Conciliating syntactic and semantic constraints for multi-phase and multi-channel region segmentation ☆

Vincent Israel-Jost [a],[*],[1], Jérôme Darbon [b], Elsa D. Angelini [a], Isabelle Bloch [a]

[a] Institut Mines-Telecom, Telecom ParisTech – CNRS LTCI – Paris, France
[b] CMLA, ENS Cachan, CNRS, PRES UniverSud, France

ARTICLE INFO

ABSTRACT

In this paper, we propose a method to extend the multi-phase piecewise-constant segmentation method of Mumford and Shah to the multi-channel case. To this effect, we show that it is crucial to find an agreement between the syntactic constraint of obtaining regions that form a partition of the image space and the semantic constraint that attributes a formal meaning to the segmented regions. We elaborate from the work of Sandberg et al. that addresses the same problem in the binary (2-phase) case and we show that the agreement principle presented there, based on De Morgan's law, cannot be generalized to the multi-phase case. Therefore, we base the agreement between syntactic and semantic constraints on another mathematical principle, namely the fundamental theorem of equivalence relation. After we give some details regarding the implementation of the method, we show results on brain MR T1-weighted and T2-weighted images, which illustrate the good behavior of our method, leading to robust joint segmentation of brain structures and tumors.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

One of the most striking feature of today's practice of data acquisition is the possibility to produce images of a given scene with several different devices, associated with the detection of a variety of physical properties. In medical imaging for example, it is common to acquire relatively low-resolution functional images together with higher resolution morphological images that provide structural information in which objects can serve as landmarks. Another example is the use of photographs (in the visible domain) in combination with infrared images which often show different objects ("hot" objects typically). Similar examples can be found in aerial and satellite imaging.

Once registered to correct for possible mis-alignments, deformations and differences in resolution, these multi-protocol images can be seen as several channels of a vector-valued image to be segmented. The different channels are then images of the same scene, at the same time. And because they show various properties of the scene, these images also show objects differently. An illustration of such situations is provided by multi-protocol magnetic resonance (MR) images. In this paper, we address the question of how to find objects in such multi-channel images, that is, we wish to propose a method of multi-channel segmentation for multiple objects. We refer to this problem as that of multi-phase and multi-channel segmentation.

This problem however, derives from another, more fundamental, one, which is to define what a multi-channel object is. *What* is it, exactly, that we wish to segment in multi-channel images? In single-channel imaging, the answer is rather straightforward. Depending on the type of image one considers, different segmentation models exist. One can use gradient-based methods that draw closed contours along high values of the image gradient, thus defining objects as that which are enclosed by a sudden change in image intensity. Another class of segmentation methods relies on the identification of objects with regular functions. For instance, in the case of an MR image with strong field heterogeneity (that has not been corrected for), objects may be mathematically described by regions associated with linear functions or even more general smooth functions. Yet, in a variety of applications, the so-called piecewise-constant model is satisfying. It corresponds to the situation where objects are identified with relatively homogeneous patches. The segmentation model therefore depends on the choice of a formal semantics to determine what objects are. In this paper, we will focus on the case where images show objects as relatively homogeneous patches.

One popular segmentation model to detect homogeneous objects is the piecewise-constant model of Mumford–Shah, that aims to partition an image into regions that are as homogeneous as possible. More specifically, in the particular case of the restriction of

---

the general Mumford and Shah functional to piecewise constant regions, for a given image u defined in Ω (a subset of $\mathbb{R}^2$ or $\mathbb{R}^3$), the model consists in finding P regions R^j such that $\cup_{j=1}^{P} R^j = \Omega$, with associated contours ∂R^j and real values c^j, minimizing the energy:

$$E((R^j, c^j)_{j=1,\ldots,P}) = \sum_{j=1}^{P} \int_{R^j} (u - c^j)^2 \mathrm{d}x + v \cdot \mathrm{Length}(\cup_{j=1,\ldots,P} \partial R^j), \qquad (1)$$

which is the functional associated with the "minimal partition problem" proposed in [1] in the piecewise constant case, with u being the image intensity function, and v being the regularization parameter, tuning the influence of the length term with respect to the data term. The minimization of this energy can be achieved by many kinds of algorithms, based on different approaches which will not be debated at length here. Our main goal will be to provide a similar energy-based model of segmentation that adapts to the vectorial case (i.e., for multi-channel image data).

An important feature of the Mumford and Shah segmentation model is that the regions R^j are created according to two constraints that do not conflict:

(i) they form a partition of Ω (each point of Ω belongs to exactly one such region). This is a structural or syntactic constraint.

(ii) each R^j has a clear meaning. Formally, they are patches that are relatively homogeneous and smooth to some extent (depending on the parameter v). This is a semantic constraint.

The only conflict that could occur in the single-channel case is when the intensity of a point of the image u falls right in the middle between the mean values of two regions and can therefore be associated with either one of them. However, taking the decision to put this point in one of these equally distant regions is harmless, because it does not significantly change the meaning of the regions.

As it will appear in this paper, the generalization of the piecewise-constant model of Mumford and Shah to the multi-channel case easily breaks the harmony between these two constraints. Hence, we will focus on the problem of finding an agreement between them, so enforcing one does not work against the other. This approach has been used in the *logic framework for active contours on multi-channel images* of Sandberg et al. [2] in which multi-channel segmentation is considered, but only in a 2-phase framework. The extension that we offer here parallels the approach taken by these authors in that we give a principle to make the syntactic and semantic constraints compatible, therefore avoiding what they call *conflictive objectives*. This extension differs from the one recently proposed in [3], where the authors guarantee the partition constraint based on an ordering of the objects to be segmented (but the result highly depends on this ordering).

