

# A Hybrid Approach to Detect Graphical Symbols in Documents

Salvatore Tabbone, Laurent Wendling, and Daniel Zuwala

LORIA-UMR 7503,  
Campus Scientifique,  
B.P. 239,  
54506 Villers-lès-Nancy Cedex, France  
{tabbone,wendling,dzuwala}@loria.fr

**Abstract.** We propose to combine a feature descriptor method with a structural representation of symbols. An adaptation of the Radon transform, keeping main geometric transformations usually required for the recognition of symbols, is provided. In order to improve the recognition step we directly process on the grey level document. In this perspective, a three-dimensional signature integrates into a same formalism both the shape of the object and its photometric variations. More precisely the signature is computed within the symbol following several grey levels. Additionally a structural representation of symbols allows to localize into the document candidate symbols.

## 1 Introduction

On the early years, the powerful development of the Internet and the large increasing of storage capabilities have given rise to high interests for the symbol recognition problem [1]. Usually documents are scanned in grey levels and afterwards a binarization step is carried out to extract symbols from the background.

A lot of approaches have been proposed during the last years for recognizing symbols. Among these approaches some are based on features descriptors [2–5]. Two kinds of descriptors are often encountered: those that work on an object as a whole and those that work on the contours of the objects. In general such approaches based on feature descriptors are robust against noise and occlusions. Region-based descriptors are less sensitive to noise and contour-based descriptors to occlusions. Nevertheless, an important drawback of all these methods is that complete symbols in a document must be clearly segmented, which is in itself an ill-posed problem. Typically, symbols are quite always embedded with other graphics layers. Moreover they can touch, partially overlap each other or intersect other lines.

Other approaches focused on the structural shape representation [6–10]. Among data structures, graphs are usually suitable to support the structural representation of symbols. More generally, a symbol is described into a set of segments and the whole segments and the spatial relations between them are represented by an attributed relational graph. The computational complexity of

graph matching techniques is rather large even though efficient algorithms have been proposed for some classes of graphs. Moreover these approaches suffer from the noise introduced by the segmentation process even if statistical assumptions can be made on that noise [8]. However structural approaches are powerful in terms of its representational capabilities by spotting target symbols.

In this paper, we propose to combine a feature descriptor method with a structural representation of symbols. Documents are often perturbed by noise (digitalization, document quality...) and a binarization step gives rise to artifacts or some parts of the symbols can be removed following a threshold set. In order to improve the recognition process we directly work on the grey level document. We make here an original adaptation of the Radon transform keeping main geometric transformations usually required for the recognition of symbols. A three-dimensional signature integrates into a same formalism both the shape of the object and its photometric variations. More precisely the signature is computed within the symbol following several grey levels. So noise effects on the boundary of the processed object are insignificant considering its whole shape. Additionally on the document we compute the skeleton obtained by the medial axis [11]. Key points (junction points) on the skeleton are organized into a graph where the graph edges describe the link between junction points. From this representation, candidate symbols are selected and validated with the signature descriptor.

The paper is organized as follows. In section 2 and 3 we recall the definition of the Radon transform and its extension to recognize grey level symbols. Then, we describe our symbol localization process using a structural representation based on a decomposition of the document into its skeleton (section 4). Examples and experimental results are given in section 5 and future investigations are provided in section 6.

## 2 Signature of Grey Level Symbols

The Radon transform  $T_{Rf}$  of a function  $f$  is defined as follows [12]:

$$T_{Rf}(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos(\theta) + y \sin(\theta) - \rho) dx dy . \quad (1)$$

where  $-\infty < \rho < \infty, 0 \leq \theta < \pi$  and  $\delta$  is the Dirac function. Intuitively, that remains to integrate the function  $f$  along a line for any parameters  $(\rho, \theta)$

