Symbol and character recognition: application to engineering drawings
— **Source link** ↗

Sébastien Adam, Jean-Marc Ogier, Claude Cariou, Rémy Mullot ...+2 more authors

Institutions: University of Rouen, Orange S.A.

Related papers:

- Visual pattern recognition by moment invariants

- A robust algorithm for text string separation from mixed text/graphics images

- Invariant image recognition by Zernike moments

- Text/Graphics Separation in Maps

- Text/Graphics Separation Revisited

Symbol and character recognition: application to engineering drawings

Sébastien Adam, Jean-Marc Ogier, Claude Cariou, Rémy Mullot, Jacques Labiche, Joël Gardes

Symbol and Character Recognition : Application to Engineering Drawings

*/** S. Adam, **J.M. Ogier, ***C. Cariou, **R. Mullot, **J. Labiche, *J. Gardes

**Laboratoire PSI - Université de Rouen - 76 821 Mont Saint Aignan - France

Tél : 02 35 14 68 75 Fax : 02 35 14 66 18

{Sebastien.Adam, Jean-Marc.Ogier}@univ-rouen.fr

***LASTI - ENSSAT Lannion - 6, rue Kerampont - BP 447 - 22305 Lannion - France

Tél : 02 96 46 66 39 Fax : 02 96 46 66 75

Claude.Cariou@enssat.fr

*France Télécom – DVSI DES/MBL - 6, Avenue des Usines - BP 383 - 90 007 Belfort - France

Tél : 03 84 54 42 35   Fax : 03 84 54 43 93

Address for correspondence:

Jean-Marc Ogier

Université de Rouen,

UFR des Sciences,

Laboratoire PSI

76821, Mont Saint Aignan, FRANCE

email: Jean-Marc.Ogier@univ-rouen.fr

Abstract: In this paper, we consider the general problem of technical document interpretation, applied to the documents of the French Telephonic Operator, France Telecom. More precisely, we focus the content of this paper on the computation of a new set of features allowing the classification of multi-oriented and multi-scaled patterns. This set of Invariant is based on the Fourier Mellin Transform. The interests of this computation rely on the excellent classification rate which is obtained with this method, and also on the possibility to use this Fourier Mellin transform within a "filtering mode", that permits to solve the well known difficult problem of connected character recognition.

Keywords: Invariant Moments, Character and Symbol Recognition, Feature Extraction and Classification, Engineering Drawing.

Symbol and Characters Recognition : Application to Engineering Drawings

1. Introduction

The current improvements of intranet structures allow large companies to develop internal communications between services. Nevertheless, the representation of the heritage of huge companies like network managers firms is often represented through paper documents, which can be either graphic or textual. As a consequence, the sharing of these kind of information will stay very difficult as long as the storage format will not be digital. This explains the current development of studies concerning the automatic analysis of cartographic or engineering documents which comes as a result of the growing needs of industries and local groups in the development and use of maps and charts. The aim of the interpretation of technical maps is to make the production of documents easier by proposing a set of stages to transform the paper map into interpreted numerical storage [Boatto 1992a, Joseph 1992, Ogier 1998b, Vaxivière 1992].

In this paper, we focus our attention on an original technique which takes place among these stages since it allows the recognition of multi-oriented and multi-scaled objects. The application which is considered in this paper is the automatic analysis of a French Telephonic operator documents, France Telecom. The paper will be organized as follows.

Although this paper does not deal with a global interpretation strategy, in the second section, we will give a short description of the general CAD conversion problems in order to justify the development of this new multi-oriented and multi-scaled pattern recognition tool.

Then, in the third section, we will focus on the pattern recognition problem and we will give in this context a synthesis of the bibliographic sources dealing with multi-oriented and multi-scaled pattern recognition problem. In the fourth part, we will then propose our new pattern description tool, based

on a set of invariant issued from the Mellin Fourier Transform. This part will also include the description of different modes of utilization of the Mellin Fourier Transform, that permits to solve the difficult problem of connected patterns recognition.

In the fifth section, we will then present how this Multi-oriented and multi scaled OCR, based on two utilization modes, is integrated in our global system. Then, in the sixth part, we will present the strategy that we have adopted for implementing, testing and comparing our method with classical tools. Statistical results will also be provided in this part. Finally, in the conclusion, we will try to have a critical point of view of our approach and we to propose our ideas in terms of future works.

2. Usual CAD conversion devices

2.1. General CAD conversion process

The CAD conversion process is quite the same in all the cases and generally consists of 5 successive stages (see Figure 1). At first, it consists in scanning the black and white images (stage 1) at a resolution which is generally within a range going from 400 to 1200 dpi, depending on the richness and on the complexity of the document.

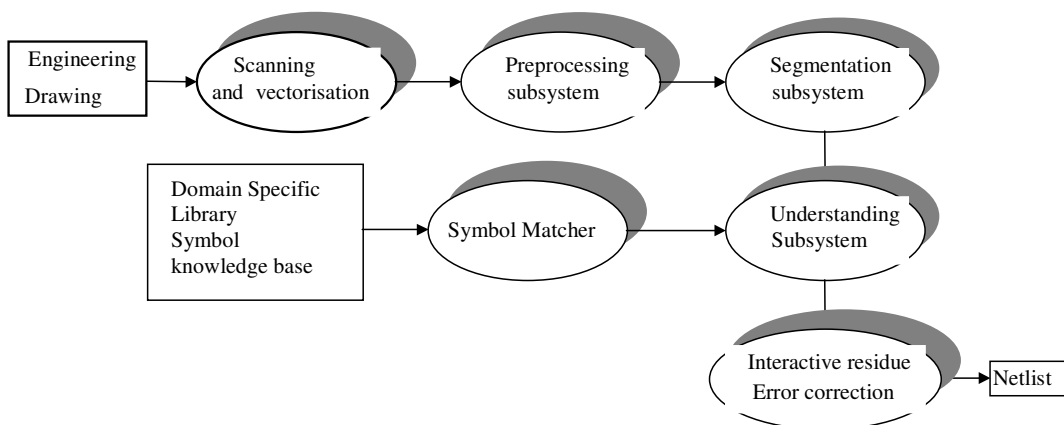

Figure 1 : *Interpretation device Yu [Yu 1997]*

Then, a vectorization process is implemented, allowing to transform the "image" or "raster" representation into a "vector" or "spaghettis" one. At the issue of this process, a post-processing of

the vector's data (stage 2) allows to suppress the different artifacts that appeared during the scanning and vectorization processes. Then, a "segmentation subsystem"(stage 3) allows to process vectors and symbols by dissociating them on different layers, on the basis of a set of rules. Actually, this subsystem allows to separate the symbols/characters which are connected to linear objects, as far as it is possible. The "interpretation subsystem" (stage 4) provides a high level of interpretation of the document, relying on a basis of the objects described by the system. In order to perform this stage, a software components library is very often used in order to find a good matching between the processed objects and some "model" objects. Finally, at the end of the global process, a checking stage of the results of the interpretation is generally performed by a human operator. This user correction cycle can become a very time consuming problem when the size of the document is large or when the document is semantically rich or dense. Indeed, in this correction cycle, generally, the user has not any clue allowing him to locate easily the mistakes made by the system. As a consequence, this checking stage must generally be performed systematically on the whole image. In some cases, this "heavy" task can be made easier when the system has at its disposal a set of external data (alphanumeric data base, for instance), that permits to help the user by verifying the relevance of the extracted data and thus to pilot the correction.

This kind of "traditional" approach is put into question by some works integrating some notions of construction cycles by using some interactions between the subsystems. One can find this kind of approach in the works of : Joseph [Joseph 1992], Den Hartog [Den Hartog 1996a], Ogier[Ogier 1998b] and Pasternak [Pasternak 1995].

2.2. Low level software components in CAD conversion devices

Even if this paper does not deal with the large set of low level software components, it appears important to us to provide to the reader a short survey of the previous and current studies, the results of some of these works being integrated in our system. From our point of view, this subject constitutes the richest bibliographic source, since many authors propose algorithmic solutions for

the extraction and the recognition of graphical entities.

Extraction and recognition processes of these graphical entities are generally applied on black and white images. Nevertheless, some authors propose some solutions performed on Grey level images [Den Hartog 1996b, Kamel 1993, Trier 1995b]. Indeed, such processing of Grey level images allows particular supports such as noisy or deteriorated documents or calques to be interpreted. Indeed, for this kind of document, the thresholding techniques, which are directly integrated in the digitization device (scanner), do not produce images of a quality sufficient to allow reliable interpretation. On this point, Sahoo [Sahoo 1988] proposes a good state-of-the-art review of the thresholding technique in a general context. Concerning the processing of graphical entities, many bibliographic references have been presented since 1980. These papers deal with the separation of different graphical entities [Fletcher 1988], the vectorization of linear objects [Kasturi 1990, Lam 1992, O'Gorman 1997], the segmentation of complex objects such as arcs and circles [Davies 1988, Liao 1990, Thomas 1989], dashed lines [Lai 1991], textures [Antoine 1991, Kasturi 1990, Ogier 1993]. Many of these processing techniques can be considered as mature, as Tombre indicates in [Tombre 1998b][Dori 1995b]. Nevertheless, the most recent literature regularly proposes some relevant improvements dealing with residual problems [Tombre 1998a]: dashed lines [Dori 1996a, Kong 1996], vectorization [Di Zenzo 1996, Dori 1999b, Janssen 1997], dimension sets [Das 1997, Lai 1994], arcs and circles [Dori 1995a, Wenyin 1998b], and global systems to convert scanned engineering drawings into vectorized file [Chen 1996,Wenyin 1998a]

On symbol and character recognition, many works can be found in the literature [Mori 1992, Chhabra 1998], some of them processing classical problems (structured documents) while others deal with specific constraints : maps, industrial documents, … However, some aspects of this pattern recognition problem are still the object of intensive research activities because of their specificity : multi-oriented and multi-scaled characters [Deseilligny 1995, Trier 1996], connected characters [Deseilligny 1995, Trier 1995a], connected symbols [Yu 1994]. This kind of problem is crucial since characters and symbols carry very important information about the contents of a

document : poor detection/recognition of such patterns leads to some very important problem in the interpretation process. In the next section, we focus on this problem of characters and symbols recognition.

