HAL
open science

Enforcing Monotonous Shape Growth or Shrinkage in Video Segmentation

Yuliya Tarabalka, Guillaume Charpiat, Ludovic Brucker, Bjoern H. Menze

To cite this version:

Yuliya Tarabalka, Guillaume Charpiat, Ludovic Brucker, Bjoern H. Menze. Enforcing Monotonous Shape Growth or Shrinkage in Video Segmentation. BMVC - British Machine Vision Conference, Sep 2013, Bristol, United Kingdom. hal-00856634

HAL Id: hal-00856634

https://inria.hal.science/hal-00856634v1

Submitted on 2 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Enforcing Monotonous Shape Growth or Shrinkage in Video Segmentation

Yuliya Tarabalka
yuliya.tarabalka@inria.fr
Inria Sophia-Antipolis, AYIN team
Sophia Antipolis, France

Guillaume Charpiat
guillaume.charpiat@inria.fr
Inria Sophia-Antipolis, STARS team

Ludovic Brucker
ludovic.brucker@nasa.gov
NASA GSFC & Universities Space Research Association
MD, USA

Bjoern H. Menze
bjoern@ethz.ch
Computer Vision Laboratory
ETH Zurich, Switzerland &
Inria Sophia-Antipolis, Asclepios team

Abstract

We propose a new method based on graph cuts for joint segmentation of monotonously growing or shrinking shapes in time series of noisy images. By introducing directed infinite links connecting pixels at the same spatial locations in successive image frames, we impose shape growth/shrinkage constraint in graph cuts. Minimization of energy computed on the resulting graph of the image sequence yields globally optimal segmentation. We validate the proposed approach on two applications: segmentation of melting sea ice floes from a time series of multimodal satellite images and segmentation of a growing brain tumor from sequences of 3D multimodal medical scans. In the latter application, we impose an additional inter-sequences inclusion constraint by adding directed infinite links between pixels of dependent image structures.

1 Introduction

One of the great challenges in computer vision is automatic segmentation of objects in videos. This task becomes more difficult when image sequences are subject to low signal-to-noise ratio or low contrast between intensities of neighboring structures in the image scene. Such challenging data are acquired routinely, for example, in medical imaging or satellite remote sensing.

While individual frames can be analyzed independently [3], temporal coherence in image sequences provides a lot of information not available for a single image. Two categories of spatio-temporal video segmentation techniques can be distinguished. Causal methods consider only past data for segmenting each next frame [16, 19]. Omniscient techniques use both past and future data by treating the video as a 3D space-time volume, so that segmentation of the entire image set supports each of the individual segmentations [7, 10].

In this work, we focus on segmenting shapes which only grow or shrink in time, from sequences of extremely noisy images. Examples of growing shapes are forest fires in satellite

© 2013. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

images and organ development in medical imaging. We consider image sequences, where both foreground and background intensity distributions can vary significantly over time, foreground can be heavily occluded or undistinguishable from a part of the background. Most of previously-proposed spatio-temporal methods rely on coherence of foreground/background intensity distributions in consecutive image frames, and therefore fail when segmenting such noisy data sets. Few approaches have been specifically designed for spatio-temporal segmentation of shapes from magnetic resonance images (MRI) with low signal-to-noise ratio [17, 20]. Applied to multi-temporal sequences that show a monotonously growing or shrinking structure, however, these smoothing methods bias results towards the mean shape obtained from averaging consecutive segmentations and, hence, underestimate rapid growth or shrinkage events.

To address this issue, we propose a new segmentation framework based on graph cuts for the joint segmentation of an image sequence. It *introduces growth or shrinkage constraint in graph cuts* by using directed infinite links that still guarantee sub-modularity, connecting pixels at the same spatial locations in successive image frames. By minimizing an energy computed on the resulting spatio-temporal graph of the image set, the proposed method yields a *globally optimal solution* (Sec. 2). Differently from the state-of-the-art spatio-temporal techniques, it does not rely on the coherence of the intensity in time, but only on the coherence of the shape. To summarize, the main contribution of this paper is:

1. a new method for segmentation of image sequences with the constraint of shape growth/shrinkage,
2. in order to be able to segment extremely noisy data,
3. in a low computational time (see Fig. 2(b)).

