



An Overview of Advances of Pattern Recognition Systems in Computer Vision

Kidiyo Kpalma, Joseph Ronsin

► To cite this version:

Kidiyo Kpalma, Joseph Ronsin. An Overview of Advances of Pattern Recognition Systems in Computer Vision. Prof. Goro Obinata and Dr. Ashish Dutta. Vision Systems, Advanced Robotic Systems, pp.26, 2007. <hal-00143842>

HAL Id: hal-00143842

<https://hal.science/hal-00143842v1>

Submitted on 27 Apr 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

An Overview of Advances of Pattern Recognition Systems in Computer Vision

Kidiyo KPALMA and Joseph RONSIN
IETR (Institut d'Electronique et de Télécommunications de Rennes)
 UMR – CNRS 6164
 Groupe Image et Télédétection
 Institut National des Sciences Appliquées (INSA) de Rennes
 {kidiyo.kpalma | joseph.ronsin}@insa-rennes.fr

1. Introduction

First of all, let's give a tentative answer to the following question: **what is pattern recognition (PR)?** Among all the possible existing answers, that which we consider being the best adapted to the situation and to the concern of this chapter is: "pattern recognition is the scientific discipline of machine learning (or artificial intelligence) that aims at classifying data (patterns) into a number of categories or classes". **But what is a pattern?**

In 1985, Satoshi Watanabe (Watanabe, 1985) defined a pattern as "the opposite of chaos; it is an entity, vaguely defined, that could be given a name." In other words, a pattern can be any entity of interest which one needs to recognise and/or identify: it is so worthy that one would like to know its name (its identity). Examples of patterns are: a pixel in an image, a 2D or 3D shape, a typewritten or handwritten character, the gait of an individual, a gesture, a fingerprint, a footprint, a human face, the voice of an individual, a speech signal, ECG time series, a building, a shape of an animal.

A pattern recognition system (PRS) is an automatic system that aims at classifying the input pattern into a specific class. It proceeds into two successive tasks: (1) the **analysis** (or description) that extracts the characteristics from the pattern being studied and (2) the **classification** (or recognition) that enables us to recognise an object (or a pattern) by using some characteristics derived from the first task.

The classification scheme is usually based on the availability of the training set that is a set of patterns already having been classified. This learning strategy is termed as supervised learning in opposition to the unsupervised learning. A learning strategy is said to be unsupervised if for the system is not given an a priori information about classes; it establishes the classes itself based on the regularities of the features. Features are those measurements which are extracted from a pattern to represent it in the features space. In other words, pattern analysis enables us to use some features to describe and represent it instead of using the pattern itself. Also called characteristics, attributes or signatures the recognition efficiency and reliability are dependent on their choice.

Pattern recognition constitutes an important tool in various application domains, but unfortunately, that is not always an easy task to carry out. Commonly, one can encounter four major methodologies in PRSs; which are: statistical approach, syntactic approach, template matching, neural networks. In this chapter, our remarks and details will be directed, principally, towards systems based on the statistical approach since it is the more commonly used in practice.

1.1 Statistical approach

Typically, statistical PRSs are based on statistics and probabilities. In these systems, features are converted to numbers which are placed into a vector to represent the pattern. This approach is most intensively used in practice because it is the simplest to handle.

In this approach, patterns to be classified are represented by a set of features defining a specific multidimensional vector: by doing so, each pattern is represented by a point in the multidimensional features space. To compare patterns, this approach uses measures by observing distances between points in this statistical space. For more details and deeper considerations on this approach, one can refer to (Jain, 2000) that presents a review of statistical pattern recognition approaches.

1.2 Syntactic approach

Also called structural PRSs, these systems are based on the relation between features. In this approach, patterns are represented by structures which can take into account more complex relations between features than numerical feature vectors used in statistical PRSs (Venguerov & Cunningham, 1998). Patterns are described in hierarchical structure composed of sub-structures composed themselves of smaller sub-structures.

As explained in (Sonka et al., 1993), the shape is represented with a set of predefined primitives called the codebook and the primitives are called codewords. For example, given the codewords on the left of figure 1, the shape on the right of the figure can be represented as the following string S , when starting from the pointed codeword on the figure:

$$S = d b a b c b a b d b a b c b a b \quad (1)$$

The system parses the set of extracted features using a kind of predefined grammar. If the whole features extracted from a pattern can be parsed to the grammar then the system has recognised the pattern. Unfortunately, grammar-based syntactic pattern recognition is generally very difficult to handle.

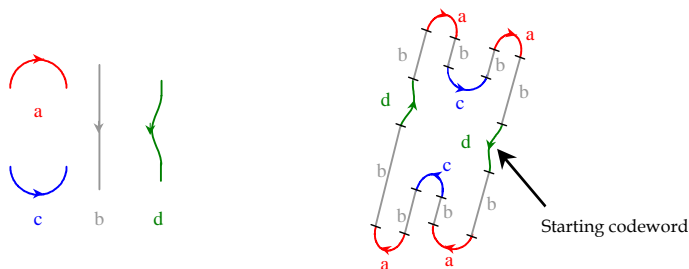


Fig. 1. Example of syntactic description features

1.3 Template matching

Template matching approach is widely used in image processing to localize and identify shapes in an image. In this approach, one looks for parts in an image which match a template (or model). In visual pattern recognition, one compares the template function to the input image by maximising the spatial cross-correlation or by minimising a distance: that provides the matching rate.

The strategy of this approach is: for each possible position (in the image), each possible rotation, or each other geometric transformation of the template, compare each pixel's neighbourhood to this template. After computing the matching rate for each possibility, select the largest one, that exceeds a predefined threshold. It is a very expensive operation while dealing with big templates and/or large sets of images (Brunelli & Poggio, 1997 ; Roberts & Everson, 2001 ; Cole et al., 2004). Figures 2 illustrate a pattern recognition based on the template matching approach. Figure 2.a is the input image I, Fig.1.b represents two templates (K representing letter 'K' and P letter 'P'). Figures 2.c and 2.d represent, respectively, the normalized cross-correlation of I with K and the normalized cross-correlation of I with P. On these two images, the cross-correlation peaks surrounded by a circle indicate the location of the most matching letter in the input image. On figure 2.e, we have superposed the templates on the input image, accordingly to the coordinates of corresponding correlation peaks. For this study, we didn't take the rotation and the scaling into account: from the result, it clearly appears that this approach retrieves only the shape that matches perfectly the model (size and rotation). This explains why only one 'K' (the rotated one) and only one 'P' (the down-scaled one) are recognised.

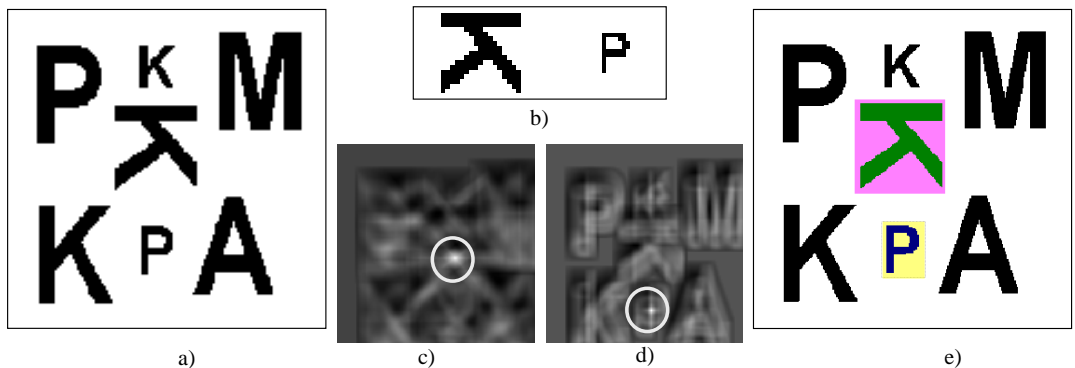


Fig. 2. Illustration of the template matching method

1.4 Neural networks

Typically, an artificial neural network (ANN) is a self-adaptive trainable process that is able to learn to resolve complex problems based on available knowledge. A set of available data is supplied to the system so that it finds the most adapted function among an allowed class of functions that matches the input.

An ANN-based system simulates how the biological brain works: it is composed of interconnected processing elements (PE) that simulate neurones. Using this interconnection (or synapse), each neurone (or PE) can pass information to another. As can be seen on figure

3, these interconnections are not necessarily binary (on or off) but they may have varying weights defined by the weight matrix W : the weight applied to a connection results from the learning process and indicates the importance of the contribution of the preceding neurone in the information being passed to the following neurone. Figure 3 shows a simple neural network representing the Perceptron as defined by Frank Rosenblatt in 1957. On this example, the output Out_j ($j=1$ or 2) is defined by a weighted combination of the inputs. In the reference (Abdi, 1994), the author presents a nice introduction to ANNs.

Besides these approaches, one can encounter other methodologies like those based on fuzzy-set theoretic, genetic algorithms. In some applications, hybrid methodologies combine different aspects of these approaches to design more complex PRSs. In (Liu et al., 2006), the authors present an overview of pattern recognition approaches and the classification of their associated applications.

