

Spatch Based Active Partitions with Linguistically Formulated Energy

Arkadiusz Tomczyk, Michał Pryczek, Stanisław Walczak,
Konrad Jojczyk, Piotr S. Szczepaniak

*Institute of Information Technology
Technical University of Lodz
Wólczańska 215, 90-924 Łódź, Poland
tomczyk@ics.p.lodz.pl*

Abstract. *The present paper shows the method of cognitive hierarchical active partitions that can be applied to creation of automatic image understanding systems. The approach, which stems from active contours techniques, allows one to use not only the knowledge contained in an image, but also any additional expert knowledge. Special emphasis is put on the efficient way of knowledge retrieval, which could minimise the necessity to render information expressed in a natural language into a description convenient for recognition algorithms and machine learning.*

Keywords: *image understanding, image segmentation, active partitions, linguistic.*

1. Introduction

In the traditional approach to image analysis, an image is first preprocessed (e.g. by means of noise elimination or edge retrieval) and then analysed by means of segmentation, which employs such techniques as thresholding, region growing and splitting, etc. Such approaches are described in [1, 2, 3]. Most of them use only the knowledge contained in the image. However, researchers e.g. in [4] have started

to notice the necessity of adding expert knowledge, the lack of which disables higher-level analysis. Therefore, the present paper is a creative development of the concept of active contours, which, in a lower-level analysis, is able to use any additional knowledge available. The present discussion proceeds in two directions. Firstly, it focuses on the idea of cognitive hierarchical active partitions that enables one to resign from a single pixel analysis (although it allows such an analysis) in favour of hierarchical analysis of semantically relevant objects. Secondly, it offers methods that allow utilisation of expert knowledge expressed in a natural language, which can be particularly helpful to specialists in domains other than computer science.

The paper is organized as follows: in section 2 the main idea of cognitive hierarchical active partitions approach as well as methods of linguistic formulation of expert knowledge are presented, section 3, section 4, section 5 describe three examples of application of the proposed methods with three different kinds of semantic objects, the last section focuses on the summary of the proposed approach.

2. Active partitions

2.1. Active contours

The term active contours was first introduced with reference to the snakes method by Kass, Witkin and Terzopoulos in [5]. In that work, contour, defined as a parametrised curve, evolved until the desired object was identified in the image. The purpose of the contour's evolution was defined by energy function, which was the objective function of the optimisation process. For the purpose of optimisation, calculus of variations was used. Its application led to iterative solution of partial differential equation set and, as a result, to iterative changes of the contour itself. The literature abounds in modifications and improvements of this basic method. Cohen in [6] introduced additional pressure forces able to compress or expand the contour, which was further elaborated by Ivins and Porrill in [7], who introduced region energy and region forces. Both of these changes allowed additional knowledge to be taken into account, in order to prevent the process from getting stuck in local minima of energy function, which happened in the case of wrong initialisation of the contour. Similar reasons can be ascribed to the introduction of gradient vector flow by Xu and Prince in [8] and distance potential by Cohen and Cohen in [9]. Other significant modifications can be found also in [10, 11, 12].

The snakes method was the first but not the only variation of active contours

to have emerged in the literature. The most significant of them are geometric active contours, which differ from the snakes method in that they do not take into account the information about parameterisation of curves representing the contour (the snakes method is regarded as a parametric active contours method). Such an approach was introduced parallelly by Caselles, Catta, Coll and Dibos in [13] and by Malladi, Sethian and Vemuri in [14] and used during the optimisation of the level-set method, earlier applied by Osher and Sethian in [15] for the solution of front propagation problem. The main advantage of this method was the possibility to change easily the contour's topology during its evolution. A variation of this technique was geodesic active contours introduced in two different ways by Caselles, Kimmel and Sapiro in [16] and Yezzi, Kichenassamy, Kumar, Olver and Tannenbaum in [17, 18]. Here, similarly to the snakes method, the purpose of evolution was formulated in the form of energy function, and the evolution was formulated in the form of forces that influence the contour. This enabled Xu, Yezzi, Prince and Hopkins [19, 20] to indicate the dependencies between those methods, which helped to share experiences.

Other types of active contours include: active shape models proposed by Cootes and Taylor in [21, 22], where the contour was described by a set of landmark points, and the appropriately tested point distribution model helped to impose additional limitations on the evolving contour, Brownian strings introduced by Grzeszczuk and Levin in [23], where the contour is described linguistically and simulated annealing algorithm is used for the contour evolution, the approaches discussed by Jacob, Blu and Unser in [24] and by Schnabel and Arridge in [25], in which the contour is represented by splines, the approach of Staib and Duncan [26], in which Fourier descriptors are used and finally the active ray approach described by Denzler and Niemann in [27]. Although substantially different in regard to contour's description model, the above methods share a few characteristics.

First of all, they all apply the notion of contour. Although it is not precisely defined in any of the works, a contour can be intuitively defined as something that can indicate those image fragments that represent visible objects. In other words, it is able to divide an image into the part representing the object and the part representing the background. For formal description, let o denote a pixel represented by point p and let \mathcal{O} and \mathcal{P} denote the set of all pixels and the set of points representing the pixels, respectively. As shown in [28, 29], the contour can be regarded as classifier k of image points, and therefore, the classifier of pixels, which assigns the pixels one of two labels: l^o or l^b . The labels denote the object and the background, respectively. The classifier partitions the set of image points and, consequently, the