We use here a shape measure, called  $\mathcal{R}$ -signature [13] which is defined from the Radon transform:

$$\mathcal{R}_f(\theta) = \int_{-\infty}^{\infty} T_{Rf}^2(\rho, \theta) d\rho . \quad (2)$$

### 2.1 $\mathcal{R}$ -Signature Properties

It is easy to show, from basic Radon properties, that the following properties are checked for the  $\mathcal{R}$ -signature:

– **Periodicity of  $\pi$ :**

$$\mathcal{R}_f(\theta \pm \pi) = \int_{-\infty}^{\infty} T_{Rf}^2(\rho, \theta \pm \pi) d\rho = \mathcal{R}_f(\theta) . \quad (3)$$

- **Rotation:** a rotation of angle  $\theta_0$  on  $f$  implies on the  $\mathcal{R}$ -signature a cyclical shift depending on  $\theta_0$ :

$$\mathcal{R}_f(\theta + \theta_0) = \int_{-\infty}^{\infty} T_{Rf}^2(\rho, \theta + \theta_0) d\rho . \quad (4)$$

- **Translation:** a  $\mathcal{R}$ -signature is invariant to a translation of  $\mathbf{u} = (x_0, y_0)$  on  $f$ :

$$\int_{-\infty}^{\infty} T_{Rf}^2(\rho - x_0 \cos(\theta) - y_0 \sin(\theta)) d\rho = \mathcal{R}_f(\theta) . \quad (5)$$

- **Scale:** a zoom of  $\alpha \neq 0$  on  $f$  implies that:

$$\int_{-\infty}^{\infty} T_{Rf}^2(\alpha \rho, \theta) d\rho = \frac{1}{\alpha^2} \mathcal{R}_f(\theta), \alpha > 0 . \quad (6)$$

## 2.2 $\mathcal{R}_{3D}$ -Signature Definition

Generally a bidimensional representation of the signature is sufficient to reach accurate results considering only the shape of the object. Nevertheless such an approach is not suitable to recognize grey level symbols. There exists two ways to manage with the photometry during the definition of the signature. First the luminance can be directly integrated onto the scan line. In this case we have noticed a smoothing of the 2D signature which is in fact less discriminate. Another idea is to consider the image as a set of level cuts performed from successive binarizations (for each grey levels contained in the image) and so to define a 3D signature. A simple approach to compute an  $\mathcal{R}_{3D}$ -signature might be to calculate for the signature associated to each cut and to append all of them. Let  $O$  be an object composed of  $\#ng$  grey levels and let  $\lambda_i$  be a binary cut of threshold  $i$  applied on  $O$ . We have  $O = U_{i=1, \#ng} \{\lambda_i\}$ , with  $\lambda_{\#ng} \subseteq \lambda_{\#ng-1} \subseteq \dots \subseteq \lambda_1$ . So,  $\mathcal{R}_{3D} = U\{R_{\lambda_i}\}_{i=1, \#ng}$  with  $R_{\lambda_i}$  the signature calculated on the binary image  $\lambda_i$  performed from a binarization of threshold  $i$  and  $\#ng$  is the number of grey levels of the image. However such successive binarizations rely to the computation of Radon transforms and should be expensive in processing time. The use of an accumulator during the definition of the Radon transform allows to take into account directly the grey levels of the image. A point with a given value appears on all the high level cuts and obviously in each associated signature. It is enough to integrate this information during the calculation of the Radon transform of this point. Then only one Radon transform is required to define an  $\mathcal{R}_{3D}$ -signature. Figure 1 presents the signature achieved on a grey level symbol.

