Using the Median Distance to compare object shapes in content-based image retrieval

Giancarlo Iannizzotto[1] Antonio Puliafito[2] Lorenzo Vita[1]

[1] Dipartimento di Matematica, University of Messina
C.da Papardo - Salita Sperone
98166 Messina - Italy
E-mail: ianni@scirocco.unime.it, vita@mat520.unime.it
[2] Istituto di Informatica, University of Catania
Viale A. Doria 6, 95025 Catania - Italy
E-mail: ap@iit.unict.it

Abstract

Turning Angles Representation (TA) is considered one of the most interesting methods for representing object shapes in content-based image retrieval systems. Nevertheless, the distance commonly used to measure the similarity between shapes represented by TAs, the Euclidean one, is generally too sensitive to small variations in shapes.

In this paper we present a new distance between shapes represented by TA, namely the *Median distance*, specially devised to minimize the effects of small variations in shapes. Its analytical properties are discussed and experimental results are provided and compared with those obtained by applying traditional techniques based on Euclidean distance. The Median distance has been implemented in the Automatic Image Storage and Retrieval (AISR) system, which allows storage and content-based retrieval of 2-D images.

Keywords: Shape distance, Turning Angles, Shape matching, Content-based image retrieval.

1 Introduction

For several years now, archives of digital images have played an important role in a large number of applications ranging from medicine (the storing of

CAT scan, magnetic resonance and x-ray images, and diagnosis by images in general) [16] [20], to chemistry and biology, cartography (telesurveying), meteorology [15] astronomy, from personnel management [11] to judicial enquiries (court files).

In the past images were stored by human operators, who characterized them with alphanumerical strings, more or less referring to the contents of the images, which were then used in subsequent search operations. Modern multimedia technology, however, suggests the development of completely different databases, in which images can be stored and subsequently sought not only by means of key words but also by graphic keys, i.e. keys that are more similar to the actual data stored; this kind of database would thus acquire certain important characteristics:

- Queries could be made by giving a description of the subjects being sought, even a complex and not necessarily accurate one, in terms of both shapes and other graphic features such as colour or texture; suitable user interfaces would allow these descriptions to be introduced directly in graphic form, thus relieving the user of the task of translating graphic concepts into key words;
- the system would be able to automatically select the images which come closer to the description provided (*nearest neighbour query*), according to a certain similarity criterion which has to be as close as possible to that used by a human being;
- alternatively, it would be possible to automatically select all the images whose similarity with the model provided falls within a certain threshold previously established by the user (*range queries*).

Obviously, in each of the search modes outlined the crucial point is measurement of the similarity between two images; this can be done by comparing them in their entirety (*whole-matching*) or by comparing parts of them (*sub-pattern matching*) [16]. A particular instance of sub-pattern matching, which is of great importance, is *object matching*, on which we will focus in this paper. An object is a portion of an image which somehow acquires significance as an entity separate from the rest of the scene; this characterization derives from research into artificial vision, the main aim of which is to identify objects present in a scene but which is also useful in image retrieval.

If we want to characterize an image according to the objects it contains (the background, if there is one, is obviously also an object), it is necessary to define a mode of representation which is capable of describing it completely in terms of objects and the relations between them; significant examples are to be found in the field of medicine, where it is often necessary to search through large databases (CAT, x-rays, etc) to find images which correspond to certain pathological conditions featuring the presence of spots and shadows in precise absolute and relative positions. A large amount of work done in recent years has not taken this need into account [2]

[14] [10], a solution to which can be provided by using relational structures such as Attributed Relational Graphs (ARG) [21], in which the objects are stored in the nodes and their relative positions in the arcs. The problem of matching ARGs is then one of isomorphism between graphs and has been investigated thoroughly in the field of computer vision [3][5]. Alternative solutions provide for use of hierarchical structures such as trees, objects with inheritance, etc.

Whatever storage and comparison techniques are used, if they consist of object-matching the crucial point is still definition of a criterion for comparison. With this aim in mind, it is necessary to identify a set of attributes, or features, such as to:

1. describe the object;
2. accept an appropriate measure of similarity, i.e. a distance between the objects;
3. act as a base for correct representation of the image.

The features used over the years by researchers in the field of both artificial vision and image databases are numerous [19], [7] [14], and several are often used together in order to obtain a better characterization of an object starting from an approximate description. Of these features, however, shape can be considered to be the most significant [22]: combinations of area, perimeter, shape factors, barycentre, main axes and/or symmetry and sets of invariant algebraic moments have been used, [24] [10] [23], alone or in conjunction with other features such as colour and texture. The simultaneous use of several features allows queries of the type "look for all images with an object having a shape similar to this one and with this composition of colour and texture" [6]; Information about the shape of an object is very often accompanied by information about its colour or texture because finding a definition for the similarity between shapes which corresponds to the human concept of similarity is an extremely difficult problem [22].

Once a representation for an object has been found by identifying an appropriate set of features, a multidimensional feature space can be identified in which each object corresponds to a point. To be able to search for an object in this space it is therefore necessary to define a distance in it: this corresponds to defining a relationship of equivalence for the set of objects, by means of which it will be possible to establish whether two elements of the set are identical or not (exact matching), and a relationship of order, by which to establish an order of similarity (nearest neighbour). This function has to correspond, in the values supplied, to the criterion of similarity chosen, in the sense that the distance between two objects has to be null if and only if they coincide, and the distance between two different objects has to be proportional to the difference between them. In addition, the measure has to avoid "false dismissals", i.e. it must not allow two objects satisfying the similarity conditions to be considered different on the basis of the distance calculated [7].

Literature provides a variety of shape representation techniques: most of them derive from research into computer vision and image analysis, while others have been developed specially for image storing. In the case of 2-D images, it is a widely accepted fact that most of the information about shape focuses on the outlines of objects; this has led many researchers to focus on the representation of outlines rather than internal points [7] [22] [1] [12]. An interesting comparison between various methods has been made on a large image database by Scassellati, Alexopoulos and Flickner [22]: according to the results they obtained, the most efficient method of representation for image retrieval seems to be the one based on Turning Angles (TA) [1].

Representation by Turning Angles has a great number of advantages over other techniques: it is more or less independent of translations of objects and can be made invariant to rotations and scale factors; it is also very natural, as it describes polygons in terms of their angles, a technique used by humans as well; last but not least, it gives a compact description in the form of a unidimensional signal which is much more simple to deal with than signals with more dimensions.

As we shall see below, however, the Euclidean distance defined in literature [1] [22] is very sensitive to small variations in outlines which cause great variations in the distance between objects: if we consider these small variations as noise superimposed on the signal representing the turning angles, we can say that the distance is highly sensitive to noise. In [1], Arkin et al. emphasize this point, stating that this sensitivity is most obvious in the presence of non-uniform noise. On account of this inherent weakness of the system, they proposed its use mainly for comparison between polygons which, having straight sides and few, well-defined angles, give better results. In [22] TAs are proposed for use with generic images, but no mention is made of the problem of noise and the images shown as examples do not seem to be affected.

This paper presents a new definition of the distance between objects represented by turning angles, specially devised to minimize the effects of additive impulsive non-uniform noise. The main analytical properties will be discussed and the results of experimental tests will be given, comparing them with the results obtained by applying traditional techniques based on Euclidean distance. The results presented, together with the demonstrated characteristics of the distance proposed, suggest use in all cases where it is necessary to compare 2-D objects, especially those belonging to real images. In our case, they have been applied to a system of Automatic Image Storage and Retrieval (AISR) currently in an advanced stage of testing at the University of Messina (Italy). The AISR System is a set of procedures for the processing, storage and retrieval of 2-D images which grants access on the basis of the contents of images (content-based image retrieval), at the same time facilitating interaction with the user by a specially designed graphic interface.

