Modern Education
and Computer Science
PRESS

Vector Image Retrieval Methods Based on Fuzzy Patterns

Yevgeniya Sulema
Department of Computer Systems Software, Igor Sikorsky Kyiv Polytechnic Institute, Ukraine
Email: sulema@pzks.fpm.kpi.ua

Etienne Kerre
Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Belgium
Email: etienne.kerre@ugent.be

Oksana Shkurat
Department of Computer Systems Software, Igor Sikorsky Kyiv Polytechnic Institute, Ukraine
Email: shkurat@pzks.fpm.kpi.ua

Abstract—In this work we present two methods of vector graphic objects retrieval based on a fuzzy description of their shapes. Both methods enable the retrieval of vector images resembling to a given fuzzy pattern. The basic method offers a geometrical interpretation of a fuzziness measure as a radius of a circle with the center in each vertex of a given candidate object. It enables the representation of uncertain information about a pattern object defined by its "fuzzy" vertices. The advanced method generalizes this approach by considering an ellipse instead of a circle. The basic method can be used for the comparison of polygons and other primitives in vector images. The advanced method can be used for complex shapes retrieval. To enable saving a "fuzzy" image as a file, the modification of the SVG format with a new attribute "fuzziness" is proposed for both methods. The advanced method practical implementation is illustrated by the retrieval of medical images, namely, heart computer tomography images.

Index Terms—Information Retrieval, Pattern Recognition, Image Processing, Fuzziness.

I. INTRODUCTION

Nowadays, the number of images created and stored in vector formats is growing dramatically. The most frequently occurring content of these images is technical drawings and clip arts. Retrieval of such graphical data enables new opportunities for patent search, plagiarism analysis, etc. Besides, there is also another important task, which can be solved with the use of vector images, namely, medical image retrieval. Despite that originally medical images are raster, their processing can require vectorization in many cases, in particular, it can be useful for the detection of a specific shape, retrieval of similar images, feature extraction, etc.

However, vector image retrieval is a complicated task because the same shape can be presented by using different combinations of basic graphical primitives (lines, rectangles, polylines, etc.). Visually, such vectorized shapes can look identically but their internal view can differ considerably because each graphical primitive has its own representation format. The problem of vector image retrieval becomes even more complex if we search not for identical shapes but for similar ones which can be of different size and their vertices coordinates can differ significantly. All these specificities of graphical data representation complicate vector image retrieval.

The research problem is to enable an object retrieval when its shape cannot be defined strictly. In this paper we attempt to solve this problem by employing a fuzzy approach to vector image representation. The main objective of this research is to elaborate a method for the retrieval of vector images described by imprecise data.

The rest of this paper is organized as follows. Section 2 presents the literature review. In Section 3, we propose two approaches to vector image retrieval. The basic method explains the general idea of imprecise graphical data representation based on fuzzy templates of shapes which enables vector image retrieval. The advanced method improves the retrieval accuracy and can be used for the analysis of complex shapes such as objects on medical images. Section 4 presents the practical results and their discussion. Finally, some conclusions are given in Section 5.

II. RELATED WORK

Since the task of vector image retrieval is topical, many researchers work in this area. In particular, Fujita and Hayashi [1] proposed a similarity retrieval system for vector images such as trademarks. The proposed system retrieves vector images from a database using a sketch

given by the user as a query. The sketch is converted into a vector image format as a pre-processing step for the similarity retrieval.

Jiang et al. [2] presented a prototype system based on SVG technology and the method of similarity measurement. They proposed to measure the similarity between a requested image and a candidate image taking into consideration shape similarity, spatial similarity, color similarity, transformation similarity, and position similarity.

Rahmani and Ansari presented their study [3] of the conventional approaches and techniques used in image retrieval systems.

Hayashi et al. [4] presented a novel method for content-based 2D vector image retrieval. The method includes two main stages. At the beginning, Weighted Feature Points (WFPs) are extracted from vector images. Next, similarity between query and candidate images by matching WFPs is evaluated by using Proportional Transfer Distance (PTD).

Pu and Ramani [5] proposed to analyze a vector image as a whole. In particular, they presented two methods to describe drawings: the first one uses the 2.5D spherical harmonics and the other uses a 2D shape histogram.

Chen and Wang [6] presented a new fuzzy logic approach called UFM (Unified Feature Matching) which enables region-based image retrieval. This approach allows associating an image with a family of fuzzy features corresponding to regions.

Chen and Weng [7] considered a measure of the similarity between two tree structures. They proposed an algorithm for a tree similarity measure based on the weighting subtree and the fuzzy inference system.

Wang and Xie [8] presented the retrieval system which uses the fuzzy language variables to describe the similarity degree of image features as well as it employs the fuzzy inference to instruct the weights assignment among various image features. The authors also proposed an improvement on the Average Area Histogram to represent color features.

