Fully Automated Brain Tumor Segmentation using two MRI Modalities

Mohamed Ben Salah[1], Idanis Diaz[1], Russell Greiner[1], Pierre Boulanger[1], Bret Hoehn[1], and Albert Murtha[2]

[1] Department of Computing Science
[2] Radiation Oncology Department
University of Alberta

Abstract. An algorithm is presented for fully automated brain tumor segmentation from only two magnetic resonance image modalities. The technique is based on three steps: (1) alternating different levels of automatic histogram-based multi-thresholding step, (2) performing an effective and fully automated procedure for skull-stripping by evolving deformable contours, and (3) segmenting both Gross Tumor Volume and edema. The method is tested using 19 hand-segmented real tumors which shows very accurate results in comparison to a very recent method (STS) in terms of the Dice coefficient. Improvements of 5% and 20% respectively for segmentation of edema and Gross Tumor Volume have been recorded.

1 Introduction

A fundamental problem in computer vision, image segmentation has been the focus of a large number of theoretic and practical studies [11]. Medical image segmentation has received a significant attention as well, due to the diverse practical applications that segmentation has. These include image-guided surgery, enhanced visualization, consistent volume measurements, and change detection in images with time. In the last decade, an increasing research interest and effort have been deployed in order to tackle various medical problems using specific image-based modalities such as brain magnetic resonance (MR) images. In this paper, we address the problem of automatically segmenting brain tumors and associated edema in MR images [3].

In practice, practitioners currently rely on experts manual delineations. Unfortunately, manual segmentation of tumor volumes through all images is prohibitively time consuming and is subject to manual variation. For instance, the new technologies for radiation therapy need very accurate automatic segmentations whereas manual delineations are known to be patient-specific and might have high variability (both between observers and within the same observer). Automatic methods do not have this variability and, as a result, any significant difference in the output (tumor shape, size, volume, etc) could be easily interpreted and assessed [2]. In addition, precise automatic segmentation methods could be used in other interesting new applications such as image retrieval

in large medical databases which enormously helps clinicians in making important decisions (e.g. tumor grading, make patient prognosis, etc). Although tumor segmentation represents an important problem and many semi-automatic approaches have been proposed, few of these meet the basic requirements in practice.

In clinical practices, any automatic approach is required to fulfill some standard features such as accuracy, speed, and minimal user intervention. Despite the extensive research focusing on brain tumor segmentation in MR images, automatic approaches with an accuracy comparable to human experts is still far from reach [6, 7]. The reasons which make this problem challenging are mainly related to two aspects: the MR imaging modalities at hand and the inherited properties from the adopted low-level imaging techniques. First, MR modalities are typically corrupted with local noise, suffer from partial volume artifacts, intensity inhomogeneity (within the set of slices at hand) and inter-slice intensity variations. Second, the elementary image processing techniques, on which a large number of the proposed semi-automatic segmentation methods are based, present various inherited weaknesses. For instance *thresholding*, typically used as an initial step for creating a rough binary segmentation of the considered image, does not take into account the spatial features which generally characterize well brain tumors from other components (connectivity property). *Edge detection* techniques are also of great importance for localization of tumors in MR images. These methods are clearly justified for segmenting objects which have a distinctive photometric profile from their surroundings. This is not always the case for brain tumor segmentation because the tumor's boundaries include some spots where the intensity gradient fades and also some neighboring tissues have very close photometric profiles. *Region growing* is also used for extracting connected regions based on some intensity-based criteria. This technique is sensitive to intensity inhomogeneity as well but, more importantly, requires at least one seed point that is manually selected by an operator. Thus, these techniques are seldom used alone but within a set of consecutive operations and generally require a lot of postprocessing. The purpose of this work is to provide a fully automatic method for segmenting brain tumor where there is no need for user intervention. The fully automatic segmentation of the enhancing tumor region has been investigated before as a simplified way to define abnormality in brain MR images. In this work we tackle the more challenging task of segmenting the Gross Tumor Volume (GTV) as well as the full tumor and edema area.

