



Representation and fusion of heterogeneous fuzzy information in the 3D space for model-based structural recognition—Application to 3D brain imaging

Isabelle Bloch^{*}, Thierry Géraud¹, Henri Maître

*Ecole Nationale Supérieure des Télécommunications, Département TSI - CNRS URA 820, 46 rue Barrault,
75013 Paris, France*

Received 2 July 2002

Abstract

We present a novel approach to model-based pattern recognition where structural information and spatial relationships have a most important role. It is illustrated in the domain of 3D brain structure recognition using an anatomical atlas. Our approach performs segmentation and recognition of the scene simultaneously. The solution of the recognition task is progressive, processing successively different objects, and using different pieces of knowledge about the object and about relationships between objects. Therefore, the core of the approach is the knowledge representation part, and constitutes the main contribution of this paper. We make use of a spatial representation of each piece of information, as a spatial fuzzy set representing a constraint to be satisfied by the searched object, thanks in particular to fuzzy mathematical morphology operations. Fusion of these constraints allows us to select, segment and recognize the desired object.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Heterogeneous knowledge representation; Fuzzy mathematical morphology; Fuzzy classification; Fuzzy spatial relationships; Fuzzy fusion; Fuzzy pattern recognition; Model-based structural recognition; Brain imaging

^{*} Corresponding author.

E-mail addresses: isabelle.bloch@enst.fr (I. Bloch), thierry.geraud@lrde.epita.fr (T. Géraud).

URL: <http://www.lrde.epita.fr> (T. Géraud).

¹ Current address of Thierry Géraud: EPITA Research and Development Laboratory, 14-16 rue Voltaire, F-94276 Le Kremlin-Bicêtre cedex, France.

1. Introduction

Structural recognition is a useful step in automated image description and interpretation. Structural recognition makes use of spatial relationships between the several components of an image to improve the recognition of every individual component and to provide a reliable global understanding of the image. We are interested in model-based structural understanding. Such systems may be of great interest in many domains, for instance in aerial imaging, computer vision or medical imaging, for all problems for which some prior knowledge about the observed scene is available. This will be illustrated here in the domain of 3D brain imagery.

Most image understanding methods are separated into two stages: (i) detection of candidate objects and (ii) recognition of these objects. The former is considered a low level process whereas the latter is performed at a high level, and as such, they are often dealt with by different layers of software, for instance the former is performed in a preprocessing algorithmic level whereas the latter is done by a rule-based system. To compensate for this stratification, in some cases, a first interpretation allows a return to the low level process to improve detection and a moderate number of iterations is performed.

Our approach is completely different. It may be seen as a simultaneous segmentation and recognition process of the scene, and the solution of the recognition task is progressive, processing successively different objects. The method starts with one object, expected to be rather easy to detect. This object is detected and recognized by gathering and comparing on the one hand information extracted from the image and, on the other hand, knowledge derived from the model or from the domain. Then the method addresses the recognition of another object using its own properties (radiometry, morphology) and also some relations (connectivity, relative position) with the previously recognized object. The process is then iterated to extract objects that are more difficult to detect, but using more structural information since the context is better known. If variability is expected between model and scene, registration can be initialized and refined at each recognition step using the newly obtained correspondence between a model object and a scene object. We suggested this approach in [1]. Here it will be more elaborated and detailed.

A second important characteristic of image understanding methods is the way spatial information is handled. Many structural recognition methods are based on graph representations. Relaxation and optimization techniques (i.e., modifying iteratively the recognition function) are often used to satisfy structural constraints. Others use constraint networks or Delaunay triangulation to capture and manipulate the spatial context. In this paper, we directly make use of the image array to gather and combine the pieces of information including structural ones which are therefore all converted into a spatially encoded structure.

The last important characteristic of image understanding methods that will be discussed here is the chosen framework for knowledge representation and management, including imprecision and uncertainty. Rule-based systems often rely on propositional logics. Probabilistic reasoning, and more specifically Bayesian reasoning, benefit from a large body of theoretical and experimental results. Bayesian networks for instance were found to be well adapted for reasoning in complex situations where many objects are present [2]. The Dempster–Shafer evidence theory [3,4] is able to represent not only uncertainty,

but also ignorance and thus better models some human deductions. In this paper we have chosen to rely on the fuzzy set framework [5,6] for two main reasons:

- (1) fuzzy sets are well adapted to provide a framework for the representation under a common language of heterogeneous pieces of knowledge about radiometry, space, morphology or structure, and for the combination of these;
- (2) fuzzy sets appropriately model information imprecision which results from image noise, unknown radiometric characteristics due to the image acquisition process, variability between model and scene or between different instantiations of the model, and intrinsically vague knowledge.

We make use here mainly of a spatial representation of fuzzy sets, which already has proved to be of use in image processing [7], and answers to both spatial information handling and imprecise knowledge representation problems.

Brain imaging has been chosen to illustrate the methodology presented hereafter because of the importance of structural information in the detection and recognition of the different components of the brain.² Segmentation of brain structures is of prime importance for many different applications: morphometry, pathology detection and measurement, diagnosis, surgery and radio-therapy planning, functional imaging, neuro-sciences and so forth. A large body of literature has been devoted to brain image segmentation (see, e.g., the syntheses in [8,9]). We will deal with magnetic resonance images (MRI). For these images, the classes that can be observed are, for the outer part of the brain: air, skin, muscle, fat and the skull. As for the brain itself, white matter, grey matter and cerebro-spinal fluid can be observed. Although the radiometry of these classes can be described by statistical distributions that significantly overlap, classifiers can separate the three main brain tissues. In the fuzzy set framework, fuzzy clustering, e.g., [10,11] has been widely used for this purpose. Unfortunately, recognition of internal structures remains difficult. For instance, the different grey nuclei which are constituted of grey matter cannot be distinguished using only radiometric information. Therefore the use of models is almost always necessary. Models used in the literature are implicit, like physics-based deformable models (e.g., [12, 13]), or explicit in atlas deformation techniques.

Atlas-based methods are generally divided into two steps. The first one consists in aligning the atlas and the 3D image using a rigid or affine matching. The second one consists of an elastic matching to achieve a better correspondence between objects. The underlying assumption is that the topological structure is the same in both volumes, and that variability is limited. Several methods have been developed, that can be classified according to four main aspects: the physical model used to model the deformations, the similarity criterion to be maximized, the anatomical structures used for this optimization, and the possible use of a multi-resolution approach to reduce the computation cost. Among the first works in this domain, it is worth mentioning the approach described in [14], which is based on local elastic deformations which are computed based on brain and ventricle

² It should be considered only as an illustration of the proposed approach, the focus being the methodological aspects, not the clinical ones.

surfaces. Extensions of this approach have been proposed, e.g., in [15]. A probabilistic model to estimate the deformation has been proposed in [16], which allows the introduction of prior information. Deformations based on viscous fluid equations have been proposed in [17] in order to better preserve the topology of structures. Most methods rely on the matching of homologous points [18], surfaces [19], or the whole volume [16,17]. Atlas-based approaches can segment all structures in a global way but have to deal with the difficult problem of anatomical variability. Contrary to the existing atlas-based methods which try to find a global deformation between the atlas and the image, our method is sequential: one step aims at recognizing one single anatomical object and then refines the correspondence between the image and the atlas. Therefore it strongly relies on the topological arrangement between structures as given by the atlas, but allows for local variations due to diversity of structures, and takes into account specificities of each of them, including their variability. It relies on both surfaces and volumes, and takes into account spatial relationships between structures, which are not explicitly included in other atlas-based approaches.

In Section 2 we discuss the types of information and knowledge that are used in model-based structural recognition, and propose an original representation as fuzzy volumes of interest in the image space. A unified approach is proposed to the representation of many spatial concepts via mathematical morphology and its fuzzy extension. Mathematical morphology is well adapted to deal with shapes, spatial representations, and spatial relations. Moreover its use guarantees good algebraic properties, and benefits from algorithmical development to compute morphological operations in an efficient way. In Section 3, we describe how this information is combined and used to drive the recognition process. In Section 4, we illustrate the proposed approach with the example of recognizing some brain structures in 3D T1-weighted MRI using an anatomical atlas.

2. Knowledge representation using spatial fuzzy sets and fuzzy mathematical morphology

This section aims at describing the type of information and knowledge used in scene recognition based on a model. We first give some general characteristics of the information. Then, because of the heterogeneity of the knowledge used for recognizing an object, we propose a common framework based on fuzzy set theory for its representation. We then detail the proposed representation for different types of knowledge and information.

2.1. General characteristics of information

There are many different properties which allow the classification of the information used in a pattern recognition task. In the case of model-based scene recognition, we consider three main families of properties, which help to answer the following three questions: What is the information about? Is it generic or factual? Under which form is it currently provided?

What is the information about? Two types of information are classically used in structural shape recognition. The first concerns information about the object to be recognized (its shape, topology, grey level, position), while the second concerns its relationships to other objects in the scene (distances, adjacency, relative directional position). Structural information is mostly encoded within this last category and in the sequel we will take care to preserve this information.

Section 2.3 describes knowledge about objects, while Section 2.4 describes knowledge about relationships between objects.

Is it generic or factual? The answer to this question mainly depends on the information source. Information extracted from the image is factual since it concerns the particular scene to be interpreted. As an example take the grey level of a region in the image: it certainly pertains to the data. However, information about the image itself (for instance about its acquisition) is more generic. It pertains to domain or contextual information. A typical example is the prior knowledge about expected grey level of a given structure given the type of acquisition. It is generally less specific than factual information and may be revised in the light of image information. Information contained in or derived from the model is generic, since it should apply, within some limits, to any image. For instance, the proposition *object A is to the right of object B in the model* is generic, and we expect the objects in the scene corresponding to *A* and *B* satisfy a similar relation.

Under which form is it currently provided? We have a great variety of answers to this question which makes the representation and combination of information difficult. Classically it can be a number (as the mean grey level of an image region, or a distance between two objects), a distribution (for instance to represent the grey levels of the pixels in a region) or a binary value (as for the inclusion relationship of an object in another one). But we will also be concerned with imprecise values and with propositional formulae which are often used by experts within a given application. Imprecise values are expressed sometimes in linguistic terms: for instance the expected grey level of a structure, which is either absolute (*bright, dark, medium, ...*), or relative (*darker than, ...*), or the expected distance between two objects (*close, far, ...*). They can also be expressed as an interval (an object thickness is between 3 and 5 mm). Propositional formulae (*object A is to the right of object B*) usually express rather complex ideas which need much prior knowledge to be correctly interpreted.

The pertinent translation of these heterogeneous pieces of information and their easy combination will dictate the choice of an adequate representation framework.

2.2. Fuzzy sets as a representation framework

In the conventional approach to pattern recognition where the two stages of detection and recognition are separated (as for instance in graph based recognition methods), what is needed is a way to assess the similarity between degrees of satisfaction of relationships between the model and image objects. In the approach proposed in this paper, the driving idea is different. The different types of knowledge have to serve as a guide for (i) exploring the image space and (ii) segmenting and recognizing a specific object. Therefore, the

assessment of similarity is far from being enough. We have to detect regions of interest, to select possible candidates and to measure their match to the model.

To detect regions of interest and to select possible candidates, we propose to translate all available knowledge into a spatial representation. Then, fusion combines all these regions of interest in order to focus attention by reducing the search space and restricting it to the area that satisfies most relationships. Since many pieces of information are delivered in an imprecise way, we make use of the framework of fuzzy sets.

This modeling is well adapted to information derived from the model. Models exhibit two different kinds of imprecision. They are generic, i.e., they do not represent every sample of the family but an “average” object which probably does not even exist with exactly the same shape and properties. They are selectively simplified and schematized to bring the essential information to the fore.