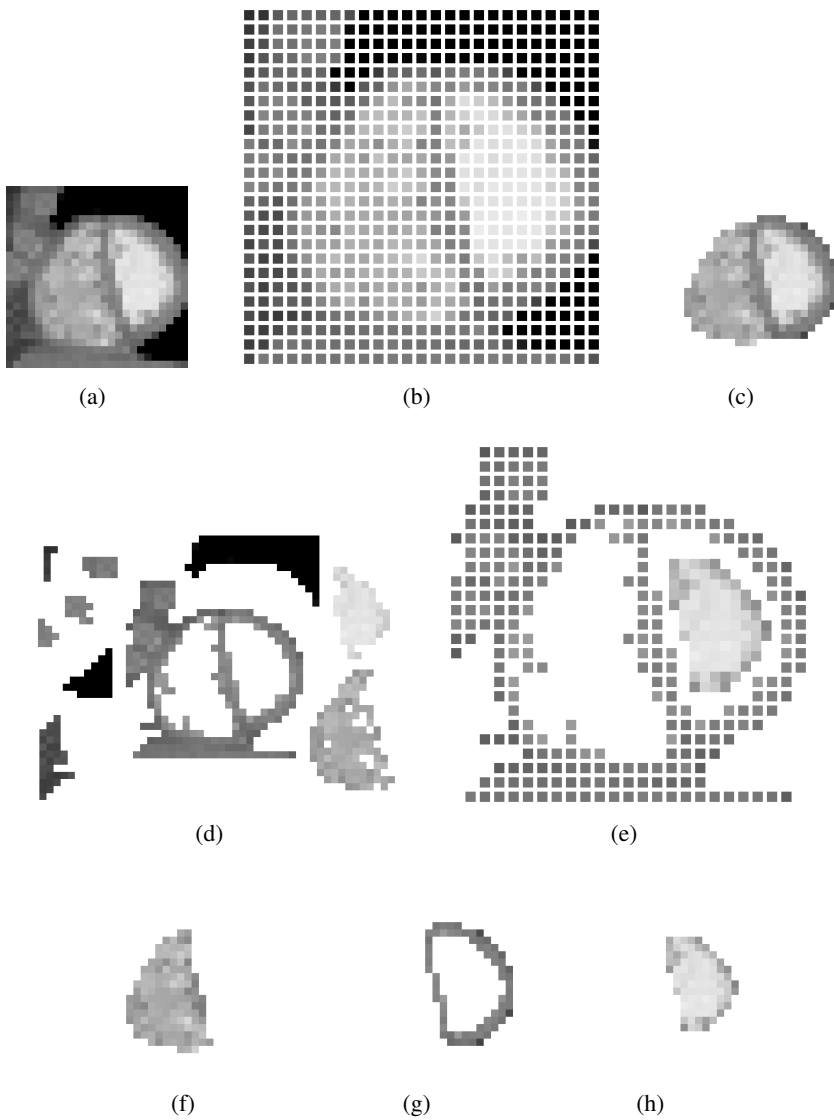


Figure 1. Localisation of heart: (a) - original image, (b) - initial knowledge (all the pixels in the image), (c) - heart that is sought, (d) - regions with similar colours, (e) - knowledge that can be used to localise cardiac muscle, (f) - right ventricle (blood with contrast), (g) - cardiac muscle, (h) - left ventricle (blood with contrast)

set of pixels, into those composing the object:

$$\mathcal{O}^{l^o} = \{o \in \mathcal{O} : k(o) = l^o\} \quad (1)$$

and those composing the background:

$$\mathcal{O}^{l^b} = \{o \in \mathcal{O} : k(o) = l^b\} \quad (2)$$

Of course $\mathcal{O}^{l^o} \cup \mathcal{O}^{l^b} = \mathcal{O}$ and $\mathcal{O}^{l^o} \cap \mathcal{O}^{l^b} = \emptyset$. The object detected in the image is represented by the set of all points corresponding to pixels labelled l^o .

Secondly, all active contour methods allow one to specify the expectations concerning the object, directly or indirectly, in the form of energy function E . This helps to include any knowledge that might be useful for pixels' classification. The knowledge can be included in the image (information about other elements of the image), which places the problem among context classification problems, or it can refer to general knowledge connected with the origin of the image and information associated with the image or coming from the experience of people who are able to analyse a particular class of images.

Finally, all the methods employ a specific way of optimal contour finding. That way, usually iterative, leads to classifier's evolution and, as a result, to further changes of the partition. The character of energy function and the method of evolution depend on the model of a classifier or, to be more precise, on the model's parameters that are optimised, and have great influence on the results.

2.2. Active partitions

The above approach can be generalised, given that image analysis generally consists in indicating those image fragments that carry a certain semantic meaning, which does not necessarily have to be (and most often is not) performed at the level of particular image points. For the purpose of generalisation, let us assume that \mathcal{P} represents a set of patches that have already been assigned semantic meaning. Individual patch $p \subset \mathbb{R}^2$ can denote, in the most simple case, a set containing a single point that has been assigned a colour - like in the active contour method, but also a division line between two regions of different characteristics, regions of uniform colour or regions representing objects of higher-level semantic meaning significant for a class of images analysed. Although these spatches correspond to separate objects, they do not have to be separate regions. Additionally, let us assume that k is not a binary classifier, but it assigns each object one of the labels

L , denoted as l_1, \dots, l_L . Then the partition defined by such a classifier consists of L parts:

$$\mathcal{O}^l = \{o \in \mathcal{O} : k(o) = l\} \quad (3)$$

for $l = l_1, \dots, l_L$, which, like previously, are separate and compose the whole set \mathcal{O} . In this case, the object corresponding to label l is naturally represented in the image by a sum of patches $p \subset \mathbb{R}^2$. If optimal partition finding process is performed in the same way as in active contours, that is, by selecting the model of the classifier and then iterative finding of its optimal parameters for a given energy function, then one can speak of a new image analysis method referred to as active partitions. The name of the method reflects evolution of the partition that results from classifier's evolution.

2.3. Cognitive hierarchical active partitions

The process of image understanding is usually more complex than direct finding of the regions that represent semantically-relevant objects. It can be regarded as a bottom-up process, in which the objects are organised in terms of increasing complexity. The process is complex, as it should include not only image information but also additional knowledge about the image analysed. What is also important for image content understanding is the experience of a person who is analysing the image, which refers not only to expert knowledge but also common life experience connected for example with the culture area the person has been raised in. Such an approach has already been advocated in [4, 30, 31].

Since the approach requires retrieval of semantically meaningful objects, one can make use of the active partitions approach described above. For this reason, the process of image understanding can be presented as iterative image partition, with each part being ascribed a semantic meaning in such a way that, after each iteration, newly understood objects constitute additional knowledge about the image. The new knowledge can be used in the next step. According to the designation used in previous chapters, one step can be represented as follows:

Algorithm 1: Cognitive hierarchical active partitions

select \mathcal{O}_i **from** \mathcal{O}
find the partition $\{\mathcal{O}_i^{l_1}, \dots, \mathcal{O}_i^{l_{L_i}}\}$ of \mathcal{O}_i
assign $\mathcal{O} = \mathcal{O} \cup \{\mathcal{O}_i^{l_1}, \dots, \mathcal{O}_i^{l_{L_i}}\}$ **and** $i = i + 1$

Of course, in the case of automatic analysis of digital images, initially set O contains just pixels – the only information available. It is worth mentioning that, at each step, every object belonging to set O is connected with a certain patch belonging to set \mathcal{P} . Due to its structure and the fact that for partition finding, especially in the more complex cases, one has to employ an algorithm that uses external knowledge and expert's experience, the method has been called cognitive hierarchical active partitions, in short - CHAP. Although it seems natural to use active partitions in the second step of the above algorithm (or active contours - at a pixel level), the partition can be performed by means of any segmentation algorithm or either supervised or unsupervised classification algorithm. These approaches can be mixed.

A possible example of CHAP in Fig. 1 is shown. In Fig. 1b the initial set, which contains all the pixels, is presented. The analysis of individual pixels helps to find only the regions of approximately the same colour, which represent structures that look the same on images. This is visible in Fig. 1d. However, these structures can represent various anatomical structures, which is impossible to determine without medical knowledge. In the example, it is particularly visible in the attempt to locate the wall of the left ventricle Fig. 1g. While ventricle cavity is clearly visible due to the contrast in patient's blood, the muscle merges with neighbouring tissues. Only with the help of additional information about the location of left ventricle shown in Fig. 1h combined with anatomical knowledge, can one arrive at proper image understanding and finding the heart depicted in Fig. 1c. The example illustrates the necessity of hierarchical approach taking into account the accumulated knowledge, that was gathered earlier, and additionally demonstrates that in one CHAP step it is sometimes necessary to use different kind of objects found so far (localisation of cardiac muscle is possible when both patches representing ventricle cavities and patches representing separate pixels allowing to divide regions of the similar colour) which is shown in Fig. 1e.

The above discussion refers to traditional 2D images, in which patches p are subsets of \mathbb{R}^2 . However, CHAP can be applied also to image sequences - both spatial and video ones. Then, only patches are subsets of a corresponding space \mathbb{R}^n for $n \in \mathbb{N}$. Therefore, instead of patches, a new term will be used, namely spatches - a blend of spatial and patch.