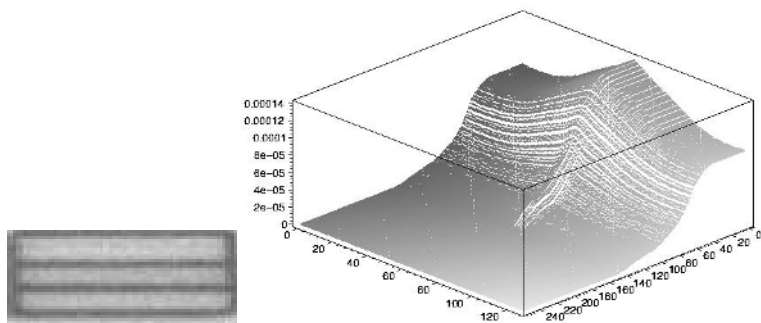


Fig. 1. Example of  $\mathcal{R}$ -signature.

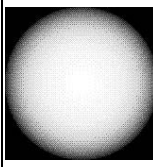
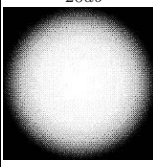
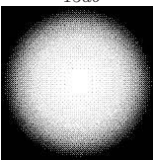
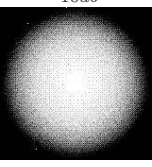
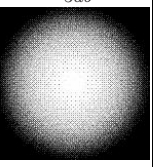
### 2.3 Properties of $\mathcal{R}_{3D}$ -Signatures

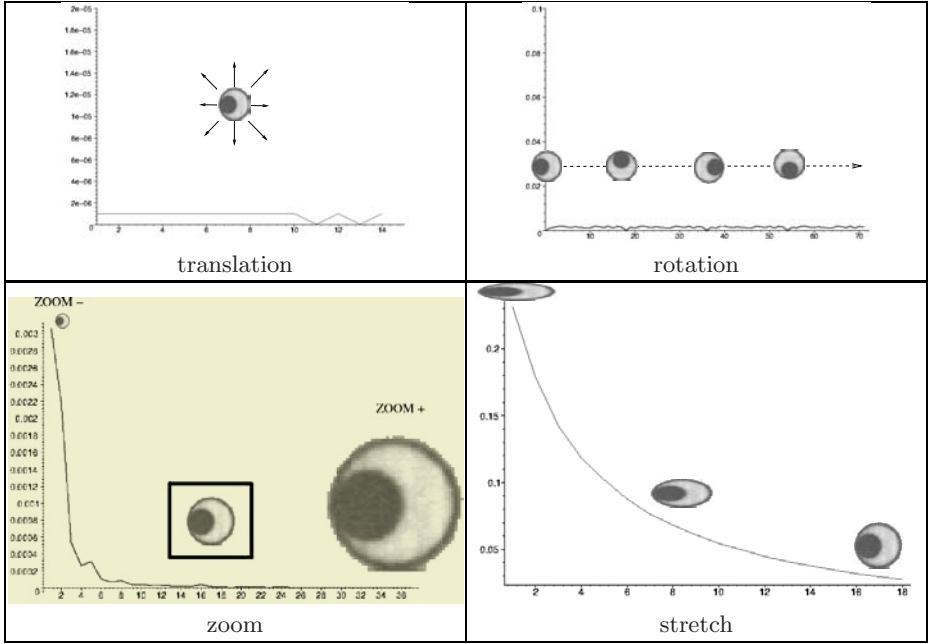
By definition an  $\mathcal{R}_{3D}$ -signature consists in a set of  $\mathcal{R}$ -signatures checking separately the properties described previously (see section 2.1). As the photometry is invariant to the rotation and the translation of the object, the invariance properties of rotation and translation of  $\mathcal{R}$ -signatures are kept for the  $\mathcal{R}_{3D}$ -signature. The scale factor is taken into account by dividing each  $\mathcal{R}_{3D}$ -signature by its volume during the matching process. Figure 2 shows error variations due to the application of basic geometric transforms on a grey level object. The values achieved are close to zero (except for high stretching values and negative zoom).

### 2.4 Noise Effect

Noise has been added to a synthetic image of a sphere. Table 1 provides the difference in term of percentage between the original signature and degraded versions. The low differences attest of the robustness of our method. Few disparities from the global structure of the object have weak influences on the global signature of the object.