We demonstrate the performance of the proposed framework for two applications with very noisy image sequences (Sec. 3): The first one addresses the segmentation of multiyear sea ice floes in a set of satellite images acquired through two different satellite sensors. The new method returns accurate melting profiles of sea ice, which is important for building climate models. The second application deals with the segmentation of brain tumors from longitudinal sets of multimodal MRI volumes, where we impose additional inter-modal inclusion constraint for joint segmentation of different image structures.

2 Enforcing shape growth/shrinkage in graph cuts

Graph cut is an optimization tool, based on the rewriting of image segmentation problems as (s,t)-min-cuts in graphs, on the equivalence of (s,t)-min-cut and max-flow problems, and on the existence of efficient algorithms to solve the latter ones [2, 4, 9]. It can be used to find the globally optimal binary segmentation of images where the segmentation criterion E is related to a Markov Random Field with submodular interaction terms:

$$E(L) = \sum_{\text{pixels } i} V_i(L_i) + \sum_{i \sim j} W_{i,j}(L_i, L_j), \tag{1}$$

where L is the binary labelling function to be found (L_i is the label of pixel i), individual potentials V_i are any binary real-valued functions measuring the disagreement between a prior probabilistic model and the observed data, $i \sim j$ denotes a pair of neighboring pixels (any neighborhood system can be used), and $W_{i,j}$ are any real-valued interaction terms between neighboring pixels expressing spatial coherency of labels.

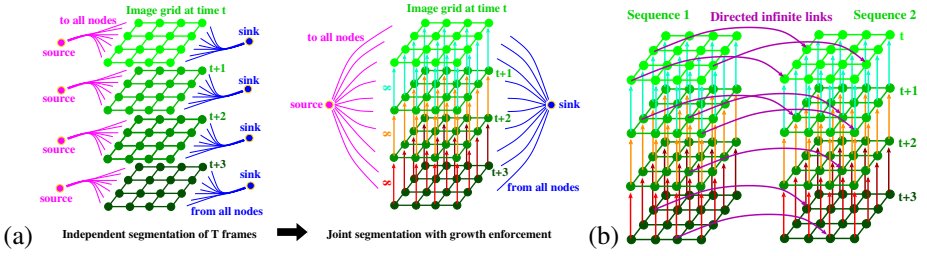

Figure 1: (a) Enforcing shape growth in an image sequence. (b) Segmenting jointly two sequences S1 and S2, by enforcing the foreground of S1 to contain the foreground of S2, with directed infinite links between all pixels of coordinates (x,y,t), from S1 towards S2.

A directed infinite link between two pixels expresses precisely the constraint that this pair of pixels cannot have the pair of labels (0,1). In the case of image binary segmentation, if 0 stands for the background and 1 for the foreground object, then this means that the second pixel may belong to the foreground only if the first one already does.

2.1 Related works

Extensions of graph cuts to multi-label problems (*i.e.* multi-class segmentation) have been proposed but generally do no guarantee optimal solutions, except *e.g.* in the case of Ishikawa's construction [12], which requires labels to be ordered and the interaction term to be a convex function of their differences, *i.e.* $W_{i,j}(L_i,L_j) = g(L_i - L_j)$ with g convex. This graph construction makes intensive use of infinite links for constraining the min-cut solutions to satisfy desired properties required to interpret them as image segmentation solutions. This was the source of inspiration for our work.

A related study to shape constraints can be found in [6], where one image has to be segmented in several possibly-overlapping objects. Infinite links are used for imposing common boundaries, inclusion or exclusion conditions between objects in a same single image.

Wolz et al. [20] used graph cuts for simultaneous segmentation of serially acquired MRI volumes. They set temporal edge weights by computing intensity differences of voxels at the same spatial locations. Thus, the same smoothness constraint was applied both in space and time, and the segmentations at different timepoints were forced to be consistent in areas where a small intensity difference between the images exist. Unfortunately, this type of temporal constraint is suboptimal in image series where intensity distributions of foreground and background vary significantly over time. To the best of our knowledge, our work is the first to use infinite links for enforcing a temporal growth constraint, and we illustrate in Sec. 3 the advantage of the new method over previous approaches such as [20].

2.2 Growth/shrinkage constraint

Shape growth in a sequence of images $I(t)$ can be easily expressed as the property that the foreground object cannot lose any pixel when time advances. Otherwise said, if a pixel belongs to the foreground object at time t_1, then it belongs also to the foreground object for all times $t_2 > t_1$. Equivalently, and simpler: a pair of pixels $((x,y,t),(x,y,t+1))$, sharing the same location and immediately successive in time, cannot have the pair of labels (1,0), with

the same binary segmentation notations as above. This can be enforced by setting directed infinite links from all pixels to their immediate predecessor in time.