2.4. Linguistic

In order to understand an image, it is essential to combine the knowledge contained in an image with external knowledge and experience of an expert. This makes methods of knowledge representation in CHAP systems important.

Knowledge about an image is contained in set O that includes semantically defined objects visible in an image (location of the objects describes the corresponding spatches from set \mathcal{P}). Because they carry semantic meaning, each of them can be described by such features as location and colour of a pixel, coordinates of circle's centre and circle's radius, coordinates of end points of a segment, coordinates of a centroid, colour and descriptors of a region's shape, etc. Such information constitutes full knowledge about the objects but might not be always useful in such a form, because it is different from human way of description. It is hard to imagine that an expert could describe an object by means of the features mentioned above, e.g. *there are two spots of a certain colour and a certain diameter, the first of which has the centroid in one point and the second – in the other*. The expert is much more likely to express his/her statement in a natural language, e.g. *there are two bright spots of average size, one of which is located to the left of the other*.

Analogously, in order to find an object in an image, it is necessary to use experience of an expert. However, to be used in a computer system, such an experience has to be expressed in an appropriate form. Unfortunately, the form of expression which is convenient to recognition algorithms or machine learning is not convenient for a human expert. The problem is usually solved in two ways. In the first one, an expert provides a set of examples together with their interpretations. Although often applied, such a solution requires accuracy, great amount of work (the set should be representative and consistent) and time (gathering a sufficient number of examples may take years). Therefore, more and more often, also during knowledge description, the other method is used, i.e. linguistic descriptions, which allows an expert to express his knowledge in a way resembling a natural language.

Thus, the present paper shows how linguistic knowledge can be used in the CHAP method. Three examples will be presented, in which:

- knowledge about an object is rendered from a natural language to mathematical formulae
- knowledge expressed by rules described in a natural language is directly used by means of fuzzy sets that represent linguistic variables and by means

of a fuzzy controller which, basing on those rules, performs reasoning

- knowledge about image and object is represented by a graph, whose vertices are representations of semantically defined objects, while the edges can reflect linguistic relations between the objects.

3. Pixel spatch

3.1. Problem

The present section shows how a linguistic description provided by an expert can be directly utilised. As a test problem we will use artificially generated black and white images depicting an elongated object which has changeable location, orientation and scale. Thanks to the method applied, the localisation of the object is not difficult. However, in practice objects usually adjoin each other or overlap and are affected by noise caused by image acquisition process. Thus, to make the example more real, random noise is added to the images. The noise has a form of randomly located additional objects. Therefore, information about a single pixel is not sufficient and it is necessary to add knowledge about the shape of the object wanted.

3.2. Spatch approach

For the localisation of the type of objects described above, we use a single step of CHAP, which considers a full set of pixels O_1 . Active partitions approach is applied, which this time is limited to active contours. As the model of classifier k , and consequently, the model of a contour, potential active contours method has been chosen, described in detail in [32, 33, 34].

Energy The research applies external energy component E_z described in the mentioned papers, which punishes for all dark pixels located outside the contour (label l^b) and all bright pixels located inside the contour (label l^o). However, in the case of noise, energy not taking into account the relations between pixels has no chance to obtain satisfactory results. Therefore, it is necessary to introduce an additional component E_f describing the shape.

For this purpose, a typical model of a fuzzy controller based on the Zadeh-Mamdani model described in [35, 36] is applied. The fuzzification using singleton

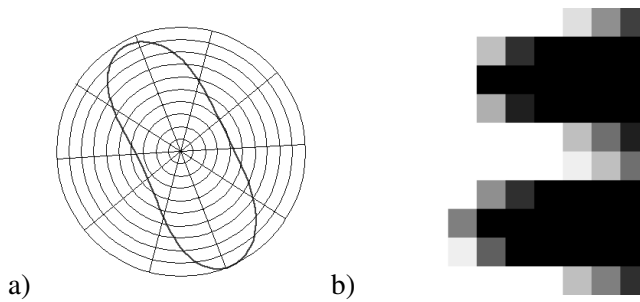


Figure 2. Shape histogram: (a) - contour with ring fragments, (b) - histogram

membership functions, reasoning with Larsen rules and center average defuzzification are used. For encoding the evaluated shape, the shape histogram, which contains information about contour's location in ring fragments of a circle centered round the centroid of the contour and with minimal radius allowing for the inclusion of the contour, is chosen. Since it is considered with relation to a circle determined for each contour separately, the description is independent of the location and scale of the contour. In order to make it also rotation-independent, it is presented starting from subsequent radii with the final result being the smallest response of the controller for those presentation. Fuzzy controller has been prepared in such a way as to evaluate the content of each ring fragment and, for this purpose, it is provided with such data as the distance from the fragment to the circle's centre, the angle between the fragment and the horizontal axis and the percentage area of the fragment occupied by the contour. As an output, the controller determines if the fragment's potential is low (as it should be) or high (as it should not be). Energy E_f is defined as a sum of potentials that are the responses of the controller for each ring fragment. All the inputs and outputs were described using linguistic variables and the proper set of rules representing the expert's knowledge in natural language was prepared.

Listing 4. Sample rules

- 1 IF dist=near AND perc=big THEN pot=low
- 2 IF dist=near AND perc=med THEN pot=high
- 3 IF dist=near AND perc=small THEN pot=high
- 4 IF angle=east AND dist=far AND perc=big THEN pot=high
- 5 IF angle=east AND dist=far AND perc=med THEN pot=high

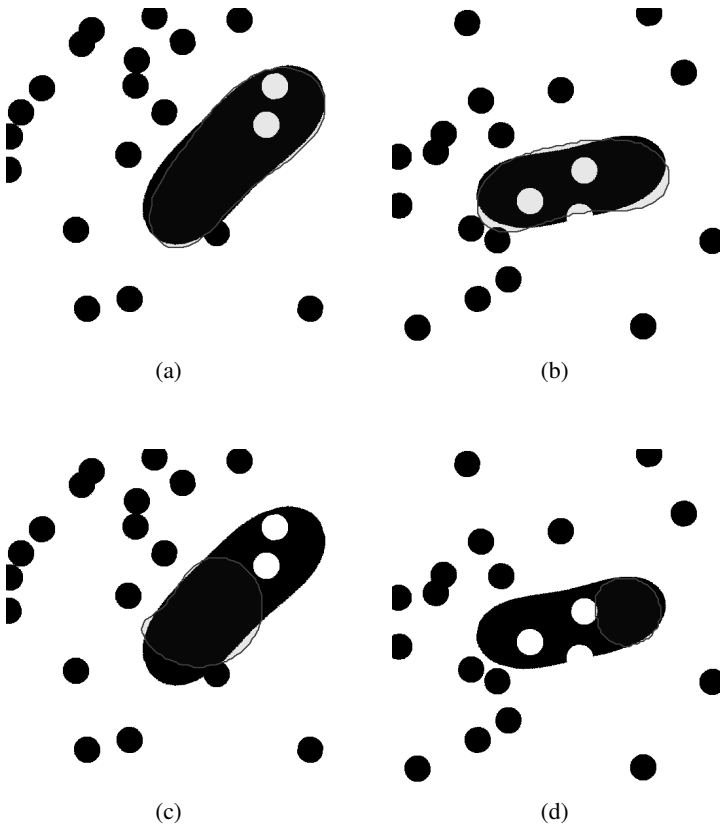


Figure 3. Sample results of potential active contour approach: (a), (b) - both components of the energy, (c), (d) - energy without E_f component

```

6 IF angle=east AND dist=far AND perc=small THEN pot=low
7 ...