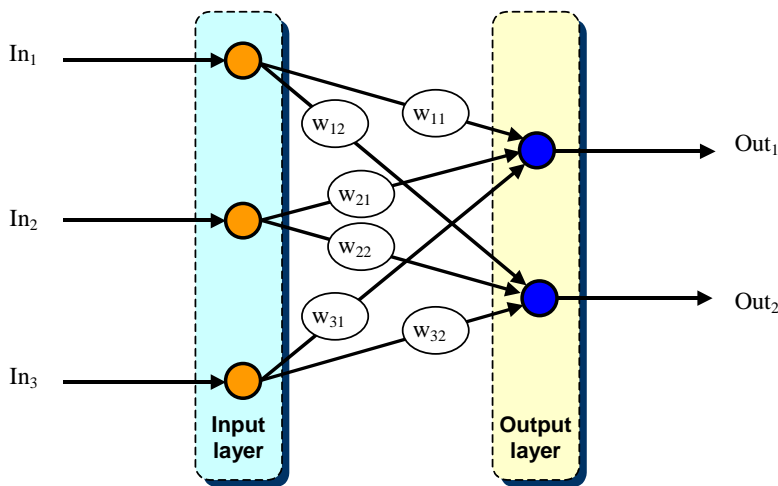


Fig. 3. Example of neural network

In the remainder of this chapter, we will develop three sections. First, we present a generic scheme of a pattern recognition system. Then we give an overview of the advances of different PRSs and some examples of their applications. Last, as an illustration, we present a specific application example based on our MSGPR (Multi-Scale curve smoothing for Generalised Pattern Recognition) description method. As presented further, MSGPR is a multi-scale method we have developed for describing planar objects by analysing their boundary.

2. A generic scheme of a pattern recognition system

From now, our concerns will be primarily focused on PRSs in computer vision. Commonly, in this field, the input is one or more images and the output is one or more images with eventually, some semantic and/or textual entities.

In figure 4, we represent a generic scheme of a (statistical) PRS. This figure summarises the principal aspects of a PRS in computer vision. On this figure the two successive tasks can be

observed: on one hand, the analysis/description task (see ❶ on figure 4) and on the other hand the classification/recognition task (see ❷ on figure 4).

After features are extracted, the features selection that may follow aims at reducing the number of features to be provided to the classification process. Features that are likely to improve discrimination are retained and the others are discarded. During this processing, higher level features can be derived by combining and/or transforming low level features, e.g. by applying the so called independent component analysis (ICA) (Roberts & Everson, 2001); this operation thus leads to the reduction of the dimension of the feature space.

These features must be as discriminative as possible to reduce false alarms due to misclassification during the second task. Efficient features must also present some essential properties such as:

- translation invariance: whichever be the location of the pattern, it must give exactly the same features,
- rotation invariance: extracted features must not vary with the rotation of the pattern,
- scale invariance: scale changing must not affect the extracted features,
- noise resistance: features must be as robust as possible against noise i.e. they must be the same whichever be the strength of the noise that affects the pattern,
- statistically independent: two features must be statistically independent,
- compact. The number of retained features is not too large. It must also be fast in extraction time and in matching,
- reliable: as long as one deals with the same pattern, the extracted features must remain the same.

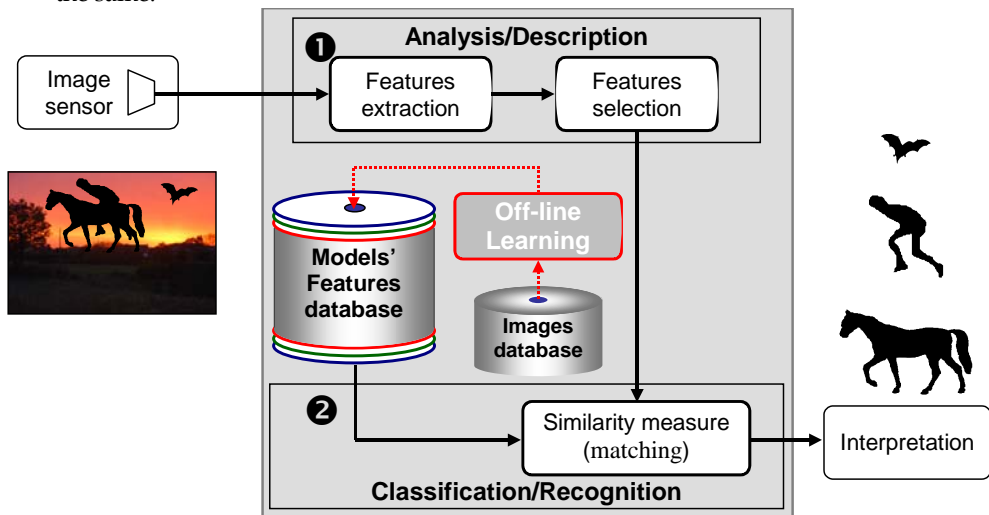


Fig. 4. A generic PRS scheme

During the classification task, the system uses the features extracted in the analysis stage from each of the patterns to compare. As illustrated on figure 4, features are extracted from the patterns of the database during an off-line learning processing. This enables to create features database before each query occurs: by proceeding this way, one doesn't need to compute features of models at each query. To compare two patterns, the system uses a metric that measures a kind of distance (the similarity or the dissimilarity) to assess how

similar are two patterns: it is an expression of the distance between the points representing the two patterns in the features space. This procedure gives the similarity index or similarity score between two patterns. In some cases (probably the most natural way), the similarity index is given in terms of a rate varying from 0% for totally different patterns to 100% for perfectly similar patterns (Kpalma & Ronsin, 2006). Some commonly used metrics are Minkowski distance, cosine distance, Hausdorff distance, Mahalanobis Distance, (Veltkamp & Hagedoorn, 2001 ; Zhang, 2002) or city block distance and Euclidian distance that are particular Minkowski distances. The following paragraph illustrates formalism of some of them.

Let $V_A(a_1, a_2, \dots, a_N)$ and $V_B(b_1, b_2, \dots, b_N)$ be the features vectors representing patterns A and B in a N-dimensional features space ; examples of distances are defined by the following expressions.

City block distance (d_1)

$$d_1(V_A, V_B) = \sum_{i=1}^N |a_i - b_i| \quad (2)$$

Euclidian distance (d_2)

$$d_2(V_A, V_B) = \sqrt{\sum_{i=1}^N (a_i - b_i)^2} \quad (3)$$

Cosine distance (d_3)

$$d_2(V_A, V_B) = 1 - \cos(\theta) = 1 - \frac{V_A V_B^T}{\|V_A\| \times \|V_B\|} = 1 - \frac{\sum_{i=1}^N a_i \times b_i}{\sqrt{\sum_{i=1}^N (a_i)^2} \times \sqrt{\sum_{i=1}^N (b_i)^2}} \quad (4)$$

where θ is the angle between the two vectors V_A and V_B .

Figure 5 shows an example of three vectors V , U and W represented in 2D space. As it can be seen on this example the value of the similarity/dissimilarity depends on the used distance (metric). In the tables on this figure, d_3 gives the same distance between U and W , on one hand, and between V and W , on the other hand, ($d_3(U,W) = d_3(V,W) = 0.15$) but it gives 0 distance between U and V . This leads to confusions, because a distance of 0 that also means vectors equality, may lead to the decision that the patterns to be compared are the same. A particular attention must be paid while choosing a distance. In (Kpalma & Ronsin, 2006) we have proposed a cosine-based distance that enables to remove the ambiguity of the distance between collinear vectors. Since the obtained distance varies from a metric to another, one must be very careful and be sure to use the same metric during all the procedure.

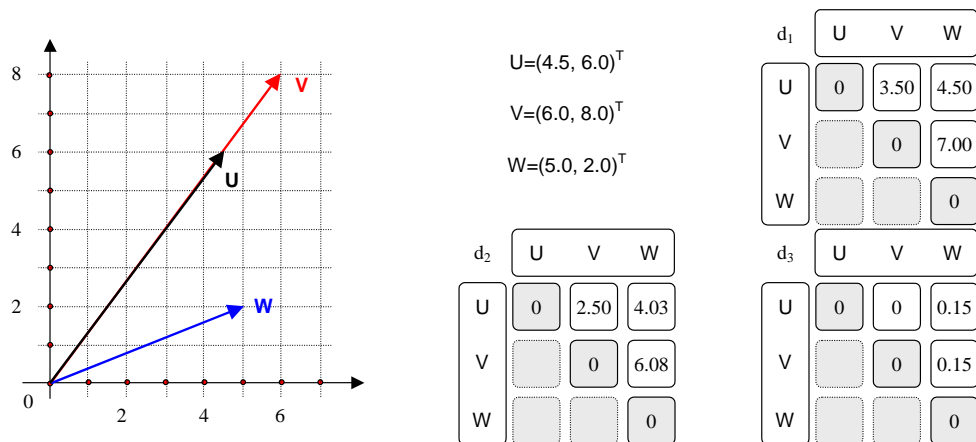


Fig. 5. Examples of similarity measures between two vectors depending on the chosen metric

3. Pattern recognition applications and an overview of advances

Pattern recognition is studied in many fields, including psychology, ethnology, forensics, marketing, artificial intelligence, remote sensing, agriculture, computer science, data mining, document classification, multimedia, biometrics, surveillance, medical imaging, bioinformatics and internet search. Pattern recognition helps to resolve various problems such as: optical character recognition (OCR), zip-code recognition, bank check recognition, industrial parts inspection, speech recognition, document recognition, face recognition, gait recognition or gesture recognition, fingerprint recognition, image indexing or retrieval, image segmentation (by pixels classification)...