Given T images $I(t), t \in [1,T]$, and as many associated submodular segmentation criteria E^t, we transform the problem of segmenting independently each image $I(t)$ according to its criterion E^t, into a joint segmentation of all images together, by enforcing the shape growth constraint with directed infinite links (see Fig. 1(a)). Thus, instead of applying graph cut T times independently to planar grids of the size of the images $W \times H$, we apply graph cut once to a 3D grid $W \times H \times T$, consisting of the same nodes and edges, but with additional directed infinite links in time. The criterion to be minimized is then $E = \sum_t E^t$ under the constraint of shape growth, and the solution found by graph cut is globally optimal, since the problem is binary and submodular.

One can enforce shape shrinkage instead of shape growth, by reversing the direction of infinite links. Another straightforward extension consists in applying this approach to the case of sequences of 3D images. The directed infinite links are then set for all pairs of voxels of the form $((x,y,z,t),(x,y,z,t-1))$ to enforce 3D shape growth.

In some applications, it may happen that growth (or shrinkage) is only very probable, but not with probability 1, *i.e.* growth should be considered as a probable hint but should not be enforced strictly at all locations at all times. In that case, one may replace directed infinite links by directed finite links: the weights of these links will encourage growth (more or less strongly depending on the weight), but sufficiently disagreeing potentials V_i may make the shape locally shrink instead. Thus, shrinkage would be discouraged but not forbidden.

The precise theoretical worst case complexity of the proposed method depends on the max-flow algorithm used, and ranges from quasi-quadratic to cubic. However, the computational time observed in practice is known to be much faster on typical image segmentation problems. We used the graph-cut algorithm of Boykov and Kolmogorov [2], and we report a linear observed complexity with the total number of pixels $T \times W \times H$ (Fig. 2(b)).

In case of long sequences of big images, the memory space required may exceed computer's capacities. This is however not an issue, as there exist graph cut implementations for massive grids [5] meant for such cases, where all information is not stored in the memory at all times. This was not required for the experiments in this article though.

2.3 Inter-sequences inclusion constraint

It is also possible to segment jointly several images sequences $I(s)(t)$ with the constraint that the foreground object in some sequences should be included in the foreground object of some other sequences (this will be useful in Sec. 3.2). This can be done similarly by considering together the graphs associated to all sequences, and, for each inclusion constraint, adding directed infinite links between pixels of the desired sequences s_1 and s_2, sharing same location and time: such links from (s_1,x,y,t) to (s_2,x,y,t) for all x,y,t will force the foreground object in sequence s_1 to contain the one of sequence s_2 (see Fig. 1(b)).

Naturally, instead of imposing an inclusion constraint over the whole time span and the whole image space, it is possible to specify spatio-temporal domains of constraints, for instance to express that the inclusion property between two sequences has to be satisfied inside a pre-defined region and/or during a pre-defined time span $[t_1,t_2]$ only, by adding directed infinite links in these sets only.

2.4 Rewriting as a multi-label problem

We express here an alternative point of view on sequence segmentation with growth constraint. The successive labels $L_i(t)$ of a given pixel i over time might change only once, and only from 0 (background) to 1 (foreground object). Hence, this vector of labels $L_i(t)$ is of the form $(0,0,\ldots,0,1,\ldots,1,1)$ and can be represented by just the time index of the first 1. This way, our initial binary optimization problem on a sequence of images with shape growth constraint can be seen as a multi-label problem defined on one single image, without any constraint. It turns out that this new problem can be expressed in the Markov Random Field form (1) and satisfies the conditions needed to apply Ishikawa's construction [12] to be solved globally. The graphs built by the two approaches are actually similar. Our initial formulation is however more flexible, in that interaction terms can depend on t, and more natural, in that inclusion constraints can easily be enforced in spatial or/and time subregions only, while this would not be expressible with the multi-label formulation.

3 Experimental results

We applied the proposed method for two applications: (a) segmentation of a melting multiyear ice floe from a sequence of satellite measurements and (b) segmentation of growing brain tumors from multimodal MRI volumes. The performance of our framework with monodirectional links, *[Mono=const]*, is compared with other graph-cut-based methods:

• *[w/o]*: Graph cut with no temporal links (independent segmentation of each frame).