Martins et al. [9] studied the benefits of combining information extracted from vector and raster images to retrieve clip arts. The authors proposed to extract color and texture features from raster images while geometry and topology features were extracted from vector images.

Bhute and Meshram [10] presented the content-based image retrieval systems which employ the color, texture and shape information of images.

Banerjee and Kundu [11] presented a robust technique for content-based image retrieval based on using fuzzy edge map of an image.

Krishnapuram et al. [12] described an image retrieval system called FIRST which uses fuzzy attributed relational graphs to represent images.

Uchida et al. [13] proposed the application of the Fisher vector representation to binary features to improve the accuracy of binary feature-based image retrieval.

Fonseca et al. [14] presented a generic approach to content-based retrieval of structured graphics and drawings. The authors proposed to index drawings using a topology graph which describes adjacency and containment relations for parts and subparts in an image.

Li et al. [15] proposed an approach to describe and to extract fuzzy color semantics. The authors applied the feedforward neural network to model the vagueness of human color perception and to extract the fuzzy semantic feature vector.

The analysis of these and other papers allows us to conclude that a fuzzy approach is mostly employed for raster image retrieval. In the next sections, we present our vision on how a fuzzy approach can be applied to vector image retrieval.

III. IMAGE REPRESENTATION AND RETRIEVAL

A. Vector image format

The analysis of these and other papers allows us to conclude that a fuzzy approach is mostly employed for raster image retrieval. In the next sections, we present our vision on how a fuzzy approach can be applied to vector image retrieval.

One of the most popular vector image formats is SVG (Scalable Vector Graphics) [16]. SVG allows describing two-dimensional graphics by using graphical primitives such as lines, circles, polygons, etc. It supports interactivity and animation as well as raster effects applied on the client side when the SVG is rendered. In fact, SVG is a language for describing two-dimensional graphics in XML [17].

The example of SVG code fragment which draws a blue triangle defined by its vertices P_1 (100,100), P_2 (300,400), P_3 (400,200) is given in Fig. 1. This simple SVG code illustrates the main problem of vector image retrieval mentioned in above: to find the image which contains a specific object (e.g. a triangle of a certain size and shape) we need to know crisp coordinates of this object. However, in most cases, we do not have this information and, therefore, we mostly look for objects which are similar to some pattern, but generally not equal to it.

```
<svg width="15cm" height="15cm">
  <g    style="fill:none;
        stroke: #0000ff;
        stroke-width:2px">
<polygon points="100,100 300,400 400,200"/>
    </g>
</svg>
```

Fig. 1. SVG code example.

Note that in further explanations we refer to such similar objects as resembling objects as well as we use the term "resemblance" instead of "similarity" in order to avoid misinterpreting this term used in geometry (similar triangles, similar solids, etc.). On the other hand, resemblance of two objects can be modelled by a resemblance relation introduced in [18, 19].

In particular, in [19], the resemblance relation is defined as follows: a fuzzy relation E on a pseudo-metric space (X, d) is called a resemblance relation (w.r.t. d) if for all x y z and u in X:

(R.1) $E(x, x) = 1$,

(R.2) $d(x, y) \leq d(z, u)$ implies $E(x, y) \geq E(z, u)$.

Thus, transition from crisp values of vertex coordinates to fuzzy values can enable effective vector image retrieval.

B. Basic Method

The idea of the proposed method is based on a geometric interpretation of a fuzziness degree: a "fuzzy" vertex is a circle with the center in this vertex defined by its crisp coordinates and a crisp radius of the circle which limits the neighborhood of the "fuzzy" vertex. Both a pattern image and a candidate image are analyzed in object-by-object manner. Every object is a single shape (e.g. polygon, circle, line, etc.) located at a separate layer of the vector image. Objects in the candidate image are defined by their crisp coordinates as in the example in Fig. 1. Objects in the pattern image are defined by fuzzy coordinates of their vertices. Since a SVG file [16] operates with crisp values only, we propose to add a new attribute "fuzziness" intended to define a fuzziness degree of vertex coordinates (Fig. 2).

```
<svg width="15cm" height="15cm">
  <g    style="fill:none;
        stroke: #0000ff;
        stroke-width:2px">
<polygon points="100,100 300,400 400,200";
        fuzziness="0, 20, 10"/>
  </g>
</svg>
```

Fig. 2. SVG code with 'fuzziness' attribute.

A value of the attribute "fuzziness" can be defined in millimeter, inch, pixel, etc. As it is mentioned above, this value corresponds to a radius of the circle around the vertex which limits the neighborhood of this "fuzzy" vertex. This idea is described by Fig. 3 where a fuzzy pattern (dotted triangle), vertices fuzziness (grey circles), and a crisp result of image retrieval (solid triangle) are shown. Other conventional signs in Fig. 3 are explained below. The number of possible retrieval results is limited by the value of the neighborhood circle radius.
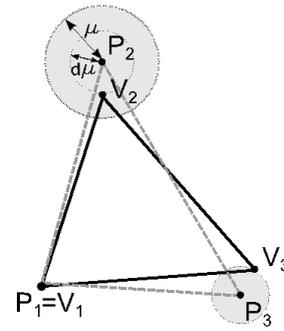
Fig. 3. A fuzzy pattern and a crisp result of image retrieval.

