this work, we extend and further develop our initial proposal and we face the important issue of matching initialization. We also show more extended results that reinforce the feasibility of the application of this approach to the recognition of hand-drawn symbols.

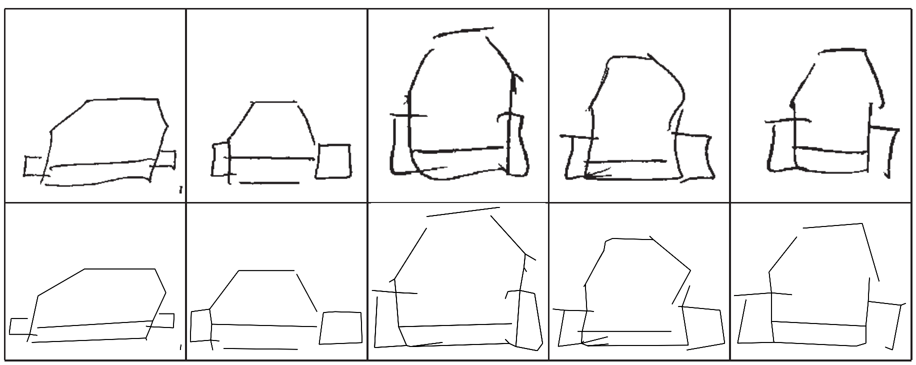

Fig. 1. Some examples of images of symbol *Sofa*. Original image in the top and vectorization in the bottom

In section 2 we explain the general framework of Bayesian inference applied to symbol recognition. In section 3 we describe the application of this general framework to the recognition of hand-drawn lineal symbols. In section 4 we discuss the problem of finding a good initialization of the symbol to get an accurate convergence of the matching algorithm. Section 5 shows some of the experiments carried out and, finally, in section 6, we state the conclusions from our work.

2 Bayesian Formulation of Symbol Recognition

The general problem of symbol recognition can be stated in this way: given an input image I and a set of predefined symbols, $\{S_1, \ldots, S_n\}$, represented by their ideal shapes, symbol recognition provides the symbol S_i that can be best identified in image I.

In a probabilistic framework the correspondence between an image and a symbol can be expressed as $P(S_i|I)$, i.e., the probability that given image I, we can identify symbol S_i in it. Then, symbol recognition consists in finding the symbol S_i which maximises the conditional probability $P(S_i|I)$. Applying Bayes' rule, we can express $P(S_i|I)$ in the following way:

$$P(S_i|I) = \frac{P(I|S_i)P(S_i)}{P(I)} \ . \tag{1}$$

Assuming that all symbols have the same prior probability, $P(S_i)$ and verifying that $P(I)$ is constant for all symbols, we can deduce that finding S_i which maximizes $P(S_i|I)$ is equivalent to finding S_i which maximizes $P(I|S_i)$.

As a hand-drawn symbol can take many different shapes, we must search over all its possible and valid shape variations when looking for the correspondence between an image and the symbol. We let D_i be any possible deformation of the ideal shape of symbol S_i. Then, $P(I|S_i)$ can be expressed as the marginal probability summing up for all deformations D_i the joint probability of the image and each of the deformations, $P(I, D_i|S_i)$:

$$P(I|S_i) = \int P(I, D_i|S_i)dD_i = \int P(I|D_i, S_i)P(D_i|S_i)dD_i \ . \tag{2}$$

Then, the probability to be maximized, $P(I|S_i)$, is expressed as the combination of the prior probability of the deformations, $P(D_i|S_i)$, and the likelihood between the image and each of the deformations, $P(I|D_i, S_i)$. Prior probability is the probability that deformation D_i is still a valid representation of symbol S_i. It penalizes excessive distortions by giving them lower probability. Likelihood is the probability that image I corresponds to deformation D_i. It is usually measured by computing the distance between the image and the deformation.

Expression (2) is solved using Laplacian approximation yielding the following expression:

$$P(I|S_i) = k \cdot P(I|\hat{D}_i, S_i) \cdot P(\hat{D}_i|S_i) \ . \tag{3}$$

where k is a constant and $\hat{D}_i$ is the deformation of the symbol that maximizes $P(I|D_i, S_i) \cdot P(D_i|S_i)$. Usually $\hat{D}_i$ is found searching for the minimum of the negative log of this expression:

$$\begin{aligned}\hat{D}_i &= \arg\max_{D_i} P(I|D_i, S_i)P(D_i|S_i) \\ &= \arg\min_{D_i} \left(-\log P(I|D_i, S_i) - \log P(D_i|S_i)\right) \ . \end{aligned} \tag{4}$$