• *[Feedforward]*: Images are segmented successively by graph cut. After segmenting the first frame, foreground/background pixels are marked as seeds with infinite unary costs in the next frame for enforcing growth/shrinkage.

• *[Bi=const]*: Smoothing by introducing bidirectional temporal links, *i.e.* in both directions (from t to $t+1$ and from $t+1$ to t), with a constant weight w (finite or infinite).

• *[Bi=variable]*: As proposed in [20], temporal links are computed based on intensity differences between pixels in successive image frames, *i.e.* in the same way as spatial links.

To compare the methods we used the Dice score [8], $D = (2|\hat{M} \cap M|)/(|\hat{M}| + |M|)$, where $\hat{M}$ and M are manually and automatically segmented foreground regions, respectively.

3.1 Application 1: melting sea ice in satellite images

Sea ice is both an active participant of the Earth's climate and a sensitive climate indicator. Therefore, it is important to automatically monitor how rapidly sea ice floes melt. Previous studies attempted to consider temporal information for ice floe segmentation, by using ice percentages, area and shape parameters of ice floes at the previous time moment as priors for segmenting a floe at the next time moment [11, 19]. While a multiyear ice floe can only melt in the summer period, these feedforward approaches were unable to accurately estimate melting ice profiles because of the low signal-to-noise ratio and lack of contrast in satellite data.

We aimed at segmenting a multiyear ice floe from a 45-day sequence (summer period of 2008) of measurements over the polar regions by two Aqua satellite sensors: Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E, 6.25 km spatial resolution) and Moderate-Resolution Imaging Spectroradiometer (MODIS, band 1, 250 m spatial resolution). Fig. 2(a) shows an example of the MODIS and AMSR-E images. The floe was tracked from the AMSR-E data, where a multiyear ice has a low microwave emissivity (dark

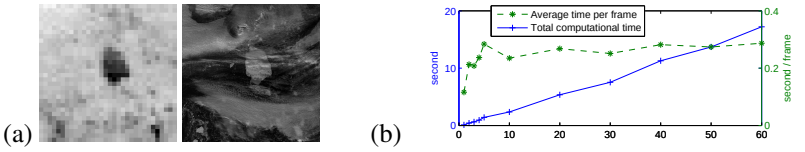

Figure 2: (a) AMSR-E and MODIS images captured on the 244th day. (b) Computational time for the proposed segmentation of the ice floe set as a function of the number of frames.

area in Fig. 2(left)), and is in this way distinguishable from clouds and younger ice which has a higher emissivity (white area in Fig. 2(left)). However, the low spatial resolution of these data does not allow to accurately quantify the ice floes areas. In accordance with tracking measurements, a time series of $T = 75$ MODIS and upscaled AMSR-E images with the ice floe was built, with spatial dimensions of 800×800 pixels. We denote each MODIS image by I^t, $t = 1,...,T$.

In order to apply the proposed method with a shrinkage constraint to the selected time series, the images must be aligned first, so that the property that the floe in the image I^{t+1} is included in the floe of the image at the previous time moment I^t can be expressed directly in terms of pixel locations. For this purpose, we detected a reliable region of the foreground, R_F, and a reliable region of the background, R_B, from the AMSR-E images, and computed histograms of the intensities I^t of the floe, $p^t(I|F)$, and of the background, $p^t(I|B)$, respectively. We calculated T maps of foreground probabilities as

$$p_i^t(F|I) = \frac{p^t(I|F)P_i^t(F)}{p^t(I|F)P_i^t(F) + p^t(I|B)P_i^t(B)}, \quad P_i^t(B) = \frac{A_i^t}{\max_j A_j^t}, \quad P_i^t(F) = 1 - P_i^t(B),$$

where $A^t, t = 1,...,T$ are AMSR-E images smoothed by Gaussian filter. The images I^t were aligned by exhaustive searching over rigid motions (rotations and translations) to maximize the correlation between maps of floe probabilities at the current and previous time moments. We then computed potentials and interaction terms between neighboring pixels as:

$$V_i^t(1) = -ln[p_i^t(F|I)], \quad V_i^t(0) = -ln[p_i^t(B|I)], \quad W_{i,j}^t = \delta_{L_i \neq L_j}\, \beta \exp\left[-\frac{(I_i^t - I_j^t)^2}{2\sigma^2}\right], \quad (2)$$