**Table 1.** Differences between a sphere and its degraded versions. *max* : maximum error.  $\bar{\mu}$  : mean error. *RMS* : root min square error.

$I$	$I_{20db}$	$I_{15db}$	$I_{10db}$	$I_{5db}$
				
max	0,043	0,054	0,073	0,11
$\bar{\mu}$	0,035	0,044	0,060	0,091
<i>RMS</i>	$6 \cdot 10^{-6}$	$8 \cdot 10^{-6}$	$11 \cdot 10^{-6}$	$18 \cdot 10^{-6}$



**Fig. 2.** Basic geometric transforms. X-axis: Transforms; Y-axis: Error variations. The original symbol is defined into the square area.

## 2.5 Distance Between $\mathcal{R}_{3D}$ -Signatures

Consider two discrete  $\mathcal{R}_{3D}$ -signatures  $\mathcal{R}_{3D}^A$  and  $\mathcal{R}_{3D}^B$ . Let  $\bar{\mathcal{R}}_{3D}^A$  and  $\bar{\mathcal{R}}_{3D}^B$  be the normalized signatures and  $p$  be the number of orientations. The similarity ratio  $SR$  (expressed as a percentage) between  $\mathcal{R}_{3D}^A$  and  $\mathcal{R}_{3D}^B$  is defined as follows:

$$m^{AB}(\theta, x) = \sum_{i=1}^n \min (\bar{\mathcal{R}}_{\lambda_i}^A(\theta), \bar{\mathcal{R}}_{\lambda_i}^B(\theta + x \% p)) . \quad (7)$$

$$M^{AB}(\theta, x) = \sum_{i=1}^n \max (\bar{\mathcal{R}}_{\lambda_i}^A(\theta), \bar{\mathcal{R}}_{\lambda_i}^B(\theta + x \% p)) . \quad (8)$$

$$SR = 100 \max_{x \in [0, p]} \left\{ \frac{\sum_{\theta=0}^p m^{AB}(\theta, x)}{\sum_{\theta=0}^p M^{AB}(\theta, x)} \right\} . \quad (9)$$

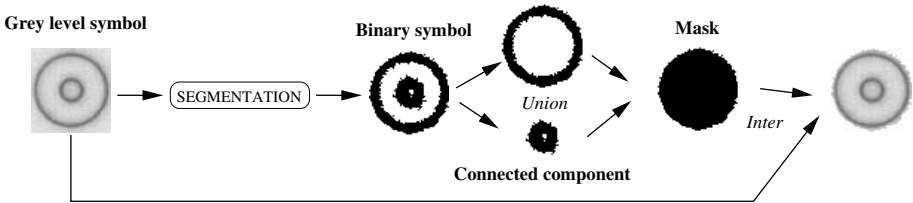
Hence  $SR$  is not sensitive to symbol scaling. Moreover, note that circular shifts  $\theta$  are applied to  $\bar{\mathcal{R}}_{3D}^B$ , and  $SR$  is obtained by maximizing the classical Tanimoto index (min over max). In that way,  $SR$  is not sensitive to object rotation either.

### 3 Symbol Extraction

We consider a set of symbols to be found in a graphical document. Two cases are studied: either the symbols are disconnected from the graphic layer network or linked to the network with only trihedral junctions (order 3).

#### 3.1 Disconnected Symbols

In this section, we focus on disconnected symbols obtained after a rough segmentation of noisy documents. Each grey level symbol is localized using a support (mask in figure 3) defined from its maximal connected component. We consider here components having a consistent size to not study too weak (assimilated to noise) or too large structures (greater than the symbols contained into the database). The selected components are filled to take into account the global shape of the symbol. Figure 3 presents the processing performed to extract one symbol in a line-drawing document.



**Fig. 3.** Grey level symbol extraction.

Finally, the support region is directly applied on the grey level document and the extracted area is matched with each model belonging to the database using the similarity ratio (see section 3). The target symbol corresponds to the maximal similarity ratio.