3. Characters and symbols recognition : classical approaches

As we have said previously, analysis of textual information and graphic symbols is a fundamental step in the context of document interpretation. Indeed, this information is semantically rich and represents an important part of the sense of the document (quotation values, river names, equipment identifier…). This point explains the fact that the recognition of this textual information represents an important stake for interpretation devices.

From a definition point of view, we will consider as "symbols", the shapes the size of which is similar to the characters' one. This definition takes into account the "alphanumeric characters" as well as the graphic symbols identified by A.Chhabra in [Chhabra 1998]

Literature and commercial products let a wide place to Optical Character Recognition (O.C.R.) systems in a classical context : horizontal characters and known fonts. However, if we consider the problem of generalization to different orientation, the number of propositions decreases considerably. As far as we know, it is even possible to say that there is no reliable industrial product being able to recognize characters submitted to intrinsic variability. Nevertheless, in the context of text recognition on technical document, this kind of tools is really necessary. Indeed, in most of the cases, characters are written, with constraints and with an orientation which is the same as the object they describe. For instance, a road name will have the same orientation as the road. As a consequence, the orientation of graphical lines near a text string can constitute a very strong cue for estimating the orientation of the text string. The works integrating particular constraints like orientation or scale changes are quite numerous in the literature. Let's note the works of Deseiligny [Deseilligny 1995] in the particular case of maps, and the one's of Dori [Dori 1999a] based on NETS [Baffes] system and Trier [Trier 1996].

Concerning the problem of text recognition in the case of an easy separation of each character, different techniques are generally used in a sequential order. These techniques concern respectively segmentation of characters, individual recognition of these characters and grouping them into words or texts. In regard to classical and recent bibliography, these different points are presented in the next section the name of which is "classical approaches.

Nevertheless, a difficult point remains for the inter-connected characters recognition and for the management of characters touching graphical parts of the document. Dealing with this set of problems, a bibliographic synthesis is proposed in section 3.2., the name of which is "Connected Characters Management"

3.1. Classical approaches

Generally, from the system point of view, the processing of textual information is performed through 3 stages

3.1.1. Characters segmentation

This stage aims at isolating characters from other elements of the document. Most of the authors propose to detect characters by extracting connected components the size of which is pre-defined [Trier 1996, Fletcher 1988, Lu 1998, Shimotsuji 1994, Lai 1994, Langrana 1997]. Let us note here the original work of Dori's team [Dori 1996b, Dori 1999a] who proposes a segmentation based on parallelograms, called "Charbox" the orientation of which is relative to the character's one. This process allows to filter candidates as a function of precise dimensions without considering the orientation of the character. Furthermore, this principle is an interesting cue for the characters research ("text box"). However, these processes seem to be very sensitive to the variability of the shapes and can be applied only under very constrained characters representations (essentially mechanical charts).

After this textual information segmentation, some "clustering" algorithms are applied in order to

link the characters into strings without any recognition. In this domain, a good reference is the work presented by Fletcher [Fletcher 1988] who proposes to detect these strings on the basis of Hough Transform and by using a set a simple heuristics. This algorithm requires three co-linear connected components in order to validate the presence of a string. In particular conditions, this methodology can lead to a set of confusion between dashed lines and text strings. More recently, Lu [Lu 1998] proposed a clustering methodology allowing to link a set of new connected components (NCC) based on a set of four heuristics (four predefined thresholds). Text orientation is estimated by minimizing the "bounding rectangle" through successive rotations of the image by step of 10 degrees.

3.1.2. Character recognition

Problem dealing with recognition of multi-scaled and multi-oriented shapes constitutes a particularly difficult point in the design of robust technical document interpretation device. In this field of research, we may distinguished three main approaches:

- The first suggests a preliminary computation of the shape's orientation, and tries, through a normalisation and a rotation step, to obtain a pattern in a reference position which can be introduced into a classical OCR (Optical Character Recognition) system. However, methods based on such a strategy are not frequently used because of the lack of methods that enable a reliable computation of the orientation to be obtained. Moreover, distortions due to sampling errors appear during the geometric transformations that are needed to normalise the pattern.

- A second approach consists in using a multi-layer feed forward classifier, fed by the original image of the pattern. In this kind of context, the classifier renders the problem invariant with respect to the desired transformations [Fukumi 1992]. However, this kind of strategy rises different problems, dealing with the loss of information during the normalisation stage and the constitution of representative data base [Lecun 1995]. As a consequence, the proposed

contributions are limited in regard with the range of angle variation. [Lecun]

- The last approach, which is probably the most frequently used, consists in extracting from the shape a set of descriptors, which are invariant to the desired transformations. An excellent description of the state of the art in this domain can be found in Trier [Trier 1996]. Generally, it is possible to note that the features used to describe patterns independently from their position, size, and rotation, can be split up into two groups, as shown in the following.

  ❑ Descriptors based on the global aspect of the pattern

Many features can be used to describe the global aspect of a shape. Since the works of Hu in 1961, invariant moments [Hu 1962], which are based on combinations of regular moments, have been very frequently used [Rothe 1996, Reiss 1993]. Among them, one can thus cite Zernike moments [Teague 1980, Khotanzad 1990a, Khotanzad 1990b, Liao 1990] which consitute a reference in the domain, pseudo-Zernike moments [Teague 1980], Bamieh moments [Bamieh 1986], and Legendre moments[Chen 1996].

These invariant moments, which can be extracted from a binary or a grey-scaled image, generally offer properties of reconstructibility, thus ensuring that extracted features contain all the information about the shape under study. Good comparative studies about moment invariants can be found in [Teh 1988] and [Belkasim 1991], both showing the superiority of Zernike moments in terms of recognition accuracy. Among these results, Belkasim [Belkasim 1991] concludes that 6th order Zernike normalized invariant moments provide the best recognition performance. Indeed, good recognition rate reach 94,9 % whereas misclassifications rate just exceeds 5 % by using a KNN classification process. The used database contains only 320 prototypes distributed in 10 classes of digits. Let's note that a mathematical transformation is indispensable for obtaining rotation invariance. Another interesting comparative study between Lengendre, Zernike and pseudo-Zernike moments is led in [Bailey 1996]. This study shows that Zernike moments of 7th

order, introduced in a KNN classification process provide the best results (about 91,7%) on a handwritten characters database. Although, these rates can not be considered as a reference, given the size of databases. Moreover, these studies also proved that moment-based approaches are sensitive to noise and that they are time-consuming, even if complexity optimisation methods can be found in the literature [Dai 1992, Lin 1991].

❑   Descriptors based on a local approach

Besides the approaches presented above, a geometric invariant description can also be obtained by using features which are supposed to contain most of the pattern information. For example, outlines are commonly used in order to obtain patterns invariant descriptions. The most frequently used descriptions deal with Fourier descriptors [Pei 1992] or elliptic Fourier descriptors [Lin 1987]. Taxt [Taxt 1990] proposed a comparative study between these descriptors. This study highlights their potential interest, in terms of simplicity and robustness.

Especially, Taxt [taxt 1990] proposed to use elliptic Kuhl moments [Kuhl 1982] when characters orientation is known. This kind of technique is also used by Trier [Trier 1996] for character recognition on hydrographic maps. Authors claim to obtain a good recognition rate of 78 % for only 2,3% of misclassification on a test set of 1760 hand-written characters. On the other hand, structural invariant features can also be extracted from characters or thinned characters [Boatto 1992a, Shimotsuji 1992, Shimotsuji 1994]. For instance, one can thus cite the number of occlusions, the number of T-joints or X-joints, the number of bend points. However, it has been shown that such features used alone do not lead to robust recognition systems [Mori 1992].

Circular primitives, which are, by definition, well adapted to rotation invariant recognition, have been used in [Kita 1992]. These are based on the analysis of the shape through a set of concentric circles. In this kind of context, Lefrere [Lefrère 1993] proposed a set of circular strobes which permit to describe the shape around its centroïd. A comparative study, available in [Dudani 1977], shows that they yield better results than Hu's moments. Nevertheless, this kind of approach stays

very sensitive to variability of the shapes [Lefrère 1993].

3.1.3. String construction

The last step concerning recognition of textual information consists of the constitution of words. Heuristics are sometimes necessary in order to detect rupture in characters string. Let's note there the work of Deseilligny [Deseilligny 1995] whose algorithm, based on dynamic programming, permits to construct optimal strings by using constraints on orientation homogeneity. After this string reconstruction, the author proposes a syntactical verification in order to improve the robustness of the results.

3.2. Connected Characters Management

In relation with textual information analysis, another important set of problems deals with the characters/symbols which are connected together or connected to linear objects [Dori 1998]. Most of the proposed methods rely on the analysis of connected components. As a consequence, if we consider these kind of approach, an isolated character which is line connected can not be localized. From this point, the lack of literature highlights the difficulty of the problem.

The processing of characters which are connected together seem to be the less complex problem. Actually, the proposed techniques generally rely on the successive stages. At first, they rely on a detection of the connected characters sets, on the basis of the analysis of the connected components sizes. Then, classical segmentation/recognition techniques are applied, with or without a first righting stage of the shape [Casey 1996]. Casey proposed in his paper a synthesis of the characters segmentation methods dealing with horizontal texts. These kinds of techniques are not applicable when the characters are connected to graphical objects of the document : lines, textures, and so on. In that kind of context, the detection of characters which are connected with large connected components is much more complex, because of the difficulty to distinguish the features characterizing the text and the element on which they are connected. In order to separate textual

information, some research teams try to implement some image processing algorithms either based on morphological techniques [Lu 1998] or on local computation of binarization thresholds [Trier 1995a, Trier 1997].