```

Additionally, E_f has been enriched by an element that evaluates the size of a contour, in order to avoid contour's reduction to a very small size.

Evolution Simulated annealing has been chosen as an evolution method as, with proper selection of parameters, it has a chance to avoid local minima of energy function E . Its application (solution generators modifying parameters of sources of potentials) was described in the works mentioned above.

3.3. Sample results

Sample results in Fig. 3 show that application of additional component E_f allows one to obtain satisfactory results even if an object is largely distorted by noise. At the same time, it is visible that the lack of this component brings unsatisfactory results (the contour is matched only with the information contained in single pixels, in this case - their colour, while their location against each other and spatial relations are also important). Of course, when the noise is big, even additional knowledge may prevent a satisfactory solution.

4. Line spatch

4.1. Problem

The present chapter presents the application of active partitions method to detection of spicular lesions in mammograms. Spicular lesions are pathological changes in breast with irregular centres and numerous fibrous spicules. Due to the spicules, a lesion is shaped like a star in radiological images, which is shown in Fig. 4 and in Fig. 5. Detection of those changes is crucial, since they are often indicative of a breast cancer.

4.2. Spatch approach

In the proposed method we look for the segments creating star-shaped forms. For this reason, in the mammogram all visible lines should be detected. The attempts to automate this process have not brought satisfactory results due to the character of mammographic images. The segments detected are numerous and very small. Therefore, the first step of CHAP for O_1 that represents a full set of pixels has been performed manually and, as a result, set O contains pixels described by their coordinates and lines described by the coordinates of their end points. Of course all those objects have their reflection in spatches, which, for a line, is shown in Fig. 4a. Further in this chapter, the second step of CHAP is described, assuming that O_2 consists only of segments detected in the first step:

$$o = \overline{AB} \tag{4}$$

where $o \in O_2$ while $A, B \in \mathbb{R}^2$ define end points of a segment.

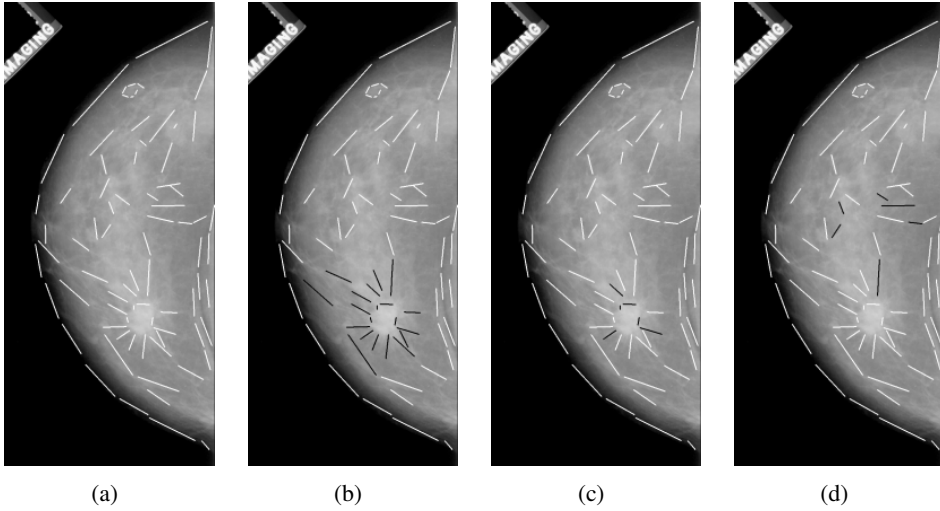


Figure 4. Sample results for line spatches: (a) - all the segments, (b) - segments of spicular lesion (black), (c) - segments of result for both energy components, (d) - segments of result without E_t component

Energy In active partitions, it is necessary to use energy function E , which applies the knowledge about the object that we look for. In the example presented, the function evaluates classifier k by taking into account two kinds of knowledge, described linguistically as follows:

- E_s - shape energy reflecting the following statement: *segments composing a spicular lesion should create a star-shaped form, and they should intersect at right angle a circle, the centre of which is the centroid of those segments, and the radius of which equals average distance between the circle's centre and the segments' midpoints*. To find its value for those segments in O_2 that are assigned the label l^o by the current classifier, the circle with centre C being the arithmetic mean of coordinates of end points A and B of and with radius equal to arithmetic mean of the distances of the centroids of those segments to centre C , is identified. Next, the punishment equal to the length of segment is counted when the segment does not intersect the circle. The reward, depending also on the length of the segment, is counted when the segment intersects circle but in this case it is the bigger the more approximate is the angle of that intersection to the right angle. The lengths of the

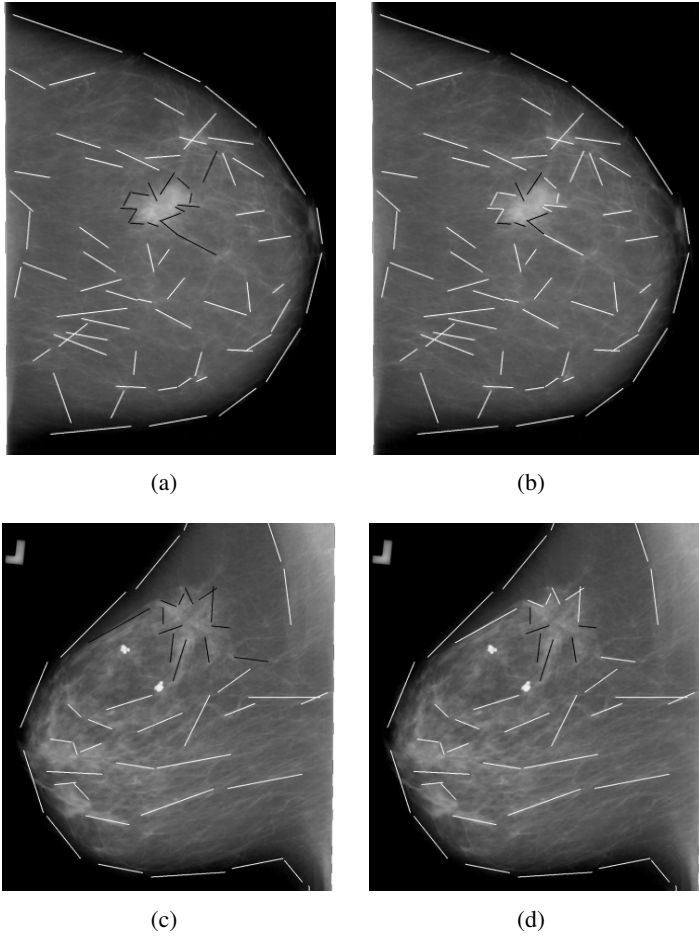


Figure 5. Sample results for line spatches: (a), (c) - segments of spicular lesion (black), (b), (d) - segments of result (black)

segment are normalised using the radius of the circle so that the object of different sizes have equal chances.

- E_t - centre energy reflecting the following statement *the centre of the spicular lesion in the image is an area of high brightness*. Its value reflects the reward equal to the square of average brightness of the pixels inside the circle found in the above energy component.