Making the following equivalences:

$$E_{ext} = -\log P(I|D_i, S_i) \ . \quad (5)$$

$$E_{int} = -\log P(D_i|S_i) \ . \quad (6)$$

$$E = E_{ext} + E_{int} = -\log P(I|S_i) \ . \quad (7)$$

the problem of symbol recognition is reduced to the problem of minimizing an energy function, E, composed of two terms: external energy, E_{ext}, which is related to likelihood and internal energy, E_{int}, which is related to prior probability. External energy plays the role of a force which tries to deform the shape of the symbol as much as possible to get the best match to the input image. On the other hand, internal energy acts like a force that prevents high deformations keeping the shape of the symbol as close as possible to the ideal shape. The minimum of the energy function is the equilibrium point between these two opposite forces. It corresponds to the shape of the symbol that best fits the image with the minimum amount of deformation. The final value of the energy function at this point is related by Eq.(7) to the maximum of $P(I|S_i)$ and by Eq.(1) to the maximum of $P(S_i|I)$. Thus, it is a measure of the degree of correspondence between the input image and the symbol. Then, image can be identified with the symbol with the lowest final energy value.

3 Deformable Template Matching for Lineal Symbols

In this section we describe the application of the general framework introduced in the previous section to the recognition of lineal symbols. Using that framework, there are three main components involved in a deformable template matching approach which need to be defined: first, prior information, i.e., the representation for the symbol and deformations and the definition of the prior probability; secondly, the likelihood between the image and deformations and finally, the matching procedure used to find the minimum of the global energy function.

3.1 Prior Information

As symbols that can be found in drawings are basically composed of lines, we define each symbol as a set of straight lines, not necessarily connected. Each line is represented by the position of its midpoint, its orientation and its length. Deformations of the symbol are generated by translating, rotating and scaling each of the lines. This is an intuitive and natural way to represent symbols and their variations and it yields a set of shapes very close to those produced by handwriting.

We define two kinds of deformations that can be applied to a symbol: global and local deformations. Global deformations apply the same transformation to all the lines of the symbol. Therefore, they do not change the global shape of the symbol. On the other hand, local deformations apply different changes in position, orientation and length to each of the lines. Thus, they can be used to change the global shape of the symbol. Prior probability has only to penalize local deformations while global deformations are assumed to have the same probability.

Prior probability and internal energy are defined assuming that each possible transformation (translation, rotation or scaling) of a line follows a gaussian distribution of zero mean and that all transformations applied to each line are independent. These assumptions allow to express prior probability and internal energy in this way:

$$P(D_i|S_i) = \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}\sigma_{t_{x_i}}} e^{-\frac{t_{x_i}^2}{2\sigma_{t_{x_i}}^2}} \frac{1}{\sqrt{2\pi}\sigma_{t_{y_i}}} e^{-\frac{t_{y_i}^2}{2\sigma_{t_{y_i}}^2}} \frac{1}{\sqrt{2\pi}\sigma_{\theta_i}} e^{-\frac{\sin^2\theta_i}{2\sigma_{\theta_i}^2}} \frac{1}{\sqrt{2\pi}\sigma_{s_i}} e^{-\frac{s_i^2}{2\sigma_{s_i}^2}} \right) . \quad (8)$$

$$E_{int} = \sum_{i=1}^{n} \left(\frac{t_{x_i}^2}{2\sigma_{t_{x_i}}^2} + \frac{t_{y_i}^2}{2\sigma_{t_{y_i}}^2} + \frac{\sin^2\theta_i}{2\sigma_{\theta_i}^2} + \frac{s_i^2}{2\sigma_{s_i}^2} \right) + K . \quad (9)$$

Internal energy is derived, as it is shown in expression (6), from the negative log of prior probability, which is defined as a product of gaussian distributions. There is one distinct gaussian for each kind of transformation applied to a specific line. Each line has its own distribution for translation, rotation and scaling. n is the number of symbol lines; t_{x_i}, t_{y_i} are the translations in x and y directions applied to the midpoint of line i; θ_i is the change in orientation applied to the line; and s_i the scaling applied to the length of the line. ; $\sigma_{t_{x_i}}$, $\sigma_{t_{y_i}}$, σ_{θ_i} and σ_{s_i} are the standard deviations for each of the gaussian distributions.

3.2 Likelihood

Likelihood must be a measure of similarity between a given deformation of the symbol and the input image. We have defined it from the distance between the deformation and the pixels of the binary input image. Working over the binary image allows to avoid errors and loss of information induced by skeletonization and vectorization and illustrated in Fig. 1.

The distance function is defined by summing up for all the lines of the deformation, the distance between the line and the closest pixels of the image. For each line of the deformation, we take a regular sample of points along it and we find, for each point, the closest image pixel. Each pair composed of a line point and an image pixel contributes to the distance with two values: the first one is based on the distance between them and the second one on the difference between the orientation of the line and the orientation of the pixel. This orientation is computed through the analysis of a window centered on the pixel. The proportion of points along the line overlapping with inked image pixels is also taken into account. In this way, we promote deformations of the symbol having lines close to the image, with the same orientation of image pixels and overlapping with them.

