Medical Imaging Report

Article: Image Segmentation with One Shape Prior A Template-Based Formulation
Siqi Chena, Daniel Cremersb, Richard J. Radkea

Author:
Jules Scholler
Adele Courot

Professors:
H. Delingette
X. Pennec

December 7, 2016

Introduction

We are going to present an article focused on image segmentation for medical imaging. In many applications, we may have a prior knowledge about the shape of the object we are looking for, and incorporating such knowledge into the segmentation problem can drastically improve the segmentation performance. This paper first reviews representative regularization-based approaches and show how they can all be formulated in the same framework. Then it points out the unavoidable drawbacks of this framework which was the motivation of this paper. In a second part, the authors describe their template-based formulation along with its advantages over the standard regularization-based formulation. Thanks to this new formulation, they designed an extremely efficient and powerful approach for image segmentation and give experimental results that we will present in the last part. In this article, the authors focus on the case when a single shape prior template is available. They chose to represent the shape prior as a characteristic function, in other words they use a binary representation of the wanted object.

1 Common framework: regularization-based formulations

For the problem of image segmentation with one shape prior, most of the previous works define an energy function in the following way:

$$E(C,T) = E_{data}(C,I) + \lambda D(C,T(C_{ref})) \tag{1}$$

$$with \quad \begin{cases} I : \Omega \to \mathbb{R} & \text{Input image} \\ C & \text{Shape representation} \\ T & \text{Shape transformation} \\ D & \text{Shape distance measure} \end{cases}$$

The different existing methods vary by their representation of C, the number of transformations taken into account and the shape distance measure.

Shape representations C The most common shape representations C used in the literature include parametric shape representations: $[0, l(C)] \to \mathbb{R}$, signed distance functions (SDF) $\phi : \Omega \to \mathbb{R}$, binary characteristic functions $u : \Omega \to \{0, 1\}$, or very recent work on relaxed characteristic functions $u : \Omega \to [0, 1]$.

Shape transformation T Typical transformations T are limited to parametric global transformation, including rigid, similarity or more general projective transformations. Rigid transformation is the simplest model as it considers only 3 parameters (t, r, s) which are respectively a translation, a rotation and a scaling.

Shape distance measure D Various shape distance measures D have been proposed and mostly they depend on the shape representation. Most approaches use a region-based Maximum Likelihood (ML) model or a boundary-based gradient model. It is also possible to combine these two models together.

The typical optimization strategy employed to minimize (1) is to optimize alternatively over C and T:

1. Fix T and update C. This becomes a standard image segmentation problem and common optimization methods can be used.

2. Fix C and update T . This is typically solved by gradient- based optimization methods such as gradient descent on the parameters of the transformation.

3. Go to Step 1 and iterate until convergence.

The main drawbacks of this method is the non-optimal and slow convergence. Indeed, due to the non convexity of the second step problem, the overall energy minimization cannot guarantee global optimality and the alternating optimization strategy can be slow to converge. Besides these optimization problems (mainly due to the form of (1)), the authors observed in practice that the regularization-based framework often needs careful tuning (e.g., gradient descent step size, order of parameter update), which makes the approach difficult to be applied in a general setting. We present an example of results obtained with the previous method on Fig. 1 where we can clearly see the

limitation of such methods.
Because a much higher precision is absolutely necessary in medical imaging, the authors developed a different method that is extremely faster to converge and which gives very good results without the tuning complication.

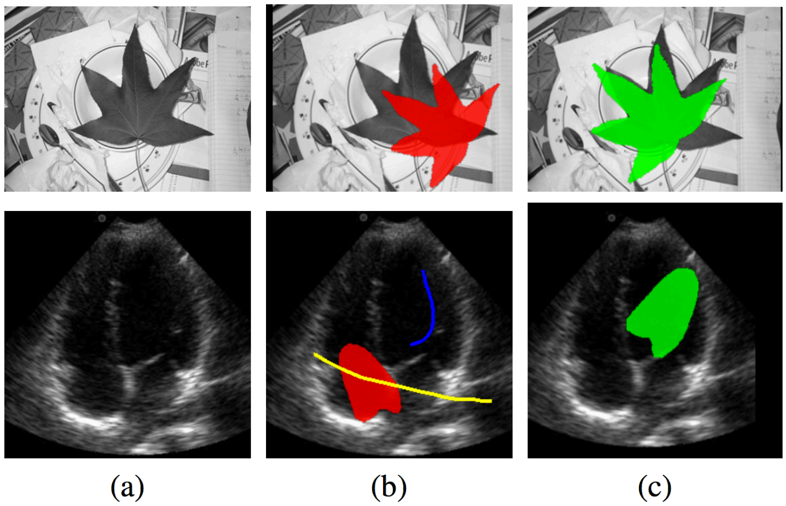

Figure 1: Segmentation results taken from the article. (a) Input image. (b) Initial position of shape template (in red) overlaid on the original image. (c) The segmentation result (in green).