The closest work to our approach is [1] which finds edema and GTV based on an automatic thresholding process followed by some morphological operations. However, our approach diverges from [1] in two important aspects. First, we use only two standard clinical MRI modalities: T1-weighted with gadolinium contrast agent (T1C) and Fluid Attenuated Inversion Recovery (FLAIR) contrarily to [1] which uses, in addition to T1C and FLAIR, T1-weighted (T1) and T2-weighted (T2). This is a major advantage because acquiring these MRI modalities is both time consuming and costs a lot of money. Second, we use active contours in order to remove the skull which is conceptually different from

the skull stripping process in [1] which is based on thresholding and a succession of morphological steps. In fact, skull stripping is not specific for the problem of tumor detection but is rather an indispensable tool for large-scale studies. A precise and automatic tool is highly wanted (preferably with no user intervention) because any inaccuracy in the skull-stripping step systematically and negatively influences the following, yet crucial, processes such as cortical thickness estimation, brain atrophy estimation, volumetric analysis, and tumor detection and growth prediction. The skull stripping methods reported in the literature can roughly be classified into three categories: 1) morphological operation based methods, 2) deformable surface based methods, and 3) hybrid methods [8, 9].

Morphological operation based methods extract brain via applying a series of thresholding and morphological dilation and erosion. Generally, these operations are repeated until the brain is isolated from the extracranial tissue based on certain criteria. Human intervention is often required due the aforementioned intensity inhomogeneity and also the apparent connectiveness of the brain tissue and the skull in certain cases.

Deformable surface based methods mainly rely on image gradient and surface internal forces to move the deformable surface towards regions of interest, the brain tumor in this case. The image gradient information is typically extracted from the MRI modalities used for the task at hand which are T1C and FLAIR in our case. In addition, internal forces are used to guide the evolution of the deformable surfaces and impose on them certain topological constraints such as smoothness, convexity or shape prior. Criticisms of this type of approaches include that the initial surface position should be close enough to the tumor volume in order to avoid local minima and failure to converge [8, 9].

Hybrid methods combine more than one of the aforementioned skull stripping techniques for better results. The effectiveness of such approaches would depend heavily on the component methods.

Our skull stripping component adopts the surface-based method formulated in the variational framework. In this framework, segmentation is formulated as an optimization problem of a given objective functional which embeds the different constraints of the problem at hand. Variational formulations are known to be more principled, easily generalized and especially more flexible. Thus, we propose in this paper an approach, operating on scalar 3D MRI data of the brain, which locates both edema and GTV using only two MRI modalities: T1C and FLAIR. The approach is based on three main steps: 1) alternation of different levels of automatic histogram-based multi-thresholding procedure coupled with sequences of morphological operations, 2) an effective and completely automatic procedure for skull stripping by evolving deformable contours, and 3) a GTV and edema final segmentation step.

2 Method

The algorithm is designed to operate on two registered 3D volumes of the same patient compound of two MR sequences T1C and FLAIR. The first step, thresh-

olding and morphological operations, produces initial masks of GTV, edema and skull. The skull mask together with the original T1C and FLAIR are fed to the proposed skull-stripping algorithm which refines the brain extraction task and produces a final skull mask. After removing the skull, the last step consists in segmenting GTV and edema. In the following, we explain in details each one of these steps.

2.1 Histogram-based operations

To provide an automatic segmentor and before applying a histogram-based technique, one should verify that some basic assumptions hold. First, the radiological images should contain most of the information necessary for identifying the abnormal anatomical regions. Second, the assumption stating that a good contrast between structures of interest and the surrounding structures should hold. In our case, the structures of interest are mainly edema and GTV; then with less importance the skull. Fortunately, histograms of brain MRI scans present a typical shape that allows thresholding using parameters that would not change considerably from one patient to another. Indeed, it turns out that brain MRI histograms are typically bimodal: a first well-defined mode which represents the most common intensity values corresponding to the image background (values close to zero) and a second mode which includes all grayish values corresponding to brain tissues like white and gray matter (refer to Fig. 1). In our case, for each patient two MRI modalities (T1C and FLAIR) are used to extract the two corresponding histograms. To simplify the modes' (peaks of the histogram) search, we smooth the histograms using Savitsky-Golay FIR filter [5][3], and then localize the two modes (μ_1 and μ_2 in Fig. 1) following the method in [4]. Generally, edema region is brighter (higher intensity) on FLAIR than on T1C and the gadolinium-enhanced lesion has the opposite behavior. The skull has typically high intensity values on both T1C and FLAIR. Based on these observations, we determine three thresholds (τ_{Fgr} τ_A and τ_B) that would distinguish the different anatomical regions within the brain.