In Section 2, we describe the syntactic and semantic aspects of a segmentation and we present the problem of conflictive objectives that is created when extending a segmentation method from single- to multi-channel image data. In Section 3, we recall how the problem is solved in the 2-phase case in [2] and we suggest that, even if this solution does not apply for more than two phases, it can be understood more generally as a demand for a compatibility principle between syntactic and semantic constraints. Since we do not wish to change the fact that the segmented regions form a partition of the image, we rather study how to choose an interpretation for these regions that is compatible with the partition constraint. In Section 4, we illustrate on a few examples the behavior of our segmentation model, both with simulated and real images and conclude in Section 5.

2. Syntactic and semantic aspects of segmentation

When we consider an energy-based model such as (1) for segmentation, we actually try to combine two sets of constraints which are equally important for our understanding of the result. The first set, that we call *syntactic constraints*, aims to drive the computed regions to take a certain form. Here, we define two of them by setting the number of regions to P, which means that our final segmentation will be composed of P regions or less, and we want these regions to form a partition of the image domain (i.e. be non-overlapping while entirely covering the image space Ω). Let us call these two syntactic constraints the *maximum number of regions constraint* and the *partition constraint*. To say that these constraints are purely syntactic means that simply constructing our regions according to them has no *a priori* reason to provide us with a segmentation that is representative of the objects seen in the image (since these constraints do not explicit any link between object features or image information and the P regions). They only help clarifying the structure of a given segmentation by reducing the number of gray levels to a small number of labels (usually, P is significantly lower than the original number of gray levels) and by placing each and every pixel in one of the regions so as to avoid ambiguity: we generally do not want a set of non-classified pixels in our segmentation and we do not want pixels to be part of two or more regions at the same time either.

Thus, we need other constraints that link the segmented regions to the actual content of the image, and we call this second set of constraints *semantic constraints*. Our choice to focus on a region-based segmentation method is guided by two considerations. Firstly, it is a remarkably simple (and yet, in many cases, efficient) model that is based on one single criterion (that the heterogeneity measure between a point and the mean value of a region be low). This permits us to remain with a rather light mathematical formulation which helps generalizing to the multi-channel case. Secondly, because regions are associated with distinct mean values, these values can be used as labels for the possibly many different objects that will be segmented. The piecewise-constant model of Mumford and Shah that we use thus accommodates well to (and is actually designed for) the multi-phase case.

This distinction between syntactic and semantic constraints of a segmentation, while commonly found in the large literature dedicated to segmentation, rarely leads to discussions regarding the mutual agreement between these constraints. As we suggested in the introduction, this is because they do not really conflict in single-channel segmentation. When constructing regions that are characterized by a certain average value, deciding in what region a pixel associated with a given value should be placed is rather straightforward: we put it in the region associated with the closest average value, provided that it does not dramatically increase the regularization term. But the same procedure cannot be applied for general multi-channel segmentation by simply considering a vector of average values, as proposed in [4]. Suppose for example, in a 2-channel situation, that we want to group together the pixels that have a similar value in at least one channel. It it easy to think of examples in which, for a given pixel p with value (p_1, p_2), two distinct regions R^1 and R^2 already exist at a given iteration, with average values (v_1^1, v_2^1) and (v_1^2, v_2^2) with the property: $p_1 = v_1^1$ and $p_2 = v_2^2$ (see Fig. 1). We are then left with two possibilities: we can either choose to enforce the semantic constraint, and accept that the pixel p be placed in *both* R^1 and R^2, therefore breaking the partition constraint, or we can decide to place p in one of these two regions to enforce the partition constraint, but the resulting segmentation will not be interpreted as "the set of regions composed of pixels with similar value in at least one channel".
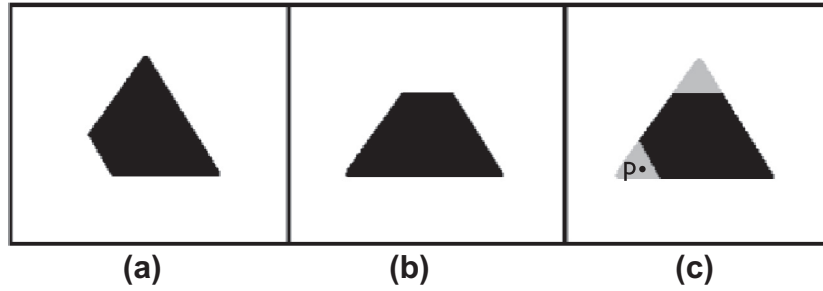
Fig. 1. Example of the use of multi-channel imaging. The two channels are shown in (a) and (b) with different appearances of the object. In (c), we show a synthetic view of the different parts of the vectorial image: in black, what is in the object in both channels, in white, what is in the background in both channels and in gray, "conflictive parts". If at some point during the iterations, two phases R^1 and R^2 represent approximately the background and the object, the point p taken in the conflictive part can legitimately be placed in either R^1 or R^2 (see text).

Sandberg et al. in [2] refer to this problem as a *conflictive objective*, that makes multi-channel segmentation harder to grasp, conceptually, because the construction of segmented objects must be understood with an additional dimension, that of the "depth" of the multi-channel image. We will now turn to the solution adopted by these authors in the 2-phase and multi-channel case to properly define objects and extend it to the multi-phase case.