The image also suffers from imprecision for several reasons, some related to the observed phenomenon, others to processing artifacts. For instance, a soft transition between tissues (e.g., healthy and pathological tissues) is surely a cause of classification imprecision inherent to the nature of the observed objects. It may also happen that for some modalities, tissues have similar characteristics. Thus the images obtained with this modality will poorly discriminate between the tissues, resulting in uncertainty on the belonging of a pixel to one or the other class. Another cause of imprecision comes from the discrete nature of digital images, resulting in a delocalization of information contained in a small volume at only one point. The partial volume effect (presence of several tissues in one pixel or voxel) also participates in this type of spatial imprecision. Other image imperfections can be caused by numerical reconstruction algorithms in computed imaging (for instance the Gibbs effect that may appear in MRI around sharp transitions), or by processing algorithms (e.g., filtering, contour detection, registration between images, etc.) which all suffer from false alarms and delocalization.

In this context, the theory of fuzzy sets appears to be well suited. Indeed, it provides a good theoretical basis to model the imprecision of the information at different levels of representation. It constitutes a unified framework for representing and processing both numerical and symbolic information. Structural information (constituted mainly by spatial relationships in image processing) is well represented by it. Moreover, fuzzy set theory has benefited from the many recent developments in information fusion, in the definitions of combination operators, of similarity measures, and in decision tools [20]. This will be used in Section 3.

The numerical representation of membership values assumes that we can assign numbers that represent degrees of satisfaction of a relationship for instance. These numbers can be derived from prior knowledge or learned from examples, but usually there remain some quite arbitrary choices. This might appear as a drawback in comparison to propositional representations. However, it is not necessary to have precise estimations of these values, and experimentally we observed a good robustness with respect to these estimations, in various problems like information fusion, object recognition and scene interpretation. This can be explained by two reasons: first, the fuzzy representations are used for rough information and therefore do not have to be precise themselves, and second several pieces of information are usually combined in a whole reasoning process, which decreases the influence of each particular value (of individual information). Therefore the

chosen numbers are not crucial. What is important is that ranking is preserved. For instance if a region of the space satisfies a relationship to some objects to a higher degree than another region, then this ranking is preserved in the representation, for all relationships described in the following sections. Assuming the existence of ranking is reasonable for the type of relations we consider.

In the rest of this paper, the image to be processed will be denoted by I (here we consider the case of 3D images, the most general case for medical imaging), and a point (volume element or voxel) in this image, by v . For each piece of knowledge, we consider its “natural expression”, i.e., the usual form in which it is given or available, and translate it into a spatial fuzzy set in the image space, the membership of which is denoted by $\mu_{\text{knowledge}}$ [21]. This membership function assigns to each voxel of I a degree in $[0, 1]$:

$$\mu_{\text{knowledge}} : \begin{cases} I \rightarrow [0, 1], \\ v \mapsto \mu_{\text{knowledge}}(v). \end{cases} \quad (1)$$

For instance, if the knowledge expresses some constraint, $\mu_{\text{knowledge}}(v)$ is the degree to which this constraint is satisfied at point v . In this representation, each piece of knowledge becomes a fuzzy region of the image space, which bridges the gap between linguistic expressions and numerical representations. If the knowledge is considered as a constraint to be satisfied by the object to be recognized, this fuzzy region represents a search area or a fuzzy volume of interest for this object, where this constraint is satisfied (to some degree). Several such regions, representing different available pieces of knowledge, have then to be combined in order to restrict this search area (see Section 3).

Although several works in image processing, robotics, etc. make use of spatial fuzzy sets to represent objects, to our knowledge, very few such representations have been proposed previously for relationships. In [22], fuzzy areas are defined for representing directional relative position, but only on one axis, on which projections of objects are considered. In [23,24], for applications in robotics, lines are represented as spatial fuzzy sets to account for uncertainty, and distances between objects expressed as linguistic variables are represented as fuzzy sets on each axis. Here we propose spatial representations in the same space as the objects themselves. Similar representations are used for instance in [25], based on simplified representations of the objects. The fuzzy spatial fuzzy sets we propose are also close to the notion of potential used in [26] for sizeless objects in a two dimensional space. Here the objects can have any dimension and any shape, even complex ones, and are processed without simplification.

2.3. Information on the object itself

In this Section we describe in detail the representation of knowledge about the object itself. One part concerns the geometry of the object, the other its radiometry.

2.3.1. Shape and localization

The model used may be heterogeneous. Typically for applications in image processing in various areas, it has two distinct parts:

- one part is iconic, and may be represented as a labeled image, where each region having a unique label is an object or a structure (this is typically the case for digital maps as used in aerial and satellite imaging, or 3D anatomical atlas as used in medical imaging, or environment maps as used in robotics, or some views of the scene in computer vision); regions can be crisp or fuzzy, and the rest of this paper applies in both cases;
- a propositional part, expressing expert knowledge, as linguistic terms, logical propositions, possibly including numerical, qualitative or imprecise values.

The iconic part provides a geometrical description of the objects in terms of shape and localization.

A first step is to perform a registration between the iconic model and the image using any available information. When no object has yet been detected, this may only be done with some prior knowledge we have on the positioning of both information sources. But, as soon as some objects have been detected, a classical way is to minimize the distance between surfaces of the detected objects and the corresponding models.

The second step is to transfer from model to image the shape and positioning information it contains. The model provides these basic pieces of information, but with imprecision due to variability and to the imperfect correspondence between instance and model.

The way to do this is to extend the region given by the model in a fuzzy manner, in order to take into account these imprecisions. An appropriate tool for this is morphological dilation. We use a fuzzy morphological dilation [27], defined as:

$$\forall v \in I, \quad D_v(\mu)(v) = \sup_{v' \in I} t[\mu(v'), v(v' - v)], \quad (2)$$

where μ denotes the object to be dilated (here a model object), v denotes a fuzzy set (also defined in the image space) called a structuring element, $D_v(\mu)$ denotes the dilation of μ by v , and t is a t-norm. Other forms of dilation are possible. Eq. (2) applies in both crisp and fuzzy cases (for objects and structuring elements as well).

Fuzzy dilation satisfies a set of properties, some of which are important (and even mandatory) for its use here, i.e.:

- it is extensive if $v(O) = 1$, where O denotes the center of the structuring element:

$$\forall v \in I, \quad D_v(\mu)(v) \geq \mu(v), \quad (3)$$

which guarantees that the shape provided by the model is actually extended in order to account for imprecision and variability (fuzzy set inclusion is defined in a classical way using \leq on membership functions);

- it is increasing with respect to both the structuring element and the set to be dilated:

$$v \subset v' \Rightarrow D_v(\mu) \subset D_{v'}(\mu), \quad (4)$$

$$\mu \subset \mu' \Rightarrow D_v(\mu) \subset D_v(\mu'), \quad (5)$$

this property guarantees that the larger the structuring element, the more imprecision is introduced, and that the larger the object in the model, the larger the volume of interest

in the image (in all these equations, subethood of fuzzy sets is defined as above, using \leq);

- dilation commutes with union, which guarantees that objects can be processed indifferently globally or by parts.

Also the dilation roughly preserves the shape of the object. Therefore, the dilated object carries information mainly about localization, but also approximated information about the shape (see for instance Fig. 2, top left). It provides a focus of attention. In this fuzzy area, a segmentation will be performed, that provides the derived shape. A segmentation step is included in all approaches to this problem. Here it is constrained by the relationships.

The important choice to be made here concerns the structuring element which represents the spatial imprecision. When all the existing sources of imprecision are taken into account into the fuzzy volume of interest, this volume should contain the object we are looking at. The choice of the structuring element reflecting the possible imprecisions depends on the application at hand. Its extent can be defined from prior knowledge, or learned from a set of representative images.

As the obtained fuzzy volume represents prior information about both the morphology and the localization in I of the object to be recognized, without any reference to the actual presence of the object in the image, we denote this information by μ_{prior} :

$$\mu_{\text{prior}} = D_v(\mu), \quad (6)$$

where μ is a model object.

2.3.2. Radiometry

The second important type of information that has to be taken into account is the radiometry or grey level of the object. For the sake of simplicity, we assume here that we are looking at homogeneous objects without shading or texture as is the case in brain imaging. Information about radiometry can be divided into two classes, each having different origins:

- the first class is generic knowledge, attached to the domain and the context, and in particular to the type of acquisition (for instance, *internal nuclei in T1-weighted MR images have an intermediate grey level*);
- the second one is derived from the data and has to be found in the image to be processed.

In the first class, knowledge is always approximate, since it has to take into account at least the inter-individual variability and the sensor calibration. A small number of values of a linguistic variable often represents this information adequately, for instance the set $\{\text{dark, intermediate, light}\}$. The semantics is given by fuzzy sets having membership functions defined on the radiometry range. Typically this range can be $L = [0, 255]$. The translation as fuzzy sets in the image space is made by a simple mapping. Let $\mu_r^L(l)$ denote the membership of a grey level l to the fuzzy set “ r grey level”, where r is one of the possible

values of the linguistic variable. Let $\mu_r(v)$ be the membership degree of a voxel v to the region of “ r grey level”. This membership value is defined as:

$$\mu_r(v) = \mu_r^L(l(v)), \quad (7)$$

where $l(v)$ denotes the grey level of voxel v in the image. This definition may be directly extended to multi-spectral or multi-modal images (in such cases, l is a vector).

The second type of radiometric information is more specific and precise, since it defines the actual radiometry of a tissue or an object in a given image. It can only be obtained after some object of similar radiometric properties has already been recognized (i.e., an object constituted of the same matter). From this first object a grey level distribution can be estimated, which is expected to apply also to the object to be recognized. For instance, if we deduce the mean m and the variance σ of the matter, we may define a fuzzy region for this matter as (this is but one possible model):

$$\mu_{\text{matter}}(v) = e^{-(l(v)-m)^2/(2\sigma^2)}. \quad (8)$$

In comparison with localization and shape information, radiometric information is more spread over the image, since several objects can have a similar constitution and appear with similar grey levels in the image.

2.4. Relationships between objects

In this Section, we describe how the knowledge about the spatial relations between objects may be represented in order to be easily combined with the previous pieces of information. This knowledge concerns the position of the object to be recognized with respect to the previously recognized objects. It expresses the structural information.

2.4.1. Set relationships

Since the proposed approach is progressive and does not reconsider previously recognized objects, one important type of relationship is made up of set relationships, which specify if areas where other objects have been recognized are forbidden or mandatory. These set relationships are expressed as inclusion in objects or exclusion from objects. For instance, if we are looking for a component of an object already detected, then the search area is included in this object and limited to it. On the contrary, if the object to be recognized is not allowed to overlap with the previous object, then the corresponding area is forbidden. In this way, we define for each object to be recognized a partition of previously recognized objects in two classes: one in which the inclusion is obligatory (denoted by O^{in}), and the other where exclusion is obligatory (O^{out}). Since previously recognized objects are not reconsidered in a further recognition step, these constraints are expressed in a crisp way. The corresponding region of interest has the following membership function:

$$\mu_{\text{constraint}}(v) = \begin{cases} 1 & \text{if } v \in O^{\text{in}} \setminus O^{\text{out}}, \\ 0 & \text{elsewhere.} \end{cases} \quad (9)$$

This constraint, i.e., the definition of O^{in} and O^{out} , is defined according to the model. The assumption behind this is that the set of the objects to be recognized and the background form a partition of the image.

2.4.2. Distances

The distance between objects is important for the assessment of spatial arrangement between objects in a scene. Therefore it is widely used in structural pattern recognition. Distances between objects A and B can be expressed in different forms, as in the following examples:

- the distance between A and B is equal to n ,
- the distance between A and B is less (respectively greater) than n ,
- the distance between A and B is between n_1 and n_2 .

In the framework of our study, these expressions will be translated into spatial volumes of interest within the image, taking into account imprecision and uncertainty, since these statements are generally approximate.

Distances between sets (average, Hausdorff, minimum distances) are usually defined by analytical expressions. But they also have equivalents in set theoretical terms by means of mathematical morphology. This allows us to include imprecision, and to deal with distances between fuzzy sets and with fuzzy distances [27]. Moreover, this allows us to express knowledge about distance to an object as a spatial fuzzy set in a very simple way, while benefiting from fast algorithms developed for the computation of dilations.