The proposed basic method can be used in the following cases:

- A raster image is previously vectorized;
- A graphical object (shape) is created in a vector format.

In any mentioned case, we suppose that there are two sets of vertices coordinates and two adjacency matrices as input data. The vertices adjacency matrices enable reconstruction of shapes: triangles, rectangles, polygons.

Any object of retrieval (an element of a pattern) is defined by its fuzzy coordinates. Any object in a candidate image is defined by its crisp coordinates.

The pattern can be specified by a vertices adjacency matrix A and a set P of fuzzy coordinates defined as follows:

$$P = \{(p_i, \mu_i) | i \in [1, n]\}, \tag{1}$$

where $p_i = (\tilde{x}_i, \tilde{y}_i)$ is a vertex defined by its crisp coordinates $\tilde{x}_i$ and $\tilde{y}_i$; μ_i is the degree of fuzziness of the vertex p_i; as we mentioned it above in the geometric sense fuzziness degree is equivalent to a radius of a circle with the center in the point p_i; n is the number of vertices.

A candidate image can be specified by a vertices adjacency matrix M and a set C of crisp coordinates defined as follows:

$$C = \{v_i | i \in [1, n]\}, \tag{2}$$

where $v_i = (x_i, y_i)$ is a vertex defined by its crisp coordinates;

n is the number of vertices.

Vertices adjacency matrices A and M are composed in the same way. In particular, matrix A is composed by the following rule:

$$a_{i,j} = \begin{cases} 1 \text{ if } p_i \text{ and } p_j \text{ are connected} \\ 0 \text{ if } p_i \text{ and } p_j \text{ aren't connected} \end{cases} \quad (3)$$

where $i, j \in [1, n]$.

As the first step of the method, we analyze the vertices adjacency matrix A. As a result of this analysis, we obtain data about the shape of the template object, in particular, the vertices number.

At the next step, we begin to analyze candidate images (CIs) from image database for the presence of objects resembling to the pattern object. For this purpose, we take a random vertex in the set C as the initial one and search all other vertices, which belong to the same shape, in the adjacency matrix M.

After reconstructing this shape by the "crisp" matrix M, we need to find the corresponding "fuzzy" vertex in the set P and then the corresponding shape in the "fuzzy" matrix A. Note, values in both adjacency matrices are crisp as it is defined in (3); we use words "crisp" and "fuzzy" (in quotes) to facilitate the understanding to what kind of an object – either a pattern ("fuzzy") or a candidate image ("crisp") – this matrix belongs.

The detection of the corresponding "fuzzy" vertex is fulfilled taking into consideration the fuzziness degree of the vertex. Thus, let v_i (defined by its crisp coordinates (x_i, y_i) be a "crisp" vertex which needs a correspondent "fuzzy" vertex to be found in the set P. Then p_i (defined by its fuzzy coordinates $(\tilde{x}_i, \tilde{y}_i)$ and fuzziness degree μ_i) is a required "fuzzy" vertex if the following inequalities hold true:

$$\begin{cases} \tilde{x}_i - \mu_i \leq x_i \leq \tilde{x}_i + \mu_i \\ \tilde{y}_i - \mu_i \leq y_i \leq \tilde{y}_i + \mu_i \\ (x_i - \tilde{x}_i)^2 + (y_i - \tilde{y}_i)^2 \leq \mu_i^2 \end{cases} \quad . \quad (4)$$

When the first "fuzzy" vertex is found, we search for the next "fuzzy" vertices which belong to the same shape.

If the whole shape of the "fuzzy" object corresponds to the pattern object, we calculate the *degree of resemblance* of these two objects. This degree can be obtained as a result of multiplication of the fuzziness measure for each vertex:

$$R = \prod_{i=1}^{n} \left(1 - \frac{d\mu_i}{\max(\mu_1, \mu_2, ..., \mu_n) + K} \right), \quad (5)$$

where $d\mu_i$ is a true fuzziness degree for the vertex p_i with respect to the "crisp" vertex in the pattern object (physical meaning of both $d\mu_i$ and μ_i is shown in Fig. 3);

n is the number of vertices in the pattern object;

$K \neq 0$ is a small positive number (e.g. $K = 0.01$) which is necessary to avoid zero fuzziness measure for vertices where the degree of fuzziness μ_i has maximum value among all other vertices.

For example, the degree of resemblance of two triangles ($P_1 P_2 P_3$ and $V_1 V_2 V_3$) in Fig. 3 can be estimated as $R = 0.25$ if we define values of fuzziness degree as $\mu_1 = 0$, $\mu_2 = 20$, $\mu_3 = 10$ (see attribute "fuzziness" in Fig. 2) and $d\mu_2 = 10$. If two shapes are equal, then the degree of their resemblance $R = 1$.