3. Agreement between syntactic and semantic constraints

Since we are now considering a multi-channel segmentation model, the first question to be addressed is the selection of the most relevant level for fusing information from different channels. These possible levels of fusion can be classified into three categories: the data (fusion of the different channels), the segmentation (fusion of segmentations obtained independently for the different channels) and the fusion of a quantity that is used during the segmentation process, typically an heterogeneity measure, evaluating the distance between the grey level of a point and the average value of the region it belongs to. Arguments against the first two options can be found in [2], and it is therefore suggested to achieve the fusion on the heterogeneity measures, which are defined for each point $x \in \Omega$ as:

$$z_i^j(x) = \frac{\left| u_i(x) - c_i^j \right|^2}{K(u_i)^2}, \tag{2}$$

where $i = 1, \ldots, M$ denotes the channel, u_i the corresponding image intensity function, and $K(u_i) = \max_{y \in \Omega} u_i(y) - \min_{y \in \Omega} u_i(y)$ is the contrast of u_i, while j denotes the phase. Note that other normalization methods could be applied as well, and this is but one classical example.[2] Then Eq. (1) becomes:

$$E\left(\left(R^j, c_i^j \right)_{i=1,\ldots,M;j=1,\ldots,P} \right) = \sum_{j=1}^{P} \int_{R^j} \mathcal{F}\left(z_i^j(x), i = 1, \ldots, M \right) dx + \nu \cdot \text{Length}(\cup_{j=1,\ldots,P} \partial R^j), \tag{3}$$

where $\mathcal{F}$ denotes a fusion operator.

3.1. The 2-phase case

The logic framework of [2] was designed to segment an object in a 2-phase framework, according to several possible set operations, expressing the fusion of heterogeneity measures, applied on multiple channels. For two channels, if O_1 and B_1 are the object

and the background as they appear in the first channel, and O_2 and B_2 are the object and the background as they appear in the second channel, the final region of the object O can be defined as $O_1 \cup O_2$, $O_1 \cap O_2$ or other logic operations.

These combination rules of channel content translate into fusion rules of heterogeneity measures. For instance, the fusion of the $z_i^j(x)$ with respect to the different channels can be defined with a union function, that considers that a point x belongs to a given phase as long as the heterogeneity measure is low in at least one channel. An example of such a union function, given in [2] and denoted by $f_\cup$, is the following:

$$f_\cup\left(z_1^j(x), z_2^j(x) \right) = \left(z_1^j(x) \cdot z_2^j(x) \right)^{1/2}, \tag{4}$$

which extends to the case of M channels as:

$$f_\cup\left(z_i^j(x), i = 1, \ldots, M \right) = \left(\Pi_{i=1,\ldots,M} z_i^j(x) \right)^{1/M}.$$

Other fusion operators could also be used, and we could for instance define $f_\cup$ as the minimum of the heterogeneity measures over the channels. An intersection function can be defined as well, that considers that a point belongs to a given phase only when all heterogeneity measures are low. The intersection function $f_\cap$ proposed in [2], is as follows:

$$f_\cap\left(z_1^j(x), z_2^j(x) \right) = 1 - \left(\left(1 - z_1^j(x) \right)\left(1 - z_2^j(x) \right) \right)^{1/2}, \tag{5}$$

and extends directly to M channels. We could also define $f_\cap$ as the maximum of the heterogeneity measures over the channels.

Note that according to the usual classifications of fusion operators [5–7], these examples are actually not a conjunction and a disjunction, respectively, but rather compromise (or mean) operators, since for any $z_i^j(x), i = 1, \ldots, M$ and for any j it holds:

$$\min\left(z_i^j(x), i = 1, \ldots, M \right) \leqslant f_\cup\left(z_i^j(x), i = 1, \ldots, M \right)$$
$$\leqslant \max\left(z_i^j(x), i = 1, \ldots, M \right)$$

and

$$\min\left(z_i^j(x), i = 1, \ldots, M \right) \leqslant f_\cap\left(z_i^j(x), i = 1, \ldots, M \right)$$
$$\leqslant \max\left(z_i^j(x), i = 1, \ldots, M \right).$$

This means that the expected behavior of a conjunction or a disjunction is actually not achieved with these two operators, since the combination of several heterogeneity measures always leads to an intermediate value, and not to a lower or higher one. The information fusion theory offers several operators that are better suited. Typically t-norms and t-conorms are suitable conjunctive

---

[2] Note also that this model does not require intensity standardization among the channels since each u_i is compared only to c_i^j (in the same channel) and the fusion is performed on the normalized heterogeneity measures, not directly on the intensities.

and disjunctive operators, with good theoretical properties, and have been extensively studied within the fuzzy set theory.

Average fusion operators have also been implicitly used in early work on color image segmentation (see e.g. [8,9] for methods relying on a similar framework). As shown below, such operators are not suitable to match the considered constraints.

Applying the same fusion rule in both phases can provoke a conflictive objective. For example, for every point taken in the "conflictive area" of Fig. 1, applying an intersection fusion rule in both phases will result in the point being rejected from both phases (its heterogeneity measure is large with respect to both of them) while applying a union fusion rule in both phases will result in the point being accepted in both phases (as its heterogeneity measure is low with respect to both of them). But the partition constraint forces the decision to place the point in one phase only. This in turn results in an impossibility to interpret the final segmentation results.

To solve this problem, the solution offered by Sandberg et al. comes from a principle that combines the semantic constraints (i.e., the interpretation given to the different phases) and the syntactic ones (so that the two phases form a partition of the image). Here, the different possibilities rely upon De Morgan's law, which states that for two sets A_1 and A_2, $(A_1 \cap A_2)^c = A_1^c \cup A_2^c$. So within a 2-phase framework, if we wish to obtain the union of the appearances of the object in one phase, the second phase must be its complement (partition constraint) and therefore, must correspond to the intersection of the appearances of the background, which gives us the right semantics.

The energy of the minimization problem then takes the following form:

$$E\left(R^1, c_i^1, R^2, c_i^2\right)_{i=1\ldots M} = \int_{R^1} f_\cup\left(z_i^1(x), i = 1, \ldots, M\right) \mathrm{d}x$$
$$+ \int_{R^2} f_\cap\left(z_i^2(x), i = 1, \ldots, M\right) \mathrm{d}x$$
$$+ v \cdot (\mathrm{Length}(\partial R^1) + \mathrm{Length}(\partial R^2)), \qquad (6)$$

where z_i^j are given in Eq. (2).