The second component is necessary as without it the set of segments shown in Fig. 4d could be found. Its elements create a shape that resembles a star, however, the location of the centroid does not overlap with the tumour.

Evolution As an optimisation mechanism the method of simulated annealing is used. In order to do it, a proper solution generator must be defined. In the case discussed, its main task is to modify the classifier k . The modification consists in changes of label values from l^o to l^b or the other way round of the randomly chosen segment.

4.3. Sample results

Sample results of algorithm performance are depicted in 5. The results show that the method presented does not offer an ideal solution. However, it indicates quite precisely the potential location of a tumour. The level of efficiency is satisfactory, as the aim of the method is not to replace a radiologist but to indicate to those fragments of the image that he/she should pay special attention to. Further research is being conducted for the enhancement line detection automation, as well as energy function improvement so it can more efficiently reflect the knowledge linguistically expressed by doctors.

5. Circle spatch

5.1. Problem

The present chapter focuses on the problem of automatic localisation of ventricular system in CT images of the brain. Exact recognition is important from the diagnostic point of view, since changes in the system, particularly deformities of shape, asymmetry, contraction or expansion are indicative of pathological changes in the central nervous system.

5.2. Spatch approach

The CHAP approach requires the selection of components of set O , which is indirectly determined by the selection of types of classifiers k . Since the retrieval of meaningful objects causes difficulties typical of such an analysis, we look for an easier solution, which would, on the one hand, reduce the granularity problem (depart from pixel analysis) and, on the other hand, help to avoid the problem of complex retrieval of objects.

The approach described in the present chapter is based on image partition performed by connecting points of a similar colour into circular regions, which will constitute set O_2 . As a result, the number of objects is reduced drastically, which enables the application of more complex analysis methods. Moreover, the very definition of a circle carries information about the size and location of a region of approximately uniform colour, while the set of circles contains information about neighbouring regions. Knowledge, which is absent if the image has not been pre-processed, is used and structured by the methods presented below.

If there are no requirements as to semantic quality, a simple circle process is performed by means of a simple, unsupervised brute force algorithm, on the basis of O_1 which comprises all pixels. Having determined minimal and maximal radius of the circles and error bound of colour cohesion, starting with the circles of maximal radius, all possible locations, which do not overlap with the already found circles, have to be systematically checked and added to the set if they fulfil the colour uniformity criterion. Despite inconveniences resulting from computational complexity, a big advantage of the algorithm is its determinism, simplicity and, in particular, the lack of necessity to use any additional knowledge, which enables unsupervised generation of descriptions.

Similarly to pixels, in the context of image analysis, it is the neighbourhood of the circles that determines their meaning and importance. The description will be recorded in whole in the circle graph, also referred to as the graph of linguistic description. The choice of graph languages as a tool of linguistic description is not accidental. Being formal and semantically precise, it is also extendable and expressive enough to carry both information about object's structure and external knowledge.

Circle graph Let $G = (V, E)$ be an undirected graph of linguistic description. Every vertex $v \in V$ is bijectively mapped to a circle in O_2 :

$$o = K(v) \tag{5}$$

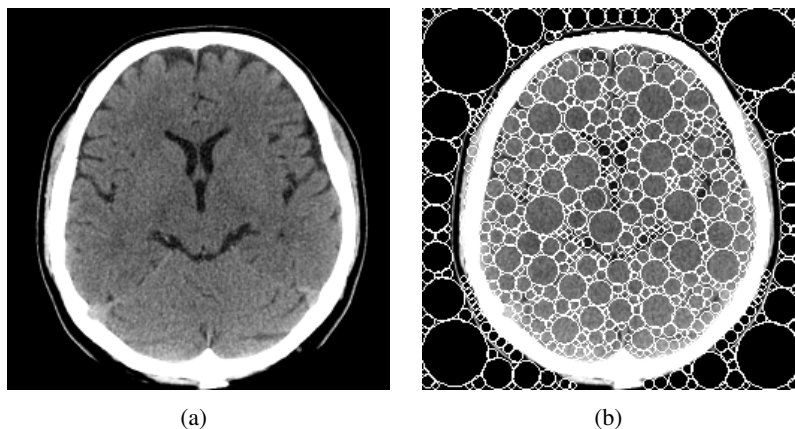


Figure 6. Description of example image: (a) - original image, (b) - circles

with centre $O(v)$ and radius $r(v)$. Two vertices $w, v \in V$ are adjacent in graph G if and only if the corresponding circles $K(w)$ and $K(v)$ are adjacent (within admissible error margin).

Semantic information A fundamental characteristic of a linguistic description is its information content. In the case of a circle graph, circles defined for each node, code information about subareas that are cohesive colour-wise. Edges carry information about the neighbourhood of such subareas and about the possibility of their composition into larger areas which may constitute potential candidates in the recognition process. Above all, however, the edges carry information about the shape of an area, supplementing surface information contained in a set of vertices and associated circles.

Task Following the denotations presented above, let G be a linguistic description of an image. Assuming that the description is a combination of descriptions of objects distinguishable in the image at a given granularity level, the task is to search for subgraph G_c , which would fulfil the criteria taken, in graph G . In the present paper, the criterion is described by a prototype graph G_p . The prototype graph carries information about an expected shape. Note that the circles are described in abstract coordinate space, which enables one to place the prototype in any point of the image in order to compare it with any candidate graph G_c .

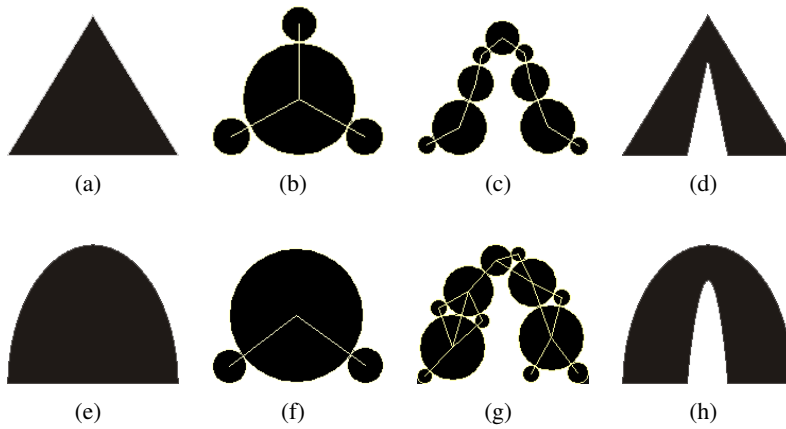


Figure 7. Vertices and edges carry spatial and shape information: (a), (d), (e), (h) - objects, (b), (c), (f), (g) - corresponding graphs

Knowledge It has to be emphasised that the graph of prototype's linguistic description carries only information about the shape of the object that is sought. Because the shapes (especially those not very complex) are not unique, this information is not sufficient for object's localisation. Thus, it is necessary to code additional knowledge in energy function E that is optimised. In the case of ventricular system image retrieval, we use knowledge of an expert, who indicated two measurable pieces of information that may characterise the lateral ventricle that we look for. The first one is blackness. It makes use of the fact that: *the object is filled with cerebrospinal fluid and in CT imaging, with appropriate data acquisition parameters, the regions filled with the fluid are dark*. The second one is centrality that: *results from supervised exposition during axial scanning*.