In (Pal & Pal, 2002; ...) number of experts address the problem of pattern recognition and present basic concepts involved. One can find the evolution of pattern recognition ; this enables the reader to establish a categorisation of the existing PRSs according to the used methodology and the application.

In (Kuncheva, 2004), the author addresses the non-trivial concept of forgetting in the challenging field of machine learning in non-stationary changing environments. This point of view is essential in on-line diagnosis when using medical imaging: indeed while dealing with PR in real world, the pattern being studied is subject to variation with respect to time. A possible solution is to continuously update the classifier. By doing so, the classifier must be able to "forget" the outdated knowledge. The idea behind this concept is to design an adaptive training system that is able to self-adapt itself accordingly to the changing of the pattern being studied.

Pattern recognition is also applied in more complex fields like data mining (DM) also called knowledge-discovery in databases (KDD). This emerging topic includes the process of automatically searching large volumes of data for patterns such as association rules. As defined in (Frawley et al., 1992), the DM "is the nontrivial extraction of implicit, previously unknown, and potentially useful information from data. Given a set of facts (data) F , a language L , and some measure of certainty C , we define a pattern as a statement S in L that describes relationships among a subset F_S of F with a certainty c , such that S is simpler (in

some sense) than the enumeration of all facts in F_s . A pattern that is interesting (according to a user-imposed interest measure) and certain enough (again according to the user's criteria) is called knowledge. The output of a program that monitors the set of facts in a database and produces patterns in this sense is discovered knowledge.

3.1 Pattern recognition in robotics

The applications of PRSs in robotics are permanent. More recently, Mario E. Munich and his co-authors (Munich et al., 2006) have presented a summary on this subject. In this paper, they show that recent advances in computer vision have given rise to a robust and invariant visual pattern recognition technology based on extracting a set of characteristic features from an image. With visual pattern recognition systems, a robot may acquire the ability to explore its environment without user intervention ; it may be able to build a reliable map of the environment and localize itself in the map: this will help the robot achieve full autonomy. Examples of robots using visual pattern recognition approaches are the Sony's AIBO ERS-7, Yaskawa's SmartPal, and Phillips' iCat.

In robotics, visual servoing or visual tracking is of high interest. For example visual tracking allows, robots to extract themselves the content of the observed scene as a human observer can do it by changing his different perspectives and scales of observation. François Chaumette (Chaumette, 1994), has addressed the problem and proposed some solutions in a closed loop system based on vision-based task. In (Chaumette, 2004), he proposes various visual features based on the image moments to characterise planar objects in visual-servoing schemes.

3.2 Pattern recognition in biometrics

The biometric authentication takes increasing place in various applications ranging from personal applications like access control to governmental applications like biometric passport and fight against terrorism. In this applications domain, one measures and analyses human physical (or physiological or biometric) and behavioural characteristics for authentication (or recognition) purposes. Examples of biometric characteristics include fingerprints, eye retinas and irises, facial patterns and hand geometry measurement, DNA (Deoxyribonucleic acid). Examples of biometric behavioural characteristics include signature, gait and typing patterns. This helps to identify individual people in forensics applications.

Reference (Jain et al., 2004a) is an interesting starting point to pattern recognition approaches and systems in biometric. This paper gives a brief overview of the field of biometrics and summarizes some of its advantages, disadvantages, strengths, limitations, and related privacy concerns. In (Jain et al., 2004b), the authors also address the problem of the accuracy of the authentication and that of the individual's right to the security, to the privacy and to the anonymity.

The reader is encouraged to have a look on the article presented in (Jain & Pankanti, 2006). The authors of this article address a problem of identity stealing through a true story and then they present some current or forthcoming systems based on biometric PRSs that will help prevent identity stealing.

3.3 Content-based image retrieval

Content-based image retrieval systems aim at automatically describing images by using their own content: the colour, the texture and the shape or their combination. As explained in (Sikora, 2001; Bober, 2001), image retrieval has become an active research and development domain since the early 1970s. During the last decade the research on image retrieval became of high importance. The most frequent and common means for image retrieval is to index them with text keywords. If this technique seems to be simple, it becomes rapidly laborious and fastidious while facing large volumes of images. On the other hand, images are rich in content so, to overcome difficulties due to the huge data volume, the content-based image retrieval emerged as a promising mean for retrieving images and browsing large images databases

With the simultaneous rapid growth of computer systems and the growing huge availability of digital data, such pattern recognition systems become increasingly necessary to help browse databases and find the desired information within a reasonable time limit. Accordingly to this observation, systems like CBIR (Content-Based Image Retrieval), QBIC (Query By Image Content), QBE (Query By Example) need more attention and take more and more place in the concerns of the researchers (Mokhtarian et al., 1996 ; Trimeche et al., 2000 ; Veltkamp & Tanase, 2001 ; Veltkamp & Hagedoorn, 2001). With query by example, the user supplies a query image and the PRS finds images of the database that are most similar to it based on various low-level features like colour, texture or shape. With query by sketch, the user draws roughly the image he is looking for and the PRS locates images of the database that match the best the sketch. In the reference (Veltkamp & Tanase, 2001), are reported various CBIR systems. After a brief description of CBIR system, the authors present different kinds of existing systems along with the features involved.

In the context of image indexing, CBIR systems use content information as summarised in figure 6. An image can then be described by using features derived from colour, texture, shape or a combination of those features.

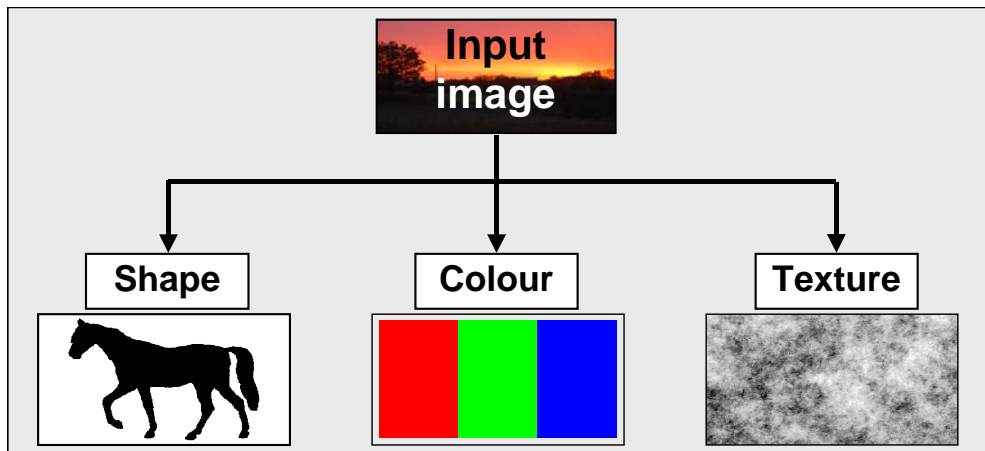


Fig. 6. Content-based image description features

3.3.1 Colour-based features

Colour features are based on colour distribution inside the image. There are many approaches to define colour-based features: colour space dominant, colour histogram or colour space. Various colour representation space exist: red-green-blue (RGB) space, hue-saturation-value (HSV) space or those based on the international commission on illumination (or CIE: *commission internationale de l'éclairage*) CIELUV space, CIELAB space, CIEXYZ space. From these representation features are defined based the colour histograms. There are different types of colour histograms depending on how the colour space is partitioned. The fixed binning for all images based on scalar linear quantisation, the adaptive binning based on an adaptive quantisation and the clustered binning based on the concept of vector quantisation. Some particular distances between histograms or main modes of histograms are used to measure the similarity/dissimilarity between colour histograms: Euclidian distance, histogram quadratic distance, histogram intersection distance (Smith & Chang, 1996), Jeffrey divergence, Kullback-Leibler divergence earth mover's distance. In the current description of the colour within MPEG-7, the following colour spaces are supported: RGB, YCrCb, HSV, hue-min-max-difference (HMMD), Linear transformation matrix with reference to RGB and monochrome (Martinez, 2004).

3.3.2 Texture-based features

For each pixel of the image, one can determine the histogram of grey levels in predefined neighbouring region centred on that pixel. Distribution of pairs of grey levels for a given spatial relation on pixels can be observed in co-occurrence matrix $M(i,j)$ (Haralick, 1973). Examples of various grey level co-occurrence matrices (GLCM) features defined by Haralick are based on these co-occurrence matrices. In Table 1, by considering a textured image with grey levels ranging from 0 to $L-1$, we present some of these texture features.

| | |
|---------------------------|---|
| Angular Second Moment | $ASM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} M(i, j)^2$ |
| Contrast | $C = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - j)^2 M(i, j)$ |
| Inverse Difference Moment | $IDM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{M(i, j)}{1 + (i - j)^2}$ |
| Homogeneity | $H = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{M(i, j)}{1 + i - j }$ |
| Entropy | $E = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} M(i, j) \ln(M(i, j))$ |

Table 1. Example of textures features

3.3.3 Shape-based features

There are many approaches (Coster & Chermant, 1985 ; Kpalma, 1994 ; Sossa, 2000), to estimates some properties of the shapes. We present, below, some samples of these properties. Figure 7 shows various shapes and the corresponding measures of their properties.