Then, external energy is made equivalent to this distance function. Likelihood is defined from external energy using a Gibbs distribution:

$$P(I|D_i, S_i) = \frac{1}{Z} e^{-E_{ext}(I,D)} \ . \tag{10}$$

$$E_{ext} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{n_i} \sum_{j=1}^{n_i} \left(1 - e^{-\lambda \cdot d(p_j, q_j)} \cdot e^{-\mu |\sin(\theta_i - \alpha_j)|} \right) + \gamma \frac{n_i - \hat{n_i}}{n_i} \right) \ . \tag{11}$$

where Z is a normalizing constant; n is the number of lines in the symbol; n_i is the number of points sampled along line i; p_j are each of the points sampled along the line; q_j is the image pixel closest to point p_j; θ_i is the orientation of line i; α_j is the orientation of pixel q_j; $\hat{n_i}$ is the number of points of line i overlapping with image pixels; and λ, μ and γ are weighting factors. The use of exponential functions in the factors related to the distance and the difference of orientation allows to control (changing the value of weighting factors λ and μ) how the distance increases when the lines in the deformation moves away from the image pixels, both in distance and in orientation.

3.3 Matching

Matching is defined as a procedure for finding the minimum of the global energy function. It corresponds to the deformation of the symbol that keeps the best compromise between minimum deformation from ideal shape and maximum fit to the input image. The value of the energy function for this deformation gives a measure of the correspondence between the symbol and the image. Therefore, symbol recognition is performed applying matching between the image and all the symbols and selecting the symbol yielding the lowest energy value.

Combination of internal and external energy is a complex, non-linear energy function, with many local minima. Complex algorithms must be used to find a solution close to the global minimum. We have employed a simulated annealing algorithm [7]. This is a well-known general optimization algorithm that allows to avoid local minima searching randomly over the space of parameters. As the algorithm runs, this random search is directed towards low energy areas. In this way, and if good parameter initialization is achieved, convergence to a point close to the global minimum can be reached. The main drawbacks are the unstability of the solution and the computational cost. Due to the random nature of the algorithm, we cannot guarantee that two distinct runs of the algorithm yield exactly the same solution. It depends very much on a good initialization and on a good setting of all the parameters involved in algorithm performance. On the other hand, many iterations are needed to guarantee an stable and robust convergence. Then, computational cost tends to be high. Currently, our attention is not focussed on computational issues but in showing the feasibility of bayesian deformable template matching to recognize symbols with high distortions. Further considerations can be found in sections 5 and 6 where alternative solutions to these problems are discussed.

The matching algorithm starts from an initial representation of the symbol and, at each step, it randomly generates a new deformation by applying local transformations to every line. This new deformation is accepted or rejected depending on its energy value. If energy is lower than energy of previous deformation, it is always accepted because we are moving towards the global minimum. If energy increases, it can also be accepted in order to jump over areas of local minima. Its acceptance depends on the change in energy and on the value of a temperature parameter. First, the following expression is evaluated:

$$e^{-\frac{E_{k+1}-E_k}{T_k}} \; . \tag{12}$$

where E_{k+1} is the value of energy for the deformation at step $k+1$, E_k is the energy for the deformation at step k, and T_k is the temperature at step k. Then, a random number u between 0 and 1 is generated. The new deformation is accepted only if the computed value in (12) is lower than u. In this way, at the beginning of the algorithm, with high temperatures almost every new state is accepted and we can randomly move over all the space of deformations. As the algorithm runs and we get focussed to areas close to the global minimum, the temperature decreases and the probability of accepting deformations moving towards higher energy values is lower.

Accurate convergence of the algorithm depends very much on two factors: first, the initial value of temperature and its decreasing rate; and secondly, the election of a good starting representation of the symbol. In the next section, we explain how this latter issue is achieved by finding the best global transformation of the symbol that best fits to the image.

4 Initialization of Matching

The goal of this step is to find the global deformation of the symbol closest to the image. This is the optimal point to start the matching between the symbol and the image. Using only global transformations we can get the best possible fit to the image with no change in the global shape of the symbol. Then, matching starts from a shape closer to the image. As a result of that, convergence is easier and smaller changes have to be applied to symbol lines.

This initialization step is performed on the lineal representation of both the image and the symbol in order to speed up the process. We have defined a metric to measure the distance between two lines and the distance between two lineal symbols. Then, initialization consists in finding the global deformation of the symbol yielding the lowest distance to the lineal representation of the image.

First, we will introduce the distance between two lineal symbols and then, we will describe the procedure for finding the best global deformation.