For each modality $m \in \{$ T1C, FLAIR$\}$, let's define the three thresholds (refer to Fig. 1)

$$\tau_i(m),\ i \in \{Fgr, A, B\}, \tag{1}$$

identified as the inflection points (slope changes sign) of the smoothed histogram located on the right-hand side of μ_2 (τ_A and τ_B) and on the left-hand side of μ_2 (τ_{Fgr}). The threshold τ_{Fgr} is extracted from FLAIR histogram and serves to select voxels belonging to the brain tissue and skull[4] (no ventricles, sinuses and sulci). The threshold τ_A serves to extract skull from T1C and FLAIR and gadolinium-enhanced lesions on T1C. Finally, τ_B is applied to FLAIR in order to extract skull and, more importantly, edema. Fig. 2 shows an example of the

[3] This filtering process preserves higher-order moments while approximating the data within a window to a higher-order polynomial using a least-square procedure.

[4] The skull extracted in this step is a rough mask. The skull-stripping method we detail in the next section provides the refined and final skull mask.

results obtained by applying these thresholds on a T1C and FLAIR scans by respectively using τ_A and τ_B.

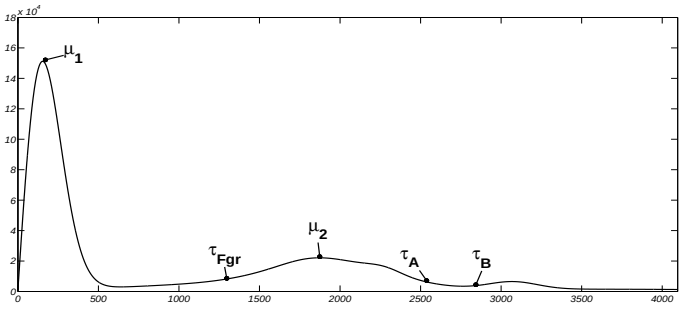

Fig. 1: Smoothed histogram of brain MRI: bi-modal structure with modes μ_1 and μ_2 and multiple thresholds: τ_{Fgr}, τ_A and τ_B.

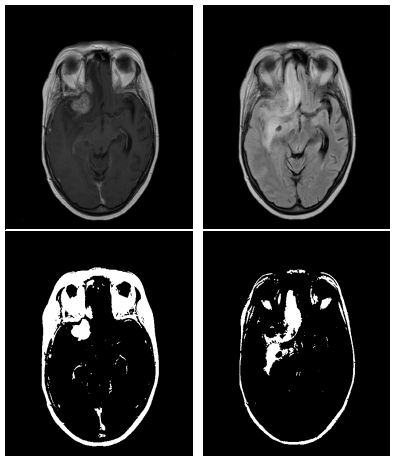

Fig. 2: Thresholding effect on one slice. Top row from left to right: T1C, FLAIR. Bottom row from left to right: thresholding results for τ_A(T1C) and τ_B(FLAIR).

2.2 Skull-stripping

The skull-stripping step is very important and it influences heavily the results in some real cases (refer to section 3). It is very important to have an automatic and yet very accurate skull-stripping component in order to expect acceptable

tumor segmentation results. We formulate our skull-stripping technique using the variational framework as an optimization problem and implement it via active contours and level sets.

Let $I : \mathbb{R}^3 \rightarrow \mathbb{R}$ be the image function (mapping voxels in 3D space to scalar gray levels (T1C and FLAIR)). Given the assumption that the objects of interest can be characterized by their boundaries, we explicitly define an external energy whose minimum corresponds to objects boundaries (or regions of high intensity gradient) based on this function $g = \frac{1}{1+|\nabla G_\sigma * I|^2}$, where G_σ is the Gaussian kernel with standard deviation σ and g is called the *stopping function*. In level sets formulation of active contours, we evolve a contour $C(t) = \{(x, y)|\Phi(t, x, y) = 0\}$ (level zero of the higher dimensional function Φ) towards the boundaries of the object of interest. The evolution equations of Φ arise from this general form

$$\frac{\partial \Phi}{\partial t} + F|\nabla \Phi| = 0, \tag{2}$$

where F is called the speed function and is obtained from the minimization of the objective functional which describes the problem at hand. In this case, we consider the following objective functional