Alternatively, the degree of resemblance can be estimated according to [18] as:

$$R(\text{object1}, \text{object2}) = \max(0,$$
$$\min(1, \alpha - d(\text{object1}, \text{object2}))), \quad (6)$$

where $\alpha \geq 1$.

Let us consider two vertices v_i and p_i defined by their coordinates (x_i, y_i) and $(\tilde{x}_i, \tilde{y}_i)$ respectively as object1 and object2. Then $d(\text{object1}, \text{object2})$ can be defined as follows:

$$d(v_i, p_i) = \frac{\max(|x_i - \tilde{x}_i|, |y_i - \tilde{y}_i|)}{\mu_i}, \quad (7)$$

where μ_i is a fuzziness degree.

Let $\alpha = 1 + \frac{1}{\mu_i}$ ($\mu_i > 0$). Then if two triangles $P_1 P_2 P_3$ and $V_1 V_2 V_3$ in Fig. 3 are defined by coordinates $P_1(14,12)$, $P_2(32,86)$, $P_3(80,10)$, $V_1(14,12)$, $V_2(32,76)$, $V_3(86,18)$ and $\mu_1 = 0$, $\mu_2 = 20$, $\mu_3 = 10$, then the degree of resemblance of $P_1 P_2 P_3$ and $V_1 V_2 V_3$ can be estimated in the following way:

$$R(V_1, P_1) = 1,$$

$$R(V_2, P_2) = \max\left(0, \min\left(1, 1 + \frac{1}{20} - \right.\right. \tag{8}$$

$$\left.\left.\frac{\max\left(\left|32 - 32\right|, \left|76 - 86\right|\right)}{20}\right)\right) = 0.55,$$

$$R(V_3, P_3) = \max\left(0, \min\left(1, 1 + \frac{1}{10} - \right.\right.$$

$$\left.\left.\frac{\max\left(\left|86 - 80\right|, \left|18 - 10\right|\right)}{10}\right)\right) = 0.30.$$

Then $R = 1 \times 0.55 \times 0.30 = 0.165$. Thus, this estimation is stricter than the degree of resemblance defined by (5).

The basic method can be used for the comparison of polygons and other primitives in vector images.

C. Advanced Method

Now let us generalize the proposed geometric interpretation of fuzziness from a circle to an ellipse. This approach gives more flexibility in search of complex shapes.

An array of fuzzy coordinates $\tilde{x}, \tilde{y}$ and an array of crisp coordinates x, y are used as initial data for resembling image retrieval in this case. Vertices with fuzzy coordinates $\tilde{x}, \tilde{y}$ define a pattern image. Vertices with coordinates x, y define a candidate image which is considered as resembling the pattern image if coordinates x, y get into the elliptic fuzziness area. The fuzziness area for a vertex with coordinates $\tilde{x}, \tilde{y}$ geometrically is an ellipse with radiuses μx and μy. These radiuses can be either set manually or calculated. In the latter case, fuzziness values μx and μy can be defined in two steps. As the first step, we set them as:

$$\mu x_1 = p_x \times D_{x1}, \quad \mu y_1 = p_y \times D_{y1}, \tag{9}$$

where μx_1, μy_1 are the "first-step" elliptic fuzziness area radiuses for coordinates $\tilde{x}_i, \tilde{y}_i; \tilde{x}_i, \tilde{y}_i$ are "first-step" fuzzy coordinates, $\tilde{x}_i, \tilde{y}_i \in [0,1]$, $i = 1, ..., K_1$; D_{x1}, D_{y1} are variance values for coordinates $\tilde{x}_i, \tilde{y}_i$ of K_1 vertices; K_1 is a quantity of vertices for the first step of the retrieval procedure.

Variance values D_{x1}, D_{y1} can be calculated as follows:

$$D_{x1} = \frac{1}{K_1} \times \sum_{i=1}^{K_1} \left(M_{x1} - \tilde{x}_i\right)^2, \tag{10}$$

$$D_{y1} = \frac{1}{K_1} \times \sum_{i=1}^{K_1} \left(M_{y1} - \tilde{y}_i\right)^2, \tag{11}$$

where M_{x1}, M_{y1} are mean values for coordinates $\tilde{x}_i, \tilde{y}_i$ of K_1 vertices:

$$M_{x1} = \frac{1}{K_1} \times \sum_{i=1}^{K_1} \tilde{x}_i. \tag{12}$$

$$M_{y1} = \frac{1}{K_1} \times \sum_{i=1}^{K_1} \tilde{y}_i. \tag{13}$$

Then the resembling images retrieval accuracy can be defined by coefficients p_x, p_y:

$$p_x = \text{int}\left(\frac{M_{x1}}{D_{x1}}\right), \tag{14}$$

$$p_y = \text{int}\left(\frac{M_{y1}}{D_{y1}}\right), \tag{15}$$

where int is an operation of rounding by truncation of a number fractional part.