4.1 Distance between Lineal Symbols

A line is defined by the position of its midpoint, its orientation and its length. These three features allow to specify any line in a very natural and intuitive

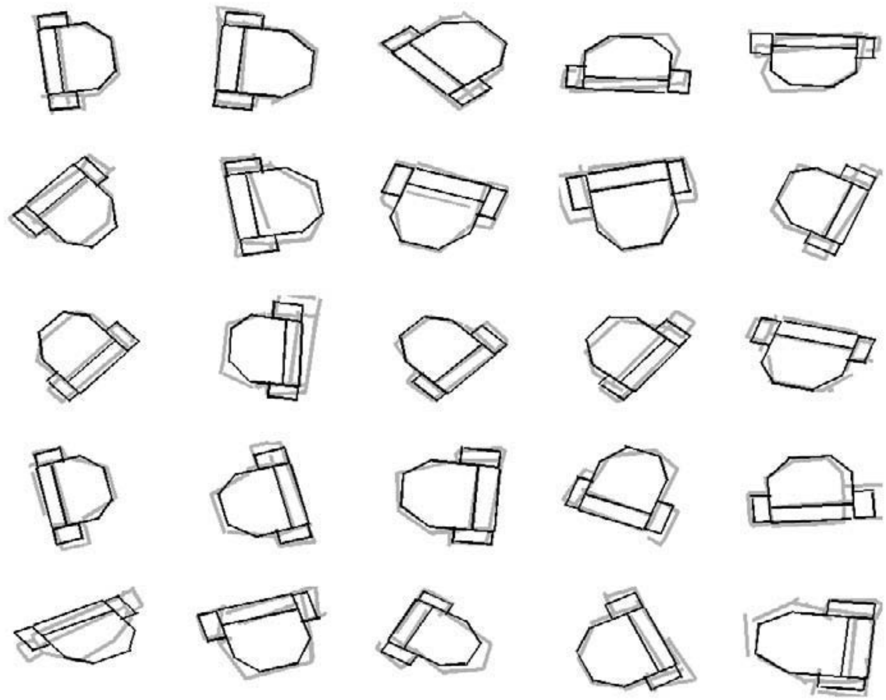

Fig. 2. Initialization of matching for 25 images of symbol *sofa*. Original image in gray and best global deformation in black

fashion. Then, the distance between two lines is based on the difference in midpoint position, orientation and length between them. A line is represented by a vector $L = (P, \alpha, l)$ where P is the position of the midpoint, α is the orientation, and l is the length. Given any two lines $L_1 = (P_1, \alpha_1, l_1)$ and $L_2 = (P_2, \alpha_2, l_2)$, the distance between them is expressed by:

$$d^2(L_1, L_2) = \omega_1 \cdot \|P_1 - P_2\|^2 + \omega_2 \cdot \sin^2(\alpha_1 - \alpha_2) + \omega_3 \cdot \frac{(l_1 - l_2)^2}{(l_1 + l_2)^2} \,. \qquad (13)$$

This measure satisfies the properties of a metric. Moreover, it yields results very close to our visual idea of similarity between lines. The distance increases as lines become more separate and more distinct in orientation and length, factors which contribute to produce more visually distinct lines. Finally, it is an easily computable function, which makes it suitable to derive other measures from it.

The distance between two lineal symbols is then deduced from this definition of distance between two lines. It is defined as the weighted sum of distances between each pair of lines in both symbols. The weighting factor for each pair of lines must be an estimation of the correspondence between the two lines. A priori, we cannot know which line of one symbol is the corresponding line of every line of the other symbol. We estimate this correspondence with a probability distribution based on the distance between the two lines. Closer lines will have higher probability of correspondence. These criteria can be expressed in the following way: given two lineal symbols S_1 and S_2, each of them composed of a set of lines, $S_1 = \{L_1, \ldots, L_{n_1}\}$ and $S_2 = \{L_1, \ldots, L_{n_2}\}$, the distance between the two symbols is defined by:

$$d(S_1, S_2) = \frac{1}{n_1} \sum_{L_i \in S_1} d(L_i, S_2) = \frac{1}{n_1} \sum_{L_i \in S_1} \sum_{L_j \in S_2} P_{ij} \cdot d(L_i, L_j) \ . \tag{14}$$

$$P_{ij} = \frac{e^{-\frac{d_{ij}^2}{2\sigma^2}}}{\sum_{k=1}^{n_2} e^{-\frac{d_{ik}^2}{2\sigma^2}}} \ . \tag{15}$$

d_{ij} is the distance between line L_i in S_1 and line L_j in S_2. P_{ij} is the weighting factor for the distance between L_i and L_j and it corresponds to the probability of correspondence between both lines assuming a normal distribution based on the distance between them. It is normalized so that global probability of correspondence of a line in S_1 sums up to 1 for all lines in S_2.