$$E(\Phi) = P(\Phi) + \alpha A(\Phi) + \beta R(\Phi), \tag{3}$$

where α and β are weighting constants. The term $P(\Phi)$ is defined by

$$P(\Phi) = \int g\delta(\Phi)|\nabla \Phi| d\mathbf{x} \tag{4}$$

and

$$A(\Phi) = \int gH(-\Phi) d\mathbf{x}. \tag{5}$$

δ is the Dirac function and H is the Heaviside function [10]. Minimizing the energy functional $P(\Phi)$ drives the zero level set towards the object boundaries (where the stopping function g is almost null). The energy term $A(\Phi)$ serves to speed up the evolution of the active contours (zero level set of Φ). The energy term $R(\Phi)$ is called the *regularization term* and it is used to impose some prior constraints on the evolving contour (shape priors, smoothness, etc). In this case, $R(\Phi)$ is simply the length of the evolving contour which serves to smooth the contour and bias the segmentation against small disconnected regions.

In this work, we initialize the evolving contour so that it encloses the whole brain region. This could be achieved using the rough foreground mask provided by the previous thresholding step. By initializing close by the external boundary of the brain, we guarantee that the contour converges very fast and also that it does not get stuck at a local minimum. Once the evolving contour reaches the external boundary of the brain, we evolve a second contour starting from the position of the first contour and going inside towards the inner boundary of the skull. This would prevent the evolving contour from reaching the inner brain unwanted components such as the tumor and other brain tissues.

2.3 Edema and GTV segmentation

Let's define the mask functions obtained using the threshold τ at a voxel $\mathbf{x}$ as follows:

$$M(m,\tau)[\mathbf{x}] = \begin{cases} 1 & \text{if } I(m)[\mathbf{x}] > \tau, \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $I(m) : \mathbb{R}^3 \rightarrow \mathbb{R}$ is the 3D image function corresponding to the MRI modality m. Then, $M(T1C, \tau_A(T1C))$ is the mask obtained after applying the threshold τ_A to a MRI modality T1C and it is a binary image where voxels labeled 1 are those whose T1C intensities are greater than $\tau_A(T1C)$. For simplicity, we will refer to this by $M(T1C, \tau_A)$. In the following we first explain how edema is segmented based on the obtained thresholds and then do the same for GTV.

As depicted by Fig. 2, $M(FLAIR, \tau_B)$ and $M(FLAIR, \tau_A)$ represent tumor edema detected in the MRI modality FLAIR at hand. From the same figure, we can notice that thresholding is not sufficient because edema might be less defined in some masks (e.g. $M(FLAIR, \tau_B)$ in comparison to $M(FLAIR, \tau_A)$). For this reason, other morphological operations are typically used such as the geodesic dilation (for this case $M(FLAIR, \tau_B)$ is the marker and $M(FLAIR, \tau_A)$ is the geodesic mask). The output resulting from FLAIR, contains generally edema with few other small regions which have the same intensity profile (recall that any very low intensity is already removed after thresholding with τ_{Fgr}).

Similarly, Fig. 2 shows an example which explains how the GTV is detected based on T1C. This procedure is performed using the threshold $\tau_A(T1C)$ which separates the enhancing rim from the rest of the brain. In case more MRI modalities are available (especially T1), this procedure could be enhanced and better detection could be achieved [1]. Finally, we can notice the presence of the skull in all the obtained masks (refer to Fig. 2) which emphasizes the importance of the skull-stripping step. Skull has a photometric profile similar to edema and GTV and is present in T1C and FLAIR. Thus, the accuracy of the skull-stripping technique influences systematically and heavily the tumor segmentation performance.

3 Experimental results

The proposed approach was evaluated based on a real dataset of patients having glioblastoma (the most common and most aggressive malignant primary brain tumor in humans) at different stages treated at the Cross Cancer Institute (CCI), Alberta. Our dataset contains nineteen patient cases, each of which has 2 sequences (between 21 and 25 slices of T1C and FLAIR) acquired with a 1.5T MR Phillips Intera Achieva scanner. All images were resampled to be of dimensions 512×512 and resolution $1 \times 1 \times 5$ mm^3.