Other combinations can be derived to obtain the segmentation of the intersection or union of all observations A_i of a given object or other possible bitwise logic operations, such as $A_1 \cap \neg A_2$. It is also worth noticing that the De Morgan's law generalizes for more than two channels, and leads to the following prescription in our context: if we choose a union fusion rule (over all the channels) for one phase, an intersection fusion rule must be applied to the other phase.

Therefore, this framework can handle the problem of multi-channel segmentation in a 2-phase framework, in the sense that conflictive objectives are avoided. Let us now see how we can extend this to the multi-phase case.

3.2. Multi-phase case: a new model

In most cases, two regions are not enough to represent all the different salient objects that are present in an image. The multi-phase framework provides us with the possibility to partition the image into an arbitrary number of phases, according to what we think is the right number of objects. Thus the constraint on the maximum number of regions is somewhat relaxed, since we can always adapt this number to match our needs.

Let us start with a remark on the possible fusion rules. In the 2-phase case, we only have considered the most obvious cases that correspond to the union and intersection of the appearances of an object over the different channels. However, other possibilities exist. For example, we could attach a point to a phase, if its associated value is close to the average of this phase in at least half of the channels. For the sake of simplicity, we do not wish to consider these possibilities and therefore restrict the discussion to union and intersection fusion rules.

Now that we consider a multi-phase framework, the De Morgan's law can no longer be used because it is binary by nature. It permits us to consider one set and its complement which only makes sense for two classes. So the only way that we could use the 2-phase case directly in the multi-phase case would be to iteratively apply the 2-phase case in a hierarchical scheme that divides an image into two parts, and then considers each of the two parts as a new image domain. But this approach seems tedious to manage in practice, as it requires the user to determine which phase is associated with the union fusion rule (the other being necessarily associated to the intersection fusion rule) for the segmentation of each subdomain. We will then consider the real multi-phase case and compare the two most simple possibilities, namely applying the same fusion rule in all the phases, i.e., either union or intersection.

As mentioned before, the two considered interpretations and channel combinations can be translated into a mathematical formulation using a fusion operator. Semantically, if a phase has been segmented with a union fusion rule, it means that "*all the points it contains are similar in at least one channel*", while the intersection fusion rule is associated with a phase such that "*all the points it contains are similar in all the channels*". Similarity is of course a fuzzy relation, which is necessary to reduce the quantity of information in the image. But it is built upon a crisp relation, that of identity, and this will be our starting point for the analysis of the two cases that are under consideration. The two rules of union and intersection can be expressed as relations that write as follows:

$$R_\cup = \{(x,y) | x \text{ and } y \text{ have a similar value in at least one channel}\} \qquad (7)$$

and

$$R_\cap = \{(x,y) | x \text{ and } y \text{ have a similar value in all the channels}\}. \qquad (8)$$

These two relations are based on crisp relations that write as

$$R_\cup^c = \{(x,y) | x = y \text{ in at least one channel}\} \qquad (9)$$

and

$$R_\cap^c = \{(x,y) | x = y \text{ in all the channels}\}. \qquad (10)$$

These two relations differ in an important aspect. The first one, $R_\cup^c$ is reflexive, symmetric, but it is not transitive. Indeed, in a 2-channel framework, we can have $x R_\cup^c y$ and $y R_\cup^c t$ if pixel x has values ($u_1(x)$, $u_2(x)$), pixel y has values ($u_1(y)$, $u_2(y)$), pixel t has values ($u_1(t)$, $u_2(t)$) and we have for example $u_1(x) = u_1(y)$ and $u_2(y) = u_2(t)$. In this case, there is no reason why $u_1(x) = u_1(t)$ or $u_2(y) = u_2(t)$ should hold, and therefore, $x R_\cup^c y \wedge y R_\cup^c t \nRightarrow x R_\cup^c t$. On the other hand, $R_\cap^c$ is reflexive, symmetric and transitive. It is therefore an equivalence relation.

The application of this to our context is obvious using the fundamental theorem of equivalence relation: the only way that we can obtain a partition of the image space is to group its points according to an equivalence relation. It is therefore this principle that we use to establish a link between the syntactic and the semantic constraints.

Concerning the union fusion rule of the heterogeneity measures $f_\cup$, we see that even its crisp version leads to a semantically irrelevant grouping of the points, since two points can be put in the same phase, even if none of their values are identical. In practice, this can result in the computation of phases that are not well interpretable since they do not correspond to the desired characterization of containing points that are all similar in at least one channel.

The fact that the intersection fusion rule is associated with an equivalence relation in the crisp case makes it a much better candidate for use in a multi-phase and multi-channel segmentation method. However, we are left with two questions: (1) how does this reasoning translate to the fuzzy case, when we replace the identity with a similarity operator? and (2) what is the role of the other syntactic constraint, namely, the number of phases?

We will not give a full answer to the first question here. A similarity relation can be defined as a *fuzzy equivalence relation*. It can be associated with a fuzzy partition of the space as shown in [10] for example. However, since our algorithm will ultimately deliver a crisp partition of the image space, what is needed here is a proof that this partition can be computed from the fuzzy partition that corresponds to the similarity relation, for example in the form of a set of α-cuts.

The second question is related to the importance of the user's input for the practical use of the method. We said at the beginning of the section that the multi-phase framework provided by the Mumford and Shah model for example permits one to anticipate the appropriate number of phases needed to catch the different objects of the scene. This possibility goes very well with the standardization of imaging techniques and protocols, from which users can generally predict what these different objects are. In multi-channel imaging, this task is slightly more complicated, because the semantics that we want to suggest implies that an object be defined by a relative homogeneity in all the channels. This means that "conflictive parts" should constitute objects on their own, even if they appear to be part of a well-defined object when considering one single channel (e.g. the conflictive parts in Fig. 1). So the task of the user is to predict, at least approximately, what the number of *vectorial* objects is.