The paper is organized as follows:

- Section 2 describes the techniques used to represent the shape of objects, from the image to representation by TAs;
- Section 3 introduces the problems relating to measuring the distance between objects represented with TAs, presenting the Euclidean distance as an initial solution;
- Section 4 introduces the median distance as the best solution to the problems discussed and demonstrates its main analytical features;
- Section 5 presents the experimental results obtained when the retrieval system was implemented, comparing them with those obtained using the Euclidean distance;
- Section 6 draws some conclusions on the work done.

2 Object Shape Description

The problem of reconstructing and representing the shape of the objects contained in an image was first dealt with several years ago in research into artificial vision, and since then a large number of solutions have been presented, mainly in the context of the research itself. In very general terms, reconstruction of the shape of a 2-D object comprises three basic steps which are intercalated, in the various methods, by intermediate processing operations to enhance their efficiency:

1. Application of a low-level extraction operator to the image, mostly one based on gradient operators; the most popular is probably the Canny operator [4], but there are many others. The result of this phase is a binary image whose points are only contour points.
2. Segmentation of the image: the contour points are grouped into chains, each representing the outline of a single object. This is certainly the most critical operation because it is by no means simple to discern the dividing line between two close objects whose shape is not known beforehand. The result of this phase is a sequence of chains of points, one for each object, representing the outlines in the form of closed lines.
3. The points making up the outline of an object are processed to obtain a compact and often analytical representation.

As far as low-level processing is concerned, the Canny operator is an excellent basis: its popularity makes it a point of reference in the sector and its performance is excellent. The phase in which the image is segmented into chains of points representing closed lines, on the other hand, needed an ad hoc algorithm to be developed [13], as those available did not give the required performance, i.e. closed lines each outlining an object, with the exception of objects having inner outlines (hollow), which can comprise more than one chain.

The technique used, when applied to an image processed by the Canny operator, uses a chain of points which initially occupy the outermost outline of the image, thus leaving all the objects visible in the image on the inside. The points in this chain can move, and feature "social" behaviour such as to give rise to generally coherent movement of the chain: it contracts until it directly and completely surrounds each object in the image, *wrapping it up*. Subsequently, at the divide between the various objects, the chain breaks and reproduces, creating as many chains as there are objects; this process repeats recursively. Any inner outlines (e.g. holes) are also detected by this technique: when a chain perfectly surrounds an outline it stops, but it sows an offspring chain inside the outline and the latter starts to contract, behaving just like the parent chain. A chain which does not encounter any outlines when it moves contracts on itself and disappears. This *reproductive* behaviour is also recursive.

The result of the algorithm described above is a sequence of closed chains of points, representing the inner and outer contours of the objects present in the image.

The outlines extracted are then scaled to normalize their area: the aim of this operation is to obtain a representation independent of the size of the object. Then the outlines are "cleaned", as they are still quite jagged and would not provide a good basis for the subsequent procedures. The strategy we chose was to apply a fitting technique developed by IBM in the QBIC project [9] [8]. It is based on fitting points with a type of spline called a Q-spline, which is not an interpolation but an approximation to the least squares; this allows us to obtain reasonably uniform outlines (see Figure 1). In addition, if the Q-spline coefficients are known we can reconstruct the

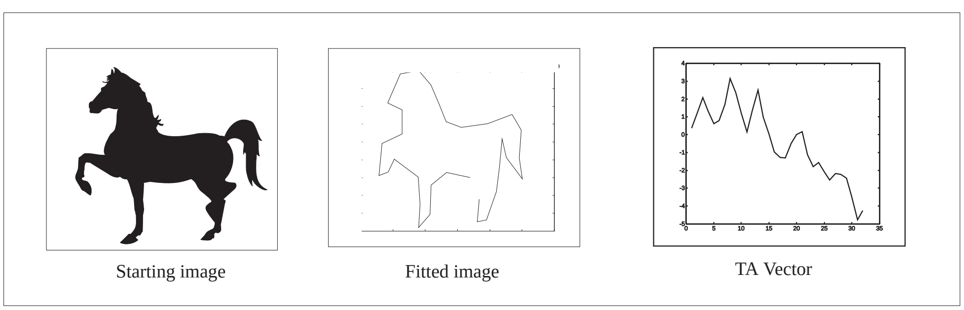

Figure 1: Starting figure, Q-Spline representation, TA representation.

curves with an arbitrary number of points which can be chosen to be the same for all the curves according to our requirements. The result of this fitting process is two series of Q-spline coefficients for each outline - one to represent the x coordinates and one for the y coordinates. The algorithms developed by the QBIC project researchers operate at an optimized speed and at the same time give excellent results as far as reproduction accuracy is concerned.

Finally, it is in this phase that a representation in the form of a list of Cartesian coordinates is transformed into a vector of the turning angles of the outline of an object. The outlines are reconstructed, starting from the Q-spline coefficients with a number of points chosen according to the desired level of accuracy, and then the turning angle value for each point is calculated.

The turning angles method for the description of shapes is based on a very simple principle [1]: let us imagine that along the curve representing the outline of an object we place a curvilinear abscissa, s and a reference point which takes the abscissa value 0. For each point of the outline we define an abscissa function, $\theta(s)$, as the angle made with a reference axis, for example the x-axis, by the counter-clockwise tangent to the trajectory which goes from the current point to the next one (see Figure 2). In other

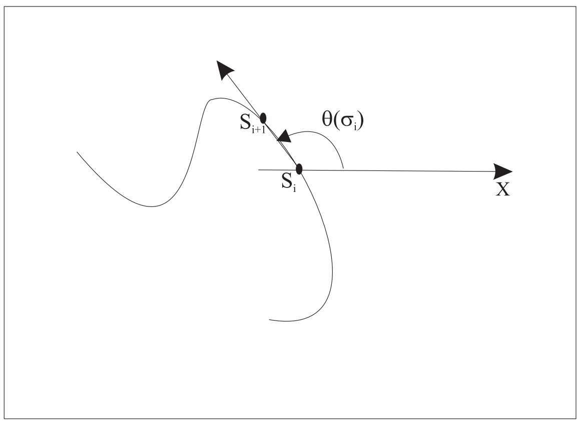

Figure 2: TA definition.

terms, $\theta(s)$ measures the turning angle from the direction of the x-axis needed to pass from the current point to the next one [22]. Formally, if $k(s)$ is the curvature of the curve, $k(s) = \theta'(s)$. As we are dealing with discrete rather than continuous digital images, the function $\theta(s)$ is represented by a vector TA with as many elements as there are points in the curve, and the curvilinear abscissa s has nonnegative integer values.

3 Shape distance

In this section we analyze the specific characteristics of a distance function which is capable of quantifying the degree of similarity between the shapes of two objects represented by means of turning angles. More specifically, we introduce the concepts of vertical and horizontal translation of a TA vector and discuss the inherent limits of the Euclidean distance, leaving the presentation of a new distance called *median distance* to the following section.