The proposed method has been compared to the method in [1] (referred to by STS) when the input is similar to ours (T1C and FLAIR). We also considered STS when taking as input four modalities (T1, T1C, FLAIR and T2) in order to asses the behavior of the proposed method even though we are given less input information. To evaluate the final results, we were given manual delineations of edema and GTV performed by experts and we adopted the Dice metric as a similarity measure between the automatic segmentation and the manual delineation. Let P and Q represent the automatically detected tumor and the manually delineated tumor, respectively. The Dice coefficient is defined as follows:

$$D(P,Q) = \frac{2 \times TP}{2 \times TP + FP + FN}, \tag{7}$$

where TP, FP, and FN are the true positive, false positive and false negative voxels.

In Figs. 3 and 4, we report the obtained results for our method (in terms of Dice coefficient) against STS (using two modalities and 4 modalities). Using four modalities is obviously better (bias in favor of STS) but we include this case for two reasons: 1) explore how much STS looses in accuracy by considering only two modalities and 2) compare to STS when it is given only two modalities (same input we are given). In comparison to STS when it is given two modalities, our method behaves better for segmenting both edema and GTV. The blue bars in Figs. 3 and 4 correspond to STS taking four modalities as input (this obviously would behave better than with only two modalities), the red bars correspond to STS given two modalities (the same we are given, T1C and FLAIR) and the green bars correspond to our method. Overall, we recorded a mean Dice coefficient of 0.76 ($\pm$0.09) and 0.64 ($\pm$0.25) against 0.72 ($\pm$0.12) and 0.53 ($\pm$0.26) respectively for the segmentation of edema and GTV. This corresponds to an improvement of 5% and 20% respectively in the segmentation of edema and GTV. From Fig. 4, one can notice that over the nineteen cases, the proposed method (green bars) performs better (or is worst cases comparable results) than the STS given two modalities (red bars). Interestingly, we perform quite good especially in the hard cases where STS with two modalities fails compared to the same method with four modalities (STS 4Mod). Also, it is worth mentioning that we even outperform STS with four modalities in few cases. This can be explained by the fact that our skull-stripping method is behaving very well which allows our method to compensate the lack of input information. From Fig. 3, similar behavior is recorded except for one case where STS with two modalities performs slightly better than us. The cases where the Dice coefficient is zero (or almost zero) correspond to MRI volumes which do not present any GTV.

In order to have a visual idea about the obtained results, we show a sample of the obtained results (GTV and edema) as depicted in Fig. 5. The top row images (of the left half) are the original T1C and FLAIR inputs fed to the proposed method and to STS (2Mod). The second row (of the left half) in Fig. 5 shows the segmentation results obtained by the proposed method where the GTV is colored in red and edema in blue. This is one of the challenging cases (check

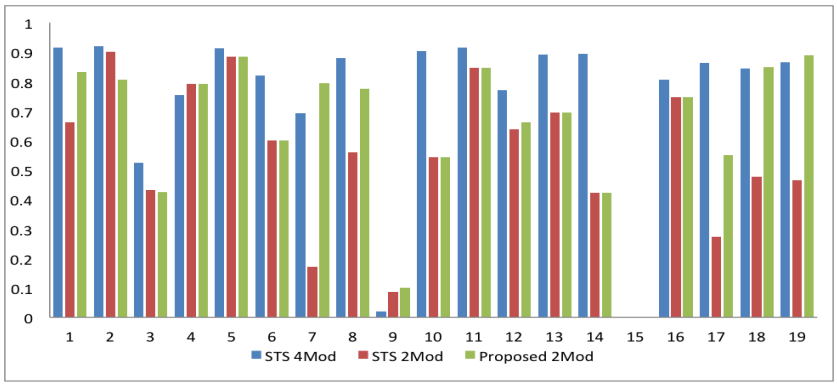

Fig. 3: Dice coefficients obtained for GTV segmentations.