Finally, the present argument seems to contradict what is proposed by Sandberg et al. in the 2-phase case, since their suggestion is to combine two different rules in the two phases, while ours is to apply an intersection fusion rule everywhere. But our respective approaches actually do not cover the same cases, since, as we just observed, our multi-phase framework is aimed to detect conflictual parts of the image, and to consider them as new objects. This means that even in the most simple case of an image with one object and background (e.g. Fig. 1), if the appearances of the object vary over two channels (i.e., if there is a conflictive part between two channels), we will need at least three phases to apply our recommendation to use intersection fusion rules in all of them when the principle that we rely on is the fundamental theorem of equivalence relation (in the example shown in Fig. 1, four phases would be necessary since the two components of the conflictive part have different values in each channel). On the other hand, if we wish to work with only two phases, then the principle that applies is the De Morgan's law and we will choose the approach described in Section 3.1.

Provided that we are able to decide for an appropriate number of phases $P > 2$, we can finally formalize our energy-based model for multi-phase and multi-channel segmentation, with phases R^j, $j = 1, \ldots, P$ such that $\cup_{j=1}^{P} R^j = \Omega$.

The multi-phase segmentation model, posed as a minimization problem in Eq. (1), can be straightforwardly extended to the multi-channel case, by fusing heterogeneity measures with a disjunctive fusion operator. Using the maximum as fusion operator, the energy to minimize is the following:

$$E\left(\left(R^j, c_i^j\right)_{i=1,\ldots,M;j=1,\ldots,P}\right) = \sum_{j=1}^{P} \int_{R^j} \max_{i=1,\ldots,M}\left(z_i^j(x)\right)\mathrm{d}x$$
$$+ \nu \cdot \mathrm{Length}(\cup_{j=1,\ldots,P}\partial R^j), \qquad (11)$$

where ∂R^j denotes the contour of the phase R^j, M is the number of channels, ν is a non-negative real number and the heterogeneity

measure z_i^j is given by Eq. (2) for each channel i with respect to the phase j.

The form of this energy is very close to that of the single-channel model seen in Eq. (1), and this is due to the fact that we have argued in favor of using the same fusion rule in all phases R^j. Hence this energy contains fewer terms than the 2-phase model of Sandberg et al. who proposed to use different fusion rules associated with the two phases. Extending this approach to the multi-phase case by assigning different fusion rules to the different phases would be possible (for example, by applying the 2-phase model iteratively), but it would be very tedious for the user who would have to supervise the fusion rules assignments, whether it be in a hierarchical scheme based on the 2-phase model or in a truly multi-phase formulation.

Even if our multi-channel segmentation model seems to be a trivial extension of that of Mumford–Shah, replacing the measure of heterogeneity of (1) by the maximum of the heterogeneity measures over all the channels, the reason why this model works – and not other models that would make use of conjunctive fusion operators for example – is profound. Hence, it has been the main goal of this section to exhibit a principle that connects the syntactic and semantic constraints, and since the main syntactic constraint is to obtain a partition of the image space, we have used the fundamental theorem of equivalence relation. We have then deduced that only operators based on an equivalence relation could be used for the fusion of the heterogeneity measures over the channels, an example of which is the maximum operator.

Therefore, our multi-phase and multi-channel segmentation model is neither a straightforward extension of that of Mumford–Shah, nor of Sandberg et al. While we share with Sandberg et al. the concern of finding an agreement between syntactic and semantic constraints in our segmentation, there are two important differences between our respective frameworks. The first one is that the work of Sandberg et al. does not apply to multi-phase partition. It is binary by nature and will therefore always lead to the segmentation of an image into foreground and background. The second one is that, in order to handle the multi-phase case, we demonstrated the need to turn to a completely different agreement principle between syntactic and semantic constraints, since De Morgan's law can only be used for binary partitions. The agreement principle based on De Morgan's law can only be straightforwardly extended to deal with more than two *channels*, as was already pointed out in Sandberg et al.'s paper, but not to handle more than two *phases*, unless one decides to implement a hierarchical scheme as was recently done in [3]. This strategy, however, must be associated with an ordering of the objects to be segmented, and a subsequent dependence of the result on this ordering. In order to treat the multi-phase case directly (that is, not by successively applying binary partitions), we formulated a new agreement principle, based on the fundamental theorem of equivalence relation. Completely departing from the original agreement principle based on De Morgan's law, it remains mathematically supported by linking the necessity to ultimately obtain a partition of the image with the semantics of the regions.

In the next section, we provide details for a simple and fast implementation and show some illustrations in medical imaging.

4. Segmentation implementation and illustration

4.1. ICM algorithm

The implementation of the minimization of E in Eq. (11) can be performed in two different ways when working in a discrete space (as for digital images).

The first approach consists in deriving the Euler–Lagrange equations, which leads to partial differential equations (PDE) governing the evolution of the c_i^j, R^j and their contours, according to an artificial time t. Using a level set formulation leads to useful expressions, avoiding a parameterization of the contours. In a second step, these PDE are solved numerically, hence the discretization of the image space and the temporal iterations are performed in the very last step, after all formal derivations have been done in the continuous domain. This approach is detailed in [4] for the multi-phase case, and in [11] for the 2^N-phase multi-channel case.