Energy The above characteristics have been coded by the following numerical features that evaluate the classifier k which currently defines a candidate graph G_c :

- E_b - blackness energy; its value is an average of the values of colours of pixels situated in the centres of the circles of G_c , divided by 255. Note that the image is an 8-bit greyscale image, where black corresponds to 0.
- E_c - centrality energy; its value is an average distance between centres of circles of G_c and the centre of the image divided by half the value of the image's diagonal.

Proper energy E function optimised during examination of subdescriptions is the sum of linguistic description's matching cost E_m and the values E_b and E_c . With a function constructed in this manner, one should remember that both blackness and centrality have supplementary meaning which promotes well-located solutions.

Prototype matching The key role in the search for solution is played by energy component that formulates the candidate-prototype similarity. What is important, the prototype may not exist in a literal sense (e.g. in the form of G_p). In this context, the term prototype should be applied to the knowledge of the system's designer.

Non-linguistic prototype matching There is no need to use the whole knowledge encoded in a linguistic description. Let us consider similarity function that measures the candidate-prototype surface matching. Let P_c be a set of image points that belong to the sum of all circles of a candidate's description, while P_p – a set of points of a prototype after it has been placed on a candidate. Minimal rectangles describing both sets are found: R_c i R_p , respectively. Next, the prototype is placed on the image in such a way that the upper left vertices of the rectangles overlap. The other option is to scale the prototype in such a way that the rectangles' diagonals are equal. However, scaling eliminates from the prototype the information about the size of the object and such knowledge does not find reflection in the resultant energy, which usually makes it difficult to find the right set of circles. Nevertheless, the size parameter can be included in function E , which allows one to avoid the necessity to create a large number of prototypes. The last element needed is the measure of surface similarity between sets P_c and P_p . If we regard both of them as the sets of sets of points belonging to the circles that constitute those sets, it is possible to apply the commonly used measures for supervised and unsupervised classification, in particular those presented in [37]: purity, entropy, precision, recall and their combinations (e.g. F1 measure), mutual information, etc.

Linguistic prototype matching In the previous subchapter the information carried by sets of graph edges G_p and G_c was ignored. However, Fig. 7 highlights important description content that they carry. In particular, image imperfections can lead to changes in number and size of circles and, as a result, surface changes. However, similar shapes are likely to have similar circle structures and spatial relations between circles and their groups. In the present work, knowledge is extracted

and utilised by calculating similarity between two graph linguistic descriptions: G_c and G_p . Because graph descriptions are usually non-identical, the evaluation algorithm is based on the search for a homeomorphism between the graphs described in [38, 39]. The algorithm employs the notion of matching operations that aim at possibly most exact reflection of G_p by the components of G_c . Every matching operation is assigned an edit cost, which helps evaluate the similarity of those shapes which are not homeomorphic in the definition sense. In search for a homeomorphism, two basic operations are performed, namely path substitution and path deletion. These are multi-stage operations which use path approximation by segments, node substitution, node deletion and segment deletion in a path. Because every such operation results in graph's modification, it is assigned a cost estimated in the context of graph descriptions. Note that there are many graph matching possibilities that differ in cost. Only minimum matching can be regarded as a measure.

It is impossible to search through the whole space of possible matches of candidate and prototype. The search for optimal matching is performed by the algorithm presented below. It is based on heuristic iterative expansion strategy for best possible graph matching.

Evolution This approach is also based on the algorithm of simulated annealing. Both observations of the algorithm and theoretical considerations prove the role of movement generator in the process of energy function optimisation.

In the case of candidate graph evolution, two strategies are applied. While using non-linguistic energy, which does not impose any additional limitations on a candidate, generation of a neighbouring solution is parameterised by certain parameter $\mu \in [0, 1]$. Let G_c be evolving candidate, the resultant solution is obtained by searching all circles $K(v) \in \mathcal{O}_2$. Each $K(v)$ has a chance to be included into generated solution with a probability π being decreasing function of distance ρ between $K(v)$ and G_c :

$$\pi = \mu e^{-\rho} \tag{6}$$

This distance can be measured by the length of the route connecting the circle's centre and the closest point being the sum of circles corresponding to vertices from G_c . However, more successful results have been obtained by measuring the distance between a closest pair of points, where one of them belongs to $K(v)$, and the other one to the sum of all candidate's circles. In this way spatial cohesiveness of a candidate graph is preserved.

Procedure GraphSimilaritySearch(G_p, G_c) $\in \mathbb{R}$

```

/* Initial match is 1-1 matching between one node from  $G_p$ 
and one from  $G_c$ ; while computing costs prototype is
considered to be placed on the candidate so that
centers of initially matched nodes overlap.          */
M - set of all initial 1-1 matches
while true do
    /* remove worst partial matches; heuristically reduces
memory consumption ant time complexity          */
    trim( $M$ )
     $best \leftarrow$  match from  $M$  having smalles cost
    remove(  $best$  from  $M$ )
    /* Extending paths has been described in [38]. In
short, searching homeomorphism between  $G_p$  and  $G_c$ 
requires searching for mappings between features of
both graphs (here edges).                          */
    if  $best$  has extending paths then
        foreach not previously matched  $path_p$  extending  $best$  do
            /* it is possible to leave  $path_p$  unmatched, or to
check all possibly matching paths in  $G_c$           */
             $best_{np} \leftarrow$   $best$  with  $path_p$  matched to nothing
             $cost(best_{np}) \leftarrow$   $cost(best) +$  penalty for unmatched  $path_p$ 
            add( $best_{np}$  to  $M$ )
            foreach  $path_c$  in  $G_c$  do
                 $best_+ \leftarrow$   $best$  extended with  $path_p \rightleftharpoons path_c$  match
                 $cost(best_+) \leftarrow$   $cost(best) + (path_p, path_c)$  match cost
                add( $best_+$  to  $M$ )
        else
             $cost_{old} \leftarrow$   $cost(best)$ 
             $cost(best) +=$  unmatched prototype edges penalty
             $cost(best) +=$  unmatched candidate edges penalty
            if  $cost_{old} = cost(best)$  then
                /*  $cost(best)$  didn't change, so it smaller than
any other processed matching, finishing          */
                return  $cost(best)$ 
            else
                add( $best$  to  $M$ )

```

Linguistic energy requires a candidate to be a connected graph. Therefore, the process of new solution generation is divided into two stages. In the first phase, the entire set of current candidate's vertices is copied to the new, generated solution candidate. Then, with a certain probability, vertices are deleted from this set provided such a removal will not lead to disconnectivity of new candidate. In the second phase, each circle adjacent to at least one corresponding to any vertex that remained in the generated candidate is subjected to addition with certain probability. Due to computational complexity of the matching algorithm, the size of candidates vertices set has been limited to a few nodes.

5.3. Sample results

Sample results of algorithm's performance for both non-linguistic and linguistic energy function are depicted in Fig. 8, Fig. 9 and Fig. 10. The solution found depends on the activation of particular components of energy function. Activation of prototype scaling allows a dynamic matching of the prototype's size with the size of a candidate set of circles. However, this is not the best matching. The same applies to the very mechanism of prototype shifting. Errors occurring at region's edges may shift the prototype and thus reduce the degree of matching, while any arbitrarily selected reference point (including centroid) will have its pessimistic examples. Although, from the theoretical point of view, the measure is well defined, there are some difficulties caused by the application of simulated annealing algorithm.