where σ is a standard deviation of I^t, β is a parameter that controls the importance of the spatial interaction energy term. We found experimentally that setting $\beta = 2$ yields robust results. The proposed method was applied with monodirectional temporal links to enforce floe shrinkage, as described in Sec. 2, with the fixed edge weight w varying from 0.25 to ∞. The results *[Mono=0.25...∞]* are compared with those obtained with other graph-cut-based approaches (listed above) in Table 1 and in Fig. 3-4. Both graph-cut with no temporal links and feedforward approaches show the worst performances, and prove to be not well suited for segmenting such noisy data sets. When a feedforward method encounters a frame with the part of the floe obscured by clouds and thus undistinguishable from the background, it segments only the visible part of the foreground, and then is trapped in a non-sense segmentation for the rest of future times. The method using gradient-based temporal links *[Bi=variable]* [20] also yields poor segmentation accuracies, because it is sensitive to both noise and variation of intensities of foreground/background in consecutive frames.

We explain in Fig. 3-4 the advantage of using monodirectional infinite links *versus* bidirectional links in the temporal dimension. Bidirectional edges with low values of w enforce

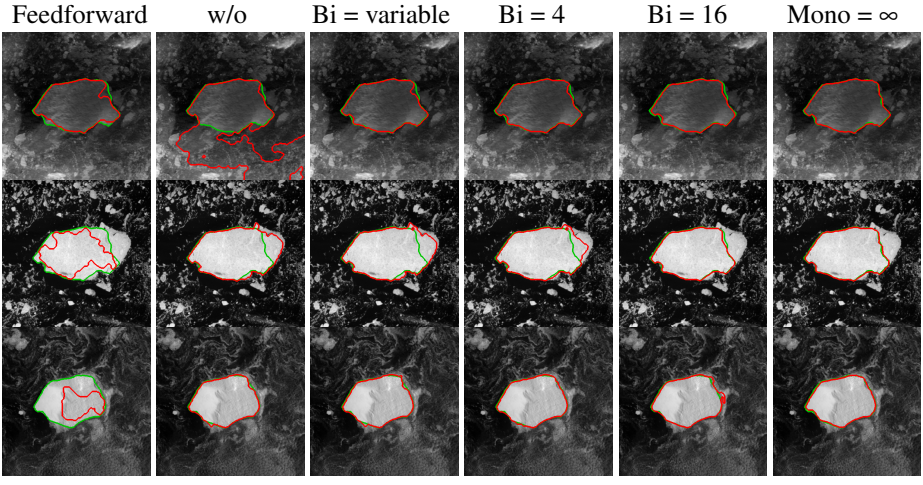

Figure 3: (From top to bottom) Segmentation contours (red) for three time moments (days 230, 235 and 267) computed by the graph-cut methods: (From left to right) *[Feedforward]*, *[w/o]*, *[Bi=variable]*, *[Bi=4]*, *[Bi=16]*, proposed method with monodirectional infinite links. Manual segmentation is shown in green. The rightmost part of the white area in the middle row is not part of the object, but another ice floe who temporarily collided.

Table 1: Mean and standard deviations of the Dice scores for the proposed method *[Mono = ∞]* and graph-cut-based approaches used for comparison.

Method	Feedforward	w/o	Bi = variable	Bi = 16	Mono = ∞
Dice score	.554 ± .128	.933 ± .099	.958 ± .048	.978 ± .007	.980 ± .007

only smoothness of variation of the contour in time, and yield segmentation errors in the case of low foreground/background contrast. For instance, in the second image of Fig. 3, the floe of interest collided temporarily with another ice floe. When using a weak smoothness constraint (see segmentation contour *[Bi=4]*), the small encountered floe collided with the floe of interest during a certain number of consecutive frames would be considered as a part of the foreground. Enforcing more smoothness in space-time to avoid this has the undesirable effect of smoothing the foreground shape, so that the segmented foreground area is lower (underestimated) than the ground-truth for the first frames, and higher (overestimated) for the last frames (see results *[Bi=16]* in Fig. 3 and-4(right)). With the increase of w, the estimated foreground tends to the constant shape for all time moments, and the Dice score decreases.