If we decrease these coefficients, the accuracy increases. It means that for the second step of the retrieval procedure the values of fuzziness should be decreased according to (16) and (17) as well as the quantity of vertices to be analyzed should be increased.

$$\mu x_2 = \left(p_x - \lambda\right) \times D_{x2}, \tag{16}$$

$$\mu y_2 = \left(p_y - \lambda\right) \times D_{y2}, \tag{17}$$

where μx_2, μy_2 are the "second-step" elliptic fuzziness area radiuses for coordinates $\tilde{x}_i, \tilde{y}_i$; D_{x2}, D_{y2} are variance values for coordinates $\tilde{x}_i, \tilde{y}_i$ of K_2 vertices used in the second step of the retrieval procedure; $\lambda = 1, ..., \min\left(p_x, p_y\right)$.

The resemblance degree of the candidate image and the pattern image for two steps of the retrieval procedure is defined as a product of the resemblance degrees for all vertices:

$$R = \prod_{i=1}^{K} R_i = \prod_{i=1}^{K} \max\left(0; 1 + \right.$$

$$\left. \varepsilon - \left(\frac{(x_i - \tilde{x}_i)^2}{\mu x_l^2} + \frac{(y_i - \tilde{y}_i)^2}{\mu y_l^2} \right) \right), \quad (18)$$

where R_i is a resemblance degree of i-vertex, $i=1,...,K$; K is a quantity of vertices; ε is an arbitrary small positive value (for example, $\varepsilon = 0.0001$); x_i, y_i are crisp coordinates, $x_i, y_i \in [0,1]$; $\tilde{x}_i, \tilde{y}_i$ are fuzzy coordinates, $\tilde{x}_i, \tilde{y}_i \in [0,1]$; $\mu x_l, \mu y_l$ are fuzziness values, $l = \overline{1,2}$.

If vertices are resembling, then $\varepsilon \leq R_i < 1 + \varepsilon$; if vertices are coincident, then $R_i = 1 + \varepsilon$; otherwise $R_i = 0$.

The advantage of this two-step retrieval procedure is evident when an object to be find has a complex shape and a large number of vertices: the first step enables excluding both dissimilar images ($R = 0$) and images with negligible resemblance ($R \rightarrow 0$) by analyzing a limited number of vertices; the second step enables finding resembling images by analyzing a full set of vertices.

To store a pattern image, which is defined by its vertices with an elliptic fuzziness area, we propose to use the extended attribute "fuzziness" with two radiuses defined for every vertex in a modified SVG format (Fig. 4).

```
<svg width="15cm" height="15cm">
  <g    style="fill:none;
        stroke: #0000ff;
        stroke-width:2px">
        <polygon points="100,100 300,400 400,200";
        fuzziness="(0,5), (20,15), (10,15)"/>
  </g>
</svg>
```

Fig. 4. SVG code fragment with "elliptic" fuzziness.

IV. RESULTS AND DISCUSSION

The objective of the experimental part of this research is to prove that the proposed methods ensure retrieval of resembling objects in a set of images. To achieve this goal, the proposed methods have been realized in the MatLab environment and tested on real images. In particular, the proposed retrieval procedure based on an elliptic fuzziness area (the advanced method) has been tested on medical images, namely, heart computer tomography images. The test data set included 100 candidate images. The retrieval request was defined by using two pattern images (Fig. 5 and Fig. 6).

(a)

(b)
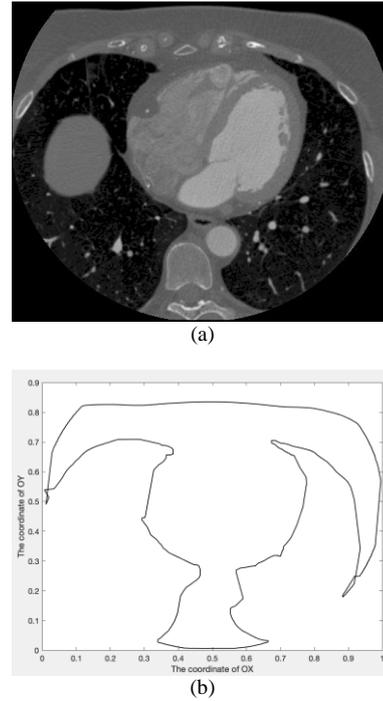
Fig. 5. An example of a heart computer tomography pattern image from the Test Set 1: (a) the original raster image, (b) the vectorized image (img1_01.bmp).