The **elongation** (EL) indicates how long is the pattern relatively to its width. It is defined by the following expression:

$$EL = 100 \frac{\lambda_m}{\lambda_M} \quad (5)$$

λ_m and λ_M being, respectively, the smallest and the largest eigenvalues of the inertia matrix of the shape. Also called elongation factor or elongation coefficient, this parameter varies from 0% for long but thick shapes to 100% for isotropic shape (see Fig. 7.e and Fig. 7.f).

The compactness measures how branchy or how tortuous is the shape. For a given 2D shape, let A be the enclosed area and P the perimeter ; the compactness is defined by:

$$CO = 100 \frac{4\pi A}{P^2} \quad (6)$$

The compactness varies from 0% for very branchy or very tortuous shapes to 100% for compact shapes like a circle (see Fig. 7.a and Fig. 7.c).

The **mass deficit coefficient** (MD) measures the area variation between the shape and the minimum enclosing circle centred on the centre of gravity of the shape. For a shape with area A, let S_C be the area of the circumscribed circle, then the mass deficit area is defined as follows:

$$MD = 100 \frac{S_C - A}{S_C} \quad (7)$$

The **mass excess coefficient** (ME) measures the area variation between the shape and the maximum enclosed circle centred on the centre of gravity of the shape. For a shape with area A, let S_I be the area of the inscribed circle, then the mass deficit area is defined as follows:

$$ME = 100 \frac{A - S_I}{A} \quad (8)$$

The two previous parameters, give another estimation of the compactness: they vary from 0% for compact shapes (e.g. a cercle) to 100% for spread out tortuous patterns (see Fig. 7.a and Fig. 7.d)

The **isotropic factor** (IF) tells how isotropic is the pattern: it indicates how regular is the shape around its centre of gravity. For a given 2D shape, let R_m be its minimal radius and R_M its maximal radius then the IF parameter is defined by:

$$IF = 100 \frac{R_m}{R_M} \quad (9)$$

The isotropic factor varies from 0% for anisotropic shapes to 100% for isotropic shapes like a circle (see Fig. 7.a and Fig. 7.d).

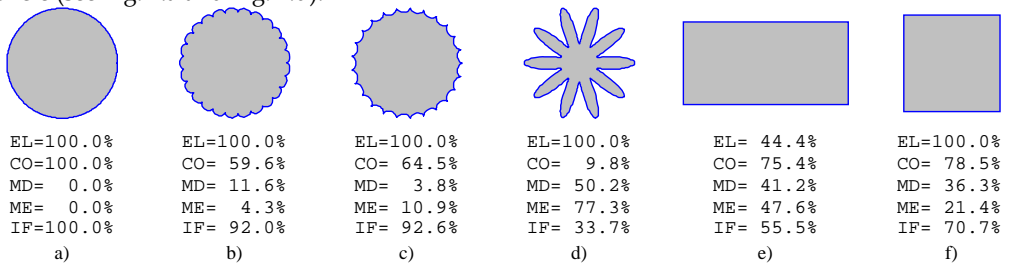


Fig. 7. Various shapes and examples of shape-based features

In the context of shape description, D. Zhang summarized very well the situation (Zhang & Lu, 2004). Figure 8 shows the flowchart of shape description approaches in a pattern recognition system. Typically, there are two kinds of approaches in shape description: the contour-based approach and the region-based one.

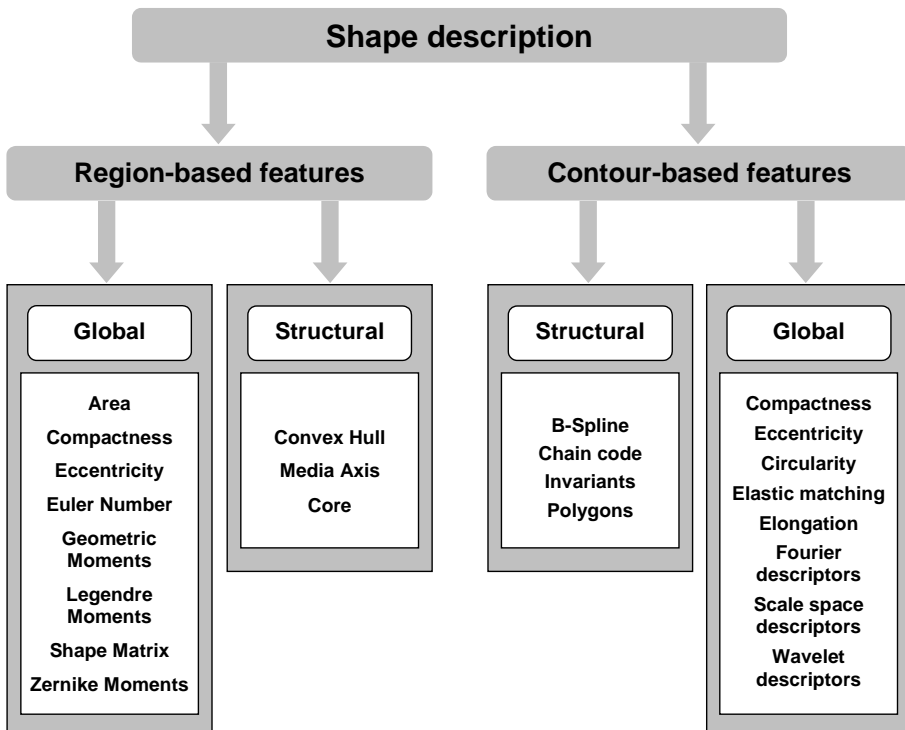


Fig. 8. A classification of shape description approaches

Contour-based approach

Contour-based approaches extract shape features from the only contour in two possible ways: structural or global. In the structural approach, the contour is divided into sub-section

to generate strings or trees according to a particular syntax. The similarity between two shapes is then measured by matching their strings or their trees.

While dealing with the contour in the global way, an appropriate technique is used to extract primitive features from the integral contour: eccentricity, perimeter, circularity... From these basic features, one defines a multidimensional vector representing the shape in the features space. From this representation, the similarity measure or the matching of two shapes is done by directly measuring a specific distance between their feature vectors.

For contour-based shape description, MPEG-7 working group (Bober, 2001 ; Martinez, 2004) has selected the so-called Curvature Scale-Space (CSS) representation which is proved to capture perceptually meaningful features of the shape (Mokhtarian et al., 1996 ; Matusiak & Daoudi, 1998 ; Lindenberg, 1998 ; Mokhtarian & Bober, 2003).

A CSS image, represented on figure 9, is a multi-scale organization of the invariant local features of a 2-D contour: it consists of the curvature zero-crossing points recovered from the contour at multiple scales of resolution. The features extracted from the CSS image consist of the coordinates of the peaks of the CSS image. Scale decreasing is obtained through progressive low-pass filtering by convolutions of a parametric representation of the contour data with Gaussian filters of increasing width. This representation carries a number of important properties, such as:

- it captures very well characteristic features of the shape, enabling similarity-based retrieval,
- it reflects properties of the perception of human visual system and offers good generalization,
- it is robust to non-rigid motion,
- it is robust to partial occlusion of the shape,
- it is robust to perspective transformations, which result from the changes of the camera parameters and are common in images and video,
- it is compact.

Some of the above properties of this descriptor are illustrated in figure 11, each frame containing very similar images according to CSS, based on the actual retrieval results from the MPEG-7 shape database. In figure 9, we represent two shapes and their corresponding CSS images. On the CSS images (bottom row) we have superposed the peaks points that are used to generate features (Mokhtarian & Bober, 2003).

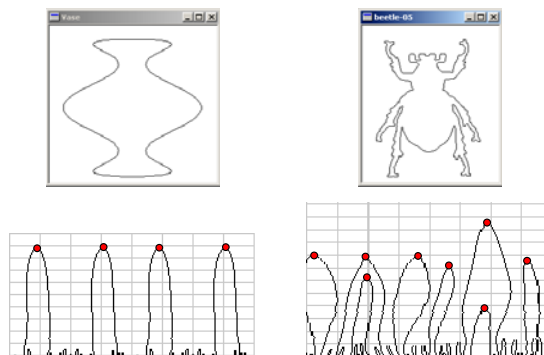


Fig. 9. Example of contours (top row) and the corresponding CSS image with the peaks points (bottom row)

Region-based approach

In region-based approaches, all pixels surrounded by the shape boundary are taken into account to generate the shape descriptor. Like in the case of contour-based approaches, we encounter the same two different ways in region-based shape description: global and structural one. In the structural approach, the shape is decomposed into sub-regions to generate a tree to represent the shape. In the global way, one computes some characteristic features to generate a vector to represent the shape. Common global features derived from a region-based approach are: geometrical moment invariants, shape matrix, area, compactness, eccentricity, Euler number, geometric moments, Legendre moments, Zernike moments... For region-based shape description, the MPEG-7 working group (Bober, 2001 ; Martinez, 2004) has selected the angular radial transform (ART). It is a moment-based approach for a 2D region-based shape description. In (Ricard et al., 2005) the authors proposed a generalization of the ART approach to describe 2D and 3D shapes for content-based image retrieval purpose.

The contour-based approaches are more appealing than region-based approaches because they involve less computation complexity, than the region-based ones, with enough discriminating efficiency. It is also demonstrated that characteristic information about a shape lie essentially on its contour features. The main drawback of contour-based descriptors is that they are more subject to noise and variations than region-based ones.

Figure 10 shows examples of shapes and illustrates situations for which the contour-based or the region-based descriptors are most suitable.