When the proposed shrinkage constraint is used, with the increase of $w \to \infty$ segmentation accuracy increases, and $w = \infty$ yields results with monotonous shrinkage of the shape size. Moreover, the proposed method copes well with rapid shrinkage events, without underestimating preceding images, or overestimating the event itself at onset. Another advantage of using directed infinite links is that there are no parameters to estimate in the temporal dimension. Fig. 2(b) depicts the computational time for the proposed graph-cut-based optimization as a function of the number of frames. The total computational time grows linearly with the number of frames, and is approximately twice the time that is taken by the independent segmentation of each frame.

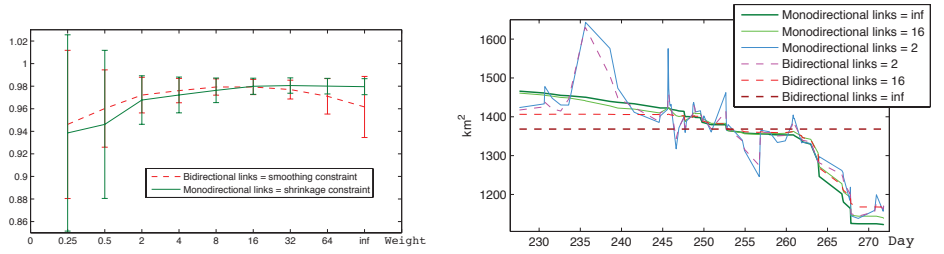

Figure 4: (Left) Mean and standard deviation for the dice score as a function of the temporal link's weight, when using mono- (green) and bidirectional (red) temporal links. (Right) Area of a multiyear ice floe as a function of time, computed by using mono- and bidirectional links with different weights.

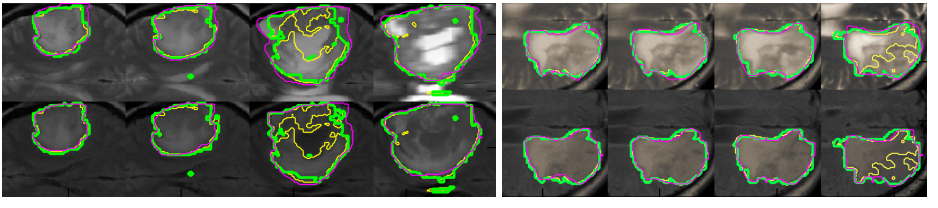

Figure 5: Two time series of T2 and FLAIR MR image volumes, each of which acquired about 3-6 months apart. The left case is rapidly growing between the second and fourth scene, the right case displays intensity modifications in the last scene, leading to a suboptimal performance of the initial multimodal segmentation (yellow) [15]. The proposed multi-temporal segmentation with growth constraints (green) delineates areas similar to the manual evaluation (magenta), being more robust against intensity variations of the MR images. It does not smooth out outlines of rapidly growing tumors as conventional bi-directional temporal constraints would do.

3.2 Application 2: growing tumor in 3D medical scans

Glioma is the most frequent primary tumor of the brain. The tumor is known to grow steadily, and lesions are diagnosed with respect to volume change in different magnetic resonance image (MRI) modalities. In our experiment we evaluate a set of 760 multimodal image volumes, each comprising T1, T1c, T2, and FLAIR MRI, acquired from ten different glioma patients and with each patient series having 3-14 time points. All image volumes are rigidly registered by using the FLIRT registration tool [13], and three orthogonal 2D slices intersecting with the tumor center are manually annotated in every volume, representing an approximate truth. Full 3D segmentations for images of each individual time point are obtained using a generative model for multimodal brain tumor segmentation [15]. This algorithm models the lesion with a latent atlas class [17] amending the tissue atlas of the standard EM segmenter. It was applied to each multimodal data set at each time independently, delineating the lesion individually in each modality. The model assumes that changes of the core (visible in T1c) will occur within the larger edema regions (visible in T2 or FLAIR) and, hence, to only have class transitions from *healthy* to *edema* and from *edema* to *core*. As the tumor grows steadily [1, 18], we can assume that negative volume changes stem from imaging artifacts. To this end we model the tumor volume to be either stable, regularizing the

segmentation along time and suppressing noise, or to expand between any two time points.