Let us detail these equivalences. We first consider the crisp case, and the minimum distance in a bounded discrete space. Let $d(A, B)$ be the distance between two crisp sets A and B , and $D^n(A)$ the dilation of size n of A (i.e., the dilation with a ball of size n as the structuring element). The following equations hold:

$$d(A, B) = n \Leftrightarrow \begin{cases} \forall m < n, D^m(A) \cap B = D^m(B) \cap A = \emptyset \\ \text{and } D^n(A) \cap B \neq \emptyset, D^n(B) \cap A \neq \emptyset, \end{cases} \quad (10)$$

$$d(A, B) \leq n \Leftrightarrow D^n(A) \cap B \neq \emptyset, \quad D^n(B) \cap A \neq \emptyset, \quad (11)$$

$$d(A, B) \geq n \Leftrightarrow \forall m < n, \quad D^m(A) \cap B = D^m(B) \cap A = \emptyset, \quad (12)$$

$$n_1 \leq d(A, B) \leq n_2 \Leftrightarrow \begin{cases} \forall m < n_1, D^m(A) \cap B = D^m(B) \cap A = \emptyset \\ \text{and } D^{n_2}(A) \cap B \neq \emptyset, D^{n_2}(B) \cap A \neq \emptyset. \end{cases} \quad (13)$$

The proof of these equations involves extensivity of dilation (for such structuring elements), and increasingness with respect to the structuring element.

We assume that A is known as one already recognized object, and that we want to detect B , subject to satisfying some distance relationship with A , as given by the model. According to the previous equations, dilations of A are an adequate tool for this. Let us consider the following different cases:

- If the model requires that $d(A, B) = n$, then the region defined by $D^n(A) \setminus D^{n-1}(A)$ is made up of the points exactly at distance n from A . Thus the border of B should intersect this region, and B should be looked for in $D^{n-1}(A)^C$ (the complement of the dilation of size $n - 1$).
- If the model requires that $d(A, B) \leq n$, then B should be looked for in A^C , with the constraints that at least one point of B belongs to $D^n(A) \setminus A$. Conversely, if the model requires that $d(A, B) \geq n$, then B should be looked for in $D^{n-1}(A)^C$.

- If the model requires that $n_1 \leq d(A, B) \leq n_2$, then B should be searched in $D^{n_1-1}(A)^C$ with the constraint that at least one point of B belongs to $D^{n_2}(A) \setminus D^{n_1-1}(A)$.

The constraints on the border may not be easy to satisfy in the recognition process. However, they can be avoided by considering both minimum and maximum (Hausdorff) distances, expressing for instance that B should lay between a distance n_1 and a distance n_2 of A . Therefore, the minimum distance should be greater than n_1 and the maximum distance should be less than n_2 . In this case, the volume of interest for B is reduced to $D^{n_2}(A) \setminus D^{n_1-1}(A)$.

In cases where imprecision has to be taken into account, fuzzy dilations are used, with the corresponding equivalences with fuzzy distances [27]. The extension to approximate distances calls for fuzzy structuring elements. We define these structuring elements through their membership function v on I . Structuring elements with a spherical symmetry are used, where the membership degree only depends on the distance to the center of the structuring element. For instance, to express a dilation of size *about* n , we define the corresponding structuring element by:

$$\forall v \in I, \quad v(v) = \begin{cases} 1 & \text{if } d_E(v, O) \leq n_1, \\ f(d_E(v, O)) & \text{if } n_1 < d_E(v, O) < n_2, \\ 0 & \text{if } d_E(v, O) \geq n_2, \end{cases} \quad (14)$$

where n_1 and n_2 are two parameters controlling the imprecision on n , such that $n \in [n_1, n_2]$, f is a decreasing function such that $f(n_1) = 1$ and $f(n_2) = 0$, O denotes the center of the structuring element, and d_E is the Euclidean distance between points in I .

The increasingness of fuzzy dilation with respect to both the set to be dilated and the structuring element guarantees that these expressions do not lead to inconsistencies.

From an algorithmical point of view, fuzzy dilations may have a quite high computational cost if the structuring element has a large support. The complexity is in $O(n_I n_S)$ where $n_I = |I|$ (size of the image) and $n_S = |Supp(v)|$ (size of the support of the structuring element v). Note that this is still less than the complexity of an exhaustive computation of distances using analytical expressions. However, in the case of crisp objects and structuring elements with spherical symmetry, fast algorithms can be implemented, in $O(n_I)$. The distance to the object A is first computed using chamfer algorithms [28]. It defines a distance map in the image, which gives the distance of each voxel v to object A corresponding to the successive dilations of A . This discrete 3D distance can be made as precise as necessary [29]. Then the translation into a fuzzy volume of interest is made according to a simple look-up table given by the function f . This algorithm has a linear complexity in the number of voxels in the image.

The membership function of a fuzzy region representing some distance information is denoted by μ_{distance} . A few examples are shown in Fig. 1.

2.4.3. Relative directional position

In contrast to the previous relationships, relative directional position (like *object A is on the right of object B*) is intrinsically vague information. The fuzzy set framework is appropriate to formally define such relationships with good properties. To the best of

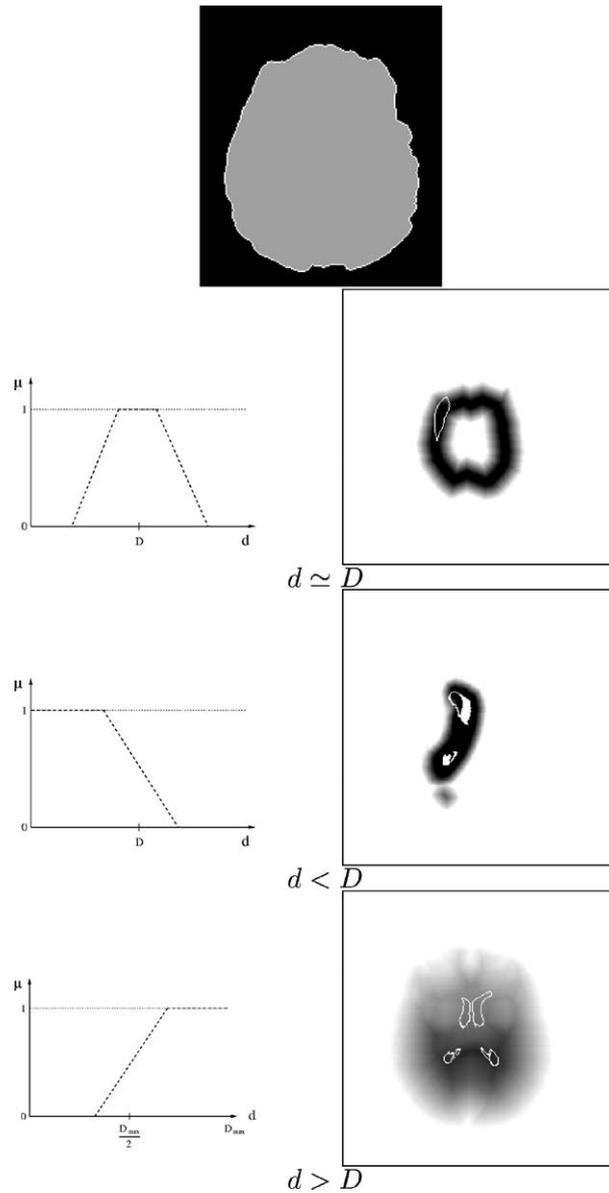


Fig. 1. Examples of representation of knowledge about distances. The first image shows one axial slice of the 3D volume (grey: brain, segmented from a 3D MRI volume, white: its surface). Left: fuzzy membership function on the distance space. Right: spatial fuzzy set representing constraint according to distance information. The second line illustrates the knowledge that the putamen has an approximately constant distance to the surface (shown on the top in white) of the brain (in grey). The third line corresponds to the knowledge that the caudate nucleus is at a distance about less than D from the lateral ventricles (in white). The fourth line corresponds to the knowledge that lateral ventricles are inside the brain and at a distance larger than about D from the brain surface. The contours of the objects we are looking at are shown in white (they are drawn just to show that they fit in the areas with high membership values). Membership values vary from 0 (white) to 1 (black).

our knowledge, almost all existing methods for defining fuzzy relative directional spatial position rely on angle measurements between points of the two objects of interest [30–32], and concern 2D objects (sometimes with possible extension to 3D). In these approaches, a fuzzy relationship is defined as a fuzzy set. More precisely, a relative position relationship is defined as a linguistic variable which is represented as a fuzzy set depending on an angle θ . On the objects, the angle $\theta(a, b)$ is measured between the segment joining two points a and b and the x -axis of the coordinate frame. Then the agreement between the relation and the measured angles is evaluated, according to three possible methods: (i) representing each object by a characteristic point as in [30,32], (ii) using an aggregation method [30, 32], (iii) using a compatibility method [31], which consists in defining a fuzzy set in $[0,1]$ representing the compatibility between the normalized angle histogram and the fuzzy relation. Another method, based on a different principle, has been recently proposed in [33] relying on a histogram of forces. Finally, the method described in [22] defines a fuzzy area, *left from A* for instance, from a projection of the object A on the horizontal axis. The degree to which B is to the left from A results from a combination of the degree of projection of B and the membership degree of B in the fuzzy area.

The approach we use is different [27]. The relationship is defined directly in the image space to be compatible with our previous developments. It is also based on a morphological approach, together with a fuzzy pattern matching procedure. It works directly in the image space, and provides the relative position between two objects in any direction.

Let us consider a reference object A and an object B for which the relative position with respect to A has to be evaluated. In order to evaluate the degree to which B is in some direction with respect to A , we use a two-step method:

- (1) We first define a fuzzy “landscape” around the reference object A as a fuzzy set such that the membership value of each point corresponds to the degree of satisfaction of the spatial relation under examination. The fuzzy landscape is defined in the same space as the considered objects, contrary to the solution proposed in [22], where the fuzzy area is defined on a one-dimensional axis. The axes of the space I are defined according to the directions of the acquisition of the volume. The direction in which the relative position has to be assessed is defined relatively to these axes.
- (2) Then we compare the object B to the fuzzy landscape attached to A , in order to evaluate how well this object matches with the areas having high membership values (i.e., areas that are in the desired direction). This evaluation is done using a fuzzy pattern matching approach, which provides as a result an interval (and not a single number).

For the application here described, the first step only is needed, which provides the fuzzy volume of interest we are interested in directly. This step is explained below.

In the 3D Euclidean space, a direction is defined by two angles α_1 and α_2 , with $\alpha_1 \in [0, 2\pi]$ and $\alpha_2 \in [-\pi/2, \pi/2]$ ($\alpha_2 = 0$ in the 2D case). We denote $\alpha = (\alpha_1, \alpha_2)$. The direction in which the relative position of an object with respect to another one is evaluated is denoted by $\vec{u}_{\alpha_1, \alpha_2} = (\cos \alpha_2 \cos \alpha_1, \cos \alpha_2 \sin \alpha_1, \sin \alpha_2)^t$. We denote by $\mu_\alpha(A)$ the fuzzy region representing the relation *to be in the direction $\vec{u}_{\alpha_1, \alpha_2}$ with respect to reference object A* . Points that satisfy this relation with high degrees should have high membership

values. In other terms, the membership function $\mu_\alpha(A)$ has to be an increasing function of the degree of satisfaction of the relation. The requirements stated above for this fuzzy set are not strong and leave room for a large spectrum of possibilities. This flexibility allows the user to define any membership function according to the application at hand and the context requirements. We propose here a definition that looks precisely at the domains of space that are visible from a reference object point in the direction $\vec{u}_{\alpha_1, \alpha_2}$. This applies to objects of any kind, in particular having strong concavities. Extensions to the case where A is fuzzy are given in [27], but are not considered here.