A shape may consist of just one single region (see Fig.10.a-c) or a set of several regions as well as regions with some holes inside them as illustrated in figures 10.d-f. Since the region-based descriptors make use of all pixels constituting the shape, they can describe any kind of shapes. They are more suitable than the contour-based descriptors to handle complex shape consisting of holes in the object or several disjoint regions (see Fig.10.d-f) in a single descriptor. Indeed, for contour-based descriptors, these shapes consist not of a single contour but of multiple contours leading, thus, to multiple descriptors.

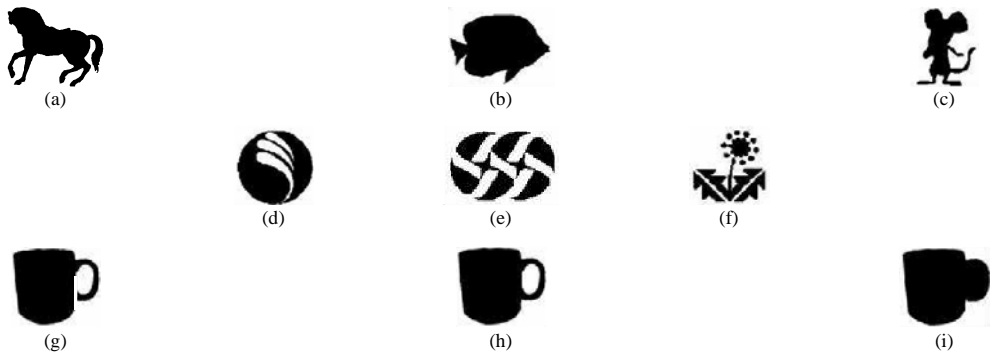


Fig. 10. Examples of various shapes

Figures 10.g-i show very similar shapes from images of a same cup. They only differ by the handle: shape 10.g has a crack at the lower handle while the handle in 10.i is filled. When comparing these shapes:

- the region-based shape descriptor will consider 10.g and 10.h similar but different from 10.i,
- the contour-based shape descriptor will consider 10.h and 10.i similar but different from 10.g.

As illustrated by MPEG-7 (Martinez, 2004), a challenge for a pattern descriptor is to enable the recognition of a pattern even if it has undergone various deformations namely partial occultation (Fig.11.a) and non-rigid deformation (Fig. 11.b).

Figure 11.a, according to (Martinez, 2004), illustrates the robustness to partial occultation: indeed, in this figure, one can note that the tails or the legs of the horses are sometimes occulted but they are recognised to be from the same class. As presented in (Mokhtarian, 1997 ; Petrakis, 2002) , this is possible because of the ability of the descriptor to handle local properties. On figure 11.c are represented various shapes are classified in the same class based on the visual perceptual similarity

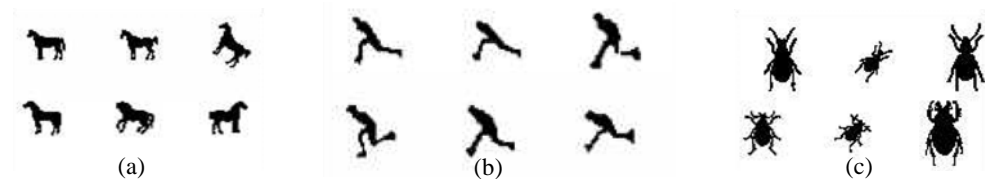


Fig. 11. a) robustness to partial occlusion, b) robustness to non-rigid deformation, c) perceptual similarity among different shapes

The choice of a description method will depend on the application so, sometimes, one needs to make a compromise. Nevertheless, MPEG-7 has set some essential principles to evaluate the suitability of shape descriptor: retrieval accuracy, compactness, generics, low computation complexity, robustness and the ability to represent a shape in hierarchical way: from coarse to fine representation.

3.4 An overview of the advances in pattern recognition

Remco C. Veltkamp and Mirela Tanase presented in (Veltkamp & Tanase, 2000) a large panel of CBIR systems. Various approaches of the state of the art in content-based image retrieval and video retrieval are explored along with the features used in each approach, they also describe the matching functions used. This overview enables to confirm, as it was said before, that commonly designed CBIR systems are generally based on visual features such as colour, texture and shape.

In (Iqbal & Aggarwal, 2002) is presented CIRES (Content-based Image REtrieval System), an online system for retrieval in image libraries. It is done to extend the retrieval paradigm, which was mostly limited to colour and texture analyses, by using image structure. Image structure was extracted via hierarchical perceptual grouping principles.

In (Mittal, 2006) the author presents an overview of the content-based retrieval along with different strategies in terms of syntactic and semantic indexing for retrieval. After an analysis of the matching techniques used and the learning methods the author addresses some directions for future research in the content-based retrieval domain.

Recently, N. Snavely and co-authors (Snavely et al., 2006) have presented a system that consists of 3D image-based modelling and representation of an unorganised images taken

by different cameras in different conditions. The challenging aim of the system is to use the content-based information to browse an image database and reply to questions like:

- "where was I? Tell me where I was when I took this picture"
- "what am I looking at? Tell me about objects visible in this image by transferring annotations from similar images"

To do this, they used the SIFT (Scale Invariant Feature Transform) keypoints detector that was shown to be transformation invariant (Lowe, 2004).

Among the various forthcoming systems, we can encounter MPEG-7. Formally named "Multimedia Content Description Interface", MPEG-7 aims at managing data in the way that content information can be retrieved easily. It is under development by the Moving Picture Coding Experts Group (MPEG) that is a working group of ISO/IEC(*) standards organization. It is in charge of the development of international standards for video and/or audio compression, decompression, processing and representation. This group has also developed well-known standards that are MPEG-1, MPEG-2 and MPEG-4. MPEG-1, MPEG-2 and MPEG-4 also make content available but MPEG-7 enables to find the desired content. MPEG-7 visual description tools consist of basic structures and descriptors that cover basic visual features: colour, texture, shape, motion, localization. Each category consists of elementary and sophisticated descriptors (Sikora, 2001; Bober, 2001). One must note that MPEG-7 addresses many different applications in various environments, thus it needs to provide a standard flexible and extensible framework for describing audio-visual data.

4. Application example based on the MSGPR method

In (Kpalma & Ronsin, 2006) we have presented an original pattern description approach based on the multi-scale analysis of the contour of planar objects. This proposed approach summarises the different presented considerations in this chapter. It is well known that some objects, especially natural ones, exists with a more or less large range of scales; and that the aspect of the object can change from one scale to another. Without a priori information about the distance of observation inside a given scene, an interesting challenge can be: to found an object without any precision about its scale of observation. Faced with this situation, it is very difficult to significantly describe a pattern using only one meaningful scale. To overcome this problem, increasingly more pattern description techniques are based on multi-scale or multiresolution representation methods (Lindeberg, 1998). Within this context, methods based on the pattern itself (Torres-Méndez et al., 2000 ; Kadyrov & Petrou, 2001 ; Belongie et al., 2002 ; Grigorescu & Petkov, 2003) exist as well as methods based on pattern contour behaviour (Matusiak & Daoudi, 1998 ; Roh & Kweon, 1998 ; Wang et al., 1999 ; Latecki et al., 2000).

This study deals exclusively with methods based on the pattern contour. Called MSGPR (A Multi-Scale curve smoothing for Generalised Pattern Recognition) this scale-space (Mokhtarian et al., 1996 ; Matusiak & Daoudi, 1998 ; Wang et al., 1999 ; Mokhtarian & Bober, 2003) method is based on multi-scale smoothing of a planar pattern contour. This method is

(*)ISO/IEC stands for International Standards Organization/International Electro-technical Committee.

totally translation and rotation insensitive and as showed in the initial studies it is also robust against scale change for a large range of scaling and resistant to additive noise.

4.1 Description of the MSGPR method

The framework of the MSGPR can be broken down into four main stages as follows (see Fig.12):

1. the input contour is separated into two parameterised functions,
2. both functions are low-pass filtered (smoothed),
3. scale adjustment is then applied to both filtered functions so that the corresponding smoothed contour has the same scale as the input one,
4. finally, the intersection points map (IPM) is generated by detecting the intersection points of the input contour and the smoothed scale-adjusted one.

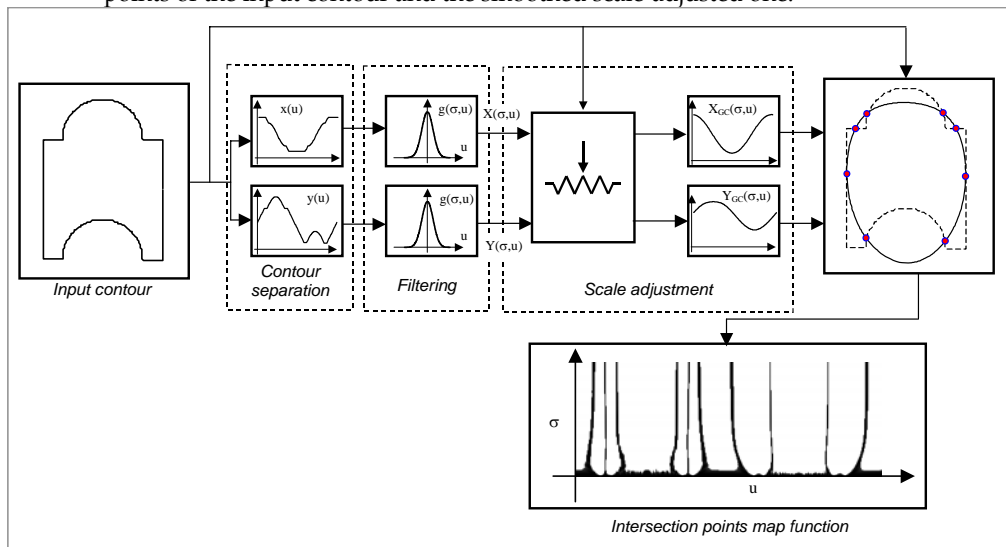


Fig. 12. MSGPR description scheme

4.1.1 Coordinate separation

The input contour is represented by a series of points defined by their (x,y) coordinates. First, the input contour is separated into two functions $x(u)$ and $y(u)$ which are functions of the normalised curvilinear u parameter that varies from 0 to 2π relative to the curve length. Each point of the curve is then represented by its parameterised coordinates $(x(u), y(u))$.