We identified the foreground label F with tumor (*edema* and *core*) and background B with healthy tissue. Then the potential $V_i^{s,t}(L_i^{s,t})$ of label $L_i^{s,t}$ at voxel i, time point t, and imaging sequence s is equal to $V_i^{s,t}(0) = p_{s,t}(F|I^{s,t})$ and $V_i^{s,t}(1) = p_{s,t}(B|I^{s,t}) = 1 - p_{s,t}(F|I^{s,t})$. The tumor probabilities $p_{s,t}(F|I^{s,t})$ were calculated from image volumes I using the generative model [15], and we identified tumor subclasses with $p(F|I^{s=T1,t})$ for *core*, and $p(F|I^{s=T2,t})$ with *edema*. We modeled the 3D spatial constraints through a 26 neighborhood ($\mathcal{N}$) linking the central voxel with all its immediate neighbors. Interaction terms $W_{i,j}^{s,t}(L_i^{s,t}, L_j^{s,t})$ between neighboring voxels in each sequence $s \in [T1, T2]$ are computed as

$$\delta_{L_i \neq L_j}\, \beta\, \frac{\alpha(i,j)}{\alpha_{\text{tot}}}\, \exp\left(-\left(\frac{I^{s,t}(i) - I^{s,t}(j)}{A}\right)^2\right)$$

with $\beta = 0.5$, $\alpha(p,q) = \frac{1}{\text{distance}(p,q)}$, $\alpha_{\text{tot}} = \sum_{q \in \mathcal{N}(\text{pixel } p)} \alpha(p,q)$ and $A = \frac{1}{3}(\max I^{s,t} - \min I^{s,t})$. We impose growth constraint in 3D+t as explained in Sec. 2.2, and inclusion constraints as in Sec. 2.3 : the foregrounds in T1 and T1c modalities are required to be included in the one of T2, which is included in the one of FLAIR.

In our test we first segment each image at each time independently by graph cut, and calculate average Dice scores for each image series as a baseline. Then we test different regularizations, i.e., *[w/o]*, *[Mono]* and *[Bi]*. Fig. 6(left) shows results for all ten time series. We find that a weak temporal regularization (both bi-/monodirectional with $w \ll 1$) outperforms a segmentation without temporal constraint. Increasing the bidirectional temporal regularization towards high values decreases the performance, while introducing monodirectional "growth" regularization through infinite links improves performance (see Fig. 5). Moreover, the log(volume)-time graph (Fig. 6) shows the exponential growth pattern that is associated with this disease [1]. It can be further analyzed, for diagnosis and treatment monitoring, *e.g.* through algorithms estimating the speed of the tumor outlines under anatomical constraints [14]. Extension of the current 5D segmentation could integrate this speed estimation, or extend the multimodal tumor segmentation [15] for longitudinal data sets.

4 Conclusions and future work

We addressed the problem of shape segmentation in 2D and 3D image time series, where shapes can only grow/shrink. In order to enforce shape growth, we proposed a new graph cut-based method for computing the globally optimal joint segmentation of an image sequence. The main idea was to introduce directed infinite links between pixels at the same spatial locations in successive image frames, which prohibit a shape to shrink or grow over time, and to perform a graph cut optimization on the constructed graph. We also demonstrated the possibility to impose inter-sequences object inclusion constraint by adding directed infinite links to the joint graph associated to all sequences. We validated the performance of the proposed approach for the segmentation of a shrinking ice floe and of growing tumor volumes from multimodal sequences of satellite and 3D MR images, respectively. The new method proved to be robust to very noisy or low-contrast images, and showed linear complexity.

We plan to apply the proposed framework for other applications, such as the segmentation of fires from satellite surface temperature measurements. We are also currently extending our work on tumor segmentation to the case of several medically-motivated classes C_i, expressing tumor evolution stages, that satisfy an inclusion constraint $C_i \subset C_{i+1}, \forall i$, considering additional relations between tumor substructures.

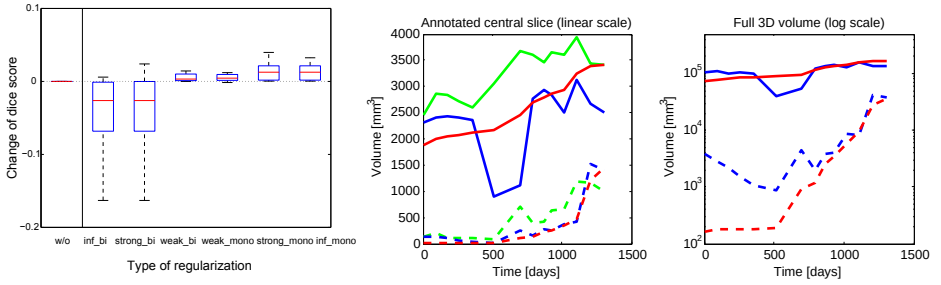

Figure 6: (Left) Boxplot changes in the average segmentation performance for the ten image sequences when testing different regularization approaches. Using a strong monodirectional regularization acting as a growth prior yields the best results. (Right) Volume-time plot for a patient with 14 observations. Solid lines indicate *edema*, dashed indicate tumor *core* that starts growing with constant rate at around day 500. The segmentation with growth constraint (red) returns results comparable to the manual segmentation (green). Segmentations obtained by evaluating image volumes of each time point individually (blue) show significant variation, even obscuring the overall trend.