Let us denote by P any point in I , and by Q any point in A . Let $\beta(P, Q)$ be the angle between the vector \vec{QP} and the direction $\vec{u}_{\alpha_1, \alpha_2}$, computed in $[0, \pi]$:

$$\beta(P, Q) = \arccos \left[\frac{\vec{QP} \cdot \vec{u}_{\alpha_1, \alpha_2}}{\|\vec{QP}\|} \right] \quad \text{and} \quad \beta(P, P) = 0. \tag{15}$$

If \vec{QP} is in the direction $\vec{u}_{\alpha_1, \alpha_2}$, we obtain $\beta(P, Q) = 0$, and $\beta(P, Q)$ increases when \vec{QP} moves apart from $\vec{u}_{\alpha_1, \alpha_2}$, until a maximum value π if \vec{QP} has exactly the opposite direction. The computation of $\beta(P, Q)$ in $[0, \pi]$ preserves the symmetry with respect to $\vec{u}_{\alpha_1, \alpha_2}$ (going apart from direction $\vec{u}_{\alpha_1, \alpha_2}$ in one sense or in the other should not change the membership values in $\mu_\alpha(A)$).

For each point P , the point Q of A leading to the smallest angle β , (denoted by β_{\min}) is determined. In the crisp case, this point Q is the reference object point from which P is visible in the direction the closest to $\vec{u}_{\alpha_1, \alpha_2}$: $\beta_{\min}(P) = \min_{Q \in R} \beta(P, Q)$. The fuzzy landscape $\mu_\alpha(A)$ at point P is then defined as: $\mu_\alpha(A)(P) = f(\beta_{\min}(P))$, where f is a decreasing function of $[0, \pi]$ into $[0, 1]$. In our experiments, we have chosen a simple linear function: $\mu_\alpha(R)(P) = \max(0, 1 - 2\beta_{\min}(P)/\pi)$.

An advantage of this approach is its easy interpretation in terms of morphological operations. It can indeed be shown [27] that $\mu_\alpha(A)$ is exactly the fuzzy dilation of A by ν , where ν is the fuzzy structuring element defined on I as:

$$\forall P \in I, \quad \nu(P) = \max \left[0, 1 - \frac{2}{\pi} \arccos \left(\frac{\vec{OP} \cdot \vec{u}_\alpha}{\|\vec{OP}\|} \right) \right], \tag{16}$$

with O as the center of the structuring element. The expression of directional relative position in terms of dilation is interesting again because of the common framework provided by mathematical morphology, which guarantees good properties. It is also a way to design faster algorithms by considering a structuring element with limited support (which limits the number of directions actually considered for β).

Among the nice properties of this definition is invariance with respect to geometrical transformations (translation, rotation, scaling), which are requirements in object recognition. Also the fact that dilation commutes with union allows to represent directly disjunctive information about directional position.

In practical situations, the knowledge of direction is used to restrict the domain of search of an unknown object B in the directions α^k of previously detected objects A^k , the α^k being given by the model.

More generally, we denote by $\mu_{\text{direction}}$ the fuzzy volume of interest representing the knowledge about direction.

2.5. Example on a brain structure

In this section we illustrate the knowledge representation method on a simple example of a brain structure (see Fig. 2).

We assume that the recognition process has already recognized and segmented three anatomical objects: the brain and the two lateral ventricles (in top right view, the black structure and its white holes respectively), and we show how information about the caudate nucleus can be represented. More details about these steps are provided in Section 4.

A region of interest of the image I is depicted in Fig. 2 (top left). It represents prior information μ_{prior} about both the morphology and the localization in I of the caudate nucleus, as given by the model. It is obtained from the crisp shape of the caudate nucleus given by the model, displaced by the elastic transformation which makes the contour of the model brain fit the contour of the already detected object brain. To express the imprecision

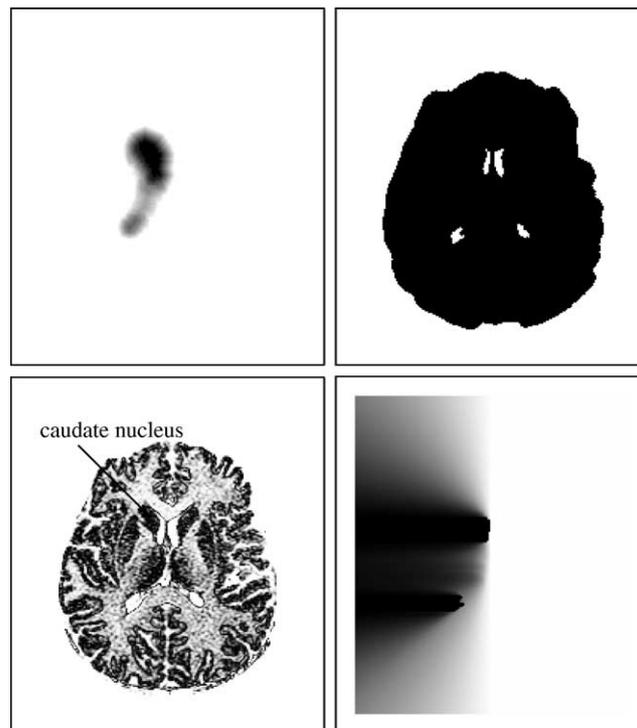


Fig. 2. Information representation in the image space (only one slice of the 3D volume is shown). This figure depicts different types of information attached to the same slice. These images are extracted from fuzzy set images built during the step of recognition of the left caudate nucleus. At this step of the recognition process, three anatomical objects have already been segmented: the brain and the two lateral ventricles (in top right view, the black structure and its white holes respectively). The prior information from the atlas (top left), the localization constraint expressing that the caudate nucleus has to be search inside the brain and outside the ventricles (top right), the a priori radiometric knowledge (bottom left) and a relative directional relationship (bottom right); white and black correspond to minimal and maximal membership values to fuzzy sets respectively.

attached to this information, a fuzzy dilation has been applied to the shape given by the model. A spherical fuzzy structuring element was used, the values of which are defined along the radius r by a trapezoidal function $e(r)$ equal to 1 for $r \leq r_k$ and to 0 for $r > r_s$, and linear in between. These two parameters define the kernel and the support of the structuring element respectively and permit us to set the degree of fuzziness of the resulting region of interest, according to the opinion of medical experts.

In Fig. 2 (top right), the binary set represents $\mu_{\text{constraint}}$. It expresses that the caudate nucleus belongs to the brain (black) but is outside of both lateral ventricles (white components inside the brain).

The result for radiometric information $\mu_{\text{medium-dark}}$ is illustrated in Fig. 2 (bottom left). It expresses that only the *medium dark* values of the T1-weighted MRI image are candidates for being pixels belonging to caudate nucleus. This knowledge also comes from medical and medical imaging experts.

Fig. 2 (bottom right) shows μ_{π} , the direction information of the caudate nucleus with respect to the lateral ventricle, an already detected object, in the π direction (to the left). It translates the knowledge from the anatomist that *the caudate nucleus is lateral to the lateral ventricle*. Such knowledge could also have been derived from the atlas.

3. Fusion and recognition

3.1. Fuzzy fusion operators

Multi-source image fusion has recently taken an important place in image processing. Most of the time, image fusion deals with the clever use of several images issued from many different sources. Here, we have to face a different situation, where we have to fuse several fuzzy images, representing different pieces of information related to the same object.

The benefit we may expect from fuzzy sets for this problem relies in the variety of combination operators [34–36]. In [20], we proposed a classification of these operators with respect to their behavior (in terms of conjunctive, disjunctive, compromise [34]), the possible variations in their behavior, their properties (mainly algebraic properties like commutativity, associativity, idempotence, etc.) and their decisiveness. Unlike other data fusion theories (like Bayesian combination), fuzzy sets provide a great flexibility in the choice of the combination operator, that can be adapted to any situation at hand.

The use of fuzzy sets in this context leads to image processing methods where the (binary) decision is rejected at the end of the processing chain. Therefore we avoid making decisions at intermediate steps with partial information only, and therefore we diminish contradictions and conflicts, which usually require a difficult control or arbitration step.

Here the problem of choice of the operator is reduced by the knowledge representation method we proposed. As imprecision is introduced directly into the representation of each piece of knowledge or information, and the obtained fuzzy regions are in general larger than the searched object, conjunctive operators are the most appropriate. Only the grey level information derived from the image (see next subsection) leads to fuzzy sets that may be slightly smaller than the searched object (because of imprecision at their

boundary) and therefore this information is more suitably combined to the others using a mean operator. In our experiments, we used mainly minimum operator (the largest t-norm), arithmetical and geometrical means. Typically, mean operators are used for pieces of information with similar spatial extensions in their representations. The specific choice of mean operators (arithmetical, geometrical means) is done experimentally. For instance, the use of a geometrical mean leads to a more severe combination, closer to a conjunction than when using an arithmetical mean. The minimum operator is used to combine these pieces of information with the binary constraint on localization (since it is a strict constraint and therefore it requires a severe operator), and with the directional relative position (since it provides very rough information and a fuzzy volume of interest that has to be revised in light of the other available information).

3.2. Determination of candidates by classification

In the proposed approach for recognition, we determine some candidate regions in the image, and try to find the one that best satisfies the constraints. In this Section, we describe a possible way to obtain candidates, that has been used in the application to detect brain structures, but which is certainly valid in other domains as well.

A major piece of information for the automatic segmentation of any brain structure is its radiometry in the image. In our method, we perform several classifications with different numbers of classes in the region of interest that corresponds to the structure. The resulting regions given by the classification are the candidates for recognition.

We have shown in [37] that two conditions are necessary for the k -means algorithm to give robust results: the number of classes must be low and several classifications with random centroid initialization should be made (so called empirical use).

Limiting the classification to a region of interest allows us to limit the number of classes and to ensure a good detection even if the radiometric distribution of the object is close to the ones of other nearby objects.

At first, the radiometry histogram of the fuzzy region of interest is calculated with the contribution of each image point weighted by its membership to the region. For each grey level l , we compute:

$$h(l) = \sum_{v \in I, l(v)=l} \mu'_{\text{prior}}(v) \quad (17)$$

where $\mu'_{\text{prior}}(v) = \min(\mu_{\text{prior}}(v), \mu_{\text{constraint}}(v))$ in order to restrict the region of interest to the area allowed according to the previously recognized objects. Using this histogram, several automatic classifications are produced by an empiric use of the k -means algorithm with different numbers n of classes (typically, $n = 2, \dots, 5$). Let us denote by $\omega_{i,n}$ the i th class in the n -class classification; its centroid and variance are:

$$c_{i,n} = \frac{\sum_{l \in \omega_{i,n}} h(l)l}{\sum_{l \in \omega_{i,n}} h(l)} \quad \text{and} \quad \sigma_{i,n} = \frac{\sum_{l \in \omega_{i,n}} h(l)l^2}{\sum_{l \in \omega_{i,n}} h(l)} - c_{i,n}^2. \quad (18)$$

Each resulting class is then translated into a fuzzy set in the image space, for instance by means of a Gaussian membership function:

$$\mu_{\text{class } i,n}(v) = e^{-((v)-c_{i,n})^2/(2\sigma_{i,n}^2)}. \quad (19)$$

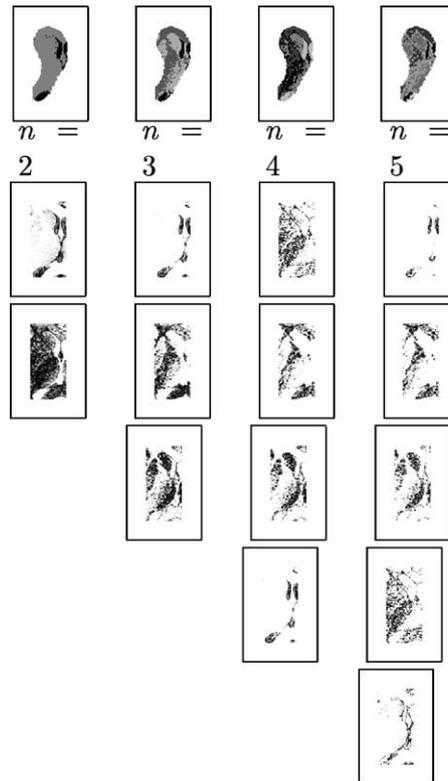


Fig. 3. Radiometric classes of a region of interest. The first line shows the results of an empirical use of the k -means algorithm processed on the histogram of a volume of interest with different numbers n of classes. The fuzzification of the classes is presented in the columns below the crisp classifications. For the caudate nucleus, the best radiometric fuzzy set is obtained for $n = 3$ (second column) and $i = 3$ (last row of this column).