4.1.2 Curve smoothing

Functions $x(u)$ and $y(u)$ are then gradually smoothed by decreasing the filter bandwidth. Similarly to the curvature scale space (CSS) method (Mokhtarian et al., 1996 ; Matusiak & Daoudi, 1998 ; Wang et al., 1999 ; Mokhtarian & Bober, 2003) or other scale-space methods, smoothing is based on the Gaussian filters $g(\sigma, u)$ with standard deviation σ :

$$g(\sigma, u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{u^2}{2\sigma^2}} \quad (10)$$

The filtered functions are then given by: $X(\sigma, u) = g(\sigma, u) * x(u)$ and $Y(\sigma, u) = g(\sigma, u) * y(u)$ so that each $(x(u), y(u))$ point on the input contour leads to the $(X(\sigma, u), Y(\sigma, u))$ point on the output smoothed contour.

Since the bandwidth is conversely proportional to σ , it is clear that the bandwidth decreases as σ increases. Thus the filter cuts increasingly lower so that the output functions move towards their mean values when σ tends towards infinity.

4.1.3 Scale adjustment

After low-pass filtering, the scale adjustment system stretches the output contour so that it reaches the same scale as the input one and so that both contours intersect at certain points. Figure 13 shows an example of a contour and two smoothed ones ($\sigma=10$ and $\sigma=180$) after they have been scale-adjusted. The input contour C_0 and smoothed scale-adjusted contours $C_{GC}(\sigma)$ are now on the same scale so that they can intersect.

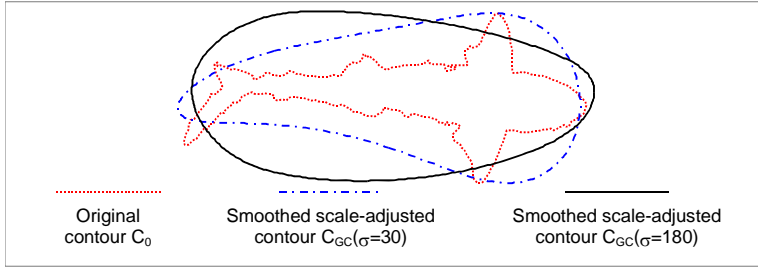


Fig. 13. Example of a contour and two smoothed scale-adjusted ones ($\sigma=30$ and $\sigma=180$)

4.1.4 Definition of the IPM function

By increasing σ , the output contour moves towards a convex curve that has some intersection points with the input contour. By marking these intersection points for each σ , we obtain the intersection points map (IPM) function defined below which characterises the pattern.

After the scale adjustment system, the IPM function is generated as follows. For each σ value, we define a function which is an image in the scale-space (u, σ) plane so that (see Fig.14):

- $IPM(u, \sigma) = 0$ (black) if the $(x(u), y(u))$ point is an intersection point between the original curve and the filtered scale-adjusted one,
- $IPM(u, \sigma) = 1$ (white) if point $(x(u), y(u))$ is not an intersection point.

Figure 14 shows examples of contours (left column) and the corresponding IPM functions (right column). On this figure, intersection points are indicated by (1) through (6) or (8), for the contour in Fig.14.a or for that in Fig.14.c, respectively. On the right column, one can see the marks corresponding to those intersection points in the IPM representation. As can be seen on this figure, the IPM function is characteristic of the contour it is derived from.

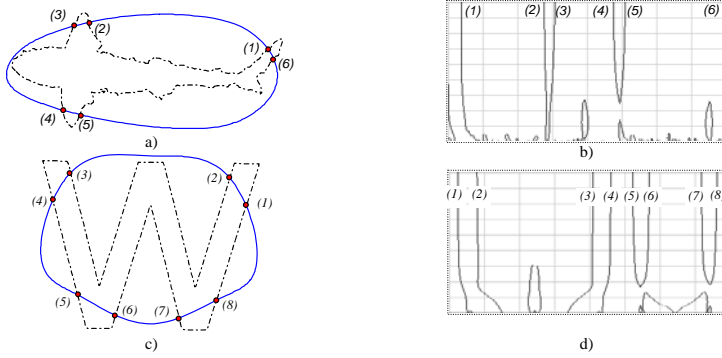


Fig. 14. Example of the IPM function

4.1.5 Features definition and selection

After generating the IPM function, the following stage, but not the least one is features definition and selection. In (Kpalma & Ronsin, 2006) we used the circular distance between IPM points at various scale values. To extract these characteristic features, we first set the scale parameter σ to σ_0 value (e.g. $\sigma_0 = 180$). Then, for each pattern:

- we consider the IPM points at the set σ_0 and select two consecutive p_a and p_b points which are, circularly, the furthest apart in the IPM function as illustrated in figure 15,
- we determine the circular distance between both points to produce the first d_1 component of the V_0 features vector,
- the next components of V_0 are distances coming after d_1 :

$$V_0 = (d_1, d_2, \dots, d_{N-1}, d_N) \quad (11)$$

To benefit from multi-scale information of the IPM function, we can define a set of M values of σ ($\sigma_0, \sigma_1, \dots, \sigma_{M-1}$) and determine the V_i feature vectors ($i=0, 1, 2, \dots, M-1$) corresponding to the σ_i scales. The global V features vector is then produced by a concatenation of the individual V_i scale vectors as follows:

$$V = (V_0, V_1, \dots, V_{M-1}) \quad (12)$$

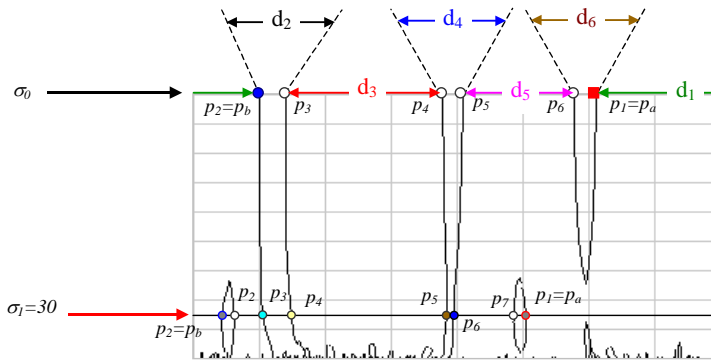


Fig. 15. Example of the IPM function

4.1.6 Similarity measure

To measure the matching rate between two V_A and V_B attribute vectors associated to two patterns, we define a similarity function as follows:

$$\text{SimScore}(V_A, V_B) = 50(1 + \cos(\gamma)) \frac{\text{Min}(\|V_A\|, \|V_B\|)}{\text{Max}(\|V_A\|, \|V_B\|)} \quad (13)$$

where γ is the angle between both vectors and where $\|\cdot\|$ indicates the module of a vector. This function ranges from 0% for very different vectors to 100% for perfectly matching vectors.

4.2 Application to car plate character recognition

In this section, we present a system we have developed to illustrate pattern recognition systems. This application can be classified into the group of the contour-based statistical approaches. Our application illustrates an automatic reading of the number plates by using their digital images. Applying the IPM-based features we carry out the automatic recognition of the characters of the number plate. Figure 16 shows two images of plates written with different fonts: the difference appears more clearly for digit '3' out of the two plates.



Fig. 16. Examples of number plates images

4.2.1 Character recognition procedure

The recognition procedure is carried out into three stages as depicted in figure 17:

- (I) **edge (or contours) detection** that will enable to obtain contours delimiting each character in the image (Fig. 18.a et Fig. 18.b). One must note that this stage is very important in our process, because, the effectiveness of character recognition will depend on it.
- (II) **contour extraction**: in this stage, one considers only the external (or the outer) boundary (Fig. 18.c), because only these contours are taken into account. As for the stage (I), one must pay particular care to the extraction of the characters so that they are continuous and closed, without self-intersection.
- (III) **character recognition**: at this last stage, we apply our IPM-based description approach to extract the features and to integrate them into the identification process to measure the similarity score between each extracted character and the models of the data base. In this application, the similarity measure is based on the *SimScore* function defined by equation (13).