#### 3.2 Symbols Linked to the Network

We present in this section a primary approach allowing to recognize basic symbols linked to a network. We consider in this application symbols having a regular shape (rectangular or circular for example). The junction points between the object and the network are assessed using common pruning and skeleton extraction steps (junctions of order 3 are processed here). The  $T_{34}$  distance transform [14] is performed on a rough binarization of the grey level document to compute the skeleton. Figure 4 shows a symbol to be found in a document and an assessment of a possible junction scheme for it.

In this case, two pairs of junctions related to a connection to two network lines have been set. A directed graph is then defined for any configuration of junctions. For example if a symbol is linked to one line the graph consists in two

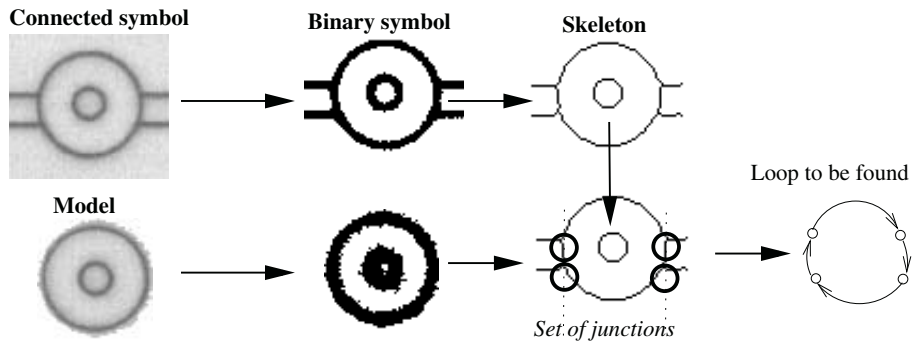


Fig. 4. Localisation of a connected symbol.

nodes, and four nodes for two lines... Such a graph is assumed to be the maximal boundary of the symbol. The retrieval of an object in a document is then guided by the detection of loops in the skeleton image following the orientations of the directed graph associated to an object. Then we search all the possible loops related to such graph. For each loop, the associated graph is extracted from the skeleton image by removing points exterior to the junctions (and by definition exterior to the current loop). It remains to fill the area delimited by such graph in order to define the mask to be applied in the document image. Nonetheless using this approach the symbol should be cropped in the document because it is limited by its exterior skeleton. So in order to grow the mask up to fully consider the symbol, the inverse distance transform is carried out onto the values of the  $T_{34}$  transform associated to the extracted loop. Such process allows to extend the area delimited by the directed graph to the boundary of the object. At last, the same processing described in the previous section are done: filling in the mask area and matching with the database of symbols.

It is easy to extend the proposed approach to the retrieval of broad symbols by considering a directed graph definition for each object. The main idea is to manage with strong graph representations defined following the set of junction points peculiar to the symbol structure. The directed loop is then set from the external hull of the skeleton representation of an object. Let  $x$  be the number of nodes of this loop. Finally, the search of symbols linked to a network by  $k$  junctions remains to the search of loops of cardinality  $x' = x + k$  in the skeleton image.

## 4 Experimental Results

In this section two examples of symbol recognition process are provided. On the one hand we show that the use of grey level information increases the accuracy of the recognition step. On the other hand, we show how a structural description of a symbol improve the symbol spotting.