The second approach consists in first discretizing the energy to be minimized, and performing all the computation and optimization steps directly in the discrete domain. A standard discretization scheme yields a discrete Markovian energy [12]. It has been shown [9] that a discrete version of the piece-wise constant Mumford–Shah model can be written in a Markovian framework. The data fidelity term, depending on the heterogeneity measure, has the same form, except that the integral over Ω is replaced by a sum over all pixels or voxels (called sites in the Markovian terminology). As for the regularization term, controlling the length of the contours, it corresponds to a Potts model [12]. Hence the discrete version of E can be written in this framework as the following:

$$E_{discr}(v) = \sum_{j=1}^{P} \sum_{s \in \Omega} \max_{i=1,...,M} \left(z_i^j(s) \right) + \frac{v}{2} \sum_{(s,t) \in \mathcal{E}} w_{st} R(v(s), v(t)), \quad (12)$$

where v denotes the labeled image, $\mathcal{E}$ defines the set of interactions (point neighborhood system), the coefficients w_{st} for any two sites s and t are some non-negative coefficients, and

$$R(v(s), v(t)) = \begin{cases} 0 & \text{if } v(s) = v(t), \\ 1 & \text{otherwise}. \end{cases}$$

The coefficient $\frac{1}{2}$ comes from the fact that a discontinuity is counted twice. The weighting factor v of the regularization term is as in Eq. (1).

In this paper, the 8-connectivity, i.e., the nearest and second nearest neighbors, is considered for the 2D case for defining the interactions, and the weights w_{st} are set to 1 for the nearest neighbors and to $\frac{1}{\sqrt{2}}$ for the diagonal ones (second nearest ones). In 3D, the 26-connectivity is considered and the weights w_{st} are respectively set to 1, $1/\sqrt{2}$ and $1/\sqrt{3}$ for the first, second and third nearest neighbors. We also refer to [13] for other definitions of discrete perimeter of discontinuities.

Numerous methods are available for minimizing this energy, such as the Iterated Conditional Modes (ICM) [14], simulated annealing [15], or graph cuts [16,17]. We have considered the ICM algorithm because of its simplicity since the goal of this paper is not to compare optimization algorithms.

In our method, we have used the second approach to minimize E (i.e., E_{discr} is actually minimized), using an ICM algorithm. The ICM picks randomly a site s and then changes the label $v(s)$ to the one that minimizes E_{discr} (note that only the terms in the sum involving s need to be considered). This operation is iterated until there is no possible label modification that decreases the energy. In our implementation, the values of the means are updated after one sweep, i.e., after all pixels have been updated once, and the ICM is relaunched again with these updated mean values. The global process stops when the means values do not significantly change over two consecutive runs, which takes about 20 iterations.

The ICM procedure decreases the energy at each iteration. It is known that the obtained solution is not necessarily globally optimal and depends on the initial guess.

4.2. Illustrative example: Brain MRI segmentation

The proposed method is a general framework that applies to all sorts of multi-channel images, whether they be natural images, medical images, astrophysical images etc., as soon as the piecewise constant Mumford and Shah model is relevant for each channel. To illustrate the behavior of the algorithm, we choose to apply it to MRI pathological brain datasets. When a patient has a brain tumor, two types of MR protocols can be used in order to gather important information regarding the patient's condition: T1-weighted images (SPGR) and T2-weighted images (FLAIR). T2-weighted images will present the tumor (here, a low-grade glioma) and surrounding edema as a single region, very contrasted from the rest of the brain. Meanwhile, T1-weighted images also show this pathological area, but not as distinctly from the other anatomical structures as with T2-weighted images. Yet, T1 images are still worth using in combination with T2 images because they provide the observer with valuable information regarding anatomical structures: white matter (WM), gray matter (GM), cerebrospinal fluid (CSF) and gray nuclei can be distinguished much better than with T2 images. Therefore, if we are able to take the best out of those two types of information, we can expect to obtain a segmentation that is informative on both sides, anatomical and pathological. T1-weighted and T2-weighted images are presented in Fig. 2. We assume here that potential field heterogeneity has been corrected in a pre-processing step.

If we run individual multi-phase segmentations of each protocol (T1 and T2) independently, we just get approximations of the original datasets, which is useful because it is easier to interpret the segmented images which only have seven (for T1) and five (for T2) phases left, as opposed to the original 256 gray levels images. But the results, shown in Fig. 3, are still unsatisfying. The segmentation of the T1 dataset is especially good at showing anatomical structures (WM, GM, CSF and gray nuclei) but the pathological area is hard to separate from those structures, especially from GM. On the other hand, the segmentation of the T2 dataset shows the pathological area very clearly, but the other phases are not meaningful. This is not due to the low number of phases that we decided to work with, there just is not enough information about anatomical structures in the original image. Note that we assume that we have a high enough number of phases to guarantee that all regions are assigned to objects. We do not deal with the
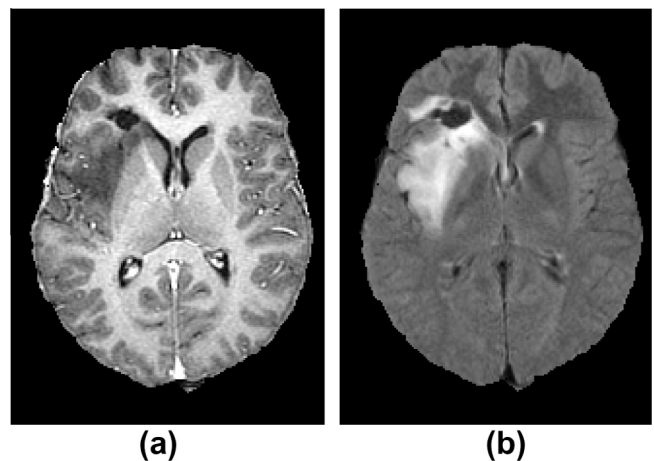