If we consider two absolutely identical objects, O_A and O_B and rotate one of them, O_B for instance, by a certain angle θ_1, the corresponding TA vector B is *vertically translated* with respect to the vector A relating to O_A, by the angle θ_1 (see Figure 3): a distance function which operates on the TAs and has to measure the similarity between the two objects without being affected by the rotations to which they are subjected has to take account of all the possible vertical translations of the TA, with θ_1 ranging between 0 e 2π. Another factor to take into account is the arbitrary

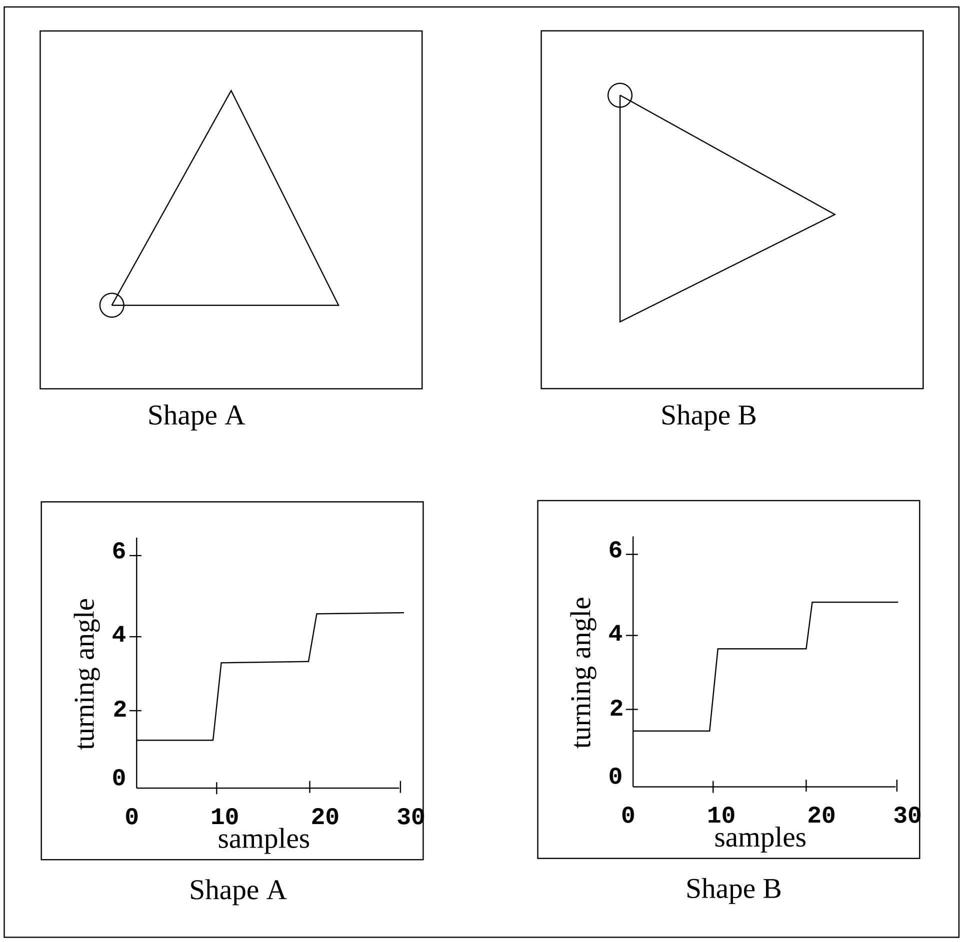

Figure 3: Q-Spline representations and TA representations for the triangle shape and for the same shape, rotated by an angle θ_1.

nature of the abscissa point 0 along the outline of the objects: on account of this arbitrariness the vectors A and B may be the same unless there is an arbitrary horizontal translation of the samples. *Horizontal translation* of a vector V corresponds to a rotation of its elements (see Figure 4), as defined by the formula:

$$\begin{array}{l} V_0 = (v_1, v_2, ..., v_{n-1}, v_n)\,; \\ V_{+1} = (v_n, v_1, v_2, ..., v_{n-1})\,; \\ V_{+2} = (v_{n-1}, v_n, v_1, ..., v_{n-2})\,; \\ ... \\ V_{+n} = (v_2, v_3, v_4, ..., v_n, v_1) \end{array} \tag{1}$$

In brief, the distance function has to take into account all the possible vertical and horizontal translations of the samples A and B [1].

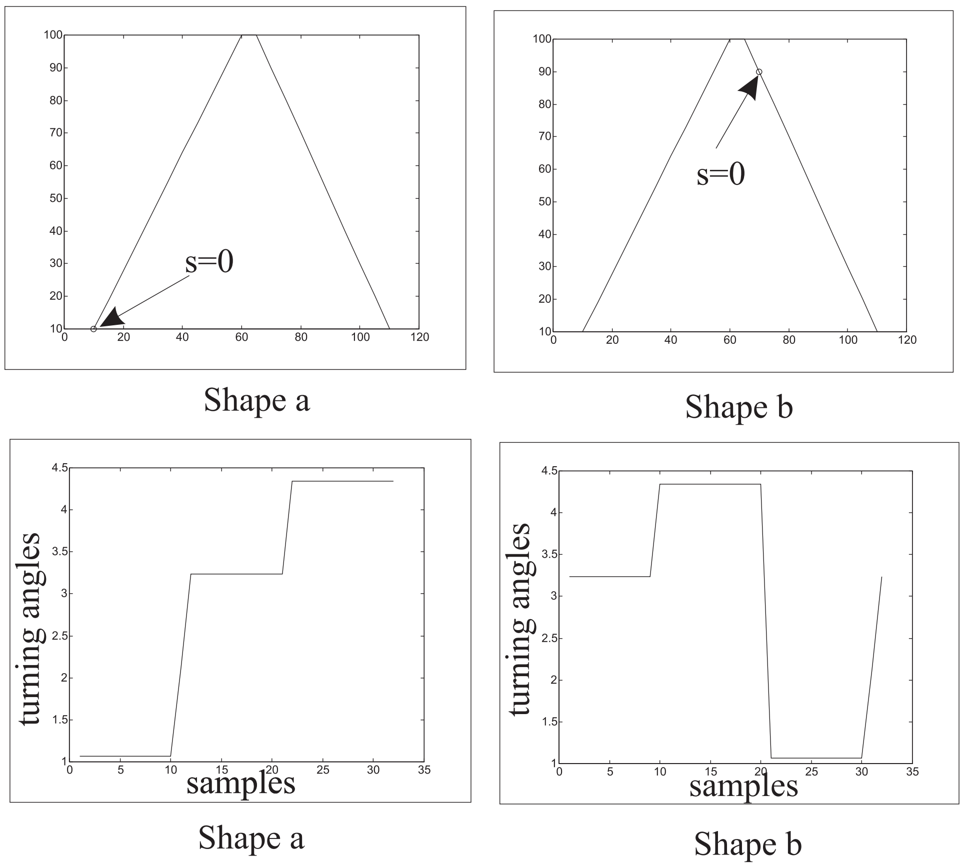

Figure 4: Q-Spline representations and TA representations for the triangle shape and for the same shape, with another choice for the starting point.

The most immediate distance function to compare the two vectors is the Euclidean distance, defined as:

$$D = \sqrt[2]{\sum_i (A_i - B_i)^2} \tag{2}$$

If we make the distance invariant to horizontal and vertical translations of the two vectors A and B, we get the definition:

$$D = \min_{r,\theta}\left(\sqrt[2]{\sum_i (A_i - B_i)^2}\right) \tag{3}$$

where $min_{r,\theta}(.)$ is the minimum of all the possible horizontal and vertical translations of B. If the two figures are the same, minimization on r (horizontal translation) makes the curvilinear abscissa points 0 on the outlines of the two figures O_A e O_B coincide, while minimization on θ (vertical translation) makes the main axes of the two figures coincide. This distance, however, is not very efficient when applied to turning angles, on account of the particular nature of the unidimensional signal deriving from the conversion of the object outline curves into turning angles.