On the other hand, some more structural techniques consist in processing the vectorized image. These kind of approach try to extract the set of vectors corresponding to potential characters or symbols [Shimotsuji 1994], or try to detect sets of vectors which do not match with the line model of the document [Nagy 1998]. Boatto [Boatto 1992b] proposed to detect the connected characters by using a technique based on "run-length" analysis. At each node, a local analysis is triggered in order to verify the line or character hypothesis. This analysis is based on a rule directed technique which integrates a set of heuristics. If a character is detected, some new constitution rules are put in place, on the basis of topological and geometrical constraints. As said in the paper proposed by Boatto, these rules are adapted to the analysis of the Italian cadaster and do not seem to be easily adaptable to other kind of documents.

An interesting approach is proposed by Dori [Dori 1996b, Liu 1997,Dori 1999a]. From the vectorized image, a recursive process allows to constitute characters on the basis of segments corresponding to potential characters. This technique tries to cluster short connected segments until the constitution of a shape which can be included in a parallelogram the size of which is pre-defined ("charBox").

Then, a combining of "charboxes" allows to validate the presence of characters strings, called "textboxes". This interesting process allows to discriminate lines from characters, but seems to be limited by the pre-defined size of the "charboxes".

At last, Deseilligny [Deseilligny 1995] initiates a systematic connected character research on each extremity of each detected strings, and when the space between two connected components permits to emit an hypothesis about the presence of a connected character presence. In this case, the research of a character with the same font, the same size, and the same orientation as the analyzed string is performed. This character research operation is performed by trying to find an optimal

matching between a set of primitives based on masks corresponding to the character references, and the image zone on which the character is researched. In fact, a normalization stage is integrated in this process, just before the matching operation. Actually, in order to simplify the matching operation, the primitives which are retained correspond to a set of sampled points on the outline. The technique, which seems to be very sound, is well adapted for the detection of regular strings without any strong variability. Furthermore, it requires to have precise information about the font which has been used during the drawing process. However, since this technique relies on the pre-detection of a string, it can not be used to detect isolated characters which are connected to the lines.

3.3. Synthesis and justifications

Given this large number of existing methods, one could argue that it is not necessary to develop new invariant features. Nevertheless, comparative evaluation studies [Belkasim 1991] have shown that the features presented above are not perfect in terms of recognition accuracy, especially when the images are noisy.

On the other hand, the documents on which we are working have specific constraints which do not allow us to use the most interesting techniques proposed by the literature [Dori 1996b], especially because of the connected characters problem. Indeed, these kinds of technique rely on different points which have been presented in the previous parts, i.e. detection of alignments between connected shapes, parallelism between strings and linear objects, pre-defined sizes of the characters to be detected, and so, on. As said in the introduction of this paper, the variability of the textual information representation on our documents do not allow us to re-use this kind of technique. Indeed, on our documents, symbols/characters can be completely isolated (i.e. not integrated in a string). Furthermore, they do not necessarily have the same orientation as the objects on which they refer. At last, their size can be variable, as well as their orientation.

Considering the set of arguments and the interest to develop a new technique based on theoretical and reliable concepts, we have decided to consider a new technique, based on Fourier Mellin

Transform.

Furthermore, many of the previously presented techniques rely on the integration of contextual information (orientation, size, and so on) which are absolutely not necessary with our approach. As a consequence, our strategy is reinforced by the development of a low level technique, which could anyway be combined with the concepts presented in the bibliography.

4 – The Fourier-Mellin Transform applied to character recognition

As we said previously, a strong constraint in the global interpretation of the document comes from the fact that characters and symbols can have any orientation and size. The consequence is that the recognition procedure to be applied must be invariant with regard to any combination of rotation and scaling of a pattern, i.e. any geometric similitude transformation.

Another strong constraint relies on the robustness of the recognition procedure. In fact, after the binarization step, many characters are still connected either together, or to the network, leaving any classical pattern recognition technique useless.

The strategy that we propose covers both constraints within a uniform framework. It is based on the application of the generalized Fourier analysis to the particular geometric group of positive similitudes. More precisely, we make use of the properties of the Fourier-Mellin transform, the properties of which are very interesting for our application. Basically, the technique developed herein is a combined use of the works of Ghorbel [Ghorbel 1994], Ravichandran [Ravichandran 1995].

First, we will recall the definition of the Fourier-Mellin transform (FMT). Then, we will recall the analytic prolongation of the FMT (AFMT) and a set of complete and stable similitude invariant features, first proposed in [Ghorbel 1994]. Next we provide the reader with some considerations about the computation of the AFMT over square lattice data, and show some early experimental results of this technique.

4.1. The Fourier-Mellin transform (FMT)

Let $f(r,\theta)$ be a real-valued function (the pattern) expressed in polar coordinates. The FMT of this function is defined as the Fourier transform on the group of positive similitudes:

$$M_f(\upsilon,q) = \int_{\rho=0}^{+\infty} \int_{\theta=0}^{2\pi} \rho^{-i\upsilon} \exp(-iq\theta) f(\rho,\theta) \frac{d\rho}{\rho} d\theta \tag{1}$$
$$\text{with } q \in \mathbf{Z}, \quad \theta \in \mathbf{R}$$

In this expression, i is the imaginary unit. It is well known that the Fourier-Mellin integral does not converge in the general case, but only under strong conditions for $f(r,\theta)$.

4.2. Analytic prolongation of the Fourier-Mellin (AFMT) and properties

In order to alleviate the above difficulty, Ghorbel [Ghorbel 1994] has proposed the use of the AFMT, defined as:

$$\tilde{M}_f(\upsilon,q) = \int_{\rho=0}^{+\infty} \int_{\theta=0}^{2\pi} \rho^{-i\upsilon+\sigma_0} \exp(-iq\theta) f(\rho,\theta) \frac{d\rho}{\rho} d\theta \tag{2}$$
$$\text{with } q \in \mathbf{Z}, \quad \upsilon \in \mathbf{R}, \text{ and } \sigma_0 \in \mathbf{R}_+^*$$

An important property of the AFMT (as well as the FMT) relies on the application of the shift theorem for the Fourier transform. Let $g(\rho,\theta) = f(\alpha\rho,\theta+\beta)$ be a scaled and rotated version of $f(\rho,\theta)$, then we have the following :

$$\tilde{M}_g(\upsilon,q) = \alpha^{-\sigma_0+i\upsilon} \exp(iq\theta) \tilde{M}_f(\upsilon,q) \tag{3}$$

Taking the modulus of both terms in Eq. (2) yields features which are invariant under any rotation of the pattern but not under scaling. To obtain scale invariance on this basis, one could use the following set of features :

$$I_f(\upsilon,q) = \left| \tilde{M}_f(\upsilon,q) \right| \left[\tilde{M}_f(0,0) \right]^{-1} \tag{4}$$

This set of invariant features provides a simple representation of shapes. However, it does not respect the completeness property, i.e. there is no bijection between the dual representations of a single pattern, since the phase information is dropped.

In [Chorbel 1994], the following set of rotation and scale invariant features was proposed:

$$I_f(\upsilon,q) = \tilde{M}_f(\upsilon,q)\left[\tilde{M}_f(0,0)\right]^{-1+i\frac{\upsilon}{\sigma_0}}\left[\tilde{M}_f(0,1)\right]^{-q}\left|\tilde{M}_f(0,1)\right|^{q} \tag{5}$$

Now if $g(\rho,\theta) = f(\alpha\rho,\theta+\beta)$, it can be easily shown that $I_g(\upsilon,q) = I_f(\upsilon,q)$. The two important properties of these *features* rely on:

(*i*) their completeness : given the quantities $\tilde{M}_f(0,0), \tilde{M}_f(0,1)$ and $\{I_f(\upsilon,q)\}$, it is possible to return to the whole set of FM coefficients $\{\tilde{M}_f(\upsilon,q)\}$, and thus to reconstruct $f(\rho,\theta)$, by using the inverse AFMT in the following way :

$$f(\rho,\theta) = \sum_q \int_{\mathbf{R}} \tilde{M}_f(\upsilon,q)\, \rho^{\sigma_0+i\upsilon} \exp(iq\theta)\, d\upsilon \tag{6}$$

(*ii*) their convergence : it is proven, in [Ghorbel 1994] that under the assumption that $\{\tilde{M}_f(\upsilon,q)\}$ is a convergent set, there exists $x \in \mathbf{R}$, $x > 1$ such that

$$\left(\int_{\mathbf{R}}\sum_{\mathbf{Z}}\left|I_f(\upsilon,q)\right|^x d\upsilon\right)^{1/x} < +\infty \tag{7}$$

One important consequence of the completeness and convergence of this set of invariant features lies in the existence of a metric in the shape representation space.

4.3. Application to 2-D square lattice images

Many difficulties and questions arise when trying to apply the AFMT to common images. The first one comes from the necessity to sample the quantity $\tilde{M}_f(\upsilon,q)$. However, this problem finds a simple solution by sampling the υ variable and computing $\{\tilde{M}_f(p,q)\}$ and $\{I_f(p,q)\}$ for $p \in \mathbf{Z}$.

Secondly, a unique pattern-invariant and pattern-representative point must be chosen as the center of development (CoD) of the AFMT. A reasonable *a priori* choice is to apply the AFMT at the centroid (center of gravity) of the pattern. For well-conditioned patterns (i.e. disconnected from any

other), the computation of the centroid is easy, and moreover it ensures that the shape will hold within a given radius, which will be of great interest in the following.

The third and major difficulty is due to the square lattice structure of images. At present, two options are available for the processing of images, namely image interpolation in polar coordinates [Derrode 1997, Ghorbel 1994], or convolution with an appropriate filter bank in Cartesian coordinates [Ravichandran 1995].

With the first option, one has to compute the AFMT from Eq. (2) at the centroid of the pattern. This assumes that the image should first be interpolated in polar coordinates in order to cover the whole range of $\rho \in \left]0, \rho_{\max}\right]$ and $\theta \in \left]0, 2\pi\right]$ ($\rho_{\max}$ being the maximum radius of the pattern). This technique may yield an important computation load, and may also introduce redundant information into the original data.