Fig. 2. T1-weighted (a) and T2-weighted (b) images of a brain with tumor and edema (one slice of the 3D volume is shown). While the pathological area, defined as the tumor and the edema, appears as one single region, very contrasted from the other tissues in the T2-weighted image (in white), the same area decomposes in the T1-weighted image into the tumor (dark gray) and the surrounding edema (lighter gray, hardly distinguishable from gray matter).
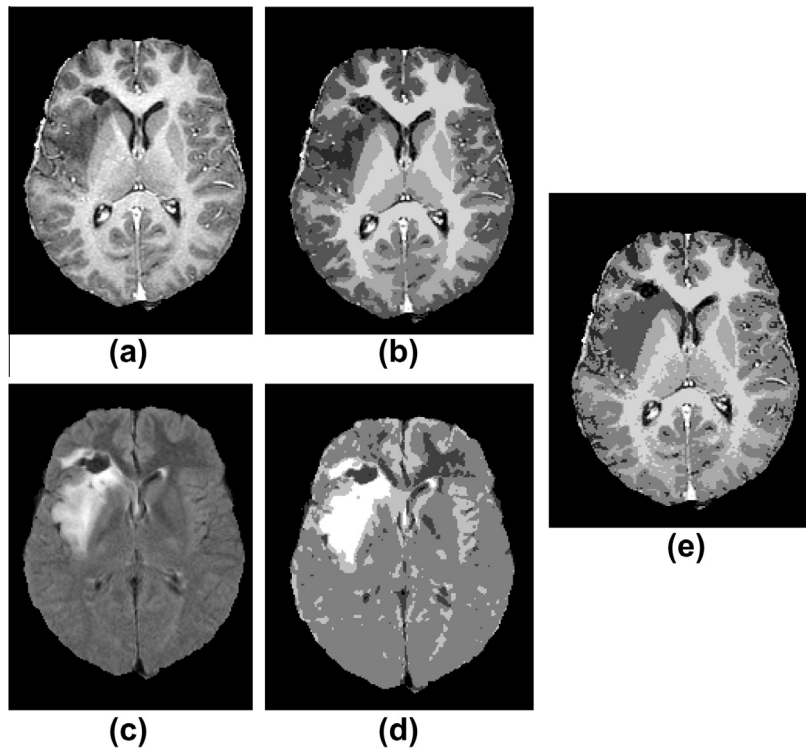
Fig. 3. MRI axial slice from the SPGR dataset (a) and FLAIR dataset (c) for Case 1. Multi-phase segmentation (ICM) of the SPGR data with seven phases (b) and of the FLAIR data with five phases (d). Multi-phase and multi-channel segmentation with seven phases (e), obtained with our method. Note that the datasets are composed of the same brain volume, acquired with different MRI protocols (SPGR and FLAIR) and that the two sets are registered. The information they provide is complementary, as FLAIR images show a well contrasted region containing the tumor and edema, while SPGR images show contrasted brain tissues (WM, GM, CSF). The final segmentation result combines what is best contrasted on each protocol, taking truly advantage of the multi-phase and multi-channel segmentation formulation. The same comment applies to Fig. 4.
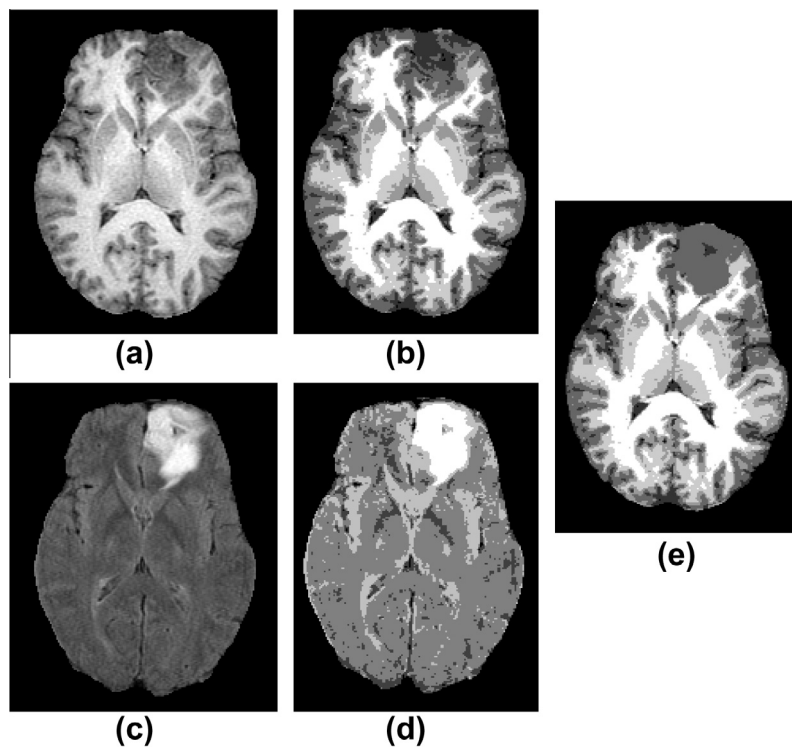

Fig. 4. One slice of SPGR dataset (a) and FLAIR dataset for Case 2 (c). Multi-phase segmentation (ICM) of the SPGR data with seven phases (b) and of the FLAIR data with five phases (d). Multi-phase and multi-channel segmentation with seven phases (e).

case of unassigned regions. This case is handled in a recent paper [3], with a different approach.

Applying our multi-phase and multi-channel segmentation method, we obtain a partition of the image that combines the most

relevant information from each channel: the pathological area is mostly segmented from the T2 dataset while anatomical structures are presented in a way that is very similar to the T1 segmentation (see Fig. 3). The number of phases that we decided to work with (seven) is of course critical, but by no means is it arbitrary or unpredictable. It simply corresponds to the number of objects that can be expected: background, WM, GM, CSF, gray nuclei and the tumor and edema (that is, the pathological area). An extra phase was added to slightly improve the results, since our semantics is rather strict. Hence, the algorithm only accepts to put a pixel in a given phase if the vectorial (multi-channel) values of this pixel are all similar to the mean values of the phase, in all the channels. Therefore, it can be useful to add an extra phase to what we predict to be the number of detected objects, that permits to better enforce the semantical constraint. But the point is that even if a little bit of experimentation is required in order to optimize the number of phases, this can be done quickly and rationally, by predicting the number of objects and then trying out this number and perhaps adding one or two phases. The rest of the parameters was set as follows, all along: 20 iterations and $\nu = 0.005$ in Eq. (11). We refer the reader to [11] for more numerical results.