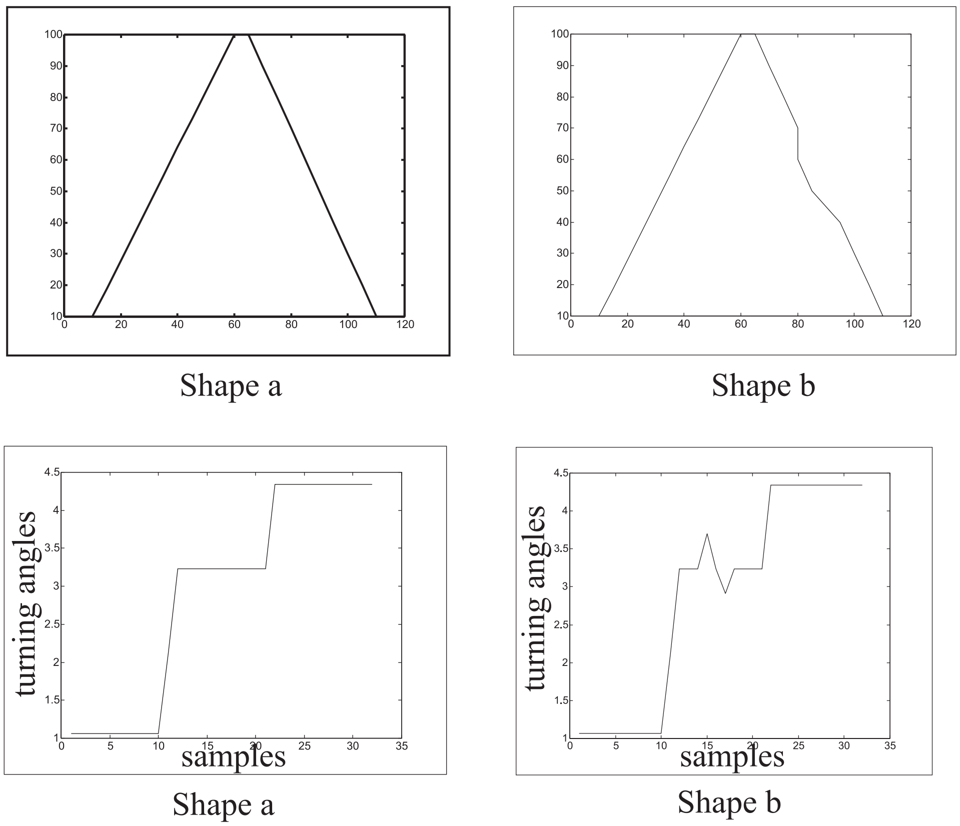

Figure 5: Q-Spline representations, TA representation for the triangle shape and the same shape, partially corrupted.

Analyzing the TA functions for two definitely similar (though not identical) objects, like those shown in Figure 5, considerable discrepancies can be seen, which generally take the form of *spikes* (see Figure 5); this happens frequently when real images are being dealt with, the outlines of which are usually not well defined. Figure 6 shows the impulsive nature of the differences between the TA functions of the two shapes shown in Figure 5.

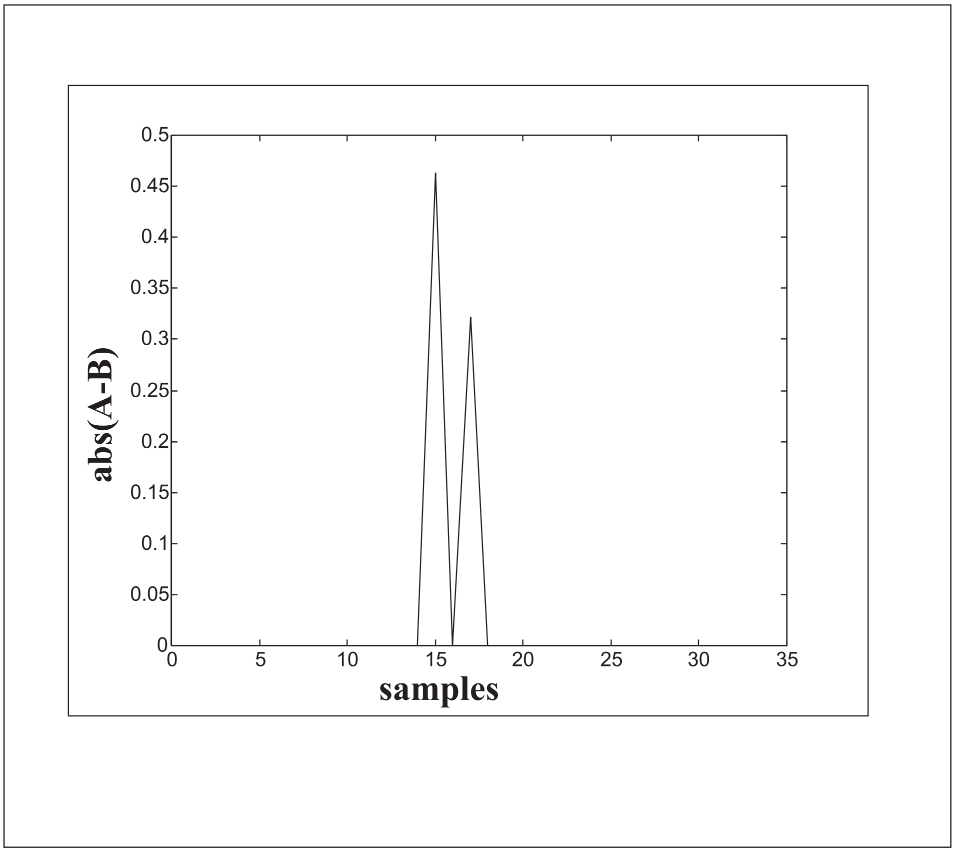

Figure 6: Abs value of the difference between the TAs of the 5.

The Euclidean distance function (2) sums the contributions made by these spikes, obtaining a high value for the distance between the two shapes, even though in reality they only differ by a few points. The problem gets worse if the outlines are affected by non-uniformly distributed noise: the Euclidean distance may grow arbitrarily even though calculated between images which are essentially similar [1]. As our aim is to define a method for representation of and comparison between shapes, whose behaviour is satisfactory in generic conditions, we felt the need to develop a new technique for turning angle comparison which would not suffer from these limits.

4 A new distance between turning angles

The impulsive nature of the discrepancies observed between the turning angle representations of essentially similar objects suggested an analogy with the problems related to the filtering of noisy signals. As is well known, one of the filters considered to be most effective on impulsive, non-uniformly distributed noise with a non-null mean is the *median filter*. The median value of an array V, of size n, is obtained by ordering the array in increasing or decreasing order and taking the value with the index $n/2$. The effect of rejection of impulsive noise is due to the fact that, by ordering the vector, values subject to strong noise peaks will be at the ends of the scale and thus far away from the central value which is chosen as the median. Starting from this simple mechanism, we continued the search for a distance which was not very sensitive to impulsive noise and finally defined the following function:

$$D_{median}(A, B) = \min_r \left(median \left(|a_i - b_i| \right) \right) \tag{4}$$

where $\min_r (.)$ is the minimum of all the possible horizontal translations of B, and both A and B have been previously normalized:

$$\begin{cases} A = A - median(A); \\ B = B - median(B); \end{cases} \tag{5}$$

The aim of this normalization (5) is to eliminate the effect of any rotation one of the two figures may have undergone: it should, in fact, be recalled that a rotation, θ_1 of the object O_A corresponds to a vertical translation (offset) of its turning angle function; so, if we suppose that the object O_B is identical to the object O_A but is rotated by the angle θ_1, as the median operator is linear, we get:

$$median(B) = median(A + \theta_1) = median(A) + \theta_1 \tag{6}$$

and recalling (5):

$$\begin{cases} A = A - median(A); \\ B = B - median(A) - \theta_1; \end{cases} \tag{7}$$

so subtracting the relative median values from both vectors A and B we eliminate the effect of rotations between the two figures. The choice of preliminary normalization instead of insertion of another minimization in (4), as was done in [1], aimed to shift part of the computational load from the search phase to the data loading phase, which is normally slower and whose service times make this normalization negligible.

It is pointed out that, before the turning angles are calculated the outlines are scaled so as to be normalized with respect to the areas: the aim of this process is to obtain representations of outlines of a similar size, thus eliminating the comparison difficulty due to scale factors.

To be applied, (4) has to be a *metric* i.e. it has to possess the following properties:

Property 1 *The distance between two vectors is always greater than or equal to zero:*

$$D_{median}(A, B) \geq 0 \;\; \forall (A, B) \tag{8}$$

Proof: Given two vectors $A = (a_1, \dots, a_i, \dots, a_n)$ and $B = (b_1, \dots, b_i, \dots, b_n)$, it follows that $|a_i - b_i| \geq 0$. According to the definition of the median operator we get: $median\,(|a_i - b_i|) \geq 0$ from which the property follows.