2 Proposed method: the template-based formulation

2.1 Template-based formulation

The authors choose the binary characteristic partition $u : \Omega \rightarrow \{0, 1\}$ as the shape representation for its simplicity. The transformation imposed on the shape template is denoted as T. The different kind of transformation considered here are similarity, rigid and deformable. They have followed the typical framework of using a Maximum Likelihood model:

$$E_{data}(u) = \int_{\Omega} (log(P_{out}) - log(P_{in}))u \ dx \quad + constant = \int_{\Omega} Q.u \ dx \quad + constant \tag{2}$$

where $P_{in}(x)$ and $P_{out}(x)$ are respectively the probabilities that pixel x belongs to the object and to the background . The log-likelihood map $Q(x)$ describes our confidence that a certain pixel belongs to the background. In this part we consider a standard two-phase Gaussian model:

$$Q(x) = log\left(e^{\frac{-(I(x)-M_{In})^2}{2\sigma^2}}\right) - log\left(e^{\frac{-(I(x)-M_{Out})^2}{2\sigma^2}}\right) = (I(x) - M_{Out})^2 - (I(x) - M_{In})^2 \tag{3}$$

The new idea is to use the reference shape as a template instead of introducing regularization as in (1). In other words, implicitly constrain that $u = T(u_{ref})$. The energy function can then writes:

$$E(u) = E_{data}(T(u_{ref})) = \int_{\Omega} Q.u(T(x)) \ dx \tag{4}$$

The image segmentation problem thus corresponds to minimizing (4) only over the geometric transformation parameters T. We will see that this makes all the difference with the methods presented in the first part. The first difference is that there is no need to find λ.

As explained in [2] image registration algorithms can be classified into intensity-based, feature-based and time-frequency domain methods.Image registration involves spatially registering the target image to align with the reference image. Intensity-based methods compare intensity patterns in images via correlation metrics, while feature-based methods find correspondence between image features . This template-based reformulation of image segmentation with one shape prior is clearly related to image registration. That is, the segmentation problem is equivalent to registering the binary shape template u_{ref} and the image log-likelihood map Q. Because image registrations problem have been heavily studied, there exists a lot of efficient algorithms to treat these types of problems which make the template-based formulation very interesting.

2.2 Similarity transformation

When the transformation is limited to a global similarity transform, the registration cost function is non-convex with respect to the parameters (translation, rotation and scaling). The optimization process is therefore complicated and the global optimum is rarely reached. The authors proposed a instead frequency-domain inspired approaches for similarity transforms.

2.2.1 Translation

If the transformation is restricted to translation, the energy function takes the following form:

$$E(t) = \sum_{x} Q(x) u_{ref}(x - t) \tag{5}$$

The Fourier shift theorem (which states that a translation in the image domain corresponds to a phase shift in the Fourier domain) shows that (5) can also be written as:

$$E(t) = \mathcal{F}^{-1} \left(\mathcal{F}_Q \mathcal{F}^*_{uref} \right) \tag{6}$$

By resolving this problem in the Fourier domain, the algorithm is much faster. Indeed, the computational complexity of the exhaustive evaluation drops from $O(N^4)$ in the spatial domain to $O(N^2 logN)$ in the frequency domain, thanks to the efficient FFT algorithm [1].

2.2.2 Rotation and scaling

In order to use the same principle as in (2.2.1) the Cartesian space (x, y) needs to be transformed into log-polar coordinates (a, b):

$$a = log \left(\sqrt{x^2 + y^2} \right) \qquad\qquad b = tan^{-1} \left(\frac{x}{y} \right) \tag{7}$$

Doing that, the rotation and scaling transformations map to simple translations in the log-polar space and it is possible to employ a similar FFT-based technique to recover the scale s and rotation r. **In order to use this trick, one must be careful because the FFT assume that the image is periodic. This is the inevitable price for using the FFT. Nonetheless, periodization artifacts could be avoided by windowing the prior shape before taking its FFT. This causes some limitations on the algorithm, such as quasi-optimal convergence.**

2.2.3 Deformable transformation

In most of the case, the prior shape does not fit exactly in the image, therefore the authors considered a deformable transformation. Such transformations are difficult to describe with global parametrized transformation because the number of parameters can be very high and the convergence of such methods can be very slow (impossible in some cases). In order to avoid those difficulties the authors decided to implement a free-form deformation (FFD) model, based on B-Splines. The algorithm is more complicated than for rigid transformation and the parameters needs to be carefully tuned. This step takes about 50 iterations, around 20 seconds to converge, which is slower than global similarity transformation estimation and is therefore the touchy step of the whole method.

2.2.4 Full transformation

In order to implement the full transformation (including both similarity transformation and deformable transformation) and because joint optimizing similarity and deformation is intractable the authors decided to compute the full transformation in two steps:

1. Alternately update the translation t and the transformation pair (s, r) using the Fourier shift theorem trick.

2. When the similarity transformation estimation converges to the optimal (t, r, s), apply the deformation estimation using FFD.

This strategy makes sense because applying first a simple similarity transformation leads to a near-optimal solution and an accurate and fast initialization. The deformation transformation is then applied to refine the solution. The previous strategy is illustrated on the following figure (the whole segmentation takes 3s. on an ordinary computer).