(a)

(b)
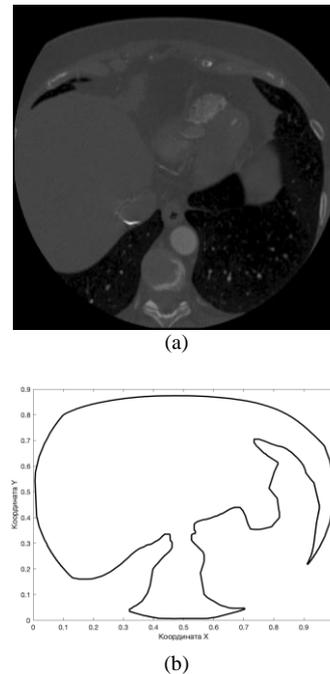
Fig. 6. An example of a heart computer tomography pattern image from the Test Set 2: (a) the original raster image, (b) the vectorized image (img2_01.bmp).

Since original heart computer tomography images are raster images, they have been vectorized by using the method presented in [20]. This method enables description of an object complex shape by any requested number of vertices. It allows us to describe objects on both a pattern image and a candidate image by the same number of vertices. Naturally, the accuracy of the object

vectorization depends on the number of vertices. The first vertex in a shape is the closest to the origin of coordinates.
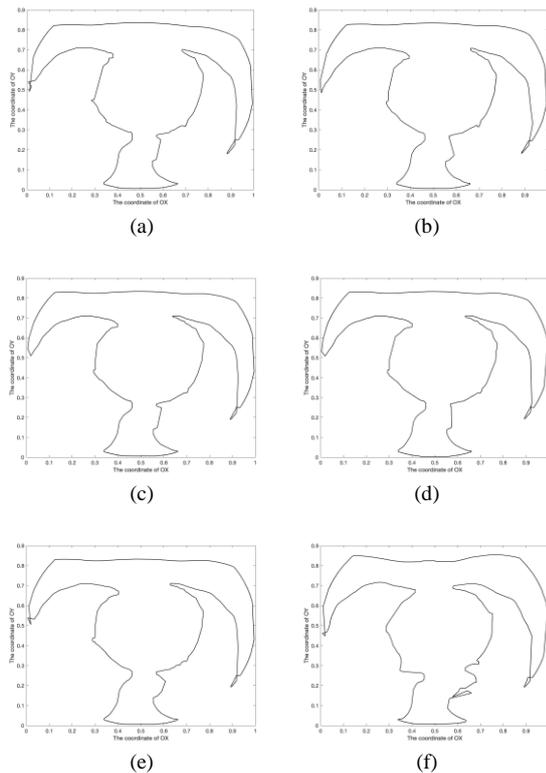
Fig. 7. The Test Set 1 of vectorized heart computer tomography images: (a) img1_02, (b) img1_03, (c) img1_04, (d) img1_05, (e) img1_06, (f) img1_07.

The number of vertices for the first step was equal to 949, for the second step it was equal to 1369. For the first step the fuzziness values were $\mu x_1 = 0.4944$, $\mu y_1 = 0.4607$; for the second step the fuzziness values were $\mu x_2 = 0.4362$, $\mu y_2 = 0.3721$. The result of the image retrieval is shown on Fig. 7. The resemblance degree of these candidate images and the pattern image img1_01 obtained at the second step is given in Table 1.

Table 1. Resemblance degree for img1_01

Image name	Resemblances Degree
img1_01	1,1467
img1_02	1,0232e-05
img1_03	1,2367e-11
img1_04	3,6728e-12
img1_05	2,2502e-18
img1_06	4,6421e-16
img1_07	9,1436e-60
img2_01	0
img2_02	0
img2_03	0
img2_04	0
img2_05	0
img2_06	0
img2_07	0

The number of vertices for the first step was equal to 949, for the second step it was equal to 1369. For the first

step the fuzziness values were, for the second step the fuzziness values were. The result of the image retrieval is shown on Fig. 8. The resemblance degree of these candidate images and the pattern image img2_01 obtained at the second step is given in Table 2.
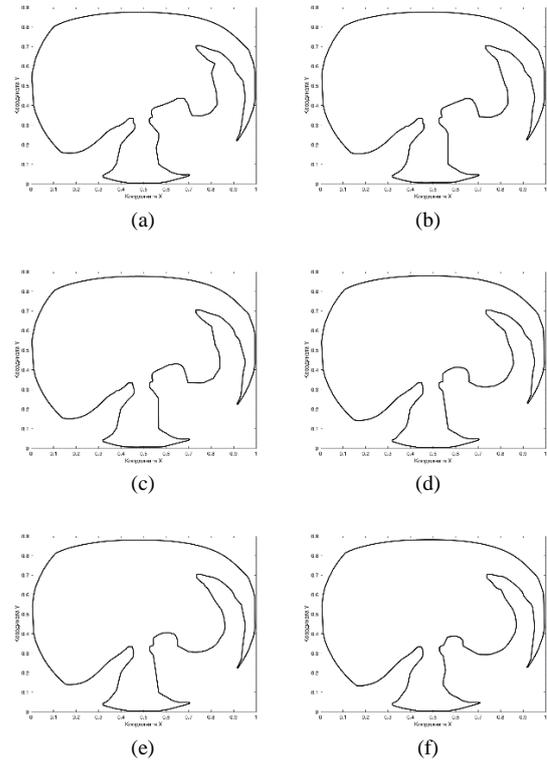
Fig. 8. The Test Set 2 of vectorized heart computer tomography images: (a) img2_02, (b) img2_03, (c) img2_04, (d) img2_05, (e) img2_06, (f) img2_07.