With the second option, it is not necessary to interpolate the pattern within the 2-D square lattice. Indeed, one can observe that the discrete version of the AFMT calculated at the CoD (k_0, l_0) of the pattern can be approximated in the following manner :

$$\tilde{M}_f(p,q) \approx \sum_{\substack{k \\ 1 \le (k^2+l^2) \le \rho_{\max}}} \sum_l h_{p,q}(k,l)\, f(k-k_0, l-l_0) \tag{8}$$

with

$$h_{p,q}(k,l) = \frac{\exp\left\{-i\left[\dfrac{p}{2}\ln\left(k^2+l^2\right)+q\tan^{-1}\left(\dfrac{l}{k}\right)\right]\right\}}{\left(k^2+l^2\right)^{1-\sigma_0/2}} \tag{9}$$

This means that the convolution of the pattern by a 2-D filter with the finite impulse response $h_{p,q}(.,.)$ at the CoD of the pattern gives the AFMT coefficient of order (p,q). Note that all filters but $h_{0,0}(.,.)$ are complex-valued. Also, it can be shown easily that the set of $h_{p,q}(.,.)$ filters is orthonormal with respect to the circular coefficient q, but not with respect to the radial coefficient p. Actually, a complete orthogonalization of the basis functions is tractable under special conditions

on the discretization of the υ variable in Eq. (2), and on the values of σ_0 and ρ_{max} ; this orthogonalization was not performed in this work. Because of its simplicity, this second option will be kept for use in our application. Note that it also offers the advantage of a possible parallel computing of AFMT coefficients.

In order to have an intuitive point of view of Fourier Mellin transform, we show on figure 2 an illustration of a filter. As Fourier-Mellin transform corresponds to the decomposition of the pattern into circular (provided by the Fourier transform) and radial (provided by the Mellin transform) harmonics, it appears that the combination of both transforms in the FMT leads to basis functions which are spiral-shaped. The number of spiral « arms » is related to the q coefficient, while the p coefficient corresponds to the spatial frequency or « contraction factor » of the spirals.
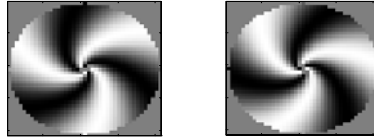

Figure2: Real and imaginary parts of $h_{p,q}$ with p = 2, q = 3, and $_0$ = 2.

Once the AFMT coefficients $\tilde{M}_f(\upsilon,q)$ are available at the CoD of a given pattern, Eq. (5) is used to derive scale and orientation invariant features $\{I_f(p,q)\}$ for this pattern. One can easily see that these features are all complex-valued, except $I_f(0,0)$ which is unity and $I_f(0,1)$ which is real-valued. Also, we have the following Hermitian property of the set of invariants:

$$I_f *(p,q) = I_f(-p,-q) \tag{10}$$

where * denotes complex conjugation. This property allows to restrict the number of non redundant features to the set of $\{(p,q) | (p \in \mathbf{N}; q=0) \cup (p \in \mathbf{Z}; q \in \mathbf{N}^*)\}$. In terms of computational complexity, such a strategy permits to reduce the number of elementary operations (complex

multiplication) to $(2PQ+P+Q+1)\pi r^2$ where $(2PQ+P+Q+1)$ represents the number of filters and r^2 the number of points per filter.

An example of invariant description $\{I_f(p,q)\}$ (which will be used as features for the future classification process) for 5 different characters is given in the next figures. A set of 13 multi-orientated digit patterns (from A to E) is presented in Figure 3. These patterns were analyzed with the filter bank { $h_{p,q}(.,.)$ }, taking $\sigma_0 = 1$, $\rho_{max} = 20$, and setting p in the range -2 to 2 and q in the
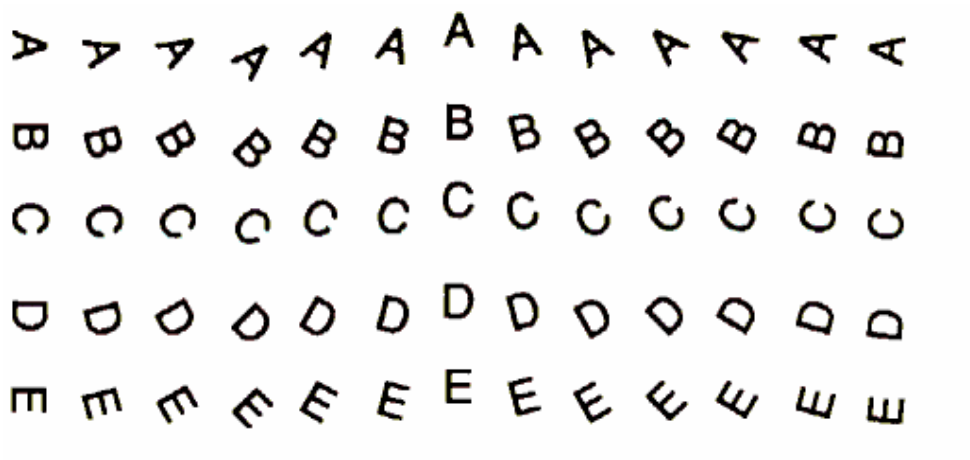
range 0 to 3. Such range values for p and q lead to 33 non redundant invariant features obtained through the use of 18 filters. The same patterns were also analyzed at a scale factor of 0.5, i.e. by simply undersampling one pixel out of two in each dimension of the original image of Figure 3. Figure 4 shows the invariant features which are obtained by using these filters: each line corresponds to the overlaying of the 33 invariant features extracted from the 13 versions of each particular shape (A, B, C, D or E). For instance, the first plot represents the overlaying of the 33 invariants for each of the 26 multi-oriented and multi-scaled "0" shape. In Figure 5, we illustrate the completeness property of the invariants, and show the reconstruction of some patterns from the analysis by a set of 62 filters, giving 119 non-redundant invariant features.

Figure 4 *: corresponding invariant features*
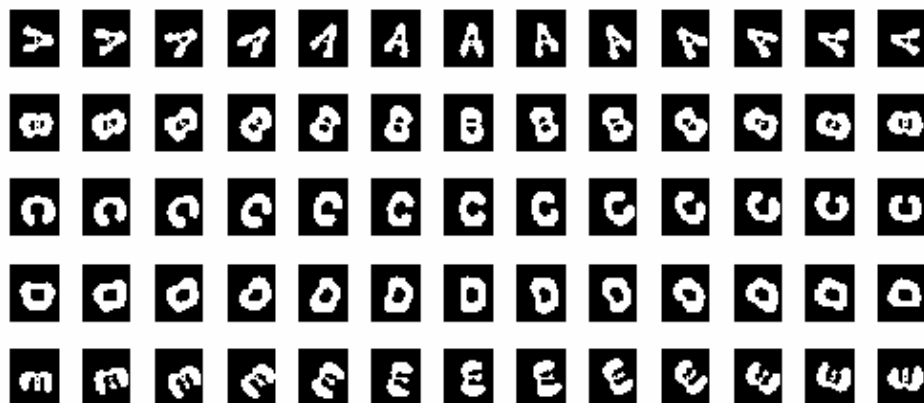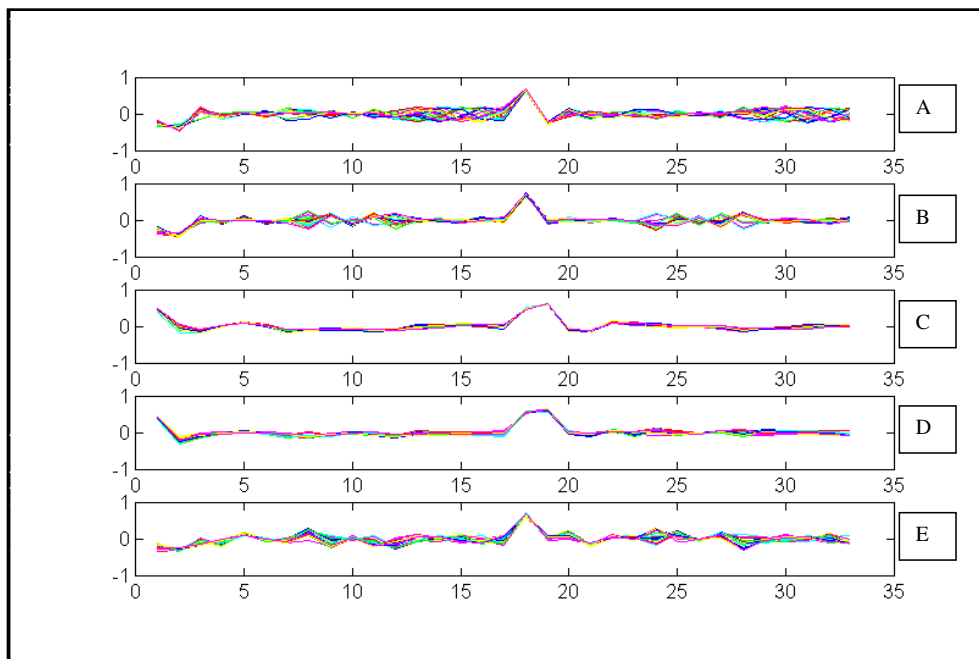
Figure 5: _Reconstruction of patterns using_ the inverse AFMT and $\left\{ \tilde{M}_f \left(\upsilon, q \right) \right\}$

5- Integration of character recognition tool in the interpretation device

The aim of this section is to give some details about the Mellin Fourier description tool, which can be implemented within two modes, that we will respectively call "recognition mode" and filtering mode". The utilization of one of these modes is triggered as a function of specific conditions detected by our intepretation device. Our technical drawing interpretation device is related to the general Artificial Intelligence problem, which can be specified by using a methodology such as object process methodology [Dori 1995c]. The aim of scene interpretation is to extract semantic information from digital images. It is difficult for an observer to recognize an object if he does not have any *a priori* mental representation of it. For an artificial interpretation system the problem is similar, and it is necessary to integrate this notion of a model if the aim is to obtain a representation close to that processed by the drawer of a map.