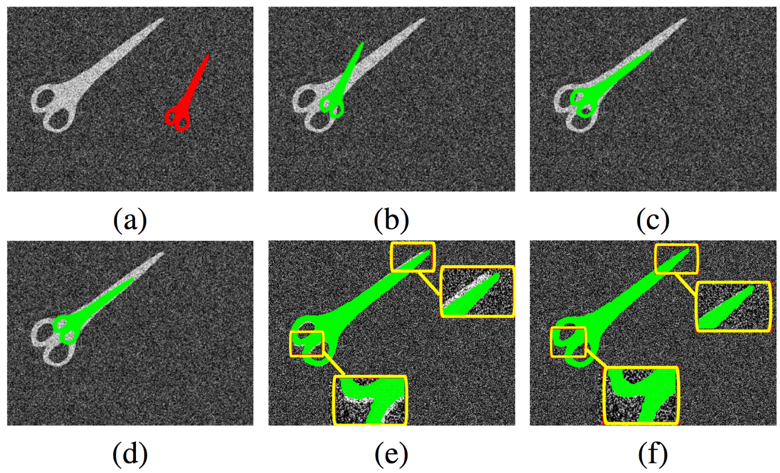

Figure 2: Full similarity transformation result. (a) Initialization. (b) Translation 1. (c) Rotation 1. (d) Translation 2. (e) Scaling 1 (f) FFD (final result).

3 Method improvements - Intensity modeling

Up to this point, the log-likelihood map has been calculated using a simple two-phase Gaussian model. This model can be refined to obtain better result.

3.1 User-provided strokes and bounding box

User-provided strokes and bounding box not only serve as interactive hard constraints that the segmentation is required to satisfy, but also provided a method to estimate the intensity distributions of both the object and the background. The idea is to consider the histograms of pixel as density probability and use this estimation to determine the log-likelihood map. **Even if this approach gives satisfying results, it requires the intervention of someone and is not fully automated. This represent a big limitation for industrialization purpose.**

3.2 Gaussian Mixture Model (GMM)

The main idea is to use two GMM, one for the object and one for the background. This model is more complex and allows to fit more general objects, not only uniform shapes. The probability model for both object and background is given below (with different parameters θ and π):

$$\mathbb{P}(x) = \sum_{k=1}^{K} \pi_k \mathbb{P}(x|\theta_k)$$

The algorithm requires now to estimate the GMMs in addition to the transformation. A classical method for the model estimation is the EM algorithm, but it is too computationally expensive. Instead the authors have chosen a simplified approach:

1. Assign the points of the object (or the background) to K clusters, by the color quantization technique.

2. From each cluster, generate one gaussian component. The weights are proportional to the cluster size.

This estimation occurs at each step, but to reduce complexity it is not done form scratch each time. Instead, the points of the background and the object are assigned to the corresponding clusters then these clusters are updated. Below are presented the algorithm and its results.

Algorithm 3 Shape prior image segmentation with GMM

Input:Shape prior template u_0 and the image I.
Initialization: Compute initial GMM parameters and Q.
while Not converged **do**
  1.Compute translation $\mathbf{t}$ from u_i and Q.
  2.Update u_i with $\mathbf{t}$ to $u_{i\mathbf{t}}$
  3.Compute s and r from $u_{i\mathbf{t}}$ and Q.
  4.Update $u_{i\mathbf{t}}$ with s and r to u_{i+1}
  5.Update the GMM paramters and Q.
end while
Deformable transformation estimation.

Figure 3: Complete algorithm

(a) (b) (c)

Figure 11: Shape-prior segmentation with GMM (a) Image with the reference shape in red. (b) The segmentation result after similarity transformation. (c)
Figure 4: Results

4 Limit of the approach

A first idea coming to mind reading this article is that the deformation model used is quite simple and therefore it seems hard to predict the behavior of the algorithm confronted to pathological cases. Indeed, will the object be recognized thanks to its correlation with the template if there is a large tumor inside? There are no test on such data, but in the conclusion of their article, the authors explain they will tackle the medical segmentation problem in future works.
Another path to explore proposed by the authors is a more complex shape prior knowledge, more precisely to study the case where a shape distribution is known and not only one realization of this distribution. This could be particularly well adapted for medical imaging where experts can provide different shape priors.

Conclusion

In this article, the authors propose a new method to address the segmentation problem. Their method is based on a mathematical simplification allowing a major gain in computational time. In the common problem formulation, an energy based on the length or the curvature of the shape is minimized subject to an equality constraint (the distance between the reference shape and the actual one should be 0). This constraint is handle through a penalization of that distance. The authors made a different choice: they integrated the equality constraint and reformulated the problem.

This method requires to estimate probability distributions of the intensity, and some models were described in the article. The results in terms of quality of the segmentation and computational time are promising.
However, there are several possible improvements to work on because as it is presented in the article, there are no guarantees that the segmentation method will work on more complex problems like medical imaging segmentation.

References

[1] James W. Cooley and John W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, Vol. 19, No. 90 (Apr., 1965), pp. 297-301(3), 1965.

[2] Bernd Fischer and Jan Modersitzki. Ill-posed medicinean introduction to image registration. *Inverse Problems*, 24(3):034008, 2008.