These disadvantages do not apply to the linguistic method that, while searching for a homeomorphism, examines all possible shifts. Scale factor cannot be selected automatically, because it is hard to predict how scaling will influence graph matching cost. On the other hand, searching through the set of scalings would result in substantial increasing of computing cost. Besides, scaling reduces information content of a candidate's description, which is undesirable.

The results produced by an algorithm that uses linguistic matching demonstrate much greater resemblance of shape between the ventricular system and the prototype. Again, emphasis is put on the shape localisation problem, which is solved by adding to energy function two components: blackness and centrality.

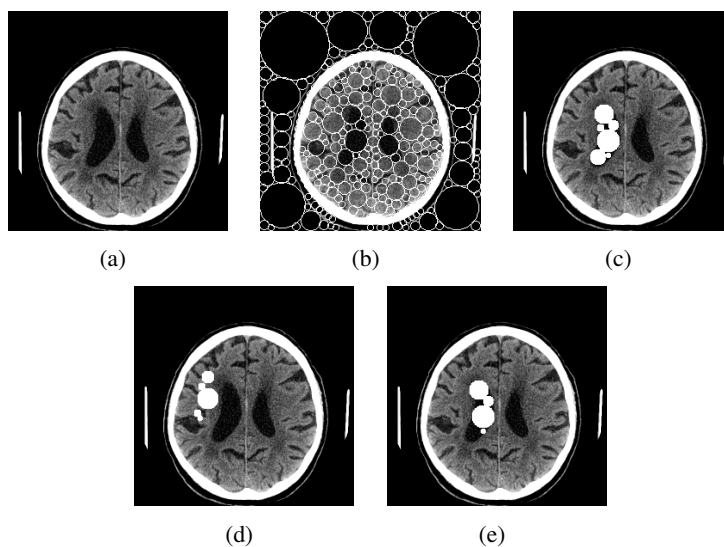


Figure 8. Non-linguistic result: (a) - original image, (b) - image description, (c) - gold standard, (d) - recognition with E_c , (e) - recognition with E_c and E_b .

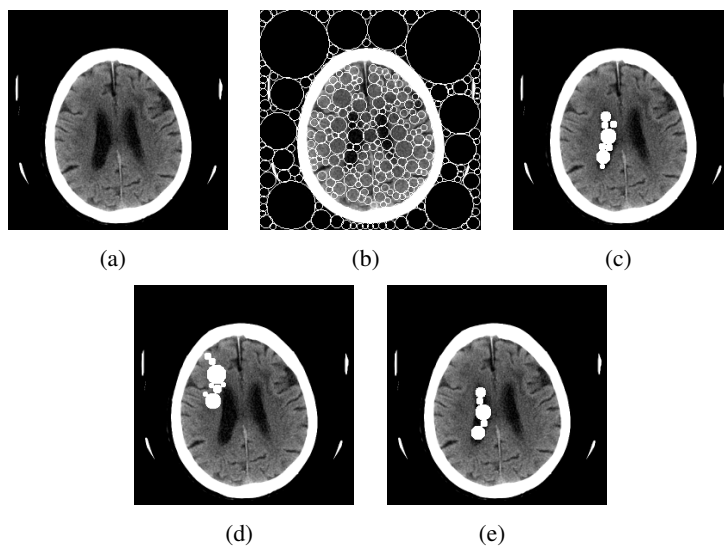


Figure 9. Non-linguistic result : (a) - original image, (b) - image description, (c) - gold standard, (d) - recognition with E_c , (e) - recognition with E_c and E_b .

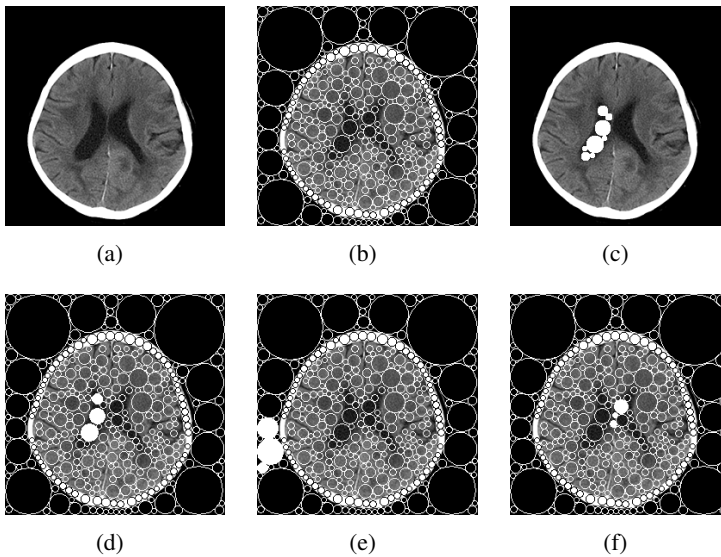


Figure 10. Sample results for graph homeomorphism approach. (a) - image, (b) - description, (c) - gold standard, (d) - $E_m + E_c + E_b$, (e) - no E_c , (f) - no E_b

6. Conclusions

The present paper has presented the CHAP method applied to automatic image understanding system development. The approach draws upon the active contours method, which employs not only the knowledge contained in the image but also any other knowledge that might be useful or sometimes indispensable for proper localisation of objects in an image. Since full utilisation of this knowledge can be complicated or even impossible, the task of image understanding has to be divided into steps, with each step contributing new semantic knowledge. This helps to avoid, if such a need should arise, low-level pixel analysis and focus on high-level spatch set analysis, which in many cases makes it easier to utilise expert knowledge. Expert knowledge utilisation, which is a key element of the approach presented, should be performed in such a way as to let an expert demonstrate their knowledge in most natural way. Thus, the present work has presented three approaches based on linguistic description: the one that requires knowledge to be rendered into the language of mathematical formulae with formal description of an image itself, the one that directly applies linguistic descriptions in the form

of rules, also with formal image description, and the one that defines linguistically both expert knowledge and image description. Sample examples prove the approach is very promising and can be applied in many domains at different analysis levels.

Acknowledgements

This work has been supported by the Ministry of Science and Higher Education, Republic of Poland, under the project number N 519 007 32/0978; decision no. 0978/T02/2007/32.

Authors would like to express their gratitude to Mr Piotr Grzelak, MD and Mr Cyprian Wolski, MD from the Department of Radiology and Diagnostic Imaging of Barlicki University Hospital in Lodz for making brain and heart images available and sharing their medical knowledge.

Authors would also like to express their gratitude to Mr Paweł Karpiński, for inspiring concepts regarding linguistic description generation.

References

- [1] Gonzalez, R. and Woods, R., *Digital Image Processing*, Prentice-Hall Inc., New Jersey, 2002.
- [2] Sonka, M., Hlavac, V., and Boyle, R., *Image Processing, Analysis and Machine Vision*, Chapman and Hall, Cambridge, 1994.
- [3] Davies, E. R., *Machine Vision, Theory, Algorithms, Practicalities*, Elsevier, Morgan Kaufmann, San Francisco, 2005.
- [4] Tadeusiewicz, R. and Ogiela, M., *Medical Image Understanding Technology*, Vol. 156 of *Studies in Fuzziness and Soft Computing*, Springer-Verlag, Berlin, Heidelberg, New York, 2004.
- [5] Kass, M., Witkin, A., and Terzopoulos, D., *Snakes: Active Contour Models*, International Journal of Computer Vision, 1988, pp. 321–331.
- [6] Cohen, L. D., *On active contour models and balloons*, Computer Vision, Graphics, and Image Processing. Image Understanding, Vol. 53, No. 2, 1991, pp. 211–218.