4.2 Finding the Best Global Deformation of Symbol

First, the input image must be vectorized to get its representation as a set of lines. Then, initialization will find the global deformation of the symbol closest to this set of lines. This is achieved looking for the minimum of the distance function between the image and each of possible global deformations. This optimal deformation is identified by the parameters of a global translation, rotation and scaling applied to all the lines of the symbol. Finding these parameters is not a straightforward task as modifying the lines of the symbol also implies modifying the probability of correspondence among symbol lines and image lines.

We have employed an implementation of the *EM* algorithm [10] to get the optimal deformation. With the *EM* algorithm we can fix the probability of correspondence in the expectation step and then, in the maximization step we can find the optimal global transformation with that estimation of the correspondence. These two steps are iterated until convergence is reached.

The algorithm starts from the ideal representation of the symbol and the lineal representation of the input image and it iteratively finds successive deformations of the symbol until the optimal global deformation is reached. At each iteration two steps are applied. The expectation step estimates the probability of correspondence, P_{ij} between each line of the image and each of the lines of the current deformation. The maximization step finds a new global deformation using these estimated probabilities. The parameters of this new deformation are found analytically deriving the expression for the distance between two lineal symbols (14) in two steps: first, the best rotation is found and secondly, best translation and scaling. The variance used in computing P_{ij} is decreased at each iteration. In this way, as the algorithm runs and the deformation is getting closer to the image, correspondence among similar lines is favoured.

Results of the application of this procedure to a set of test images can be seen in Fig. 2. In it we can see 25 images of a symbol, and superimposed in black, the starting initialization of it for every image. It can be seen that this initialization reflects orientation and scaling of the image. In this way, it will be easier to find the best fit when introducing local deformations of the symbol and applying the matching procedure described in section 3.3.

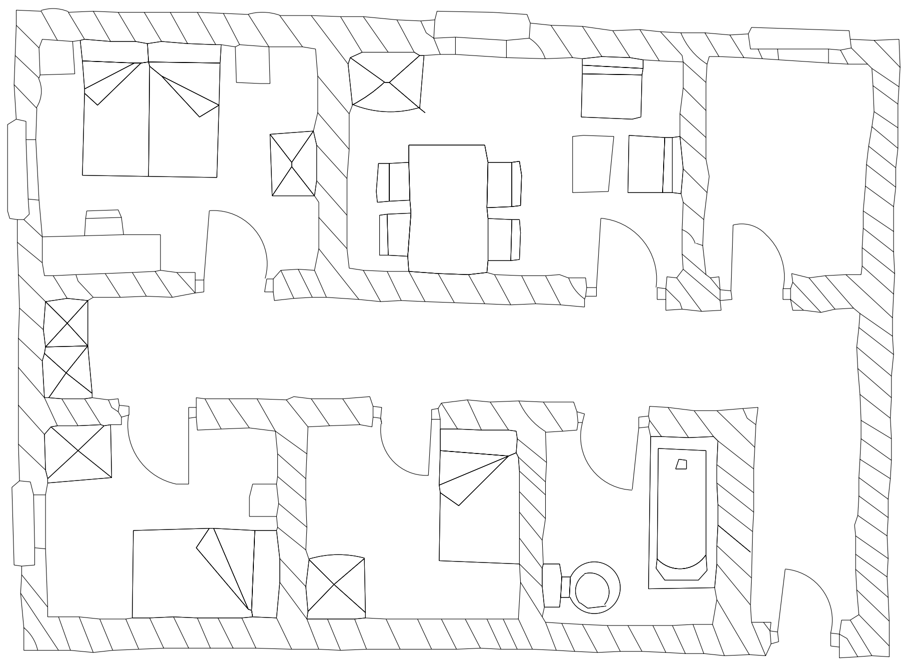

Fig. 3. Example of a hand-drawn architectural drawing

5 Results and Discussion

We have applied this method to the recognition of symbols in hand-drawn architectural drawings. Fig. 3 shows an example of this kind of drawings with the symbols to be recognized. In these drawings, the identification and the recognition of all the symbols play a very important role in the semantic analysis. We