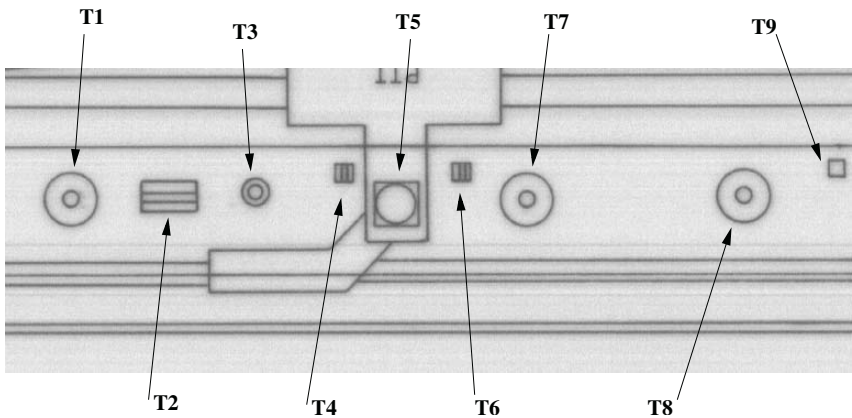


Fig. 5. Example of isolated symbols.





















										
T1	<b>99</b>	82	98	89	90	96	96	97	72	97
T2	82	<b>98</b>	82	83	83	84	83	84	81	83
T3	<b>99</b>	82	<b>99</b>	89	90	96	96	97	72	98
T4	94	86	94	92	92	<b>96</b>	95	94	75	95
T5	90	83	90	97	<b>98</b>	92	90	90	73	90
T6	97	83	97	91	92	<b>99</b>	95	96	73	96
T7	<b>99</b>	82	<b>99</b>	89	90	96	95	97	72	97
T8	<b>99</b>	82	98	89	90	96	95	97	72	97
T9	90	84	90	<b>97</b>	<b>97</b>	92	90	90	74	90

Fig. 6. Symbol ratio (*SR*: formula 9) between binary symbols.

4.1 Disconnected Symbols

A database of ten symbols having relatively close shapes is provided. Figure 6 shows an application based on binary symbols found in figure 5. In this case noise brings artifacts and hence errors to the signature of the object. As a consequence, very close clusters (both circles and squares in figure 6) give rise to miss-classified symbols. That is the scores are high for similar shapes. The use of grey level data allows to overcome such problem by improving the classification (see figure 7). In this case, near clusters are better differentiated.



										
<b>T1</b>	<b>98</b>	60	49	63	69	44	57	78	31	69
<b>T2</b>	62	<b>96</b>	74	84	79	66	70	57	44	83
<b>T3</b>	50	76	<b>97</b>	67	72	83	72	47	51	70
<b>T4</b>	44	67	83	66	65	<b>96</b>	71	43	59	61
<b>T5</b>	74	78	67	84	<b>93</b>	62	75	68	42	88
<b>T6</b>	45	69	86	69	67	<b>97</b>	73	44	58	64
<b>T7</b>	<b>95</b>	58	48	62	68	43	56	80	30	67
<b>T8</b>	<b>95</b>	58	48	63	68	44	57	81	31	67
<b>T9</b>	61	81	74	<b>97</b>	86	68	66	57	45	86

**Fig. 7.** Symbol ratio ( $SR$ : formula 9) between grey level symbols.

4.2 Connected Symbols

The figure 8 presents the main steps performed to extract a symbol linked to a network by two lines in a grey level document: that is, the skeleton computed from a rough segmentation, the areas extracted from the loops associated to the directed graph of the object and the location of the symbol (the scores in black).