Table 2. Resemblance degree for img2_01

Image name	Resemblances Degree
img1_01	0
img1_02	0
img1_03	0
img1_04	0
img1_05	0
img1_06	0
img1_07	0
img2_01	1.1467
img2_02	0.0122
img2_03	0.0066
img2_04	8.3672e-04
img2_05	1.1979e-06
img2_06	9.0422e-07
img2_07	3.3365e-09

The comparative analysis of the resemblance degree values (Table 1 and Table 2) and the object shapes (Fig. 5, Fig. 6, Fig. 7, and Fig. 8) allows concluding that the resemblance degree adequately characterizes the visual resemblance of a pattern image and a candidate image. All images which are characterized by zero resemblance degree contain artifacts such as loops and distortions. At the same time, the candidate images with non-zero

resemblance degree are visually similar to the corresponding pattern image. What is important that the more similar the pattern image and the candidate image are, the higher the computed value of their resemblance degree is. In particular, Table 1 shows that the pattern image img1_01 (Fig. 5) has the highest resemblance degree with itself and it has zero resemblance degree with the images from the Test Set 2 which have another shape. The rest of the resemblance degree values are non-zero values and the lowest value (9,1436e-60) corresponds to the image img1_07 shown at Fig.7 (f) which includes vectorization artifacts (loops) and, thus, it can be considered as the least resembling image. Similar results are shown in Table 2: the highest resemblance degree is obtained from comparison of the pattern image img2_01 (Fig. 6) with itself (1.1467). The resemblance degree of the pattern image img2_01 and the images from the Test Set 1, which have another shape, is equal to zero. The rest of images have non-zero resemblance degree.

The accuracy of the proposed methods can be improved by increasing the number of vertices in the shapes of both a pattern image and a candidate image.

The obtained experimental results allow concluding that the proposed approach based on fuzzy logic can be used for retrieval of vectorized images, including medical images. The advantage of this approach is that most medical images have large size due to their high resolution. Thus, their vectorization can be useful in cases when the images are stored to be a subject of further retrieval. In this case, the use of modified SVG format can help to decrease required data storage volumes.

V. Conclusions

The use of fuzzy values for the description of a template (a pattern image) allows to make the search more flexible and therefore to achieve better results of vector graphics retrieval.

Both proposed methods can be especially useful in cases when the shape of an object to be retrieved is described roughly. Human brain makes the search in such uncertain conditions easy, but algorithmic methods can give unacceptable results. Methods based on artificial intelligence approaches enable improving retrieval results, but they are quite complicated in their software realization [21]. Unlike methods based on machine learning which are time- and resource-consuming, the proposed methods based on fuzzy logic do not require training and high computing resources. At the same time, they allow to achieve comparative results in a simpler way. Therefore, the proposed retrieval approach based on fuzzy patterns can be used for search of resembling shapes in vector images. Initial data can be either raster or vector images and objects to be retrieved can have any complex shape. The use of the proposed methods in medical information systems enables efficient study of clinical cases. These methods can be also used for image classification. Application of the proposed methods enables development of new algorithms for data analysis of resembling image groups.

References

[1] Fujita, R. and Hayashi, T. Vector Image Retrieval Based on Approximation of Bezier Curves with Line Segments, in *IEEE Pacific Rim Conf. on Communications, Computers and Signal Processing* (2011), pp. 431–436.

[2] Kaiyuan Jiang et al. Information Retrieval through SVG-based Vector Images Using an Original Method, in *IEEE Int. Conf. on e-Business Engineering* (2007), pp.183–188.

[3] Md. Khalid Imam Rahmani and M. A. Ansari. Fuzzy Image Retrieval: Recent Trends, in *Indian Journal of Industrial and Applied Mathematics*. 4 (2013), 131–137.

[4] Hayashi, T. et al. Retrieval of 2D vector images by matching Weighted Feature Points, in *IEEE Int. Conf. on Image Processing* (2013).

[5] Pu J. and Ramani K. On visual similarity based 2d drawing retrieval, in *Journal of Computer Aided Design*. 38, 3 (2006), pp. 249–259.

[6] Yixin Chen and James Z. Wang. A Region-Based Fuzzy Feature Matching Approach to Content-Based Image Retrieval, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 24, 9 (2002), pp. 1–16.

[7] Chin-Sheng Chen and Chi-Min Weng. An Efficient Retrieval Technique for Trademarks Based on the Fuzzy Inference System, in *Applied Sciences*. 7(8) (2017), 1–24.