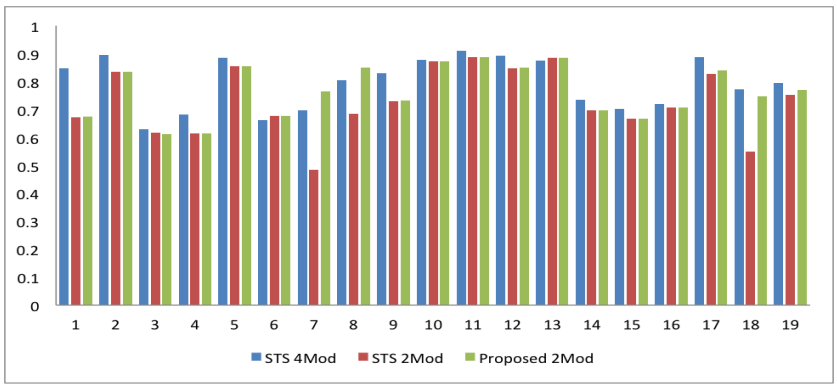

Fig. 4: Dice coefficients obtained for edema segmentations.

case 7 in Figs. 3 and 4) because the tumor is very close to the brain boundaries and is touching the skull. The STS failed dramatically in segmenting the tumor mainly because the skull-stripping method has considered part of the tumor as skull which affected the final results. When four modalities are given to STS, the results were better but still not as good as the results we obtained. This example highlights the importance of the skull-stripping step.

References

1. I. Diaz, P. Boulanger, Russell Greiner, B. Hoehn, L. Rowe, and A. Murtha, "An Automatic Brain Tumor Segmentation Tool," 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka, Japan, July 3-7, 2013.
2. Mazzara, G. P. and Velthuizen, R. P. and Pearlman, J. L. and Greenberg, H. M. and Wagner, H., "Brain tumor target volume determination for radiation treatment

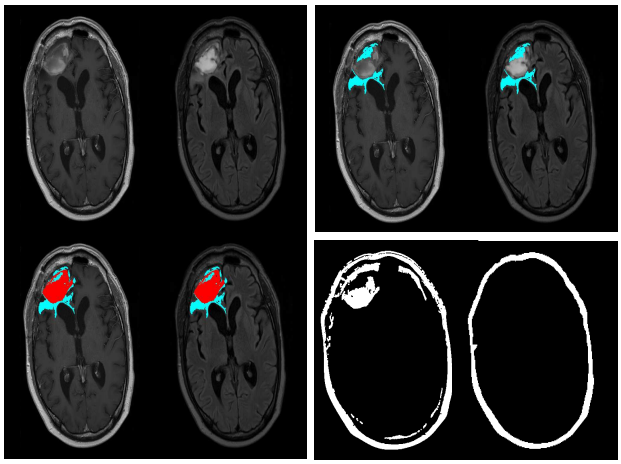

Fig. 5: Segmentation results on a challenging case. Left half: the top row depicts the original T1C and FLAIR; second row: segmentation results using our method. Right half: the top row: segmentation results using STS (2Mod); second row: the extracted skull by STS and the proposed method. Colors: GTV in red and edema in blue.

planning through automated MRI segmentation," Int J Radiat Oncol Biol Phys, vol. 59, 2004.
3. M. Schmidt, "Automatic brain tumor segmentation," M.sc. thesis, University of Alberta, 2005.
4. M. E. Brummer, R. M. Mersereau, R. L. Eisner, R. J. R. "Automatic detection of brain contours in MRI data sets IEEE Transactions on Medical Imaging," Vol. 12, No. 2. (June 1993), pp. 153-166.
5. A. Savitzky and M. Golay, "Smoothing and differentiation of data by simplified least squares procedures," Anal. Chemm. vol. 36, 1964.
6. D. Pham, C. Xu, and J. Prince, "Current methods in medical image segmentation," Annu. Rev. Biomed. Eng., vol. 2, 2000.
7. I. Bankman, Handbook of Medical image: processing and Analysis, 2008
8. Y. Wang, J. Nie, P.-T. Yap, F. Shi, L. Guo, and D. Shen, "Robust deformable-surface-based skull-stripping for large-scale studies," MICCAI, 2011.
9. X. Tao, and M.-C. Chang, "A skull stripping method using deformable surface and tissue classification," SPIE, 2010.
10. C. Li, C. Xu, C. Gui, and M. D. Fox, "Level Set Evolution without Re-Initialization: A New Variational Formulation," CVPR, 2005.
11. M. Ben Salah, A. Mitiche and I. Ben Ayed, "A continuous labeling for multiphase graph cut image partitioning," Adv. in Vis. Comp., LNCS, G. Bebis et al. (Eds.), Vol. 5358, pp. 268-277, Springer-Verlag, 2008.