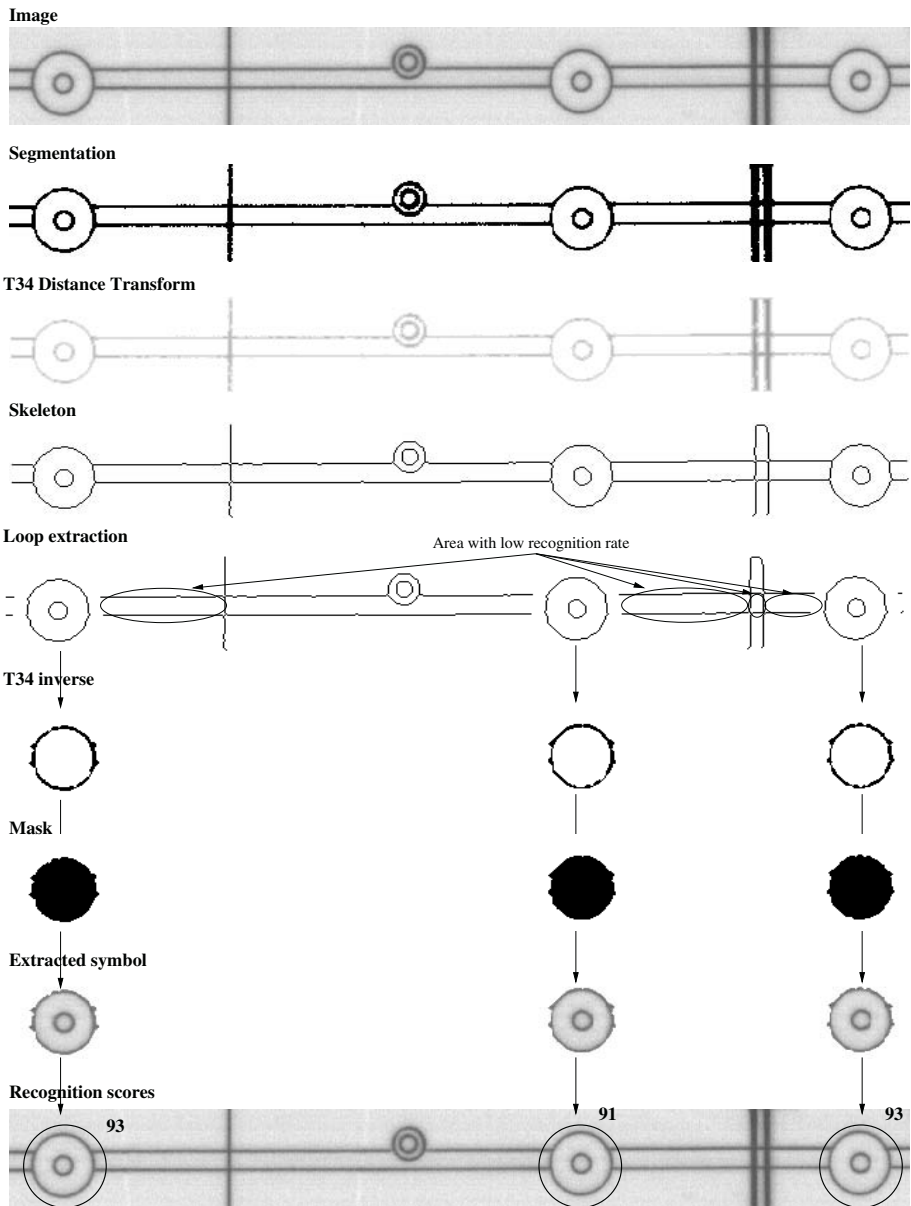
Another symbol extraction is given figure 9. Here a square symbol linked by tree lines is extracted from the network and recognized. In this case the loop consists in an odd number of nodes.

In term of complexity, the most costly is the enumeration of the all the loops. The fastest known algorithm has an upper time bound of  $O((n+e)(c+1))$ , where  $e$  is the number of edges,  $n$  the number of vertices and  $c$  the number of loops. Typically, we don't search for all the loops. We only focus the search of those with a number of edges inferior to about 10. This step is very time expensive, but it can be done offline.

5 Conclusion

The use of grey level information in noisy document allows to better recognize symbols because it limits the loss of accuracy due to the segmentation process. We have also drawn a way to handle with connected symbols in line-drawings using a structural description.

Although this method can only be applied to symbols with convex envelope, these results are very promising. However they still need further validation by



**Fig. 8.** Connected symbol recognition.

processing much larger databases of technical drawings, with more graphical symbols. Currently, we try to extent our method to the retrieval of any symbol by considering a structural representation taking into account partially occluded symbols.