We obtained essentially the same results on a second patient and in the same conditions (SPGR and FLAIR images and same parameters for the algorithm) (see [11] for more details on the results). Again, the multi-channel image combines the most relevant information taken from T1-weighted and T2-weighted images, leading to a segmentation that is accurate on both the tumoral region and the various anatomical structures (see Fig. 4).

We believe that such a behavior from the algorithm, that leads to multi-channel segmentations that take the best out of the original data, is due to the semantics chosen for all the phases. Hence, our semantics (i.e., using intersection fusion rules in *all* the phases) does not ignore large differences between objects, even when these only occur in only one of the two (or more) channels. But while such a choice for the semantics would seem illegitimate at first, especially if we follow the prescription given by Sandberg et al. to combine intersection fusion rules and union fusion rules, we also have offered an argument in favor of it, namely that it is compatible with the syntactic constraint of ultimately obtaining a partition of the image space.

While the proposed approach intrinsically sets limits regarding the fusion rules that can be used, a different approach was proposed recently in [3]. As mentioned before, it relies on a hierarchy or ordering scheme between the objects to be segmented, and the results highly depend on that. In contrast, our approach does not require such ordering and it provides a systematic approach to solve the problem of finding a compatibility principle between the two constraints.

5. Conclusion

The goal of this paper was to explore the possibility to extend multi-channel segmentation methods to the multi-phase case. In order to do so, the critical part was to identify the core idea of any such adaptation, which is that for segmented vectorial regions to make sense, one has to find principles that relate their semantics to the syntactic constraints that generally apply to such problems. In particular, we identified one such principle in the work of Sandberg et al. (De Morgan's law) but came to the conclusion that it was not applicable when more than two phases were used. We thus proposed another principle to be used instead, the fundamental theorem of equivalence relation, which in our context imposes that an equivalence relation be used to determine the content of the

different regions. One important remaining question is how exactly this reasoning translates in the case of fuzzy relations, since we were interested in using a similarity relation instead of the crisp relation of equality. But in spite of not having fully answered this theoretical question, we were able to support our new multi-phase and multi-channel segmentation method with convincing results obtained on brain MR images (T1-weighted and T2-weighted images). These suggest that our adaptation of the piecewise constant Mumford–Shah model to the multi-channel case (implemented using an ICM algorithm) works well. Indeed, we could retrieve all of the relevant objects visible on the T1 and T2 original images in the result of the multi-channel segmentation, achieving to *take the best of the two images*, all in one segmented image, which can be extremely useful for practicians. In addition, we believe that this new method has no reason to be limited in scope to MR images or to any other imaging modality since we have kept the model very general and simple.

References

[1] D. Mumford, J. Shah, Optimal approximation by piecewise smooth functions and associated variational problems, Commun. Pur Appl. Math. 42 (1989) 577–685.
[2] B. Sandberg, T. Chan, A logic framework for active contours on multi-channel images, J. Vis. Commun. Image R. 16 (2005) 333–358.
[3] M. Keegan, B. Sandberg, T. Chan, A multiphase logic framework for multichannel image segmentation, Inverse Problems Imag. 6 (1) (2012) 95–110.
[4] L. Vese, T. Chan, A multiphase level set framework for image segmentation using the Mumford and Shah model, Int. J. Comput. Vis. 50 (2002) 271–293.
[5] I. Bloch, Information combination operators for data fusion: a comparative review with classification, IEEE Trans. Syst. Man Cybernet. 26 (1) (1996) 52–67.
[6] D. Dubois, H. Prade, R. Yager, Merging fuzzy information, in: J. Bezdek, D. Dubois, H. Prade (Eds.), Handbook of Fuzzy Sets Series, Approximate Reasoning and Information Systems, Kluwer, 1999 (Chapter 6).
[7] I. Bloch (Ed.), Information Fusion in Signal and Image Processing, ISTE-Wiley, London, UK, 2008.
[8] J. Shah, Curve evolution and segmentation functionals: application to color images, in: IEEE International Conference on Image Processing, vol. 1, 1996, pp. 461–464.
[9] S. Zhu, A. Yuille, Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation, IEEE Trans. Pattern Anal. Machine Intell. 18 (9) (1996) 884–900.
[10] M. De Cock, E. Kerre, On (un)suitable fuzzy relations to model approximate equality, Fuzzy Sets Syst. 133 (2) (2003) 137–153.
[11] V. Israel-Jost, J. Darbon, E. Angelini, I. Bloch, Multi-Phase and Multi-Channel Region Segmentation and Application in Brain MRI, Tech. rep., UCLA, CAM 08-75, November 2008.
[12] G. Winkler, Image Analysis, Random Fields and Dynamic Monte Carlo Methods: A Mathematical Introduction, 3rd ed., Springer-Verlag, Amsterdam, 2006.
[13] A. Chambolle, J. Darbon, On total variation minimization and surface evolution using parametric maximum flows, Int. J. Comput. Vis. 84 (3) (2009) 288–307.
[14] J. Besag, On the statistical analysis of dirty pictures, J. Roy. Stat. Soc. Ser. B 48 (1986) 259–302.
[15] S. Geman, D. Geman, Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Machine Intell. 6 (1984) 721–742.
[16] Y. Boykov, O. Veksler, R. Zabih, Fast approximate energy minimization via graph cuts, IEEE Trans. Pattern Anal. Machine Intell. 23 (11) (2001) 1222–1239.
[17] J. Picard, H. Ratlif, Minimum cuts and related problem, Networks 5 (1975) 357–370.