□

Property 2 *Given two vectors A and B the following assertions are true:*

$$\begin{cases} A = B & \Longrightarrow D_{median}(A,B) = 0 \\ D_{median}(A,B) = 0 & \Longrightarrow \exists\, B^A \mid A = B^A, \textit{ almost everywhere.} \end{cases} \quad (9)$$

where B^A is a horizontal translation of B such as to minimize (4) [1].

Proof: The first assertion can easily be proved considering that if $A = B$ then $|a_i - b_i| = 0\ \forall i$. Applying the median operator we get $median\,(|a_i - b_i|) = 0$, from which the proof follows.

The second assertion comes from the observation that, once the B^A minimizes the function (4), then if $median\left(\left|a_i - {b_i}^A\right|\right) = 0$ it follows that $\Longrightarrow \left|a_i - {b_i}^A\right| = 0$ *almost* $\forall\, i$, from which the proof comes.

□

Property 3 *Given two vectors A and B the order of comparison does not matter. It implies that:*

$$D_{median}(A,B) = D_{median}(B,A) \quad (11)$$

Proof: It is straightforward to demonstrate that if A^B is a horizontal translation of A such as to minimize $median\left(\left|{a_i}^B - b_i\right|\right)$ and B^A is a horizontal translation of B such as to minimize $median\left(\left|a_i - {b_i}^A\right|\right)$, then

$$\exists\, A' \mid median\left(\left|a_i - {b_i}^A\right|\right) = median\,(|b_i - {a_i}'|) = m_1 \quad (12)$$

$$\exists\, B' \mid median\left(\left|b_i - {a_i}^B\right|\right) = median\,(|a_i - {b_i}'|) = m_2 \quad (13)$$

[1] The particular form of (9) is due to the minimization introduced in (4): it does not affect our definition of distance and, if we call the two objects for which A and B are the turning angle vectors OA and OB, it is equivalent to:

$$D_{median}(A,B) = 0 \iff O_A = O_B \quad (10)$$

Consequently we have:

$$median\left(\left|a_i - b_i{}^A\right|\right) = median\left(\left|b_i - a_i{}^B\right|\right) \tag{14}$$

in fact A^B is such that $median\left(\left|b_i - a_i{}^B\right|\right) \leq median\left(\left|b_i - a_i{}'\right|\right)$, so

$$median\left(\left|b_i - a_i{}^B\right|\right) \leq m_1$$

and necessarily $m_2 \leq m_1$; on the other hand B^A is such that

$$median\left(\left|a_i - b_i{}^A\right|\right) \leq median\left(\left|a_i - b_i{}'\right|\right)$$

from which we get $m_1 \leq m_2$ and therefore $m_1 = m_2$. Expression (14) thus proves the property.

□

Property 4 *Given three vectors A, B and C, the following Triangle Inequality holds:*

$$D_{median}(A,B) + D_{median}(B,C) \geq D_{median}(A,C) \;\; \forall\,(A,B,C) \tag{15}$$

Proof: By using (11), it can be observed that:

$D_{median}(A,B) + D_{median}(B,C) = median\left(\left|a_i{}^B - b_i\right|\right) + median\left(\left|b_i{}^C - c_i\right|\right)$

$= median\left(\left|a_i{}^B - b_i\right| + \left|b_i{}^C - c_i\right|\right)$.

By applying the same rotation to both vectors A^B and B, the last term is equal to:

$= median\left(\left|a_i{}^{BC} - b_i{}^C\right| + \left|b_i{}^C - c_i\right|\right) \geq median\left(\left|a_i{}^{BC} - b_i{}^C + b_i{}^C - c_i\right|\right) \geq$

$median\left(\left|a_i{}^C - c_i\right|\right) = D_{median}(A,C)$.

□

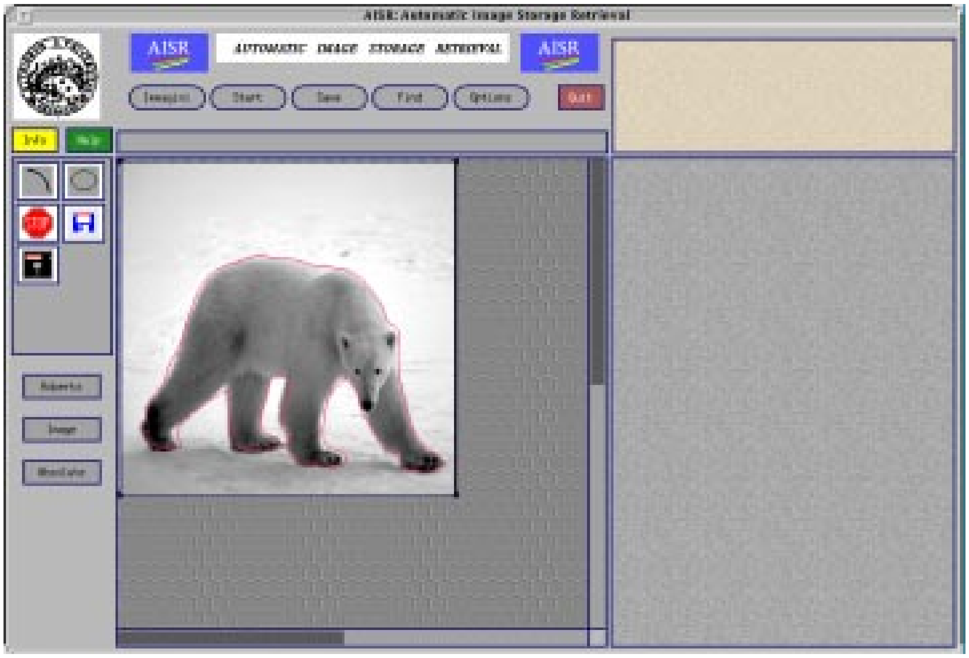

Figure 7: The "Define Query" screen of AISR: the user is defining a portion of interest of the image (see text).

5 Experimental Results

The distance defined above was tested by implementing the AISR (Automatic Image Storage and Retrieval) system, devised for accurate testing of the algorithms and techniques we have developed for image retrieval, on a SUN Sparc20 workstation. Some of the images used in the test came from commercial sources and the rest were constructed specially so as to have a suitable sample of inputs for significant testing of the system. We used a graphic tool which was capable of rotating, scaling and corrupting the outlines of objects, with which some of the samples were processed. In this section we present some of the results obtained using the median distance and compare them with the results obtained by applying the Euclidean distance.

Figure 7 shows a snapshot of the way in which the AISR query system works: the user has supplied the system with an image of his own choice and has selected a portion of interest, from which the system has started recovering the object contours. This operation is optional, as AISR can perform segmentation autonomously. After specifying the object to be sought, the user sets the query off by pressing the corresponding key. Figure 8 shows the final stage of the search process: in a special window, the system shows the stored images which most closely resemble the one being sought, ordered according to the median distance (*nearest neighbour query*). Note the presence of the scroll bars at the sides of the windows:

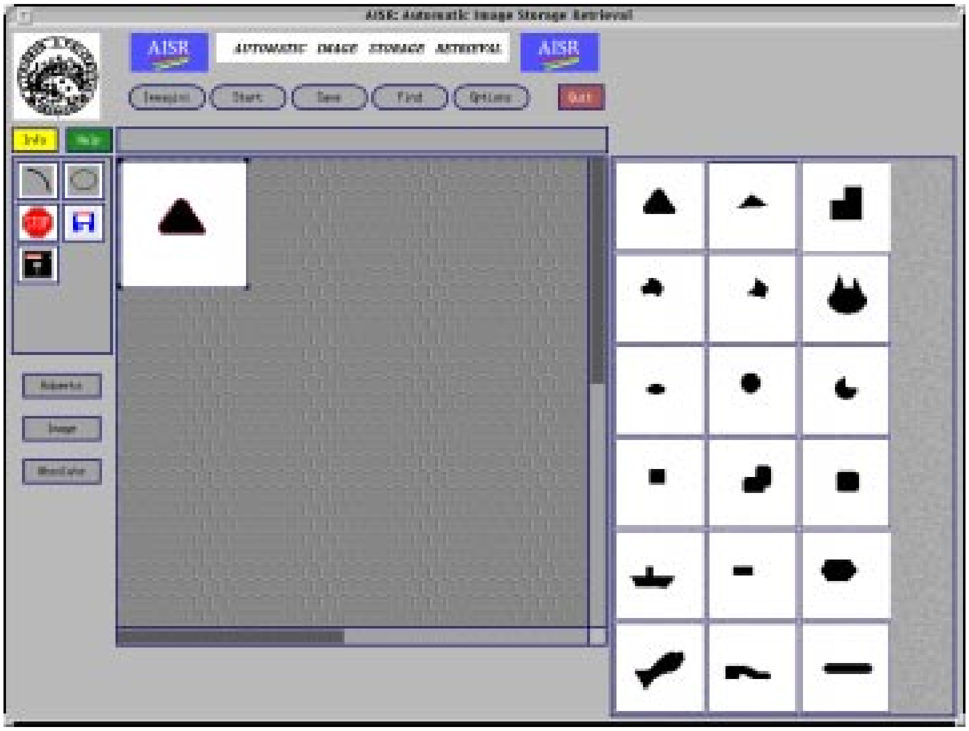

Figure 8: The "Query Results" screen of AISR: the system is presenting the results of the query made by the user (see text).

at the moment only the first 18 shapes are visible, but the system can, if necessary, order the whole archive according to similarity with the object of the query. The number of elements supplied is established by the query. The shapes displayed belong to the same series as those shown in Figs. 10 and 11 and are very simple ones in order to make the paper as clear as possible. As shown in Figure 9, the system works just as well with more complex shapes and photographic images: in this case the segmentation has been performed autonomously by AISR, and the head of the tiger has been perfectly outlined. The same figure shows how the images retrieved can be blown up at will by pressing a key, to simplify operations.

Figure 10 and 11 show the first 14 images found by querying the AISR system, filled with about 200 images, for the images similar to the "Triangle Shape". Both Median Distance and Euclidean Distance have been used, in two distinct tests. These figures provide a qualitative comparison between the Median and the Euclidean distances (the reader should note that the values reported in the axis are not relievant and so can be ignored): as can be noticed, the former distance behaves visibly better in distinguishing between similar shapes. Let us consider the third and fourth rows of both figures 10 and 11. According to the median distance, after the triangle shapes come the "twin rocks" shape and the "arrowhead" shape, followed by the "square" shape; instead, according to the Euclidean distance, the

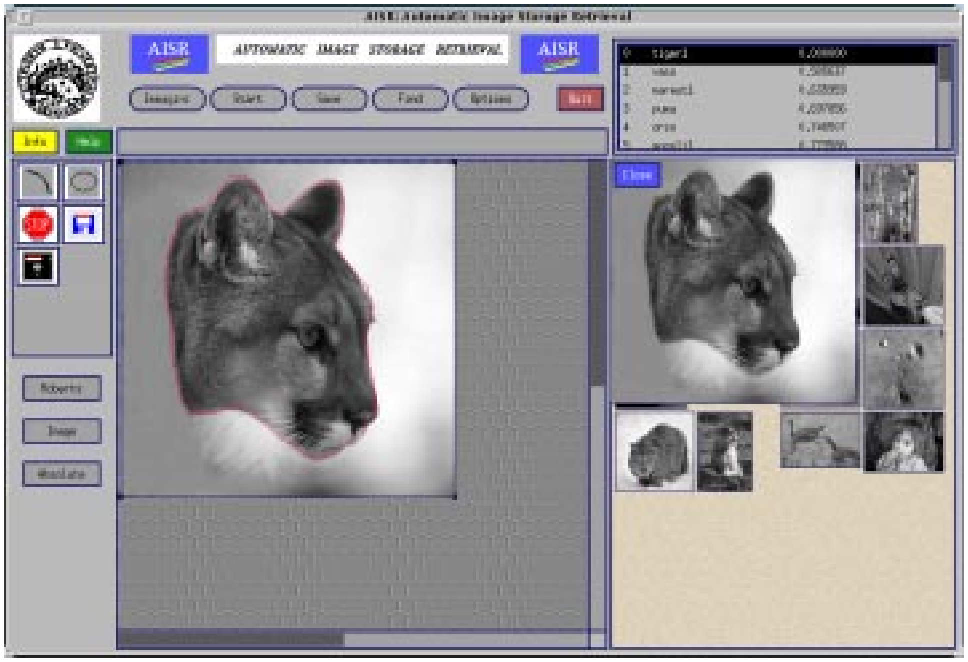

Figure 9: The AISR system can deal with general purpouse photographic images.

triangle shapes are immediately followed by the "square" shape, then by the "arrowhead", the "amorphous" shape and the "twin rocks" shape. This ordering is definitely less consistent with the human idea of similarity.

From the graphs in Figure 12 and in Figure 13 it is possible to obtain the values for the median and Euclidean distances of the first 14 shapes resulting from the nearest-neighbours query with the "Triangle" shape assumed as reference. The "Triangle" shape is numbered as zero; the 14 shapes retrieved are numbered from 1 to 14 according to the order in which they are shown in Figures 10 and 11, respectively. As can be noticed in Figure 12, the distance values are close to zero for all the triangles, jumping to 0.5 only with the 8th shape, which is the "twin rocks" one; the distance of the following shape is then similar, while another jump is encountered correspondending to the "square" shape.

According to the Euclidean distance, instead, the modified triangles are not so close to the model (see Figure 13, where shapes 3 to 7 return distance values greater than zero); moreover, both the "square" and the "arrowhead" shapes have almost the same, huge distance from the "Triangle" shape, and the "twin rocks" shape is as far as the "amorphous" shape from the "Triangle" one. In conclusion, here again the median distance behaves better, giving values close to zero or null values when the shapes are the same and returning results very close to those which a human being could return, when the shapes are similar but not identical.

6 Conclusions and discussion

In this paper we have presented a new definition of the distance between objects represented by turning angles, with the aim of obtaining a system for shape comparison that was as insensitive as possible to small variations in outlines. The distance presented is a metric for any kind of outline, both polygonal and curvilinear, convex and nonconvex, and only compares shapes, regardless of any translations, rotations, scaling or partial blurring of the outlines. The complexity obtained for the algorithm used is in the order of $O(n * o_1)$, where o_1 is the complexity of the sorting algorithm used to calculate the median value, which is, in the average case, $O(nlog(n))$ if *quick sort* is used; our distance therefore has a lower complexity than the one described by Arkin et al. [1], which is in the order of $O\left(n^2 \log\left(n^2\right)\right)$. Experimental tests have also confirmed the good behaviour of the technique presented, even as compared with the Euclidean distance; in particular, it should be noted that the capacity to recognise the similarity between objects which are essentially similar but partly affected by occlusion (covering of part of the outline of one or both by other objects or shadows) makes the system extremely interesting for application to real images, for example diagnostic images.

Interesting comparisons could be done with other works, such as Flick et alii [18]; in particular, it could be useful to introduce in the median distance some dynamic programming techniques, in order to take better account of the vertical traslations of TAs.

At this stage, no indexing issues have been introduced yet, since our main goal was to explore the quality of the results obtained by the matching algorithm; however, Some speed measurements have been made on the software and hardware architecture we used. In particular, on a SUN Sparcstation 20 with Solaris and with 500 images in the database an average response time of 60" has been experienced, using simple sequential search. Actually, the system is being widely tested with a set of about 500 images by 10 users of the University of Messina, and according to the users about the 80% of the queries returned satisfying results.