Fig. 3 shows the resulting fuzzy sets. Each fuzzy set is a candidate for recognition.

This method avoids a lot of training in order to choose the parameters, since the classification is performed with several sets of parameters. The best result is then automatically chosen according to the similarity measure presented next. This makes the classification insensitive to the choice of parameters.

3.3. Similarities and selection

The selection of the best candidate is based on a similarity computation between two fuzzy sets, one representing the candidate as previously detected, and another one given by the volume of interest representing knowledge and information about the searched structure.

A lot of similarity measures have been proposed in the literature for comparing fuzzy sets (see, e.g., [38,39] for review and classification). For their use in pattern recognition in images, it is useful to classify them according to the type of information they convey.

In particular, we distinguish between operations that compare membership functions only, and operations that also use spatial distances (similarities being derived from distances). Here objects to be compared are at similar locations in the space, and operations of the first type are sufficient.

Of the different similarity functions, we have chosen three, well adapted to recognition purposes in images, one based on the volume of intersection of fuzzy sets (it is derived from Tversky's measure [40] and has been widely used in the literature, e.g., [6,38,41]), and the two others derived from fuzzy pattern matching approaches [42,43]. For two fuzzy sets μ_1 and μ_2 , they are defined by:

$$S_1(\mu_1, \mu_2) = \frac{\sum_{v \in I} \min[\mu_1(v), \mu_2(v)]}{\sum_{v \in I} \max[\mu_1(v), \mu_2(v)]}, \quad (20)$$

$$S_2(\mu_1, \mu_2) = \max_{v \in I} \min[\mu_1(v), \mu_2(v)], \quad (21)$$

$$S_3(\mu_1, \mu_2) = \max_{v \in I} (\min \max[\mu_1(v), 1 - \mu_2(v)], \min \max[1 - \mu_1(v), \mu_2(v)]). \quad (22)$$

The first measure corresponds to the volume of intersection normalized by the volume of the union, and can be considered an average measure. The two other measures correspond to extreme values, S_2 being optimistic and S_3 being pessimistic. They will be used as such when necessary.

It is important to note that the matching does not concern only points, but candidate regions provided by the clustering algorithm. For instance, if we have information about shape and relative directional position, then we build a fuzzy set representing this shape restricted to the area satisfying the directional constraint. Then the similarity aims at finding the image region which best matches this shape. So for instance it is not sufficient that a region totally belongs to the fuzzy set representing the shape information. It has to have a high similarity with this fuzzy set, which is stronger than an inclusion, and guarantees that the chosen region has actually the right shape.

3.4. Recognition: selection of the best candidate

The recognition step consists in selecting the best candidate among those obtained by classification. We have chosen to first select the best candidate in each classification result (i.e., over all classes i for a fixed n), and then among the selected candidates, to make the final selection.

Before the selection process, an initialization step is performed, which consists in restricting all fuzzy volumes representing knowledge to the localization constraints provided by μ_{prior} (shape and localization provided by the model object) and $\mu_{\text{constraint}}$ (inclusion or exclusion with respect to previously recognized objects). This restriction is performed by a conjunctive combination operator, for instance the minimum, and expressed, for any knowledge $\mu_{\text{knowledge}}$, by:

$$\forall v \in I, \quad \mu'_{\text{knowledge}}(v) = \min[\mu_{\text{knowledge}}(v), \mu_{\text{prior}}(v), \mu_{\text{constraint}}(v)]. \quad (23)$$

The two-stage selection process is as follows.

- First, we select for each classification the most appropriate candidate mode $\mu_{\text{class } i_n, n}$, by means of a similarity measure S (chosen among the three measures described above) between two fuzzy sets $\mu'_{\text{class } i, n}$ and $\mu'_{\text{radiometry}}$, where $\mu'_{\text{class } i, n}$ is derived from $\mu_{\text{class } i, n}$ and $\mu'_{\text{radiometry}}$ is derived from μ_r or μ_{matter} according to Eq. (23). This first selection is therefore based on a rough criterion on radiometry. The information on radiometry is either generic (μ_r) or specific (μ_{matter}) depending if another structure of the same matter has already been recognized or not. For each n , i_n is the index of the candidate that maximizes $S(\mu'_{\text{class } i, n}, \mu'_{\text{radiometry}})$.
- Then, we select the best radiometric mode $\mu_{\text{class } i_{n_s}, n_s}$ over the remaining candidates. For the final selection, the similarity measure is applied between each fuzzy set $\mu'_{\text{class } i_n, n}$ and the result of fusion of prior information, inclusion/exclusion constraints, and knowledge about directional and distance relationships. This second selection is based on a criterion that includes shape and structural knowledge. This selection is application dependent. The selected candidate is combined with this knowledge.

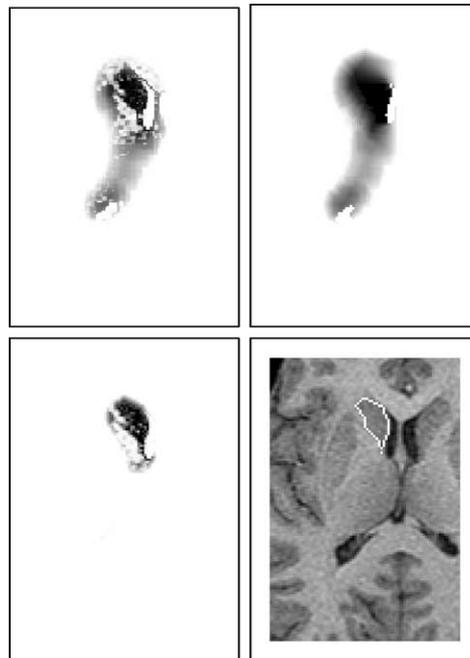


Fig. 4. Candidate selection, fusion and segmentation for recognition of one caudate nucleus. The information based on radiometry knowledge (top left) is compared to each fuzzy class resulting of a classification (for each column in Fig. 3, we get a candidate class). The information which is representative of the object localization and morphology (top right) permits finding the correct radiometric class among the different candidates. A fusion process gives a fuzzy object (bottom left) and the segmented object is deduced after regularization. Its boundary is depicted in white, superimposed on the MRI (bottom right).

In Fig. 3, the highest similarity is obtained for $n = 3$ and $i = 3$. The classification for these values is not much different from the ones obtained for $n = 4$ and $n = 5$ (same line in this figure), but still leads to a slightly higher similarity value.

Finally, a post-processing step is applied, in order to complete the segmentation of the recognized structure. It consists of a first morphological regularization using opening and closing with a small structuring element (the radius is typically chosen as the voxel size), followed by a final threshold, the threshold value being automatically given by the classification parameter corresponding to n_s and i_{n_s} .

Fig. 4 illustrates this selection process on a brain structure.

4. Application to atlas-based brain structure recognition

4.1. The proposed approach for atlas-based brain structure recognition

To guide the recognition, we make use of an atlas which is a labeled image obtained from a MRI acquisition of a normal subject. An alternative could have been to use either a probabilistic atlas or a mean atlas. A slice extracted from the atlas 3D volume is presented in Fig. 5 (left); the right view shows the corresponding slice in the 3D MRI acquisition to be processed (a different subject from the one used for building the atlas). This labeled image constitutes the iconic part of the model. The propositional part is constituted by expert knowledge about relationships between objects and expected radiometry of each structure. It will be given below for each object of interest. Note that on this example, the atlas and the 3D image to be recognized have different resolutions, and the shapes and localizations of the objects are quite different in both volumes. This example is therefore appropriate for illustrating the feasibility of the approach.

Then the objects to be detected are chosen in the order of increasing difficulty. We start with the segmentation of the brain which can easily be done from the image itself by many already existing techniques. We use the method described in [37,44], which is based on 3D mathematical morphology, and then initialize a deformation field between the atlas and the image based on the only brain surface (this deformation is composed of translation, rotation and scaling). Then we successively focus our attention on the lateral

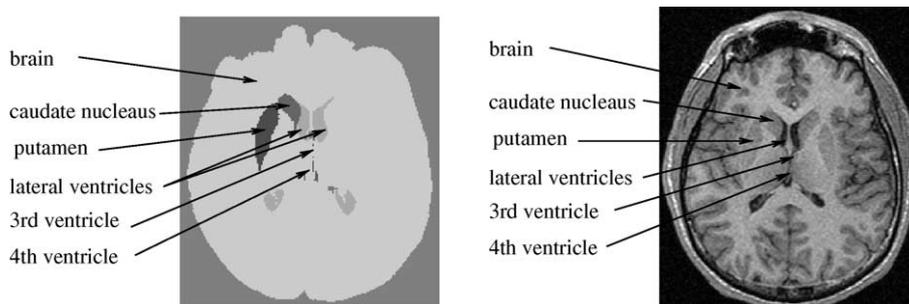


Fig. 5. One axial slice extracted from the 3D atlas and from the 3D T1 MRI image. In the atlas, each grey level represents a different object we are interested in.

ventricles, the caudate nuclei, the putamen, the fourth ventricle and the third ventricle. Each stage therefore corresponds to the detection of one structure, assuming that the detection of previous objects has been successfully obtained.

From the pairing of the previously detected objects of the image with their atlas counterpart, we deduce an elastic geometrical transformation which matches the object surfaces and interpolates the deformation to the voxels in the volumes inside and between the surfaces.

An object detection can be described by seven steps: the first five ones concern the recognition of a particular object and the last two deal with updating the correspondence to take into account the new object in the geometrical transformation which leads from the atlas to the image.

- Step 1** With the help of the correspondence field, the object shape as given by the atlas is projected in the image.
- Step 2** This binary shape is dilated with a fuzzy morphological operator in order to define in the image a region of interest that should contain the object we look at (see Section 2.3.1). This region reflects the prior information on the shape and position of the object.
- Step 3** Fuzzy classifications based on the radiometry are performed in the region of interest with different numbers of classes (see Section 3.2 for details), thus defining candidates for the searched object.
- Step 4** Each piece of symbolic information that describes the object is expressed by a fuzzy set in the image space. It can be prior radiometric knowledge either on the average grey level or on the grey level distribution, directional or distance relationships with respect to any object that has already been recognized, exclusion or inclusion from already known regions, etc. Fuzzy set construction has been presented in Section 2.
- Step 5** A two-stage fuzzy fusion process combines the prior information from Step 2 and symbolic knowledge from Step 4; two rough descriptions of the object we look at are obtained. With the help of similarity measures between these descriptions and the fuzzy sets resulting from the classifications of Step 3 (called region radiometric modes in Fig. 6), the proper candidate for the object in the image is selected (as explained in Section 3.4). A final fusion process combines this radiometric information with all pieces of knowledge about the object excluding the prior radiometric one; it leads to a fuzzy object description. A regularization followed by a binarization gives the object segmentation. This step is illustrated in Fig. 6.
- Step 6** A discrete matching to make the object definition provided by the atlas fit the segmented object is calculated with an elastic registration algorithm based on object surfaces. This step is based on the Iterative Closest Point (ICP) algorithm [45] followed by a regularization procedure [1,37].
- Step 7** A new global volume deformation field is inferred from the set of surface matching of the segmented objects. The volumic deformation is computed based on a simple mathematical model expressing that the Laplacian of the deformation field is null. The discrete deformation field is derived from its discrete values on

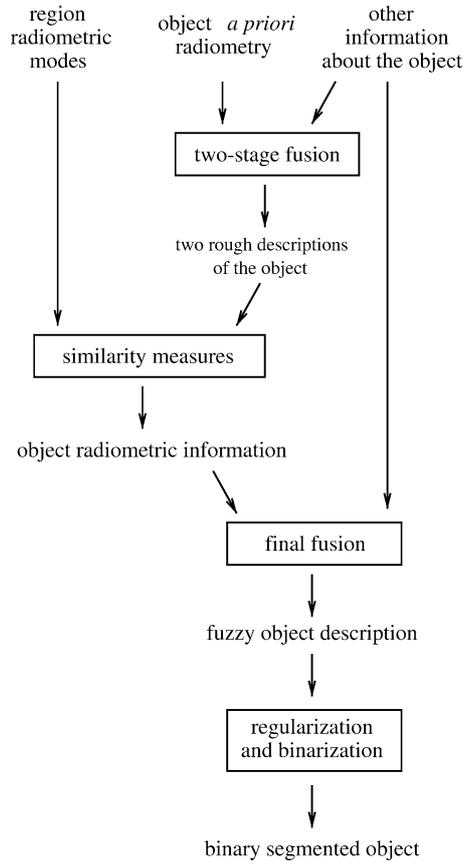


Fig. 6. Step 5: flow-chart of the object recognition using information fusion.

object surfaces under this assumption. The resolution is iterative and local to areas bounded by object surfaces [1,37]. The detailed expression of the deformation process is outside the scope of this paper.