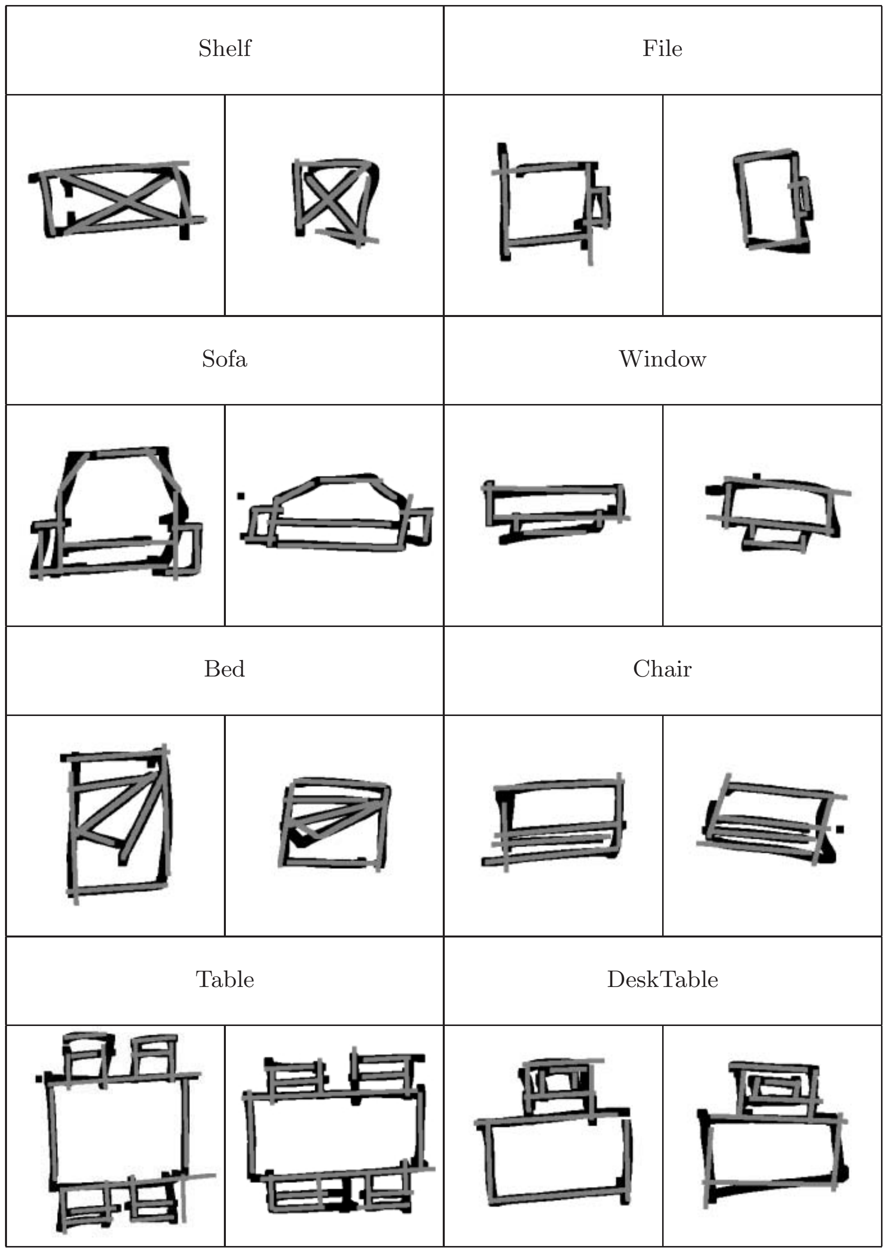

Fig. 4. Visual matching of two images of each symbol. Original image in black and final deformation in gray

have focused on the problem of being able to recognize isolated symbols. We assume that symbols have been located and segmented. Thus, our method relies on a previous segmentation step. This is an issue to be further studied and developed. We have worked with a set of symbols drawn with no constraints by ten different people. These symbols have a wide range of variations and distortions in their shape. Results show that most of the deformations can be handled and that symbols are accurately recognized.

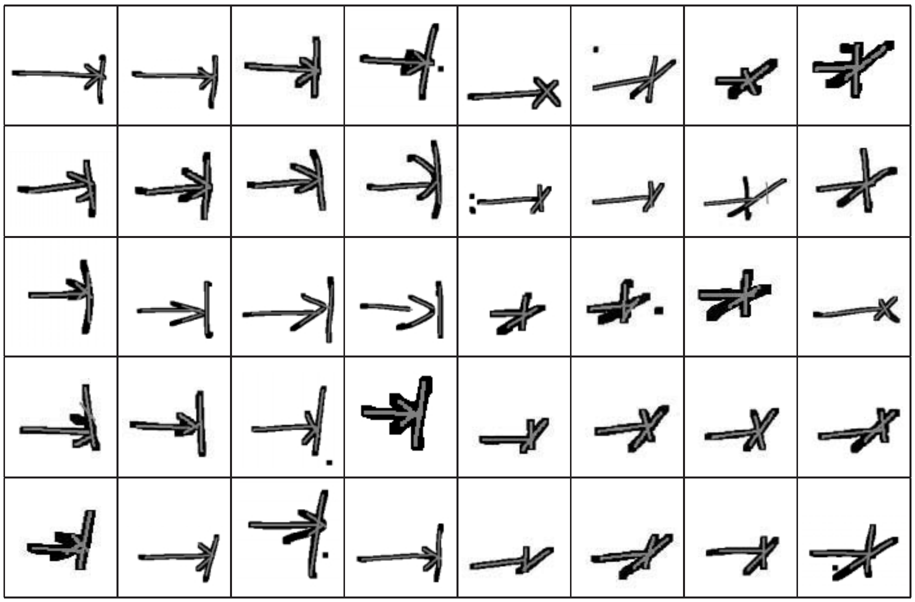

Fig. 5. Visual matching of dimension symbols.Original image in black and final deformation in gray