In order to carry out this modeling, we start from the principle that the whole of the graphic document relates to a specific organization of all the graphical primitives (strokes, dotted lines...) and to the round-up rules for these graphical entities for object representation. In the case of engineering drawings [Antoine 1992, Joseph 1992, Vaxivière 1992] or cartographic maps [Deseilligny 1995, Kasturi 1988, Suzuki 1990], the problem is particularly complicated since it often requires the expertise of a specialist to interpret the document.

In our approach, a model of the document is defined and image processing tools and interpretation operators are integrated into a general device, so that knowledge is constructed progressively, from the low level (pixel level) to the high level (semantic objects). At each level of our device, an analysis of the coherence of the data verifies that the constructed knowledge is in accordance with the defined model. When there are ambiguities between the constructed knowledge and the model, some specific operators are executed in order to solve the problem.

Concerning characters and symbols, the general scheme of the strategy applied for this part of our device is presented in Figures 6,7 and 10. First, the different overlaying layers (see Figure 9.1.) are segmented thanks to the selection of morphological criteria [Ogier 1998a]. This operation separates

the urban, the network (see Figure 9.2) and the character/symbol layers (see Figure 9.3.). The character/symbol layer contains symbols and characters that may be either isolated or connected to each other. The network layer contains the network on which some characters may still be connected because of drawing faults or segmentation (binarization) errors.

The symbol/character layer is first processed within a mode that we will define as "recognition mode" (see Figure 6). On this layer, most patterns are extracted by a shape extractor. Then, on each of these patterns, similitude invariant features are computed at the centroid of each extracted pattern, based on the Fourier-Mellin transform (FMT). This vector is then introduced into a preliminary calibrated classifier that proposes a response for unknown shapes.
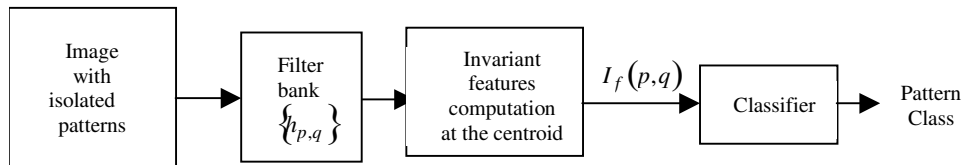

Figure 6 *: Scheme for the isolated character recognition : Recognition mode*

The second mode, called "filtering mode", deals with the characters that are connected together or the character that are connected to the network in the network layer : indeed, on the network layer, many symbols and characters may still be connected. For this layer, connected shapes are inconsistent with regard to the information that is to be stored in the Data Base. This is why detection and recognition of the characters and symbols are necessary in order to construct consistent information. This detection / recognition operation is performed thanks to the same tool as for the recognition mode, i.e. the FMT, applied in a filtering mode.

More precisely, the image is first convolved by the set of $h_{p,q}(.,.)$ filters and one tries to locate the pixels for which the response is pre-specified and may be recognized as the CoD of a known
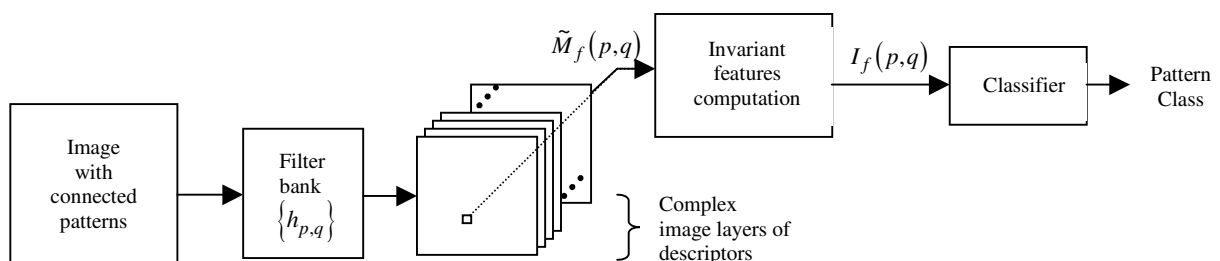
pattern. For this, a set of invariant features is computed for each pixel and enters the classification method, as shown in Figure 6 and 7. The recognition of overlapping symbols and characters is thus made possible. In fact, this is for a major part due to the decay of the magnitude of $h_{p,q}(.,.)$ as the radius from the center point of the filter increases : in most configurations where characters and symbols are connected to each other, this problem occurs only at the vicinity of the external envelope of a given pattern. In such a case, the contribution of another overlapping pattern to the computation of $\tilde{M}_f(p,q)$ at the centroid of the current pattern is small, and one can hope that the classification process will not be much affected by this contribution. Of course, in order to reduce the computational burden, and taking into account the large size of each filter, the convolution of the filter bank with the image is performed via the 2-D Fourier transform on areas of interest issued from the system approach.

__Figure 7__ : Scheme for the analysis of images with connected patterns : filtering mode

Since this filtering technique may be a time consuming operation [Tan 1998] because of a convolving process, it is not applied through a blind procedure on the whole image. In fact, the application of this technique is guided by the system approach, which identifies some areas of interest, for which it has emitted some hypotheses concerning the presence of characters in ambiguous zones [Adam 1999].
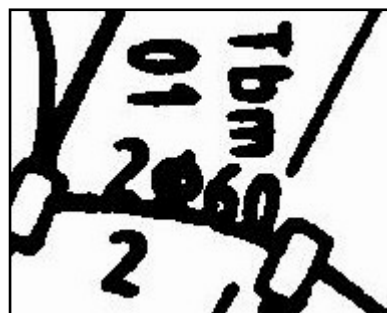

__Figure 8__ : Characters connected to the network

The hypotheses are emitted thanks to the consistency analysis of the information extracted from the considered layer. An inconsistency between the model of the document and the extracted objects is detected as soon as a character is connected to the network information layer, for instance. On Figure 8 is highlighted the problem of connected characters.
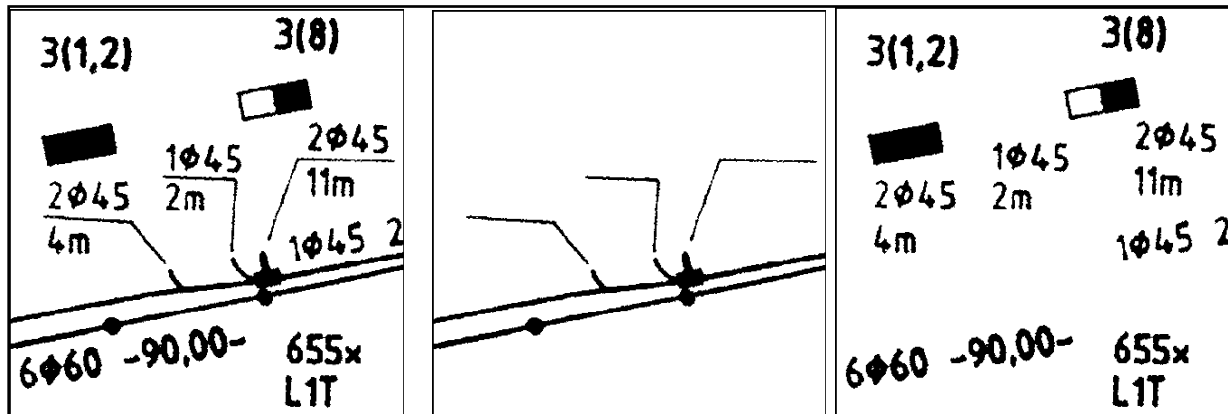

Figure 9.1 : Original image   *Figure 9.2 : Network Layer*   *Figure 9.3:Character/Symbol layer*

Figure 9 : Segmentation in layers

Actually, the global strategy consists in triggering one of these two previously presented modes, as a function of the considered layer, and as a function of the consistency analysis. Actually, as said before, the first stage of our device consists in separating the characters layer form the others, on the basis of simple geometrical features set. The Recognition mode is applied on the layer containing the resulting characters layer. Nevertheless, as said before, many connected characters stay in another layer, which generally contains linear objects. On this layer, an analysis of the set of the vectors permits to detect automatically some hypothetical connected characters. This approach is based on a coherency analysis, in regard with a model of the linear objects [Ogier 1998b]. This coherency analysis allows to detect some "area of interest" on which the filtering mode is applied. As a consequence, these two recognition process are applied on different information layers. The figure 10 is an illustration of the global process.
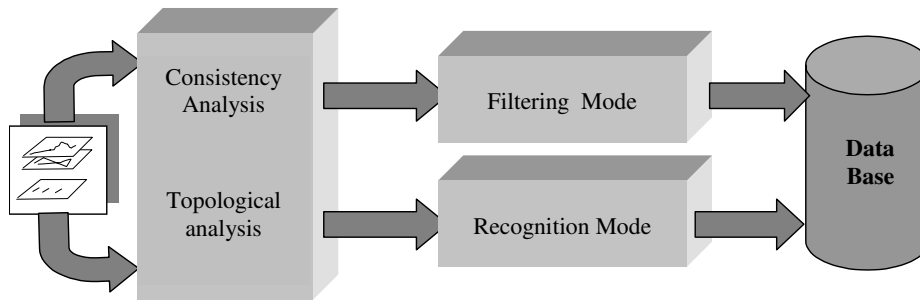
Figure 10 : *Global Recognition process*

6 - Experimental results

Significant tests have been performed in order to provide a reliable evaluation of these methods. The classification strategy relies on the tests of different classifiers. For our tests, we have implemented two different classifying techniques : the Learning Vector Quantization (LVQ) and the *K*-nearest neighbors (KNN). The choice of these classifiers was motivated by two arguments : at first, KNN's technique represents a good reference on which an evaluation could be objectively performed; secondly LVQ's approach is an interesting technique from the implementation point of view, in our industrial context : low time consuming. The classifying methodology consists in determining a distance between the vector representing an unknown shape and a set of reference vectors describing labeled patterns. The results of our tests are two-fold : in a first part we consider the case of isolated characters ("recognition mode"), and in a second part we deal with the connected components ("filtering mode"). In all the experiments, the number of considered classes to be identified equals 51 (i.e. upper case lower case letters and 0 to 9 digits, by grouping classes that have exactly the same shape in our font – 6 and 9, I and l, ….). Let's note that in this paper, tests are performed only on alpha-numeric characters, in order to have sufficiently representative databases. Nevertheless, it is possible to apply such a strategy to any kind of symbols that can be found on technical document.