References

[1] F. G. Blankenberg, R. L. Teplitz, W. Ellis, M. S. Salamat, B. H. Min, L. Hall, D. B. Boothroyd, I. M. Johnstone, and D. R. Enzmann. The influence of volumetric tumor doubling time, DNA ploidy, and histologic grade on the survival of patients with intracranial astrocytomas. *Am J Neuroradiol*, 16:1001–1012, 1995.

[2] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *TPAMI*, 26(9):1124–1137, 2004.

[3] J. Chen, S. Paris, and F. Durand. Real-time edge-aware image processing with the bilateral grid. In *SIGGRAPH*, 2007.

[4] G. B. Dantzig and D. R. Fulkerson. On the max-flow min-cut theorem of networks. *Ann. Math. Studies*, (38), 1956.

[5] A. Delong and Y. Boykov. A scalable graph-cut algorithm for n-d grids. In *CVPR*, 2008.

[6] A. Delong and Y. Boykov. Globally optimal segmentation of multi-region objects. In *ICCV*, 2009.

[7] D. Dementhon. Spatio-temporal segmentation of video by hierarchical mean shift analysis. In *Statistical Methods in Video Processing Workshop (SMVP)*, 2002.

[8] L. Dice. Measure of the amount of ecological association between species. *Ecology*, 26(3):297–302, 1945.

[9] A. V. Goldberg and R. E. Tarjan. A new approach to the maximum-flow problem. *J. ACM*, (35(4)), 1988.

[10] M. Grundmann, V. Kwatra, M. Han, and I. Essa. Efficient hierarchical graph based video segmentation. In *CVPR*, 2010.

[11] D. Haverkamp and C. Tsatsoulis. Using temporal information in an automated classification of summer, marginal ice zone imagery. In *IGARSS'96*, pages 109–111, May 1996.

[12] H. Ishikawa. Exact optimization for Markov random fields with convex priors. *TPAMI*, 25, October 2003.

[13] M. Jenkinson and S. Smith. A global optimisation method for robust affine registration of brain images. *Med. image analys.*, 5(2):143–156, 2001.

[14] E. Konukoglu, O. Clatz, B. H. Menze, B. Stieltjes, M. Weber, E. Mandonnet, H. Delingette, and N. Ayache. Image guided personalization of reaction-diffusion type tumor growth models using modified anisotropic eikonal equations. *IEEE Trans Med Imaging*, 29:77–95, 2010. ISSN 1558-0062.

[15] B. H. Menze, K. Van Leemput, D. Lashkari, M.-A. Weber, N. Ayache, and P. Golland. A generative model for brain tumor segmentation in multi-modal images. In *MICCAI*, pages 151–159, 2010.

[16] S. Paris. Edge-preserving smoothing and mean-shift segmentation of video streams. In *ECCV*, 2008.

[17] T. Riklin-Raviv, K. V. Leemput, B. H. Menze, W. M. Wells, and P. Golland. Segmentation of image ensembles via latent atlases. *Med. image analys.*, 14(5):654–665, Oct. 2010.

[18] K. R. Swanson, R. C. Rostomily, and E. C. Alvord. A mathematical modelling tool for predicting survival of individual patients following resection of glioblastoma: a proof of principle. *Br J Cancer*, (98):113–119, 2008.

[19] Y. Tarabalka, L. Brucker, A. Ivanoff, and J. C. Tilton. Shape-constrained segmentation approach for arctic multiyear sea ice floe analysis. In *IGARSS*, 2012.

[20] R. Wolz, R. A. Heckemann, P. Aljabar, J. V. Hajnal, A. Hammers, J. Lotjonen, and D. Rueckert. Measurement of hippocampal atrophy using 4d graph-cut segmentation: Application to ADNI. *NeuroImage*, 52(1):109–118, 2010.