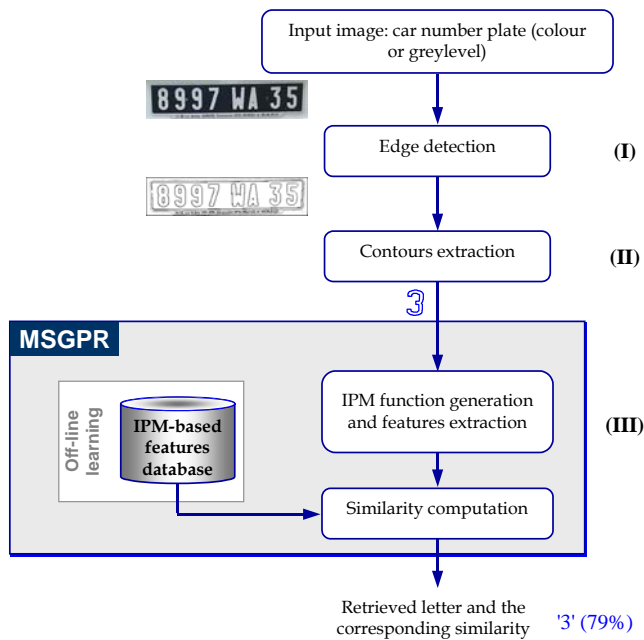


Fig. 17. An overview of an automatic number plate reading system.

4.4.2 Experimental results

Figures 18.a and 18.b represent the output images of the edge detection when applied to images corresponding to figure 16. The figure 18.c presents the set of the extracted haracters from figures 18.a and 18.b. On figure 18.d we present a sample set of characters of the database: this base consists of the character set "bold.chr" of Borland®.

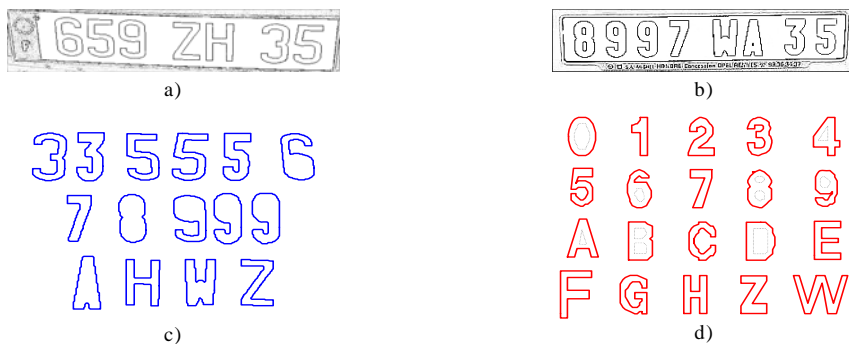


Fig. 18. a) and b) detected edges - c) extracted contours from a) and b) - d) examples of the content of the database.

It must be noted that in this study, the database is composed of only one font while the query characters come from two different fonts. In order to improve the identification

results, a possible solution would be to integrate in the database, all the possible fonts used to create car plates. Figures 19 show some results obtained from the input images presented on figure 16. On these figures, we represent some results of character recognition: on each figure, the contour on the upper left corner represents the query contour. Following contours in left-to-right and top-to-down scanning, represent eight retrieved contours which give the highest similarity scores.

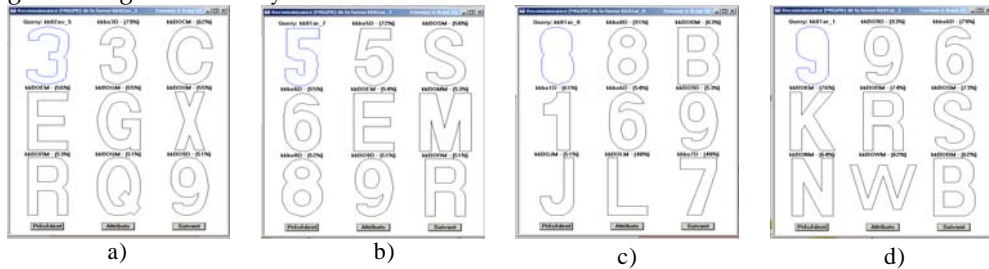


Fig. 19. Examples of the recognition output

As can be seen on these figures, the identification of different characters is effective enough: for each query, the identified character (the most similar: the character next to the query in figures 19.a-d) is exactly the required character. Thus, for the query '3', we identify the letter '3' with a similarity score of 79%. Table 2 summarises the three highest similarity scores for the contours presented on figure 19. For the contour '9' as a query, we retrieved the digit '9' with a similarity score up to 96% followed by the digit '6' with a similarity score of 79%. One can notice that the contour '6' of the used font is not other than the contour '9' which underwent a rotation of 180°: this explains that the digit '6' occupies the second position during the retrieval process. In the same way, the topological similarity between the digit '5' and the letter 'S' or between the digit '8' and the letter 'B' results in the appearance of 'S' and 'B', respectively, into the second position in the retrieval ranging. In spite of this topological similarity, specific properties of each character lead to sufficiently important variations of similarity scores to avoid mistakes.

| Query | Retrieved character (Similarity score) | | |
|-------|--|-----------|-----------|
| '3' | '3' (79%) | 'C' (62%) | 'E' (56%) |
| '5' | '5' (72%) | 'S' (58%) | '6' (55%) |
| '8' | '8' (91%) | 'B' (63%) | '1' (61%) |
| '9' | '9' (96%) | '6' (79%) | 'K' (76%) |

Table 2. Retrieved characters and the corresponding similarity scores.

5. Conclusion

As mentioned pattern recognition does not appear as a new problem. A lot of studies have been performed on this scientific field and a lot of works are currently developed. Pattern recognition is a wide topic in machine learning. It aims to classify a pattern into one of a number of classes. It appears in various fields like psychology , agriculture, computer

vision, robotics, biometrics... With technological improvements and growing performances of computer science, its application field has no real limitation. In this context, a challenge consists of finding some suitable description features since commonly, the pattern to be classified must be represented by a set of features characterising the pattern. These features must have discriminative properties: efficient features must be affined transformations insensitive. They must be robust against noise and against elastic deformations due, e.g., to movement in pictures.

Through the application example based on our MSGPR method, we have illustrated various aspects of a PRS. With this example, we have illustrated the description task that enabled us to extract multi-scale features from the generated IPM function. By using these features in the classification task, we identified the letters from a car number plate so that we automatically retrieved the license number of a vehicle.

The research topic of pattern recognition is under continuous development and in perpetual progress. With the large volumes of digital images, the challenge for pattern recognition in computer vision is now the development of a CBIR-like system: system that is able to retrieve useful information by using the only content of the input image. With the growing huge availability of digital images, pattern recognition takes more and more place in our daily life to help us while browsing and looking for the desired information in a reasonable time limit.

Pattern recognition appears as integrated into the forthcoming standard MPEG-7 via indexing approaches. Such standardization does not bring restriction to a domain: it gives synergy of best actors mixing challenge and cooperation. And moreover international standardization occurs as a requirement from different applications so it meets all conditions for large diffusion. Standards use the possibilities of last technological developments, and drive strong investments and focus research on the concerned domain. As it has been observed, for example, for coding when it was integrated inside different MPEG standards, the integration of pattern recognition inside MPEG-7 will boost its last developments.

6. References

- Abdi, H. (1994). A neural network primer. *Journal of Biological Systems*, Vol. 2, No. 3, pp. 247-281
- Belongie, S., Malik, J., and Puzicha, J., Shape matching and object recognition using shape contexts. *IEEE PAMI-24*, No 24, pp 509-522, 2002
- Bober, M. (2001). MPEG-7 Visual Shape Descriptors, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 6, pp 716-718.
- Bruckstein, A. M., Rivlin, E., and Weiss, I. (1996). Recognizing objects using scale space local invariants, *Proceedings of the 1996 International Conference on Pattern Recognition (ICPR '96)*, August 25-29, pp. 760-764, Vienna, Austria.
- Bruckstein, A., Katzir, N., Lindenbaum, M., and Porat, M. (1992). Similarity invariant signatures for partially occluded planar shapes, *IJCV*, Vol. 7, No. 3, pp. 271-285.
- Brunelli, R. and Poggio, T. (1997). Template Matching: Matched Spatial Filters And Beyond, *Pattern Recognition*, Vol. 30, No 5, pp. 751-768
- Chaumette, F. (2004), Image Moments: A General and Useful Set of Features for Visual Servoing, *IEEE Transactions on Robotics*, Vol. 20, No. 4, pp. 713-723