6.1. Results concerning the isolated characters (characters layer)

The strategy that we have adopted for our tests concerning the isolated shapes can be split in two parts. First, in order to validate the methodology, we applied our technique on a set of "clean"

characters, issued from a database especially provided for this application. In the second part, we considered a set of characters issued from real *France Telecom* documents.

- Clean database

We present on Figure 11 a sample of image which shows a set of "clean" characters, issued from a data base provided for this application. Concerning these tests, we used a set of 394 training samples and a test database composed of 412 samples.
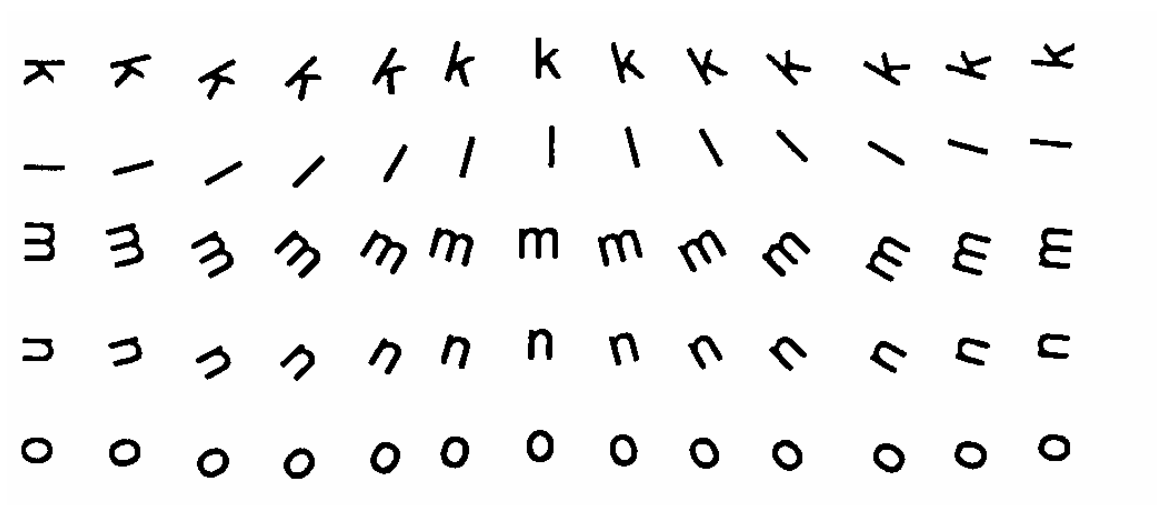

Figure 11 : a set of clean characters

The results obtained by using this method are very encouraging since we reached 97.5% of correct classification with the KNN strategy ($K = 1$) and 97% with the LVQ procedure.

The analysis of the results show that the confusions are coherent since they concern similar shapes such as "m" and "E" (which are quite similar for our fonts if we consider a rotation of 90 degrees and a scale factor) or "N" and "Z". These particular cases do not invalidate the general methodology since this kind of problem is generally solved thanks to contextual information. Within our interpretation strategy, this contextual information is to be provided by the global interpretation system.

- Real database

As explained above, the second experiment consisted in testing the method on a set of characters from *France Telecom* documentation. Actually, the problem is much more complicated than in the case of "clean" characters, since the variability of the shapes is much greater, and the shape contours are sometimes very noisy after the binarization step. To illustrate this, we present on Figure 10 an example of a processed image.

__Figure 12__ : a set of characters from real technical maps

The conditions of our experimentation were as follows. The training database was composed of 4890 samples and the testing database was composed of 14878 samples.

The classification rate is quite good with regard to our industrial constraints since the correct classification rate reaches 95,74% with the KNN approach and 95,11% with the LVQ classifier.

The confusions encountered are still due to similar shapes, like "B" and "8" (which are quite difficult to distinguish because of the noise on the image), or "5" and "S".

Finally, in order to have an objective point of view about the proposed invariant features, we have compared them with other techniques that are generally considered in the literature as being excellent. So, we compared the Fourier-Mellin features with the Zernike moments and the circular primitives. The results of all these tests, which highlight the superiority of the Fourier Mellin invariants, are presented in Table 1.

Classifier	Clean Data Base	Real Data Base

LVQ	Fourier Mellin transform	97.0 %	Fourier Mellin transform	95.11 %
	Zernike moments	75.5 %	Zernike moments	83.86 %
	Circular primitives	73.6 %	Circular primitives	82.61 %
1-NN	Fourier Mellin transform	97.5 %	Fourier Mellin transform	95.74 %
	Zernike moments	80.9 %	Zernike moments	84.55 %
	Circular primitives	77.5 %	Circular primitives	83.92 %

TABLE 1 : CLASSIFICATION RESULTS AND COMPARISONS ON CLEAN AND REAL DATABASES

6.2. Results concerning connected shapes

For the connected patterns, the strategy that we applied was the following. Firstly, a set of 257 sub-images including connected characters were selected. On these images, the connected shapes could either concern characters connected to each other or characters connected to the network as shown in Figure 13 and 14. A total number of 408 truly connected characters were analyzed on this set of images.

Two points were considered in this experiment, namely detection and classification. Firstly, the detection is evaluated according to whether the technique can or cannot detect a shape for a particular pixel. The non-detection rate is quite encouraging since only 7 of the 408 samples have not been detected as characters. Moreover, we considered the false detection rate that corresponds to pixels for which a character was detected while nothing representing a character was on the image. We had an excellent false detection rate since it happened only once on our set of 257 images.

The second evaluation deals with the correct classification, once a shape has been detected on a pixel. These results are also excellent since the rate of correct classification reached 83% for the connected components. As in the case of clean patterns, misclassifications were essentially due to similar shapes, but they could also be due to graphical representation variations induced by the connection between shapes. As an illustration, one could quite recognize an horizontal "8" on the figure 13.1, when regarding the two connected "0". On the figure 13 and 14, we present an illustration of the application of the filtering mode that permits to detect (figure 13.3 and 14.3) and recognize characters (figure 13.4 and 14.4) which were initially connected to the network or
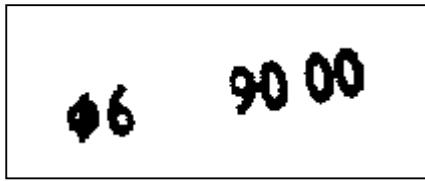
interconnected.

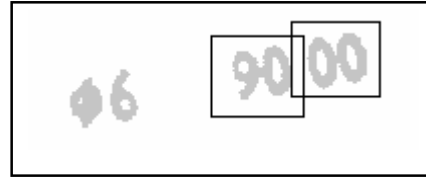*Figure 13.1 : Connected characters*　　　*Figure 13.2 : Areas of interest*
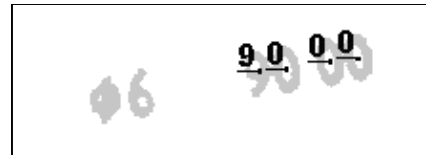

Figure 13.3 : Character detection　　　*Figure 13.4 : Character recognition*

Figure 13: interconnected characters

Figure 14.1 : Connected character　　　*Figure 14.2 : Areas of interest*

Figure 14.3 : Character detection　　　*Figure 14.4 : Character recognition*

Figure 14: character connected to the network

Even if these results concerning connected characters and characters touching graphics are not statistically sufficient, they prove that the process is viable.
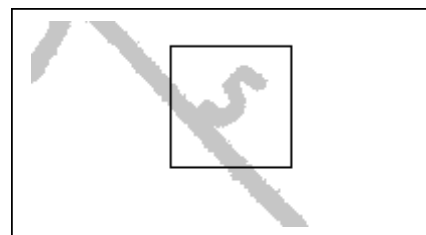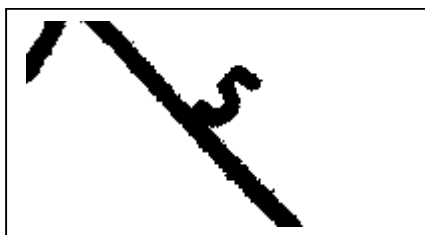
7 - Conclusion and perspectives
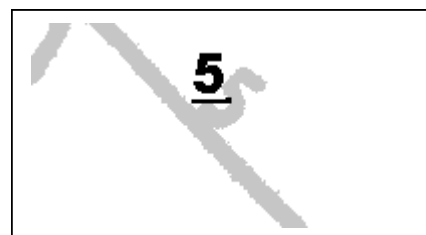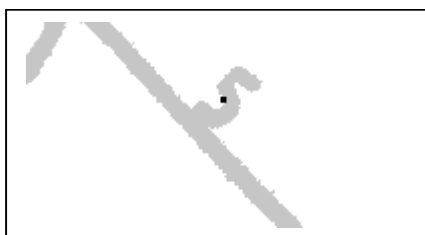
In this paper, we have proposed an original methodology allowing the detection and recognition of multi-oriented and multi-scaled shapes. The supports on which the method was

applied are technical documents representing the network of the French Telephone operator (*France Telecom*). The technique adopted, based on the Fourier-Mellin transform, is integrated in a global strategy, which permits to resolve within a same formalism, without any *a priori* knowledge, the problem of the recognition of multi-oriented/multi-scaled shapes, these one being connected or not.

The strategy, which is applied to solve the characters/symbols classification, can be divided in two stages. The first one consists in constructing a geometric invariant features vector from each shape of the characters layer which is extracted through a connected component extractor. The second one, which is probably the strongest point of this work, consists in detecting and recognizing connected shapes. The technique is applied on small images, called « areas of interest », on which the analysis of the coherence of the constructed objects permits to emit some hypothesis concerning the presence of a set of connected characters. For these areas, a filtering technique detects and recognizes connected shapes.