- [7] Ivins, J. and Porrill, J., *Active Region Models for Segmenting Medical Images*, In: IEEE International Conference ICIP, 1994, pp. 227–231.
- [8] Xu, C. and Prince, J., *Snakes, Shapes, and Gradient Vector Flow*, 1998.
- [9] Cohen, L. D. and Cohen, I., *Finite Element Methods for Active Contour Models and Balloons for 2D and 3D Images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No. 11, November 1991, pp. 1131–1147.
- [10] Amini, A. A., Weymouth, T. E., and Jain, R. C., *Using Dynamic Programming for Solving Variational Problems in Vision*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 9, 1990, pp. 855–867.
- [11] McInerney, T. and Terzopoulos, D., *Topologically Adaptable Snakes*, In: ICCV, 1995, pp. 840–845.
- [12] Delingette, H. and Montagnat, J., *New Algorithms for Controlling Active Contours Shape and Topology*, In: ECCV (2), 2000, pp. 381–395.
- [13] Casseles, V., Catta, F., Coll, T., and Dibos, F., *A geometric model for active contours in image processing*, Numerische Mathematic, Vol. 66, 1993, pp. 1–31.
- [14] Malladi, R., Sethian, J. A., and Vemuri, B. C., *Shape Modeling with Front Propagation: A Level Set Approach*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 2, 1995, pp. 158–175.
- [15] Osher, S. and Sethian, J. A., *Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations*, Journal of Computational Physics, Vol. 79, 1988, pp. 12–49.
- [16] Caselles, V., Kimmel, R., and Sapiro, G., *Geodesic Active Contours*, International Journal of Computer Vision, Vol. 22, No. 1, 1997, pp. 61–79.
- [17] Yezzi, A., Kichenassamy, S., Kumar, A., Olver, P., and Tannenbaum, A., *A Geometric Snake Model for Segmentation of Medical Imagery*, IEEE Transactions on Medical Imaging, Vol. 16, No. 2, 1997.
- [18] Kichenassamy, S., Kumar, A., Olver, P. J., Tannenbaum, A., and Yezzi, A. J., *Gradient Flows and Geometric Active Contour Models*, In: ICCV, 1995, pp. 810–815.

- [19] Xu, C., Yezzi, A., and Prince, J., *On the Relationship between Parametric and Geometric Active Contours*, In: 34th Asilomar Conference on Signals, Systems and Computers, 2000, pp. 483–489.
- [20] Xu, C., Hopkins, J., Yezzi, A., and Prince, J. L., *A Summary of Geometric Level-Set Analogues for a General Class of Parametric Active Contour and Surface Models*, In: Proc. of 1st IEEE Workshop on Variational and Level Set Methods in Computer Vision, July 2001, pp. 104–111.
- [21] Cootes, T. and Taylor, C. J., *Active shape models - 'smart snakes'*, In: Proceedings of 3rd British Machine Vision Conference, Springer-Verlag, 1992, pp. 266–275.
- [22] Cootes, T., Taylor, C., Cooper, D., and Graham, J., *Active Shape Model - Their Training and Application*, CVGIP Image Understanding, Vol. 61, No. 1, 1994, pp. 38–59.
- [23] Grzeszczuk, R. and Levin, D., *Brownian Strings: Segmenting Images with Stochastically Deformable Models*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 10, 1997, pp. 1100–1113.
- [24] Jacob, M., Blu, T., and Unser, M., *A unifying approach and interface for spline-based snakes*, In: in Proc. SPIE Med. Imaging, I. 4322, 2001, pp. 340–347.
- [25] Schnabel, J. and Arridge, S., *Active contour models for shape description using multiscale differential invariants*, In: In Proceedings of British Machine Vision Conference, edited by D. Pycock, 1995, pp. 197–206.
- [26] Staib, L. H. and Duncan, J. S., *Parametrically Deformable Contour Models*, In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1989, pp. 98–103.
- [27] Denzler, J. and Niemann, H., *Active rays: A new approach to contour tracking*, International Journal of Computing and Information Technology, Vol. 4, 1996, pp. 9–16.
- [28] Tomczyk, A. and Szczepaniak, P. S., *On the Relationship Between Active Contours and Contextual Classification*, In: CORES, edited by M. Kurzyński, E. Puchała, M. Woźniak, and A. Żołnierek, Vol. 30 of *Advances in Soft Computing*, Springer, 2005, pp. 303–310.

- [29] Tomczyk, A., Szczepaniak, P. S., and Pryczek, M., *Active Contours as Knowledge Discovery Methods*, In: Discovery Science, edited by V. Coruble, M. Takeda, and E. Suzuki, Vol. 4755 of *Lecture Notes in Computer Science*, Springer, 2007, pp. 209–218.
- [30] Tomczyk, A. and Szczepaniak, P. S., *Contribution of Active Contour Approach to Image Understanding*, In: IEEE International Workshop on Imaging Systems and Techniques, 2007.
- [31] Tadeusiewicz, R., Ogiela, M. R., and Szczepaniak, P. S., *Notes on a Linguistic Description as the Basis for Automatic Image Understanding*, Applied Mathematics and Computer Science, Vol. 19, No. 1, 2009, pp. 143–150.
- [32] Tomczyk, A., *Image Segmentation Using Adaptive Potential Active Contours*, In: Computer Recognition Systems 2, edited by M. Kurzyński, E. Puchała, M. Woźniak, and A. Żołnierek, Vol. 45 of *Advances in Soft Computing*, Springer, 2007, pp. 148–155.
- [33] Tomczyk, A., Wolski, C., Szczepaniak, P. S., and Rotkiewicz, A., *Analysis of Changes in Heart Ventricle Shape Using Contextual Potential Active Contours*, In: Computer Recognition Systems 3, edited by M. Kurzyński and M. Woźniak, Vol. 57 of *Advances in Soft Computing*, Springer, 2009, pp. 397–405.
- [34] Tomczyk, A. and Szczepaniak, P. S., *Segmentation of Heart Image Sequences Based on Human Way of Recognition*, In: Brain Informatics, edited by N. Zhong, K. Li, S. Lu, and L. Chen, Vol. 5819 of *Lecture Notes in Computer Science*, Springer, 2009, pp. 225–235.
- [35] Rutkowska, D., Piliński, M., and Rutkowski, L., *Sieci neuronowe, algorytmy genetyczne i systemy rozmyte*, Wydawnictwo Naukowe PWN, Warszawa, 1997.
- [36] Rutkowska, D., *Inteligentne systemy obliczeniowe, algorytmy genetyczne i sieci neuronowe w systemach rozmytych*, Akademicka Oficyna Wydawnicza PLJ, Warszawa, 1997.
- [37] Strehl, A., *Relationship-based Clustering and Cluster Ensembles for High-dimensional Data Mining*, Ph.D. thesis, The University of Texas at Austin, May 2002.

- [38] Rocha, J. and Pavlidis, T., *A Shape Analysis Model with Applications to a Character Recognition System*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 16, No. 4, 1994, pp. 393–404.
- [39] Rocha, J. and Pavlidis, T., *Character Recognition Without Segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, 1995, pp. 903–909.