In Fig. 4 we can see the visual result of the application of matching to two images of each symbol. Original image is shown in black while final deformation of the symbol found by the matching procedure is superimposed in gray over the image. It can be seen how the original shape of the symbol is deformed in order to fit the shape in the image, and how many different kinds of distortions can be handled, such as: changes in orientation or length of the lines, variations in relative position between the lines, spurious lines, non-touching lines at crossings, etc. Before applying matching, we have applied the initialization step to find the best global transformation of each symbol. Fig. 5 shows similar results for the recognition of dimension symbols in architectural drawings. Once again, in most cases, the ideal shape of the symbol is deformed to fit the input image. Only in some few cases, symbol cannot be deformed to adjust it to the input image. These errors can be due to a bad initialization of the symbol and the algorithm or to excessive distortions in the image.

Visual matching is an indicator of the goodness of the fit between the image and the symbol. However, the identification of an image with a symbol is done, as stated in section 2, by analyzing the minimum value of energy after the matching procedure. The graphic in Fig. 6 shows the final energy value found by the matching procedure after comparing 50 images of symbol *sofa* with each of the eight symbols taken into account. The wider line corresponds to values of matching each image with the symbol *sofa*, while thinner lines show the values of matching with the other symbols. It can be seen how, in almost all cases, energy of matching with symbol *sofa* corresponds to the minimum value. The graphic also illustrates that when the energy of matching with symbol *sofa* is too high (due to errors of matching), the image can be confused with some other symbol with lower energy.

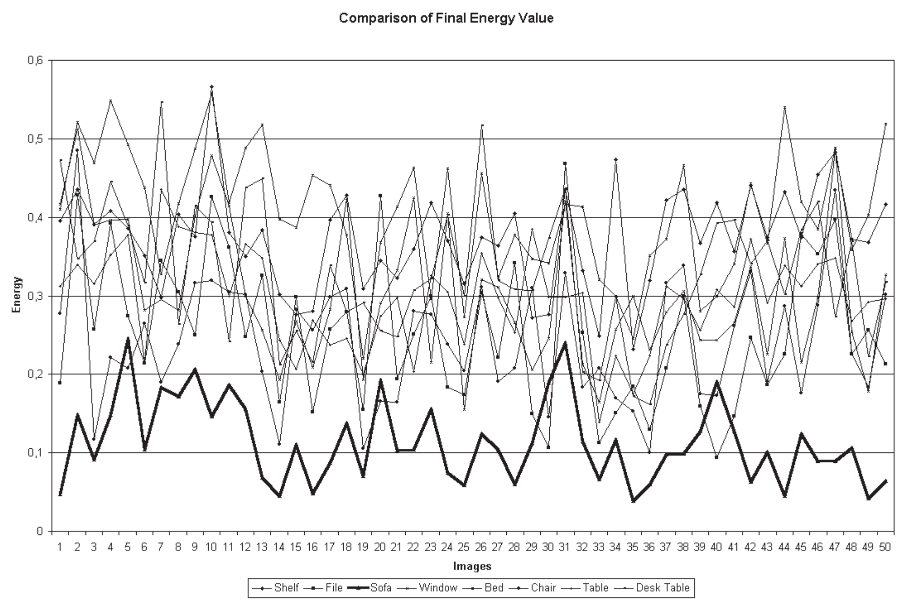

Fig. 6. Comparison of energy for matching of 50 images of a *sofa* with all symbols. The widest line is energy for matching with symbol *sofa*

Table 1 shows the recognition rates achieved with the application of this criterion. We have matched 50 images of each symbol with the model of each of eight symbols, applying first the initialization step and we have identified each image with the symbol with lower minimum energy. The table shows, for each symbol, the percentage of images correctly classified. We have got 85.25% of average accuracy for all symbols. We can see how, in some symbols, as in symbol *window*, the recognition rate is much lower. This fact is due to the confusion of this symbol with two other similar symbols: symbol *file* and symbol *chair*. Confusions are due to errors in matching (bad initialization or excessive distortion) and to the ability of confused symbols (*file* and *chair*) to deform yielding lower energy values.

Finally, Table 2 illustrates the computation time of the algorithm. For each symbol, it shows the average time (in seconds) of matching an image with the symbol. We can see how the complexity is approximately linear with the number of lines in the symbol.