The results of the application of this technique are very encouraging since the correct classification rate reaches excellent scores if we consider that contextual information, such as the orientation of strings of characters or data issued from dictionaries stored on alpha-numeric data bases, is not yet integrated in the recognition process.

The perspectives of this work are numerous and concern different points. First, concerning the Fourier-Mellin approach, a study is currently under way in order to optimize the construction of the invariant features. More precisely, it is necessary to investigate the choice of features or combination of features to keep, in order to improve the classification rate, while avoiding the "curse of dimensionality" that happens in many pattern recognition and decision-theoretic problems. This optimization can be performed through the use of genetic algorithms, for instance. Our reflections also deal with the improvement of the methodology by adding new and complementary invariant features.

From another point of view, we are currently working at the integration of contextual information, within the system approach. As an example, a set of aligned characters provides some information

concerning a string of characters, and as a consequence, guides the recognition process through the knowledge about the orientation. In these conditions, knowing the orientation of a shape, the recognition process may be coupled with other more usual techniques.

From this point dealing with orientation analysis, our works also deal with he orientation estimation through the use of Fourier Mellin descriptors. Indeed, the descriptors provided by the AFMT allow to estimate with a good reliability the orientation of the analyzed shape. As a consequence, this information should reinforce our recognition process, since the orientation estimated may be compared with the string's one.

All these considerations and the current results make us very optimistic concerning the future of our project since the results and the possible improvements of the methodology seem to indicate that the classification rate should still increase of several points. Our current studies also deal with the implementation of our pattern recognition tool in the context of the interpretation of electricity network documents. The first results of this implementation are very encouraging since the classification rate are similar to the one obtained with the interpretation of telephonic network documents. These tests seem to confirm our opinion about the stability of these invariant in regard with the graphic variability of the shapes to be recognized and also in regard with the important noise which one can find on these different supports. The consequence of the quality of this pattern recognition tool is its integration in a real industrial process.

8 – Bibliography

[Adam 1999]	S. Adam, R. Mullot, J.M. Ogier, C. Cariou, J. Gardes, Y. Lecourtier, Processing of the Connected Shapes in Raster-to-Vector Conversion Process, in *Proc of Graphic RECognition workshop (GREC*'99), Jaipur, India, pp 38-45, 1999.
[Antoine 1991]	D. Antoine, CIPLAN: A model-based system with original features for understanding French Plats, *in Proc. ICDAR'91*, St Malo, France, pp 647-655, 1991
[Antoine 1992]	D. Antoine, S. Collin, K. Tombre, Analysis of technical documents : The REDRAW system, in *Structured Document Analysis*, Baird H.S., Bunke H., Yamamoto K. Editors, Springer-Verlag,.pp 385-402, 1992
[Baffes]	P. Baffes, NETS – Neural Network Simulatot, Sofware Technology Branch NASA, Johnson Space Center, ftp://ftp.technion.ac.il/pub/unsupported/dos/simtel/neurlnet/nasanets.zip
[Bailey 1996]	R.R. Bailey, M. Srinath, Orthogonal moment features for use with parametric and non parametric classifiers, *IEEE Trans. on PAMI*, **18**(4), pp. 389-399, 1996.

[Bamieh 1986]   R. Bamieh, R. De Figueiredo, A general moments invariants/attributed graph method for the three dimensional object recognition from a single image, *IEEE Journal of Robotics Automation,* **2**, pp 240-242, 1986

[Belkasim 1991]   S.O. Belkasim, M. Shridar, M. Ahmadi, Pattern recognition with moment invariants: A comparative study and new results, *Pattern Recognition* **24**(12), pp 1117-1138, 1991

[Boatto 1992a]   L. Boatto and al., An interpretation system for land register maps, *IEEE Computer Magazine,* **25**, pp 25-33, 1992

[Boatto 1992b]   L. Boatto *et al.*, Detection and separation of symbols connected to graphics in line drawings, *Proc. of 11th International Conference on Pattern Recognition*, The Hague, Netherlands, vol. 2, pp 545-548, 1992.

[Casey 1996]   R.G. Casey, E. Lecolinet, A Survey of methods and strategies in character segmentation, *IEEE Trans. on PAMI*, **18**(7), pp. 690-706, 1996

[Chen 1996]   Y. Chen, N.A. Langrana, A.K. Das, Perfecting Vectorized Mechanical Drawings, *Computer Vision and Image Understanding (CVIU),* **63**(2), pp 273-286, 1996

[Chhabra 1998]   K. Chhabra. Graphic Symbol Recognition: An Overview, *Lecture Notes in Computer Science*, vol. 1389, pp 68-79, 1998.

[Dai 1992]   M. Dai, P. Baylou, M. Najim, A efficient algorithm for computation of shape moments from run-length codes or chain codes, *Pattern Recognition*, **25**, pp 1119-1128, 1992

[Das 1997]   A.K. Das, N.A. Langrana, Recognition and Integration Sets in Vectorized Engineering Drawings, *Computer Vision and Image Understanding (CVIU),* **68**(1), pp 90-108, 1997

[Davies 1988]   E.R. Davies, A Modified Hough Scheme for General Circle Location, *Pattern Recognition Letters*, **7**, pp 37-43, 1988.

[Den Hartog 1996a] J.E. Den Hartog , T. K. Ten Kate, J.J. Gerbrands, Knowledge-based interpretation of utility maps, *Computer Vision and Image Understanding (CVIU),* **63**(1), pp 105-117, 1996

[Den Hartog 1996b] J.E. Den Hartog, T.K Ten Kate, J. Gerbrands. Knowledge-Based Segmentation for Automatic Map Interpretation, *Lecture Notes in Computer Sciences*, vol 1072, pp 159-178, 1996

[Derrode 1997]   S. Derrode, F. Ghorbel, Transformée de Fourier-Mellin numérique – Reconstruction et estimation de mouvements d'objets à niveaux de gris », *in Proc GRETSI'97*, Grenoble, France, pp 655-658, 1997

[Deseilligny 1995] M.P. Deseilligny, H. Le Men, G. Stamon, Character String Recognition on maps, a rotation-invariant recognition method, *Pattern Recognition Letters*, **16**, pp 1297-1310, 1995

[Di Zenzo 1996]   S. Di Zenzo, L. Cinque, S. Levialdi, Run-Based Algorithms for Binary Image Analysis and Processing, *IEEE Trans. on PAMI*, **18**(1), pp83-89, 1996.

[Dori 1995a]   D. Dori, Vector-Based Arc Segmentation in the Machine Drawing Understanding System environment, *IEEE Trans. on PAMI*, **17**(11), pp 1057-1068, 1995.

[Dori 1995b]   D. Dori, K. Tombre, From Engineering Drawings to 3-D CAD Models : Are we Ready Now ?, *Computed Aided Design*, **27**(4), pp 243-254, 1995

[Dori 1995c]   D. Dori, Representing Pattern Recognition-Embedded System though Object-Process Diagrams : the Case of the Machine Drawing Understanding System, *Pattern Recognition Letters*, **16**(4), pp 377-384, 1995.

[Dori 1996a]   D. Dori, L. Wenyin, M. Peleg, How to Win a Dashed Line Detection contest, In *Lecture Notes in Computer Sciences*, vol 1072, pp 286-300, 1996

[Dori 1996b]   D. Dori, W. Liu, Vector-based segmentation of text connected to graphics in engineering drawings, P. Perner, PSP. Wang and A. Rosenfeld, eds., *Lecture Notes in Computer Science*, vol 1121, pp. 322-331, 1996.

[Dori 1998]   D. Dori, Y. Velkovitch, Segmentation of Dimensioning Text in Engineering Drawings, *Computer Vision and Image Understanding*, **69**(2), pp 196-201, 1998

[Dori 1999a]   D. Dori, L. Wenyin, Automated CAD Conversion with the Machine Drawing Understanding System : Concepts, Algorithms, and Performances, *IEEE Trans.*