This process may now be incremented with the detection of another object.

4.2. Results

We illustrate now on a few brain structures the type of knowledge that is used for the recognition of each of them and the obtained results. In all figures, the standard medical convention “left is right” is adopted (meaning for instance that on axial slices, the right caudate nucleus appears on the left part of the image).

After a rough registration between atlas and image using the brain surface (obtained as in Steps 6 and 7), the next step is to detect one of the *lateral ventricles*, e.g., the right

one. The object proposed by the atlas is quite closed to the expected result (Fig. 7 top left). Then, following knowledge is used:

- variability: the atlas region is dilated by a fuzzy structuring element with a core of 1 cm and a support of 1.5 cm, providing μ_{prior} ;
- set relationship: the lateral ventricle is included in the brain (therefore $\mu_{\text{constraint}}(v) = 1$ if v belongs to the brain, and 0 otherwise);
- generic radiometry: the lateral ventricle is filled up with cerebro-spinal fluid, which is dark in T1-weighted MR images; the corresponding fuzzy set is denoted by μ_{dark} ;
- distance: the lateral ventricle is about in the middle of the brain; the corresponding fuzzy set μ_{distance} is obtained as in the third case of Fig. 1.

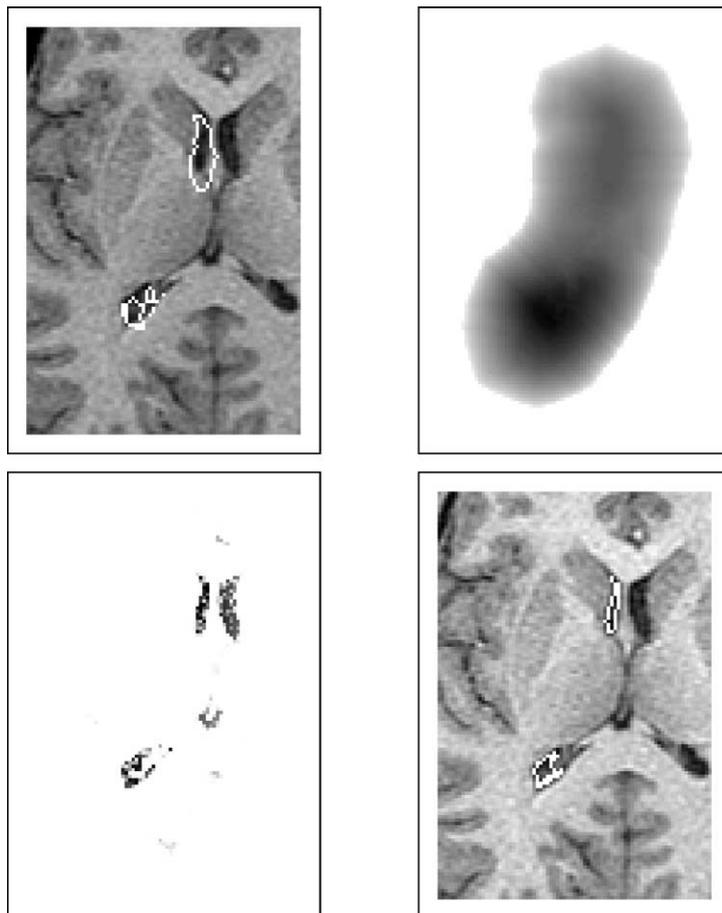


Fig. 7. Recognition of right lateral ventricle (one axial slice). Surface as given by the atlas, selection information, fusion, result.

We add to this knowledge image information provided by the classifications $\mu_{\text{class } i, n}$.

All these fuzzy sets are transformed according to Eq. (23).

The first selection is made according to radiometric information, and the second one according to distance information: for each n , i_n is chosen such that $S(\mu'_{\text{class } i, n}, \mu'_{\text{dark}})$ is maximal. Then n_s is chosen such that $S(\mu'_{\text{class } i_n, n}, \mu'_{\text{distance}})$ is maximal. The final fusion step combines the selected class with localization and distance using a mean operator. This result is then restricted to the set relationship constraint by a min operator. These operations are expressed as:

$$\min[\mu_{\text{constraint}}, m_g(\mu_{\text{class } i_{n_s}, n_s}, m_a(\mu_{\text{prior}}, \mu_{\text{distance}}))], \quad (24)$$

where m_a denotes arithmetic mean, and m_g denotes geometric mean.

The result is shown in Fig. 7. It permits us now to estimate more precisely the grey levels of cerebro-spinal fluid. The deformation field is then updated.

Then we proceed to the recognition of the *left lateral ventricle*. The knowledge used is similar as for the right lateral ventricle except for a few points, that account for previous recognition of the other ventricle:

- set relationships now also include an exclusion relationship from the right ventricle;
- radiometric information is now no more the generic and rough one, but the precise one derived from the previous step;
- an additional relationship to the right lateral ventricle is used: the left ventricle has to be searched to the left of it. This knowledge is used also in the second selection.

The results are shown in Fig. 8.

The next object is the *right caudate nucleus*. The dilation of the object atlas uses a smaller structuring element (core of 0.5 cm and support of 1 cm), since the deformation field is more precise, and the location of this nucleus is strongly constrained by the ventricles. The set relationships now include exclusion from both lateral ventricles. The radiometric knowledge is different: caudate nuclei are similar to grey matter, which appears as middle dark in these images. Here, only rough generic knowledge can be used since no grey matter object has already been recognized. Directional knowledge states that this nucleus is to the right of the right lateral ventricle. The results are shown in Fig. 9. This segmentation leads to an estimation of the grey level distribution of internal nuclei in this image. A similar process can be performed for the *left caudate nucleus*, which is not illustrated here.

The next object we look at is the *right putamen*. The previously learned radiometric information about internal nuclei is used. The results are shown in Fig. 10.

The last two objects (*third and fourth ventricles*) have been chosen to illustrate the capability of the proposed method to recognize objects that are difficult to segment directly in MRI images. They are small, and may have complex shapes. To our knowledge they have not been segmented automatically in previous works. Because of higher variability of the fourth ventricle, we used a structuring element with a larger support (1.5 cm) for the dilation of atlas object. After it has been detected, the deformation field gives a very precise localization of the third ventricle, and a very small structuring element is used (with

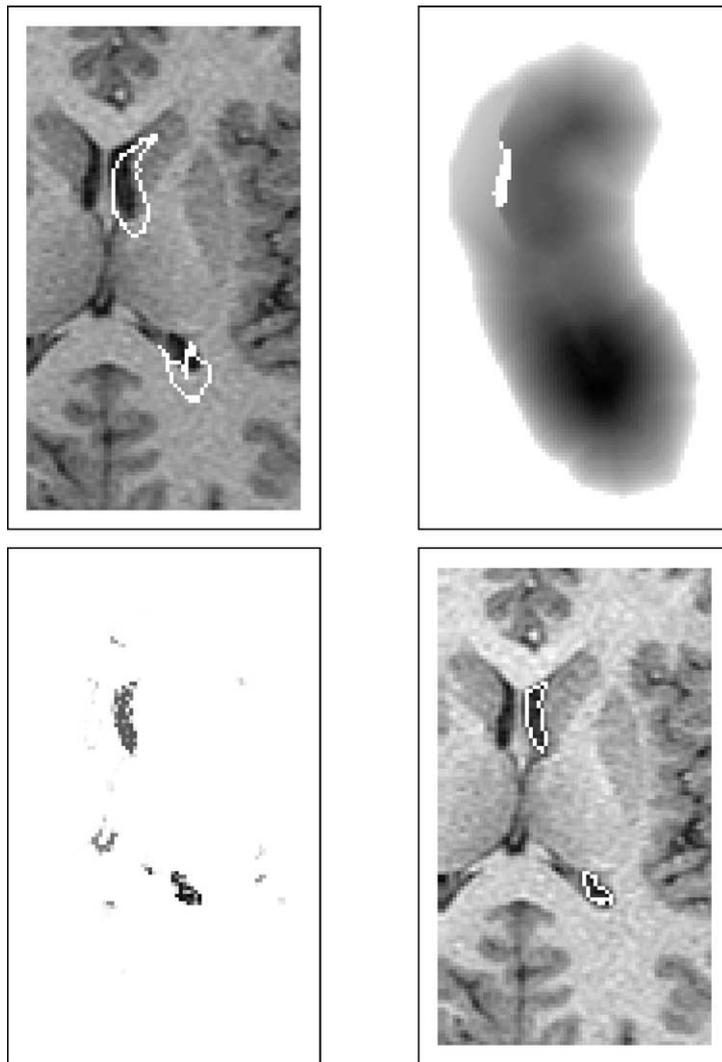


Fig. 8. Recognition of left lateral ventricle (one axial slice). Surface as given by the atlas, selection information, fusion, result.

a support of only 4 mm). Results are shown in Fig. 11 for the fourth ventricle and in Fig. 12 for the third ventricle. Note that these structures are obtained from the same volume as the previous ones. Sagittal slices (instead of axial ones as for the other structures) are presented in the figures just for better visualization purpose.

Fig. 13 shows 3D views of these objects as defined in the atlas and as recognized in an MR image with our method. They are correctly segmented although the size, the location and the morphology of these objects in the image significantly differ from their definitions in the atlas. Note in particular the good recognition of third and fourth ventricles, that are

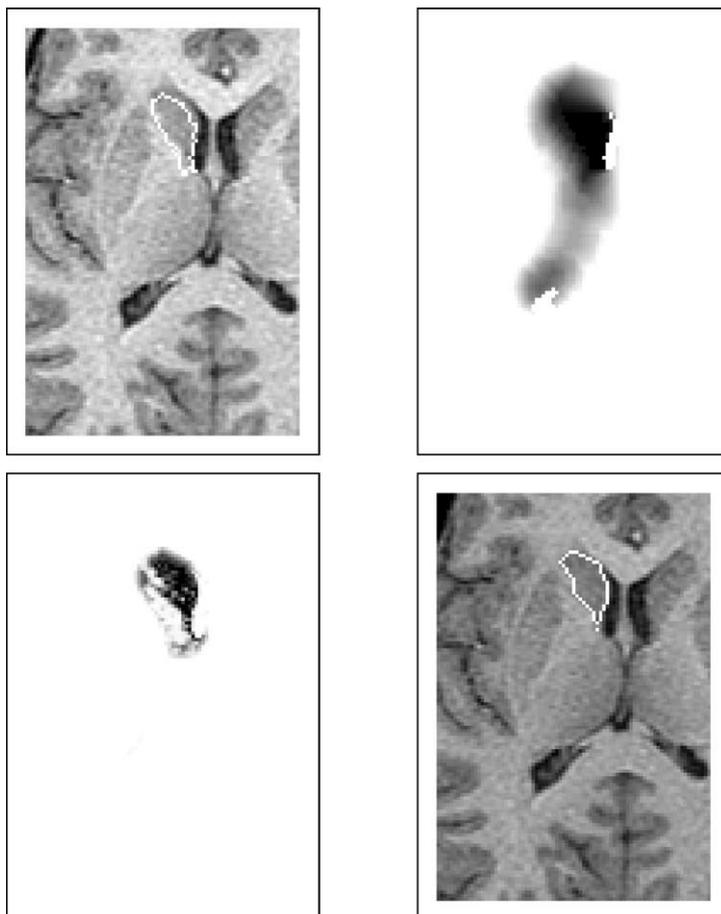


Fig. 9. Recognition of right caudate nucleus (one axial slice). Surface as given by the atlas, selection information, fusion, result.

very difficult to segment directly from the image. Here the use of relationships to other structures is very important and conditions the quality of the results.