[8] Wang Xiaoling and Xie Kanglin. Application of the Fuzzy Logic in Content-based Image Retrieval, in *Journ. of Computer Science Technology*. 5 (2005), 19–24.

[9] Pedro Martins et al. Clip Art Retrieval Combining Raster and Vector Methods, in *the 11th Int. Workshop on Content-Based Multimedia Indexing* (2013).

[10] Avinash N. Bhute and B. B. Meshram. Content Based Image Indexing and Retrieval, in *Int. Journal of Graphics & Image Processing*. 3, 4 (2013), pp. 235–246.

[11] Minakshi Banerjee Malay K. Kundu. Content Based Image Retrieval with Fuzzy Geometrical Features, in *IEEE Int. Conf. on Fuzzy Systems* (2003), pp. 932–937.

[12] Raghu Krishnapuram et al. Content-based image retrieval based on a fuzzy approach, in *IEEE Transactions on Knowledge and Data Engineering* (2004), 1185–1199.

[13] Yusuke Uchida, Shigeyuki Sakazawa, Shin'ichi Satoh. Image Retrieval with Fisher Vectors of Binary Features, in *ITE Transactions on Media Technology and Applications* (2016), pp. 1-11.

[14] Manuel Fonseca et al. Retrieving Vector Graphics Using Sketches, in *Lecture Notes in Computer Science*. 3031 (2004), pp. 66–76.

[15] Qingyong Li et al. Image Retrieval Based on Fuzzy Color Semantics, in *IEEE Int. Conf. on Fuzzy Syst.* (2007), 1–5.

[16] Scalable Vector Graphics (SVG), available at: http://www.w3.org/TR/SVG11

[17] Extensible Markup Language (XML), available at: http://www.w3.org/TR/REC-xml

[18] De Cock, M. and Kerre, E. On (un)suitable fuzzy relations to model approximate equality. *Fuzzy Sets and Systems*, 133 (2003), pp. 137–153.

[19] De Cock, M. and Kerre, E. Why fuzzy T-equivalence relations do not resolve the Poincaré paradox, and related issues, in *Fuzzy Sets and Systems*, 133 (2003), 181–192.

[20] Shkurat O., Sulema Y., Dychka A. Complicated Shapes Estimation Method for Objects Analysis in Video Surveillance System, *KPI Science News*, 3 (2018), 53–62.

[21] Karamti H., Tmar M., Gargouri F. Vectorization of Content-based Image Retrieval Process Using Neural Network, in *16th Int. Conf. on Enterprise Information Systems* (2014), pp. 435–439.

Authors' Profiles

Dr. Yevgeniya Sulema is an Associated Professor at the Computer Systems Software Department and a Vice-Dean at the Faculty of Applied Mathematics of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". She received her Ph.D. degree from the National Technical University of Ukraine "Kyiv Polytechnic Institute" in 1999. She is a member of the editorial advisory board of Systemics, Cybernetics and Informatics Journal and a member of the program committees at several international conferences. She is a member of the International Institute of Informatics and Systemics. The research in her laboratory of Multimedia, Mulsemedia and Immersive Technologies covers multimodal data representation and processing, mulsemedia and immersive applications, image and audio processing, and multimedia data protection methods. She is an author of more than 130 scientific publications. She participates in numerous European and national research projects. She teaches the course on multimedia technologies and the course on digital signal and image processing to Master program students and the course on mathematical and algorithmic fundamentals of computer graphics to Bachelor program students at her university. She also gave lectures on multimedia, mulsemedia and immersive technologies to students at several European universities as a guest lecturer within academic mobility projects.

Prof. Dr. Etienne E. Kerre obtained his M.Sc. degree in Mathematics in 1967 and his Ph.D. in Mathematics in 1970 from Ghent University. Since 1984, he has been a lector, and since 1991, a full professor at Ghent University. In 2010 he became a retired professor. He is a referee for more than 80 international scientific journals, and also a member of the editorial board of international journals and conferences on fuzzy set theory. He was an honorary chairman at various international conferences. In 1976, he founded the Fuzziness and Uncertainty Modeling Research Unit (FUM) and since then his research has been focused on the modeling of fuzziness and uncertainty, and has resulted in a great number of contributions in fuzzy set theory and its various generalizations. Especially the theories of fuzzy relational calculus and of fuzzy mathematical structures owe a very great deal of him. Over the years he has also been a promotor of 30 Ph.D's on fuzzy set theory. His current research interests include fuzzy and intuitionistic fuzzy relations, fuzzy topology, and fuzzy image processing. He has authored or co-authored 25 books, and more than 500 papers appeared in international refereed journals and proceedings.

Oksana Shkurat is a PhD candidate of the Computer Systems Software Department, the Faculty of Applied Mathematics at the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". She is member of "Multimedia Data Processing – Research Group" (MDP-RG) of the Computer Systems Software Department, the Faculty of Applied Mathematics at the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". Her research interests are technologies of medical image processing, intelligence, and expert systems.