	on SMC – Part A, **29**(4), pp 411-416, 1999
[Dori 1999b]	D. Dori, W. Liu, Sparse Pixel Vectorization : An Algorithm and Its Performance Evaluation, *IEEE Trans. on PAMI*, **21**(3), pp 202-215, 1999.
[Dudani 1977]	S.A. Dudani, K. J. Bredding, R. M. McGhee, Aircraft identification by moment invariants, *IEEE Transactions on Computers,* **26**, (1977), pp 39-45.
[Fletcher 1988]	L.A. Fletcher, R. Kasturi, A Robust Algorithm for Text String Separation from Mixed Text/graphics Images, *IEEE Trans. on PAMI*, **10**(6), pp 910-918, 1988
[Fukumi 1992]	M. Fukumi, S. Omatu, T. Takeda and T. Kosaka, Rotation invariant neural pattern recognition system with application to coin recognition, *IEEE Transactions on Neural Networks,* **3**, pp 272-279, 1992
[Ghorbel 1994]	F. Ghorbel, A complete invariant description for gray level images by the harmonic analysis approach, *Pattern Recognition Letters,* **15**, pp 1043-1051, 1994
[Hu 1962]	M.K. Hu, Visual pattern recognition by moment invariants, *IRE Transactions on Information Theory*, **8**, pp 179-187, 1962
[Janssen 1997]	R. Janssen, A Vossepoel, Adaptative Vectorization of Line Drawing Images, *Computer Vision and Image Understanding (CVIU),* **65**(1), pp 38-56, 1997.
[Joseph 1992]	S.H. Joseph, P. Pridmore, Knowledge-directed Interpretation of Line Drawing Images, *IEEE Trans. on PAMI*, **14**(9), pp 928-940, 1992
[Kamel 1993]	M. Kamel, A. Zhao, Extraction of Binary Character/Graphics Images from Grayscale Document Images, *CVGIP*, **55**(3), pp 202-217, 1993
[Kasturi 1988]	R. Kasturi, J. Alemany, Information Extraction from Images of Paper-Based Maps, *IEEE Transactions on Software Engineering,* **14**(5), pp 671-675, 1988
[Kasturi 1990]	R. Kasturi and al, A System for Interpretation of Line Drawing, *IEEE Trans. on PAMI*, **12**(10), pp 978-992, 1990.
[Khotanzad 1990a]	A. Khotanzad, Y. H. Hong,, Rotation invariant image recognition using features selected via a systematic method, *Pattern Recognition,* **23**, pp 1089-1101, 1990
[Khotanzad 1990b]	A. Khotanzad, Y. H. Hong, Invariant image recognition by Zernike moments, *IEEE Trans. on PAMI*, **12**(5), pp. 489-497, 1990.
[Kita 1992]	N. Kita, Object locating based on concentric circular description, in *Proc. ICPR'92*, Den Hague, Vol. 1, pp 637-641, 1992
[Kong 1996]	B. Kong, I. T. Phillips, R. M. Haralick, A. Prasad, R. Kasturi. A Benchmark: Performance Evaluation of Dashed-Line Detection Algorithms, *Lecture Notes in Computer Sciences*, Vol 1072, pp 270-285, 1996.
[Kulh 1982]	F.P. Kuhl, C.R. Giardina, Elliptic Fourier features of closed contour*, Computer Vision, Graphics and Image Processing*, **18**, pp. 236-258, 1982.
[Lai 1991]	C. Lai, R. Kasturi, Detection of Dashed Lines in Engineering Drawings and Maps, *in Proc. ICDAR'91*, St Malo, France, pp 507-514, 1991.
[Lai 1994]	C. Lai, R. Kasturi, Detection of Dimension Sets in Engineering Drawings, *IEEE Trans. on PAMI*, **16**(8), pp 848-855, 1994.
[Lam 1992]	L. Lam, S.W. Lee, C.Y. Suen, Thinning Methodologies – A comprehensive Survey, *IEEE Trans. on PAMI*, **14**(9), pp 869-885, 1992
[Langrana 1997]	N.A. Langrana, Y. Chen, A.K. Das, Feature identification from vectorized mechanical drawings, *Computer Vision and Image Understanding*, **68**(2), pp. 127-145, 1997.
[Lecun 1995]	Y. LeCun, Y. Bengio. Pattern recognition and neural networks. In M. A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*. MIT Press, 1995
[Lecun]	NETS-5, Image processing research Department, Research Lab AT&T, http://www.research.att.com/~yann/lenet/
[Lefrère 1993]	L. Lefrère, *Contribution au Développement d'Outils pour l'Analyse Automatique de Documents Cartographiques*, PhD Thesis, University of Rouen, 1993.
[Liao 1990]	C.W. Liao, J.S. Huang, Stroke segmentation by Bernstein-Bezier Curve Fitting, *Pattern Recognition*, **23**(5), pp 475-484, 1990.
[Lin 1987]	B.C. Lin, New forms of shape invariants from elliptic Fourier Descriptors, *Pattern Recognition,* **20**, pp 535-545, 1987
[Lin 1991]	B.C. Lin, J. Shen, Fast computation of moment invariants, *Pattern Recognition,*

24, pp 807-813, 1991

[Liu 1997]   W. Liu, D. Dori, A protocol for performance evaluation of line detection algorithms, *Machine Vision and Applications*, **9**(5), pp. 240-250, 1997

[Lu 1998]   Z. Lu, Detection of text regions from digital engineering drawings, *IEEE Trans. on PAMI*, **20**(4), pp. 431-439, 1998.

[Mori 1992]   Mori, C.Y.Suen, K. Yamamoto, Historical review of OCR Research and Development, *Proceedings of the IEEE*, **80**(7), pp 1029-1058, 1992 .

[Nagy 1998]   G. Nagy, A. Samal, S. Seth, T. Fisher, E. Guthmann, K. Kalafala, L. Li, P. Sarkar, S. Sivasubramaniam, Y. Xu, A prototype for adaptative association with streets on maps, *in* K. Tombre and A.K. Chhabra, eds., *Lecture Notes in Computer Science*, 1389, 1998

[O'Gorman 1997]   L. O'Gorman, R . Kasturi, Executive Briefing - Document Image Analysis, *IEEE Computer Society Press*, 1997

[Ogier 1993]   J.M. Ogier, R. Mullot, J. Labiche, Y. Lecourtier, Attribute Extraction for French Map Interpretation, *in Proc. ICDAR'93*, Tsukuba Science City, Japan, pp 672-675, 1993

[Ogier 1998a]   J.M. Ogier, C. Cariou, R. Mullot, J. Gardes and Y. Lecourtier, Interpretation of technical document: Application to French telephonic network, *in Proc. ISAS*'98, Orlando, Florida, pp 457-463, 1998

[Ogier 1998b]   J.M. Ogier, R. Mullot, J. Labiche and Y. Lecourtier, Multilevel approach and distributed consistency for technical map interpretation: application to cadastral maps, *Computer Vision and Image Understanding(CVIU)*, **70,** pp 438-451, 1998

[Pasternak 1995]   B. Pasternak, The Role of Taxonomy in Drawing Interpretation, *in Proc. ICDAR'95*, Montreal, Canada, pp 799-802, 1995

[Pei 1992]   S.C. Pei, C. N. Lin, Normalisation of rotationally symmetric shapes for pattern recognition, *Pattern Recognition,* **25**, pp 913-920, 1992

[Ravichandran 1995]   G. Ravichandran, M. Trivedi, Circular-Mellin features for texture segmentation, *IEEE Trans. on Image Processing,* **4**, pp 1629-1640, 1995

[Reiss 1993]   T.H. Reiss, Recognizing Planar Objects Using Invariants Image Features, in *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 1993.

[Rothe 1996]   I. Rothe, H. Susse, K. Voss, The Method of Normalization to Determine Invariants, *IEEE Trans. on PAMI*, **18**(4), pp 366-379, 1996

[Sahoo 1988]   P.K. Sahoo and al, A survey of Thresholding Techniques, *CVGIP*, **41**(2), pp 233-260, 1988

[Shimotsuji 1992]   S. Shimotsuji, O. Hori, M. Asano, K. Suzuki, F. Hoshino, T. Ishii, A Robust Recognition System for a Drawing Superimposed on a map, *IEEE Computer magazine*, **25**(7), pp 56-64, 1992

[Shimotsuji 1994]   S. Shimotsuji, O. Hori, M. Asano, Robust Drawing Recognition Based on Based-Guided Segmentation, *in Proc. of IAPR Workshop on Document Analysis Systems*, Kaiserslautern, Germany, pp 337-348, 1994

[Suzuki 1990]   S. Suzuki, T. Yamada, MARIS : MAp Recognition Input System, *Pattern Recognition,* **23**, pp 919-933, 1990

[Tan 1998]   N. Tan, Rotation invariant texture features and their use in automatic script identification, *IEEE Tran. on PAMI,* **20**, pp 751-756, 1998

[Taxt 1990]   T. Taxt, J. B. Olafsdottir, M. Daehlen, Recognition of handwritten symbols, *Pattern Recognition,* **23**(11), pp 1155-1166, 1990

[Teague 1980]   M. Teague, Image analysis via the general theory of moments, *Journal of Optical Society of America* **70**, pp 920-930, 1980

[Teh 1988]   C. Teh, R. Chin, On image analysis by the method of moments, *IEEE Trans. on PAMI,* **10**(4), pp 496-512, 1988

[Thomas 1989]   S.M. Thomas, Y.T. Chan, A Simple Approach for the Estimation of Circular Arc center and its Radius, *CVGIP*, **45**, pp 362-370, 1989.

[Tombre 1998a]   K. Tombre, A. K. Chhabra (editors). Graphics Recognition - Algorithms and Systems, *Lecture Notes in Computer Science*, vol. 1389, Springer Verlag, 1998..

[Tombre 1998b]   K. Tombre, Analysis of Drawings : State of the Art and Challenges, *Lecture Notes in Computer Science*, vol. 1389, pp 257-264, 1998.

[Trier 1995a]   O.D. Trier, T. Taxt, A. K. Jain, Data capture from maps based on gray scale

topographic analysis, in *Proc. of ICDAR*'95, Montreal, Canada, pp923-926, 1995

[Trier 1995b]   Trier, T. taxt, Evaluation of Binarization Methods for Document Images, *IEEE Trans. on PAMI*, **17**(3), pp 312-315, 1995

[Trier 1996]   O.D. Trier, A. K. Jain, T. Taxt, Features extraction methods for character recognition – a survey, *Pattern Recognition*, **29**, pp 641-662, 1996

[Trier 1997]   O.D. Trier, T. Taxt, A.K. Jain, Recognition of Digits in Hydrographic Maps : Binary Versus Topographic Analysis , *IEEE Trans. on PAMI*, **19**(4), pp 399-404, 1997

[Vaxivière 1992]   P. Vaxivière, K. Tombre, CELESTIN : CAD conversion of Mechanical Drawings, *IEEE Computer Magazine,* **25**, pp 46-54, 1992

[Wenyin 1998a]   L. Wenyin, D. Dori, A Generic Intergrated Line Detection Algorithm and Its Object-Process Specfication, *Computer Vision and Image Understanding*, **70**(3), pp 420-437, 1998

[Wenyin 1998b]   L. Wenyin, D. Dori, Incremental Arc Segmentation Algorithm and Its Evaluation, *IEEE Trans. on PAMI*, **20**(4), pp 424-431, 1998.

[Yu 1994]   Y. Yu, A. Samal, S. Seth, Isolating Symbols from Connected Lines in a Class of Engineering Drawings, *Pattern Recognition*, **27**(3), pp 391-404, 1994

[Yu 1997]   Y. Yu, A. Samal, S.C. Seth, A system for Recognizing a Large Class of Engineering Drawings, *IEEE Trans. on PAMI*, **19**(8), pp 868-890, 1997