We have carried out tests on other images, and results of similar quality have been obtained. Ten images from different subjects and different acquisition devices have been tested. The contribution of the proposed approach to the detection of brain structures actually varies depending on the structures. For instance, the ventricular system could be detected with a direct segmentation approach. The use of the atlas and of the model only makes it more robust and allows to separate different parts of this system (i.e., to distinguish between lateral, 3rd and 4th ventricles). For the caudate nucleus, we observed that all information we are using is indeed necessary to have a good detection in several cases (if we try to suppress one piece of knowledge, the method does not work well on some images). For the next structures, the fact that the registration becomes more and

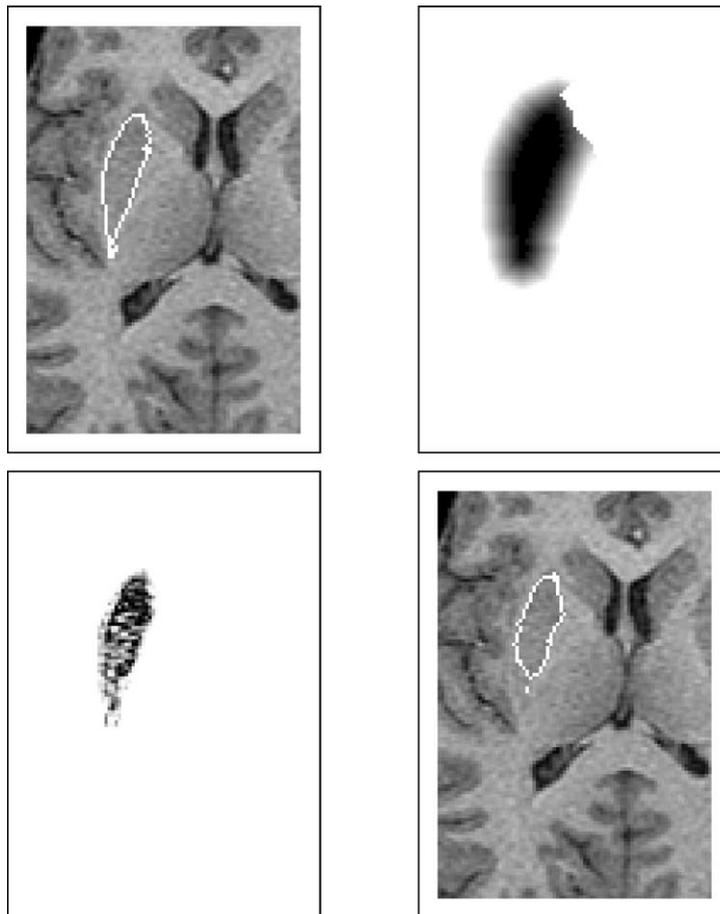


Fig. 10. Recognition of right putamen (one axial slice). Surface as given by the atlas, selection information, fusion, result.

more precise makes the use of spatial relationships less crucial, but it is still useful and results are improved, and are more robust when applied on different images.

Another reason explaining the robustness of our approach is that it relies on the segmentation of the brain as the first object, using a 3D mathematical morphology method previously developed that proved to be very robust and reliable (it is now used in routine and was evaluated on more than 30 images). Then the registration between the segmented brain and the brain in the atlas actually guarantees that the structures to be recognized are not very far from the ones of the atlas (see, e.g., the top left images in Figs. 7–12). And this fact is even improved during the subsequent steps.

If we apply the method on the image that served to build the atlas, perfect results are obtained, with a voxel accuracy. For the other images, different from the atlas, we do not

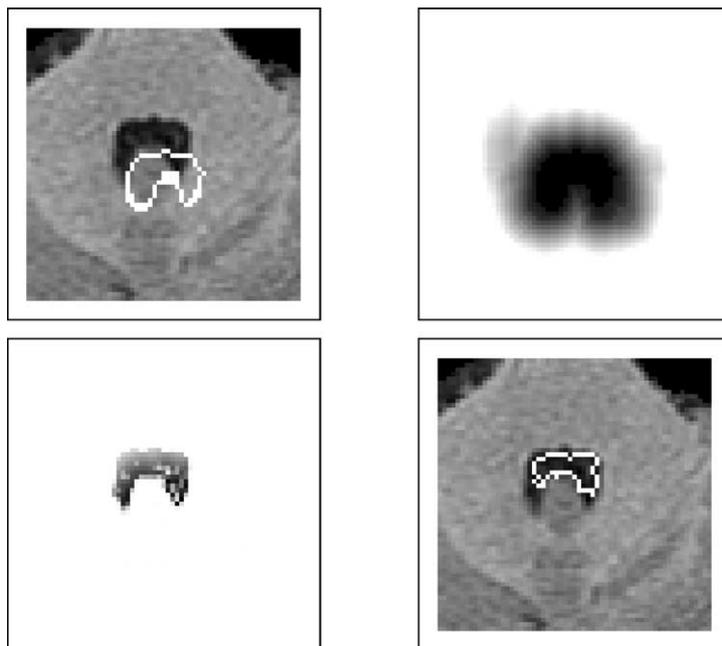


Fig. 11. Recognition of fourth ventricle (one sagittal slice). Surface as given by the atlas, selection information, fusion, result.

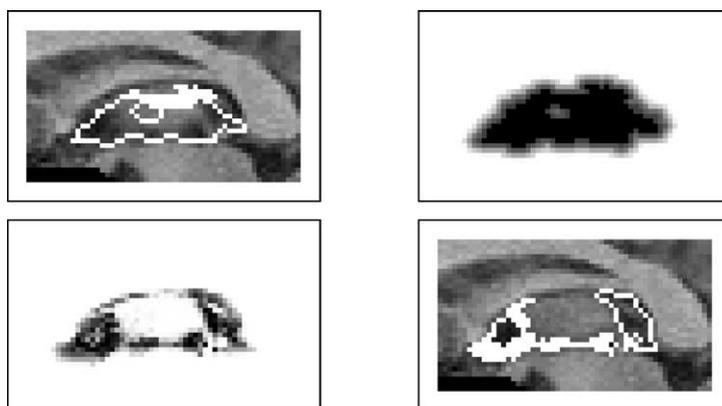


Fig. 12. Recognition of third ventricle (one sagittal slice). Surface as given by the atlas, selection information, fusion, result.

have the ground truth and it is therefore not really possible to provide quantitative results. We asked a neuro-anatomist to judge the results, and he was very satisfied. The results were even above his expectation.

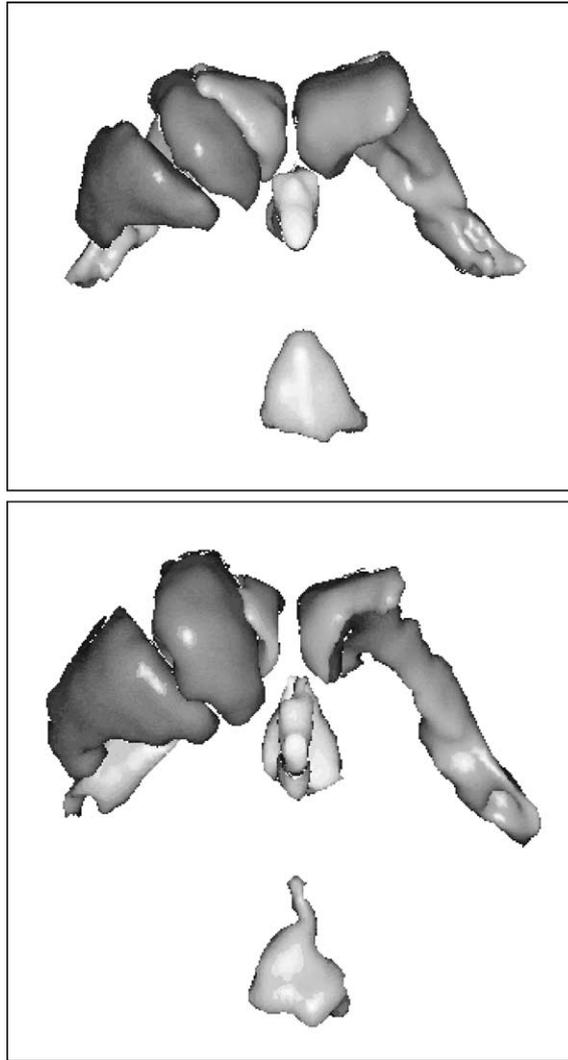


Fig. 13. Recognition results. The upper view represents six objects from the atlas: lateral ventricles (medium grey), third and fourth ventricles (light grey), caudate nucleus and putamen (dark grey). The lower view represents the equivalent objects recognized from a MRI acquisition.

5. Conclusion

We have presented an original recognition method which is atlas-guided and progressive, and which fully benefits from every piece of available structural information. A main feature of our method is that knowledge is directly expressed in the image space by the mean of fuzzy sets. Another original aspect is that it takes advantage of objects that have already been recognized. We have shown how heterogeneous knowledge can be repre-

sented in a unified framework, and combined in order to guide the recognition. The type of knowledge representation, associated with the proposed recognition method, can be used in other recognition problems using a model. Also the semi-quantitative (or semi-qualitative) interpretation of fuzzy sets bridges the gap between purely symbolic or linguistic representations and purely numerical ones. The use of morphological operators has also an interest from this point of view, since spatial knowledge can be expressed through these operators in a numeric, semi-quantitative, or logical way [27].

Most model-based approaches in medical imaging aim either at segmenting one structure based on a model of it, and then they are usually dedicated to this structure, or at performing a global registration between an atlas and a 3D image, in order to achieve segmentation of several structures. In this last case, it is difficult to account for specificities of individual structures, in particular concerning their variability, because the approach is global. Here the proposed approach overcomes both types of limitations by applying a unique method for all structures, but exploiting specific pieces of knowledge about each of them. So it is neither restricted to one particular structure, nor it has to make a global compromise. Each structure is processed according to the knowledge we have about it, while satisfying some consistency constraints with respect to the other structures.

Another advantage of the method is that no back-tracking is needed, and once an object has been recognized, it is not further considered. This may seem quite constraining, but actually the robustness of the method comes from the complete procedure, and from the fact that imprecision is explicitly represented, as well as any piece of available knowledge, no matter how heterogeneous it may be. However, at least in the considered application, the order in which structures are recognized is crucial.

This absence of back-tracking can also be a weakness of the progressive approach, due to possible wrong detection at an intermediate stage with all the consequences we may imagine. We have no solution for this problem now, but in the context of an operational brain segmentation system, we expect that the segmentation step will be under spatial supervision of an expert. In this context, the progressive detection facilitates and reduces the human intervention to the only situations where the machine is confused, and makes the human-machine interaction more efficient.

However, it should be noted that in our experiments we never found examples where one step provides a wrong result, making the whole process fail. This is due to the good initialization provided by the segmentation of the brain, and by the subsequent steps.