Table 1. Recognition rates for 50 images of each symbol

Shelf	File	Sofa	Window	Bed	Chair	Table	D.Table	**Average**
88%	88%	92%	68%	74%	88%	92%	92%	**85.25%**

Table 2. Average recognition times for each symbol (in seconds)

Symbol	Chair	Shelf	File	Window	Bed	D.Table	Sofa	Table
Time	5,62	7,31	6,04	6,64	9,84	10,79	16,34	28,93
N. of lines	6	6	7	7	8	11	13	20

6 Conclusions and Future Work

We have shown how Bayesian inference and deformable template matching can be applied to the recognition of symbols in graphic documents. This approach is more flexible and able to handle uncertainty inherent to handwriting than methods based on previous vectorization and feature extraction.

Symbols are represented as a set of lines, and their deformations are generated by geometric transformations of these lines which yields very natural distortions, close to those produced by handwriting. Bayesian formulation of matching over the binary image allows to derive an energy function, whose minimization gives the equilibrium point between low deformation and maximum fit. This minimization is not simple. We have employed a simulated annealing algorithm with a previous initialization step to facilitate convergence. The initialization step finds the best global orientation and scaling of the symbol, based on distance between lines of the symbol and lines of the image. Simulated annealing is a random algorithm, so that convergence is not always guaranteed to be stable through successive runs of the algorithm. The complexity of the algorithm is linear with the number of lines in the symbol, although computation time is high.

Our goal was mainly to show the feasibility of bayesian inference and deformable template matching to hand-drawn symbol recognition without taking into account computational issues. However, there are some relevant points to be further studied in order to be able to apply this method to real applications: first, segmentation of symbols in the drawing should be solved in order to locate candidate areas where applying the recognition. Secondly, we are investigating other ways to define internal and external energy yielding an energy function easier to minimize. In this way, we could reduce computation time and we could get more stable convergence. Thirdly, the response of the algorithm of the scalability problem should always be tested with a wider set of symbols and images. Finally, we have assumed independence in the deformations applied to each line; however this assumption is not always true, specially with complex symbols. More accurate representation of prior information about deformations and their associated cost would also allow to improve recognition rates. Generalization to other types of primitives other than straight lines should also be considered.

References

1. K. W. Cheung and D. Y. Yeung. A bayesian framework for deformable pattern recognition with application to handwritten character recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1382–1388, Desember 1998. 193
2. A. K. Chhabra. Graphic symbol recognition: An overview. In K. Tombre and A. K. Chhabra, editors, *Graphics Recognition: Algorithms and Systems*, pages 68–79. Springer Verlag, Berlin, 1998. 193
3. A. H. Habacha. Structural recognition of disturbed symbols using discrete relaxation. In *Proceedings of 1st. International Conference on Document Analysis and Recognition*, pages 170–178, Sep-Oct 1991. Saint Malo, France. 193
4. A. K. Jain, Y. Zhong, and M. P. Dubuisson-Jolly. Deformable template models: A review. *Signal Processing*, 71:109–129, 1998. 193
5. A. K. Jain and D. Zongker. Representation and recognition of handwritten digits using deformable templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(12):1386–1391, Desember 1997. 193
6. P. Kuner and B. Ueberreiter. Knowledge-based pattern recognition in disturbed line images using graph theory, optimization and predicate calculus. In *Proceedings of 8th. International Conference on Pattern Recognition*, pages 240–243, October 1986. Paris, France. 193
7. P. J. M. Van Laarhoven and E. H. Aarts. *Simulated Annealing: Theory and Applications*. Kluwer Academic Publishers, 1989. 198
8. S. Lee. Recognizing hand-written electrical circuit symbols with attributed graph matching. In H. S. Baird, H. Bunke, and K. Yamamoto, editors, *Structured Document Analysis*, pages 340–358. Springer Verlag, Berlin, 1992. 193
9. J. Lladós, J. López-Krahe, and E. Martí. A system to understand hand-drawn floor plans using subgraph isomorphism and Hough transform. *Machine Vision and Applications*, 10(3):150–158, 1997. 193
10. G. J. MacLachlan and T. Krishnan. *The EM algorithm and extensions*. John Wiley and Sons, Inc., 1997. 201

11. B. T. Messmer and H. Bunke. Automatic learning and recognition of graphical symbols in engineering drawings. In R. Kasturi and K. Tombre, editors, *Graphics Recognition: Methods and Applications, Selected Papers from First International Workshop on Graphics Recognition, 1995*, pages 123–134. Springer, Berlin, 1996. Volume 1072 of Lecture Notes in Computer Science. 193
12. M. Revow, C. K. I. Williams, and G. Hinton. Using generative models for handwritten digit recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(6):592–606, June 1996. 193
13. K. Tombre. Graphics recognition - general context and challenges. *Pattern Recognition Letters*, 16(9):883–891, 1995. 193
14. E. Valveny and E. Martí. Application of deformable template matching to symbol recognition in hand-written architectural drawings. In *Fifth IAPR International Conference on Document Analysis and Recognition ICDAR'99*, pages 483–486, Bangalore, India, September 1999. 194