Further developments could come from the application of "quick and dirty" techniques [2] to our retrieval method, in order to speed up the search phase. The aim of "quick and dirty" techniques is to preselect elements which are likely to satisfy a query from a set of large dimensions, so as to reduce the number of elements among which to conduct the search. This preselection is generally made by algorithms with a low degree of complexity, which makes this multi-step procedure more convenient than retrieval based on a single step. An interesting quick and dirty algorithm, which we are currently exploring, is based on the use of wavelett multiresolution analysis [17] applied to Turning Angles vectors.

References

[1] E.M. Arkin, L.P. Chew, D.P.Huttenlocher, K.Kedem, and J.S.B. Mitchell. An efficiently computable metric for comparing polygonal shapes. *IEEE Transactions on PAMI*, 13(3), March 1991.

[2] J. Ashley, R. Barber, M. Flickner, J. Hafner, D. Lee, W. Niblack, and D. Petkovic. Automatic and semi-automatic methods for image annotation and retrieval in qbic. Technical report, IBM Research Center of Almaden, 1995.

[3] Dana H. Ballard and C.M. Brown. *Computer Vision*. Prentice-Hall, 1982.

[4] J. F. Canny. A computational approach to edge detection. *IEEE Trans. on PAMI*, 8(6):679-698, June 1986.

[5] Shi-Kuo Chang. *Principles of Pictorial Information Systems Design*. Prentice-Hall International Editions, 1989.

[6] W. Equitz. Retrieving images from a database using texture - algorithms from the qbic system. Technical report, IBM Almaden Research Center, San Jose, CA., 1994.

[7] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz. Efficient and effective querying by image content. *Journal of Intelligent Information Systems*, 3, 1994.

[8] Myron Flickner, James Hafner, Eduardo J. Rodríguez, and Jorge L. C. Sanz. Periodic quasi-orthogonal spline bases and applications to least-squares curve fitting in digital images. *IEEE Transactions on Image Processing*, 5(1):71–88, Jan 1996.

[9] Myron Flickner, Jim Hafner, Eduardo Rodríguez, and Jorge Sanz. Fast least-squares curve fitting using quasi-orthogonal splines. In *Proceedings of the 1st International Conference on Image Processing*, volume I, pages 686–690, 1994. IBM techreport RJ 9819.

[10] Y. Gong, H. Zhang, H.C. Chuan, and M. Sakauchi. An image database system with content capturing and fast image indexing abilities. In *Proceedings of International Conference on Multimedia Computing and Systems*, Boston, May 1993.

[11] M. Hirakawa, A. Yoshitaka, S. Kishida, and T. Ichikawa. Knowledge-assisted content-based restrieval for multimedia databases. In *Proceedings of International Conference on Multimedia Computing and Systems*, Boston, 1993.

[12] D. P. Huttenlocker, G. A. Klanderman, and W. J. Rucklidge. Comparing images using the hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15, September 1993.

[13] G. Iannizzotto and L. Vita. A fast, accurate method to segment and retrieve object contours in real images. In *International Conf. on Image Processing*, Lausanne, Switzerland, 1996.

[14] H.V. Jagadish. A retrieval technique for similar shapes. In *Proc. SIGMOD 91 Conf.*, Denver, 1991.

[15] E. K. Jones and A. Roydhouse. Intelligent retrieval of archived meteorological data. *IEEE Expert*, December 1995.

[16] F. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, and Z. Protopapas. Fast nearest neighbor search in medical image databases. Technical report, University of Maryland, Dept. of Computer Science, 1996.

[17] S. G. Mallat. A theory for multiresolution signal decomposition: The wavelett representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, July 1989.

[18] W. Niblack and J. Yin. A pseudo distance measure for 2d shapes based on turning angle. In *International Conf. on Image Processing*, Washington, DC, USA, 1995.

[19] M. O'Docherty. A multimedia information system with automatic content retrieval. Technical report, University of Manchester, 1995.

[20] S. C. Orphanoukakis, C. Chronaki, and S. Kostomanolakis. I^2c: A system for the indexing, storage, and retrieval of medical images by content. Technical report, Institute of Computer Science, Foundation for Research and Technology - Hellas, 1994.

[21] E.G.M. Petrakis and C. Faloutsos. Similarity searching in large image databases. Technical report, University of Maryland, 1995.

[22] Brian Scassellati, Sophoclis Alexopoulos, and Myron Flickner. Retrieving images by 2d shape: A comparison of computation methods with human perceptual judgments. In *SPIE Conference on Storage and Retrieval for Image and Video Databases*, volume 2185, pages 2–14, February 1994.

[23] S. Sclaroff and A.P. Pentland. Modal matching for correspondence and recognition. *IEEE Trans. on PAMI*, 17(6), June 1995.

[24] G. Taubin and D.B. Cooper. Recognition and positioning of rigid objects using algebraic moment invariants. In *Geometric Methods in Computer Vision, SPIE 91*, 1991.

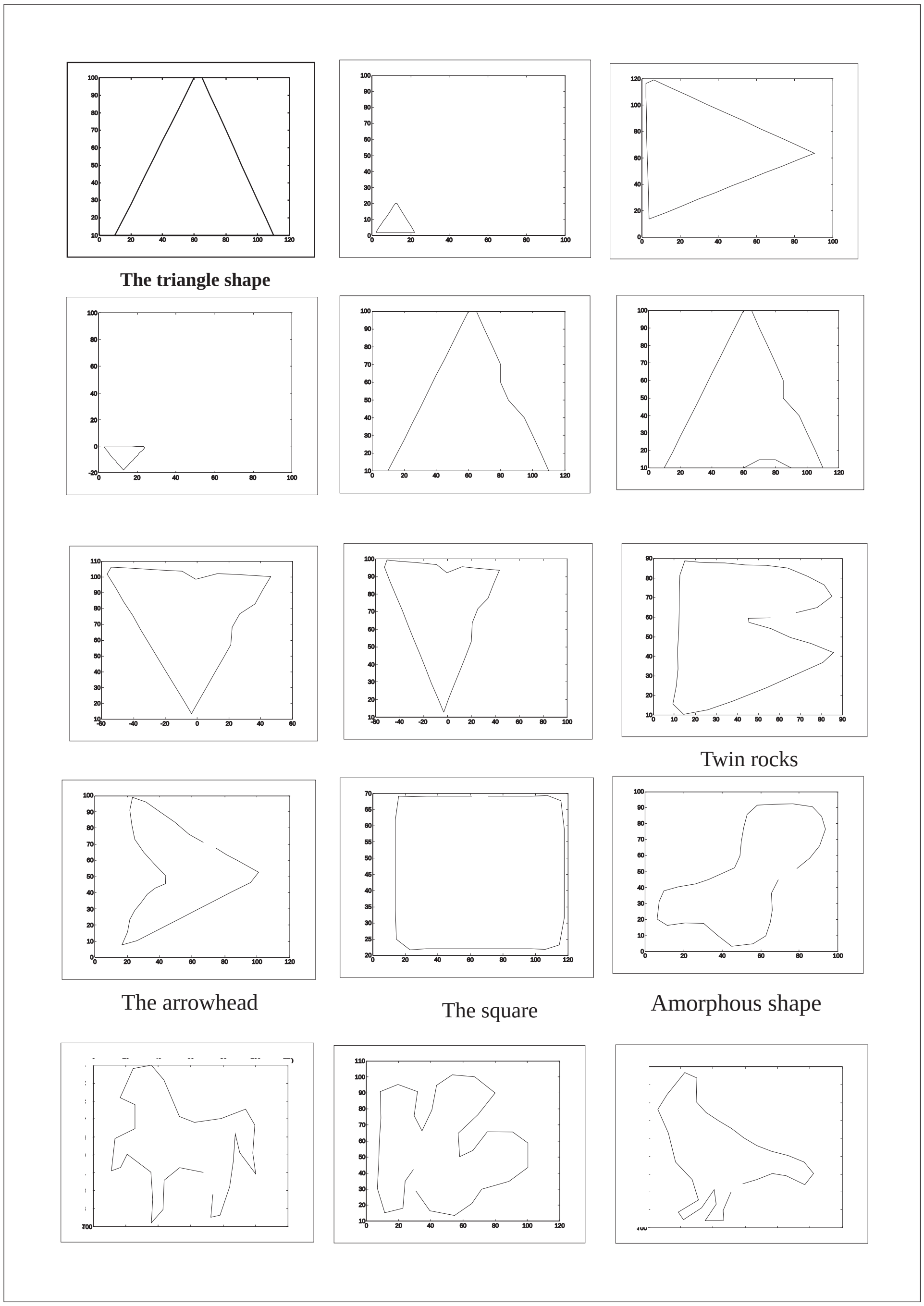

Figure 10: The Triangle shape and the other shapes sorted according to their Median distance from it.