The interest and the power of our approach appear also in the fact that it is now used for different applications (e.g., [46]), and by other teams (e.g., [47]).

We could also consider a fuzzy object at each step, and make the final decision on the precise delineation of each object at the end of the processing.

Several aspects could still be improved. It could be interesting to try to infer automatically from the iconic part of the model the most pertinent relationships. The choice of parameters (e.g., extent of fuzzy structuring element) could also be automatized, using a learning procedure from a set of representative images with sufficient variability. Until now these parameters have been set experimentally and then were not changed for all tested images. Making this method of routine use would require a larger evaluation and parameter testing.

Until now, the only morphological information we are using is provided by the a priori model object. No geometrical characteristics (like perimeter, area, symmetry, etc.) are used since such characteristics are implicitly represented in the volume of interest. In the same way, no topological information (like holes or tunnels) is explicitly introduced. For this type of information, the fuzzy dilation may change the topology, and then nothing guarantees that the recognized object will have the desired topology. This problem is left for future work, and topologically constrained dilations could be used [48]. In a similar way, constraints about adjacency between objects could be added. As shown in [27], it can be directly related to distance and extended to the fuzzy case based on fuzzy dilation.

Other selection strategies could be implemented for the choice of the best candidate. For instance, we may imagine to check each candidate with respect to each $\mu_{\text{knowledge}}$, and then fuse the degrees of satisfaction of each constraint expressed by $\mu_{\text{knowledge}}$. On the contrary, all $\mu_{\text{knowledge}}$ could be first fused, and the best candidate would be the one having the highest similarity with respect to this fused information. Here we have chosen a hybrid strategy in comparison to these two extreme ones, that provides some flexibility in the way each type of knowledge is used. It is well adapted to the problem of brain segmentation. For different image understanding problems, according to different strategies of recognition, the previous scheme may be easily modified. The knowledge representation part, on the contrary, remains general and its principle can be applied in other domains.

References

- [1] T. Géraud, I. Bloch, H. Maître, Atlas-guided recognition of cerebral structures in MRI using fusion of fuzzy structural information, in: Proc. CIMA'99 Symposium on Artificial Intelligence, La Havana, Cuba, 1999, pp. 99–106.
- [2] J. Pearl, Fusion, propagation, and structuring in belief networks, *Artificial Intelligence* 29 (1986) 241–288.
- [3] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.
- [4] P. Smets, The combination of evidence in the transferable belief model, *IEEE Trans. Pattern Anal. Machine Intelligence* 12 (5) (1990) 447–458.
- [5] L.A. Zadeh, Fuzzy sets, *Inform. and Control* 8 (1965) 338–353.
- [6] D. Dubois, H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, 1980.
- [7] R. Krishnapuram, J.M. Keller, Fuzzy set theoretic approach to computer vision: An overview, in: Proc. IEEE Internat. Conference on Fuzzy Systems, San Diego, CA, 1992, pp. 135–142.
- [8] J.C. Bezdek, L.O. Hall, L.P. Clarke, Review of MR image segmentation techniques using pattern recognition, *Medical Physics* 20 (4) (1993) 1033–1048.
- [9] L.P. Clarke, R.P. Velthuizen, M.A. Camacho, J.J. Heine, M. Vaidyanathan, L.O. Hall, R.W. Thatcher, M.L. Silbiger, MRI segmentation: Methods and applications, *J. Magnetic Resonance Imaging* 13 (3) (1995) 343–368.
- [10] M.C. Clark, L.O. Hall, D.B. Goldgof, L.P. Clarke, R.P. Velthuizen, M.S. Silbiger, MRI segmentation using fuzzy clustering techniques, *IEEE Engineering in Medicine and Biology* 13 (5) (1994) 730–742.
- [11] D. Pham, J.L. Prince, C. Xu, A.P. Dagher, An automated technique for statistical characterization of brain tissues in magnetic resonance imaging, *Internat. J. Pattern Recognition and Artificial Intelligence* 11 (8) (1997) 1189–1211.
- [12] L.H. Staib, A. Chakraborty, J.S. Duncan, An integrated approach for locating neuroanatomical structure from MRI, *Internat. J. Pattern Recognition and Artificial Intelligence* 11 (8) (1997) 1247–1269.
- [13] P.E. Undrill, K. Delibasis, G.G. Cameron, An application of genetic algorithms to geometric model-guided interpretation of brain anatomy, *Pattern Recognition* 30 (2) (1997) 217–227.
- [14] R. Bajcsy, R. Lieberman, M. Reivich, A computerized system for the elastic matching of deformed radiographic images to idealized atlas images, *J. Computer Assisted Tomography* 5 (1983) 618–625.

- [15] D.L. Collins, A.C. Evans, C. Holmes, T.M. Peters, Automatic 3D segmentation of neuro-anatomical structures from MRI, in: Y. Bizais, C. Barillot, R. Di Paola (Eds.), in: *Computational Imaging and Vision: Information Processing in Medical Imaging*, Vol. 2432, Kluwer Academic, Dordrecht, 1995, pp. 139–152.
- [16] J.C. Gee, L. Le Briquer, C. Barillot, D.R. Haynor, Probabilistic matching of brain images, in: Y. Bizais, C. Barillot, R. Di Paola (Eds.), in: *Computational Imaging and Vision: Information Processing in Medical Imaging*, Vol. 2432, Kluwer Academic, Dordrecht, 1995, pp. 113–126.
- [17] G.E. Christensen, R.D. Rabbitt, M.I. Miller, 3D brain mapping using a deformable neuroanatomy, *Phys. Med. Biol.* 39 (1994) 609–618.
- [18] A.C. Evans, D.L. Collins, C.J. Holmes, *Computational Approaches to Quantifying Human Neuroanatomical Variability*, Academic Press, New York, 1996, pp. 343–361.
- [19] P.M. Thompson, A.W. Toga, A surface-based technique for warping three-dimensional images of the brain, *IEEE Trans. Medical Imaging* 15 (4) (1996) 402–417.
- [20] I. Bloch, Information combination operators for data fusion: A comparative review with classification, *IEEE Trans. Systems Man Cybernet.* 26 (1) (1996) 52–67.
- [21] I. Bloch, Spatial representation of spatial relationships knowledge, in: A.G. Cohn, F. Giunchiglia, B. Selman (Eds.), 7th International Conference on Principles of Knowledge Representation and Reasoning (KR-2000), Breckenridge, CO, Morgan Kaufmann, San Francisco, CA, 2000, pp. 247–258.
- [22] L.T. Koczy, On the description of relative position of fuzzy patterns, *Pattern Recognition Lett.* 8 (1988) 21–28.
- [23] J. Gasos, A. Ralescu, Using imprecise environment information for guiding scene interpretation, *Fuzzy Sets and Systems* 88 (1997) 265–288.
- [24] J. Gasós, A. Saffiotti, Using fuzzy sets to represent uncertain spatial knowledge in autonomous robots, *J. Spatial Cognition and Computation* 1 (2000) 205–226.
- [25] K.P. Gapp, Basic meanings of spatial relations: Computation and evaluation in 3D space, in: *Proc. AAAI-94*, Seattle, WA, 1994, pp. 1393–1398.
- [26] A. Yamada, T. Nishida, S. Doshita, Figuring our the most plausible interpretation from spatial descriptions, in: *Proc. 12th National Conference on Computational Linguistics, COLING'88*, Budapest, Hungary, 1994, pp. 764–769.
- [27] I. Bloch, Mathematical morphology and spatial relationships: Quantitative, semi-quantitative and symbolic settings, in: L. Sztandera, P. Matsakis (Eds.), *Applying Soft Computing in Defining Spatial Relationships*, Physica Verlag, Springer, 2002, pp. 63–98.
- [28] G. Borgefors, Distance transforms in the square grid, in: H. Maître (Ed.), *Progress in Picture Processing*, Les Houches, Session LVIII, 1992, Chapter 1.4, North-Holland, Amsterdam, 1996, pp. 46–80.
- [29] J.-F. Mangin, I. Bloch, J. Lopez-Krahe, V. Frouin, Chamfer distances in anisotropic 3D images, in: *Proc. EUSIPCO 94*, Edinburgh, UK, 1994, pp. 975–978.
- [30] R. Krishnapuram, J.M. Keller, Y. Ma, Quantitative analysis of properties and spatial relations of fuzzy image regions, *IEEE Trans. Fuzzy Systems* 1 (3) (1993) 222–233.
- [31] K. Miyajima, A. Ralescu, Spatial organization in 2D segmented images: Representation and recognition of primitive spatial relations, *Fuzzy Sets and Systems* 65 (1994) 225–236.
- [32] J.M. Keller, X. Wang, Comparison of spatial relation definitions in computer vision, in: *Proc. ISUMA-NAFIPS'95*, College Park, MD, 1995, pp. 679–684.
- [33] P. Matsakis, L. Wendling, A new way to represent the relative position between areal objects, *IEEE Trans. Pattern Anal. Machine Intelligence* 21 (7) (1999) 634–642.
- [34] D. Dubois, H. Prade, A review of fuzzy set aggregation connectives, *Inform. Sci.* 36 (1985) 85–121.
- [35] R.R. Yager, Connectives and quantifiers in fuzzy sets, *Fuzzy Sets and Systems* 40 (1991) 39–75.
- [36] D. Dubois, H. Prade, Combination of information in the framework of possibility theory, in: M. Al Abidi, et al. (Eds.), *Data Fusion in Robotics and Machine Intelligence*, Academic Press, New York, 1992.
- [37] T. Gérard, Segmentation des structures internes du cerveau en imagerie par résonance magnétique tridimensionnelle, Ph.D. Thesis, École Nationale Supérieure des Télécommunications, ENST 98E012, 1998.
- [38] R. Zwick, E. Carlstein, D.V. Budescu, Measures of similarity among fuzzy concepts: A comparative analysis, *Internat. J. Approx. Reason.* 1 (1987) 221–242.
- [39] B. Bouchon-Meunier, M. Rifqi, S. Bothorel, Towards general measures of comparison of objects, *Fuzzy Sets and Systems* 84 (2) (1996) 143–153.
- [40] A. Tversky, Features of similarity, *Psychological Rev.* 84 (4) (1977) 327–352.

- [41] C.P. Pappis, N.I. Karacapilidis, A comparative assessment of measures of similarity of fuzzy values, *Fuzzy Sets and Systems* 56 (1993) 171–174.
- [42] M. Cayrol, H. Farreny, H. Prade, Fuzzy pattern matching, *Kybernetes* 11 (1982) 103–116.
- [43] D. Dubois, H. Prade, C. Testemale, Weighted fuzzy pattern matching, *Fuzzy Sets and Systems* 28 (1988) 313–331.
- [44] T. Géraud, J.-F. Mangin, I. Bloch, H. Maître, Segmenting internal structures in 3D MR images of the brain by Markovian relaxation on a watershed based adjacency graph, in: *Proc. ICIP-95, Washington, DC, 1995*, pp. 548–551.
- [45] P.J. Besl, N.D. McKay, A method for registration of 3D shapes, *IEEE Trans. Pattern Anal. Machine Intelligence* 14 (2) (1992) 239–256.
- [46] O. Camara, G. Delso, V. Frouin, I. Bloch, Improving thoracic elastic registration in oncology by using anatomical constraints, in: *Proc. Medical Image Understanding and Analysis, MIUA-2002, UK, 2002*, pp. 205–208.
- [47] E. Frenoux, V. Barra, J.-Y. Boire, Quantification of neurotransmission defects in functional imaging using information fusion: A prospective study, in: *Proc. IPMU-2002, Annecy, France, Vol. III, 2002*, pp. 1595–1600.
- [48] P. Dokladal, R. Urtasun, I. Bloch, L. Garnero, Segmentation of 3D head MR images using morphological reconstruction under constraints and automatic selection of markers, in: *IEEE International Conference on Image Processing, ICIP-2001, Thessalonique, Greece, Vol. III, 2001*, pp. 1075–1078.