- Chaumette, F. (1994). Visual servoing using image features defined upon geometrical primitives, International 33rd IEEE Conference on Decision and Control, Vol. 4, pp. 3782-3787, Orlando, Florida
- Cole, L.; Austin, D. and Cole, L. (2004). Visual Object Recognition using Template Matching, Australasian Conference on Robotics and Automation 2004
- Coster, M. and Chermant, J.-L. (1985). Précis d'Analyse d'Images, Editions du CNRS, 15, quai A. France, Paris, 1985
- Frawley, W. J.; Piatetsky-Shapiro, G. & Matheus, C. J. (1992). Knowledge Discovery in Databases: An Overview, *AI Magazine* 13(3), pp. 57-70
- Grigorescu, C., and Petkov, N. (2003). Distance Sets for Shape Filters and Shape Recognition. *IEEE Trans. Image Processing* 12(9).
- Haralick, R.M. (1979), Statistical and structural approaches to texture, *Proceedings of the IEEE*, No. 5, Vol. 67, pp. 786-804
- Haralick, R.M., Shanmugam, K. and Dinstein, I. H. (1973). Textural features for image classification, *IEEE Transaction on Systems, Man and Cybernetics*, Vol. SMC-3, n°6, pp. 610-621
- Iqbal, Q. and Aggarwal, J. K. (2002). CIREs: A System for Content-based Retrieval in Digital Image Libraries, Seventh International Conference on Control, Automation, Robotics and Vision (ICARCV), Singapore, pp. 205-210, December 2-5, 2002
- Jain, A. K. and Pankanti, S. (2006). A Touch of Money, *IEEE Spectrum*, vol. 43, no. 7, pp. 22-27, July 2006.
- Jain, A. K.; Ross, A. and Prabhakar, S. (2004a). An Introduction to Biometric Recognition, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 14, No. 1, January 2004
- Jain, A. K., Pankanti, S., Prabhakar, S., Hong, L., Ross, A., and Wayman, J. L. (2004b). Biometrics: A Grand Challenge, *Proceedings of the 17th International Conference on Pattern Recognition*, Vol. 11, August 2004, pp. 935-942.
- Jain, A. K.; Duin R. P.W. and Mao, J. (2000). Statistical Pattern Recognition: A Review, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, pp. 4-37
- Kpalma, K., and Ronsin, J. (2006). Multiscale contour description for pattern recognition, Elsevier Science Inc, *Pattern Recognition Letters*, Vol.27, No.13, pp 1545-1559, 1 October 2006
- Kpalma, K., and Ronsin, J. (2003). A Multi-Scale curve smoothing for Generalised Pattern Recognition (MSGPR), Seventh International Symposium on Signal Processing and its Applications (ISSPA), pp 427-430, Paris, France.
- Kpalma, K. (1994). Caractérisation de textures par l'anisotropie de la dimension fractale, *Proceedings of the 2nd African Conference on Research in Computer Science (CARI)*, October 1994, Ouagadougou, Burkina Faso.
- Kadyrov A., Petrou, M. (2001). Object descriptors invariant to affine distortions. *Proceedings of the British Machine Vision Conference, BMVC'2001*, Manchester, UK.
- Kuncheva, L. I. (2004). Classifier Ensembles for Changing Environments, *Proc. 5th International Workshop on Multiple Classifier Systems*, Cagliari, Italy, Springer-Verlag, LNCS, Vol. 3077, 1-15
- Latecki, L. J., Lakamper, R., and Eckhardt, U. (2000). Shape Descriptors for Non-rigid Shapes with a Single Closed Contour, *IEEE Conf. On Computer Vision and Pattern Recognition (CVPR)*, pp. 424-429, 2000

- Lindeberg, T. (1998). Principles for Automatic Scale Selection, Technical report ISRN KTH NA/P--98/14--SE. Department of Numerical Analysis and Computing Science, KTH (Royal Institute of Technology), S-100 44 Stockholm, Sweden.
- Lindeberg, T. (1994). *Scale-Space Theory in Computer Vision*, Kluwer Academic Publishers, Dordrecht, Netherlands.
- Liu, J., Sun, J. and Wang, S. (2006). Pattern Recognition: An overview, *International Journal of Computer Science and Network Security (IJCSNS)*, Vol. 6, No.6, June 2006
- Lowe, D. G. (2004). Distinctive image features from scale invariant keypoints, *IJCV*, 60 (2):91-110, 2004.
- Martinez, J. M., (editor), (2004), MPEG-7 Overview (version 10), ISO/IEC JTC1/SC29/WG11 N6828, Palma de Mallorca, October 2004
- Martinez, J.M. (2002). Standards - MPEG-7 overview of MPEG-7 description tools, part 2., *IEEE Multimedia* 9 (3), July-Sept. 2002, pp. 83 -93
- Matusiak S., Daoudi M. (1998). Planar Closed Contour Representation by Invariant Under a General Affine Transformation, *IEEE International Conference on System, Man and Cybernetics (IEEE-SMC'98)*, pp. 3251-3256, October 11-14, Hyatt Regency La Jolla, San Diego, California, USA.
- Mittal A. (2006). An Overview of Multimedia Content-Based Retrieval Strategies, *Informatica, International Journal of Computing and Informatics*, Vol. 30, No. 3, pp. 347-356
- Mokhtarian, F., and Bober, M. (2003). *Curvature Scale Space Representation: Theory, Applications and MPEG-7 Standardization*. Kluwer Academic.
- Mokhtarian, F. (1997). Silhouette-Based Occluded Object Recognition through Curvature Scale Space, *Machine Vision and Applications*, Vol. 10, No. 3, pp. 87-97.
- Mokhtarian, F., Abasi, S., and Kittler, J. (1996). Efficient and Robust Retrieval by Shape Content through Curvature Scale Space, in *Proceedings International Workshop on Image Databases and Multimedia Search*, pp 35-42, Amsterdam, The Netherlands.
- Mokhtarian, F., and Mackworth, A. K. (1992). A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-14, N° 8.
- Munich, M. E.; Pirjanian, P.; Di Bernardo, E.; Goncalves, L.; Karlsson, N. and Lowe, D. (2006). Application of Visual Pattern Recognition to Robotics and Automation, *IEEE Robotics & Automation Magazine*, pp.72-77, September 2006
- Pal, S.K. & Pal, A., (Editors). (2002). *Pattern recognition: from classical to modern approaches*, World Scientific, ISBN No. 981-02-4684-6, Singapore
- Petrakis, E. G.M.; Diplaros, A. and Milios, A. (2002). Matching and Retrieval of Distorted and Occluded Shapes Using Dynamic Programming, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 11, pp. 1501-1516
- Ricard, J., Coeurjolly, D. and Baskurt A. (2005). Generalizations of angular radial transform for 2D and 3D shape retrieval, *Elsevier Science Inc, Pattern Recognition Letters* Volume 26, Issue 14 , 15 October 2005, Pages 2174-2186
- Roberts, S. and Everson, R. (2001). *Independent Component Analysis- principles and practice*, Cambridge University Press, ISBN 0521792983
- Roh, K.-S., Kweon, I.-S. (1998). 2-D object recognition using invariant contour descriptor and projective refinement, *Pattern Recognition*, Vol. 31, N° 4, pp. 441-455.

- Smith, J. and Chang, S. F. (1996). Tools and Techniques for Color Image Retrieval. in IS&T/SPIE proceedings of Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases IV vol. 2670, pp. 1630-1639, San Jose, CA, February 1996.
- Snaveely, N., Seitz, S. M. and Szeliski, R. (2006). Photo tourism: Exploring photo collections in 3D, ACM Transactions on Graphics (SIGGRAPH Proceedings), 25 (3), pp. 835-846.
- Sonka, M.; Hlavac, V. and Boyle, R. (1993). *Image Processing, Analysis and Machine Vision*, Chapman & Hall, London, UK, 1993, pp. 193-242
- Sossa, H., 2000. Object Recognition, Summer School on Image and Robotics, INRIA Rhône-Alpes, France.
- Sun, K. B. and Super, B. J. (2005). Classification of Contour Shapes Using Class Segment Sets Full text, Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 2
- Torres-Méndez, L. A., Ruiz-Suárez, J. C., Sucar, L. E. and Gómez, G. (2000). Translation, Rotation, and Scale-Invariant Object Recognition, IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews, Vol. 30, No. 1, pp 125-130.
- Trimeche M., Alaya Cheikh F., and Gabbouj, M. (2000). Similarity Retrieval of Occluded Shapes Using Wavelet-Based Shape Feature, Proc. SPIE International Symposium on Internet Multimedia Management Systems (VV10), Boston, Massachusetts, USA.
- Vapillon, A.; Collin, B. and Montanvert, A. (1998). Analyzing and Filtering Contour Deformation, International Conference on Image Processing (ICIP), Chicago, Illinois, USA.
- Wang Y.-P., Lee, S.L., and Toraichi, K. (1999). Multiscale curvature-based shape representation using B-spline wavelets, IEEE Transactions on Image Processing, Vol. 8, No 11, pp 1586-1592.
- Sikora, T. (2001). The MPEG-7 Visual Standard for Content Description—An Overview, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 11, No. 6, June 2001
- Veltkamp, R C and Tanase M. (2001). Content-based retrieval systems: a survey, Technical Report UU-CS-2000-34, citeseer.ist.psu.edu/veltkamp00contentbased.html
- Veltkamp, R. C.; Burkhardt, H. & Kriegel, H.-P. (2001). *State-Of-The-Art in Content-Based Image and Video Retrieval*, ISBN 1-40200-109-6, Kluwer Academic Publishers.
- Veltkamp, R. C., & Hagedoorn, M. (2001). *State-of-the-art in shape matching*. In *Principles of Visual Information Retrieval*, M. Lew (editor), Springer, ISBN 1-85233-381-2, pp 87-119.
- Venguerov, M. & P. Cunningham, P. (1998). Generalised Syntactic Pattern Recognition as a Unifying Approach in Image Analysis, LNCS, Vol. 1451, pp913-920, Springer Verlag, Sydney, (Australia)
- Watanabe, S. (1985). *Pattern recognition: human and mechanical*. Wiley, 1985
- Zhang, D., and Lu, G. (2004). Review of shape representation and description techniques, Pattern Recognition, Vol.37, pp 1-19.
- Zhang, D. (2002). Image Retrieval Based on Shape, PhD dissertation, Faculty of Information Technology, Monash University, Australia
- Liu, J., Sun, J. and Wang, S. (2006). Pattern Recognition: An overview, International Journal of Computer Science and Network Security (IJCSNS), Vol. 6, No.6, June 2006