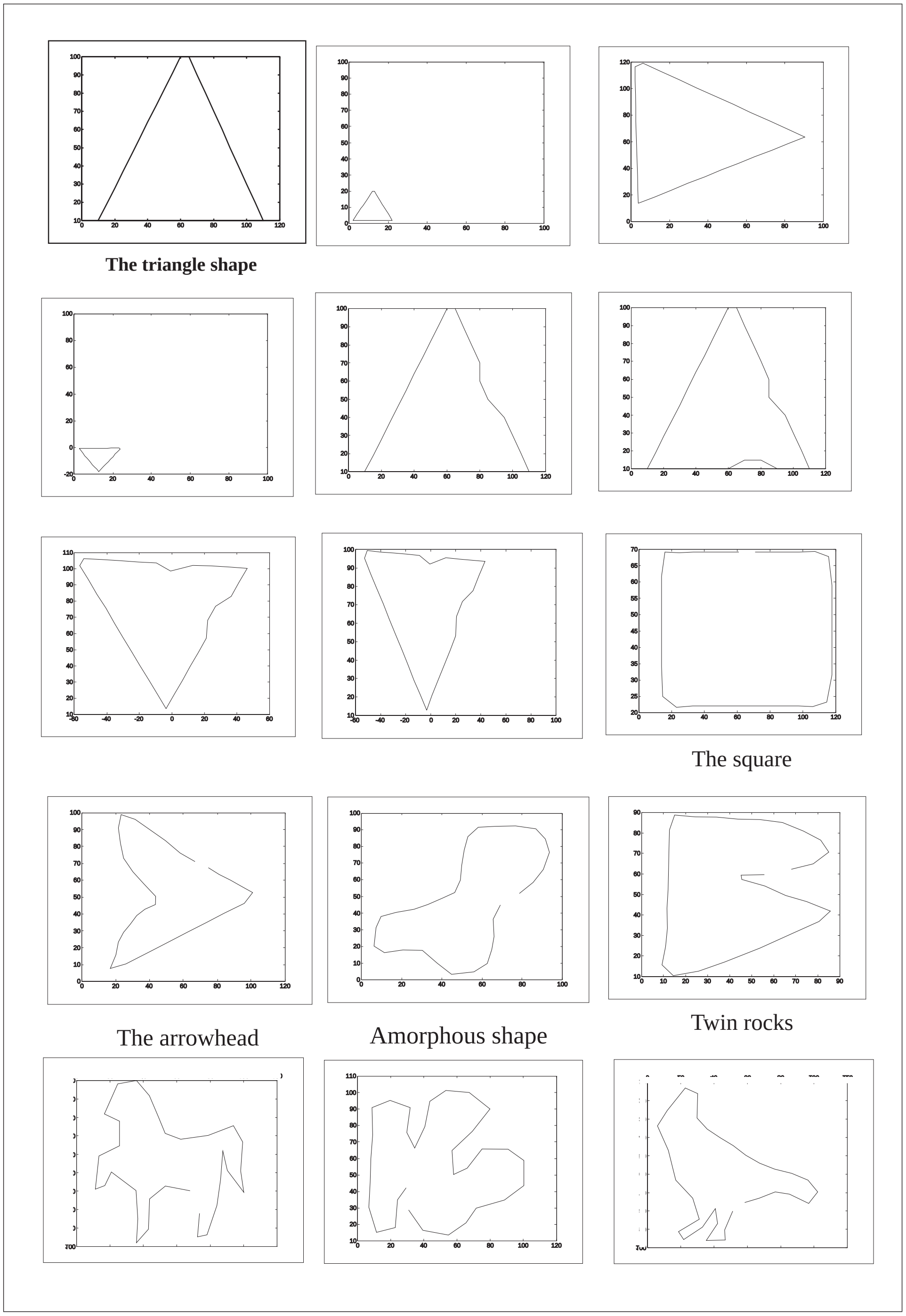

Figure 11: The Triangle shape and the other shapes sorted according to their Euclidean distance from it.

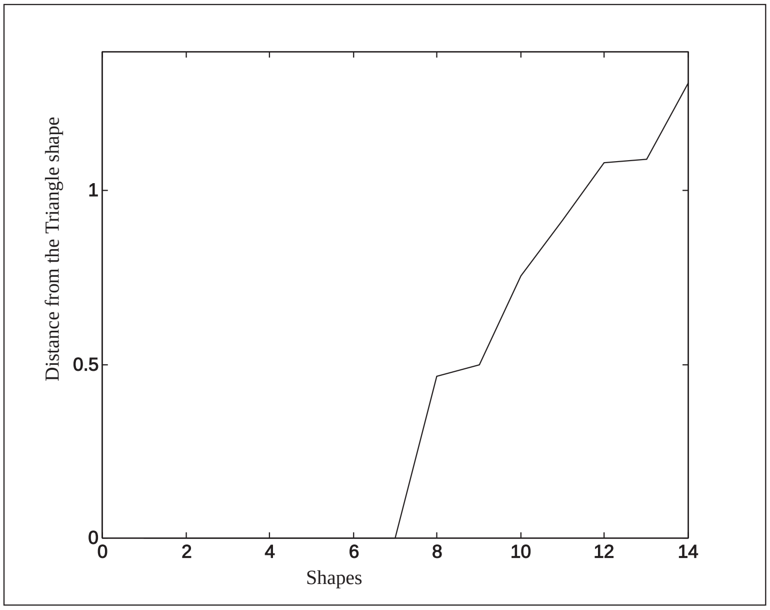

Figure 12: Graph of the Median distance between the Triangle shape and the other shapes.

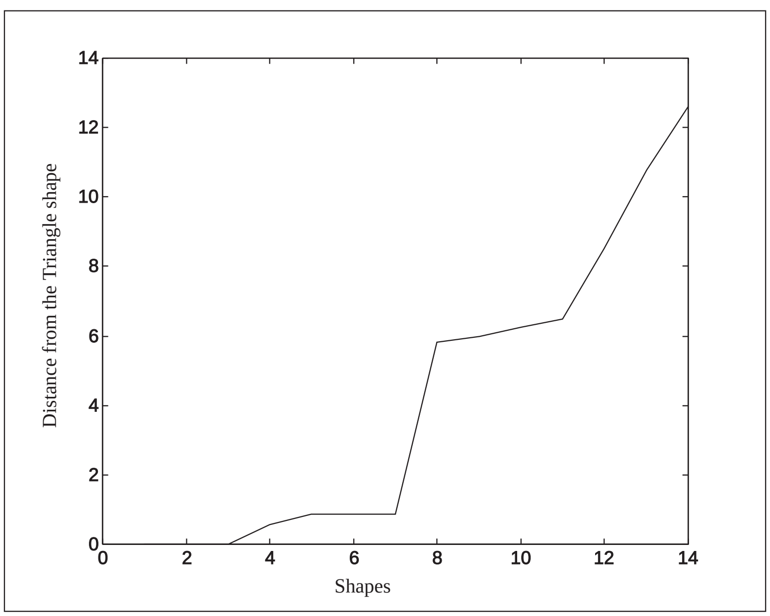

Figure 13: Graph of the Euclidean distance between the Triangle shape and the other shapes.