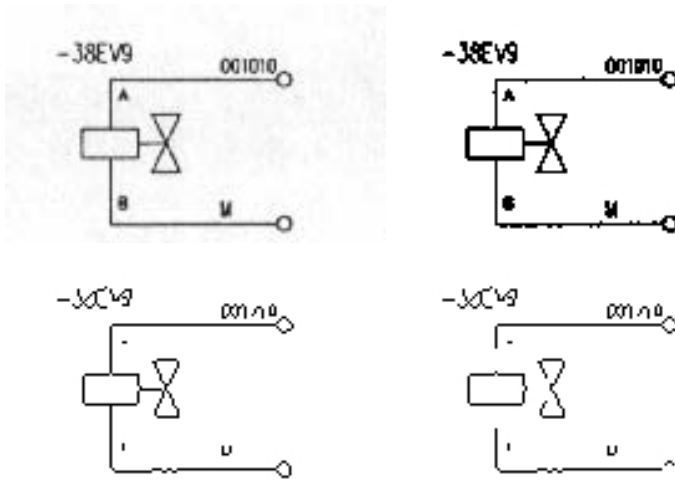


Fig. 9. Another connected symbol extraction.

## References

1. Cordella, L.P., Vento, M.: Symbol recognition in documents: a collection of techniques? *International Journal on Document Analysis and Recognition* **3** (2000) 73–88
2. Belkasim, S.O., Shridar, M., Ahmadi, M.: Pattern Recognition with Moment Invariants: A Comparative Study and New Results. *Pattern Recognition* **24** (1991) 1117–1138
3. S. Belongie, J.M., Puzicha, J.: Shape matching and object recognition using shape contexts. *IEEE Transactions on PAMI* **24** (2002) 509–522
4. Ghorbel, F.: A complete invariant description for gray level images by harmonic analysis approach. *Pattern Recognition Letters* **15** (1994) 1043–1051
5. Lin, B.C., Shen, J.: Fast Computation of Moment Invariants. *Pattern Recognition* **24** (1991) 807–813
6. J. Llado, E.M., Jose, J.: Symbol Recognition by Subgraph Matching Between Region Adjacency Graphs. *IEEE Transactions on PAMI* **23** (2001) 1137–1143
7. B. Messmer, H.B.: Automatic learning and recognition of graphical symbols in engineering drawings. In: *Graphics Recognition: methods and applications (GREC'95)*, LNCS 1072. (1996) 123–134
8. B.G. Park, K.M. Lee, S.L., Lee, J.: Recognition of partially occluded objects using probabilistic ARG (attributed relational graph)-based matching. *Computer Vision and Image Understanding* **90** (2003) 217–241
9. J.Y. Ramel, G.B., Emptoz, H.: A structural representation adapted to hand-written symbol recognition. In: *Graphics Recognition: methods and applications (GREC'99)*, Jaipur, India. (1999) 259–266
10. Yan, L., L.Wenxin: Engineering drawings recognition using a case-based approach. In: *International Conference on Document Analysis and Recognition*, Edinburgh. Volume 2886. (2003) 190–194
11. Arcelli, C., di Baja, G.S.: A One-Pass Two-Operation Process to Detect the Skeletal Pixels on the 4-Distance Transform. *IEEE Transactions on PAMI* **11** (1989) 411–414

12. Deans, S.R.: Applications of the Radon Transform. Wiley Interscience Publications, New York (1983)
13. Tabbone, S., Wendling, L.: Binary shape normalization using the radon transform. In: Discrete Geometry for Computer Imagery, LNCS 2886, Naples, Italy. Volume 2886. (2003) 184–193
14. Borgefors, G.: An Improved Version of the Chamfer Matching Algorithm. In: Proceedings of 7th International Conference on Pattern Recognition, Montréal (Canada). Volume 2. (1984) 1175–1177