A Shape Prior-Based MRF Model for 3D Masseter Muscle Segmentation

Tahir Majeed[a], Ketut Fundana[a], Marcel Lüthi[a], Jörg Beinemann[b], Philippe Cattin[a]

[a] University of Basel, Basel, Switzerland
[b] University Hospital, Basel, Switzerland

ABSTRACT

Medical image segmentation is generally an ill-posed problem that can only be solved by incorporating prior knowledge. The ambiguities arise due to the presence of noise, weak edges, imaging artifacts, inhomogeneous interior and adjacent anatomical structures having similar intensity profile as the target structure. In this paper we propose a novel approach to segment the masseter muscle using the graph-cut incorporating additional 3D shape priors in CT datasets, which is robust to noise; artifacts; and shape deformations. The main contribution of this paper is in translating the 3D shape knowledge into both unary and pairwise potentials of the Markov Random Field (MRF). The segmentation task is casted as a Maximum-A-Posteriori (MAP) estimation of the MRF. Graph-cut is then used to obtain the global minimum which results in the segmentation of the masseter muscle. The method is tested on 21 CT datasets of the masseter muscle, which are noisy with almost all possessing mild to severe imaging artifacts such as high-density artifacts caused by e.g. the very common dental fillings and dental implants. We show that the proposed technique produces clinically acceptable results to the challenging problem of muscle segmentation, and further provide a quantitative and qualitative comparison with other methods. We statistically show that adding additional shape prior into both unary and pairwise potentials can increase the robustness of the proposed method in noisy datasets.

Keywords: MRF, Markov Random, Graph-Cut, Masseter, Muscle, Segmentation, Max-Flow Min-Cut, MAP

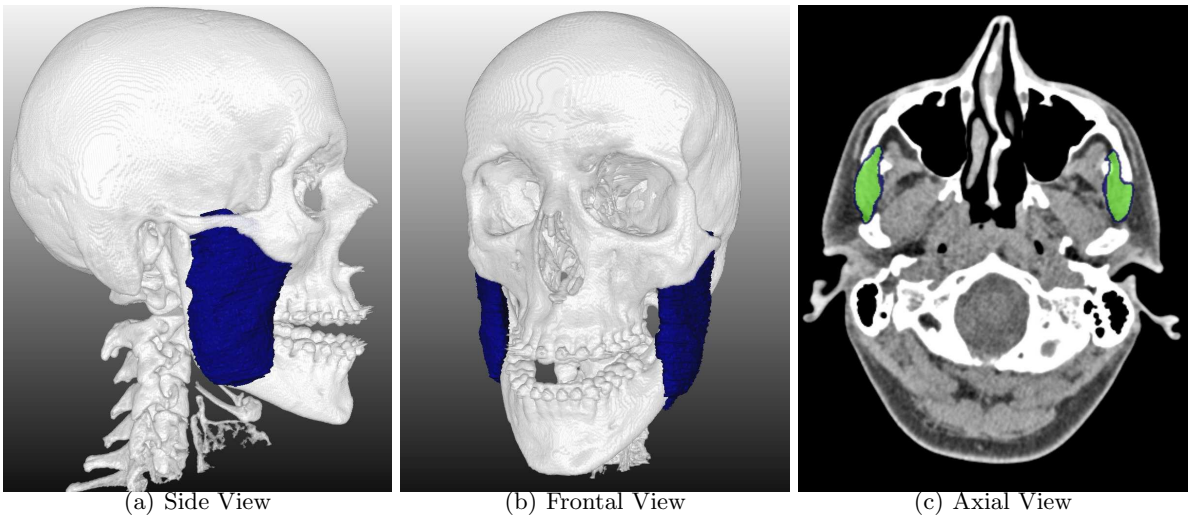
1. INTRODUCTION

Face being the unique external perceivable identity of a human is of fundamental importance for them. It represents the individual personality and traits in daily contact with other humans. Underneath the facial skin is the musculo-skeletal system consisting of bones and soft tissues. Bones and soft tissues are of major importance for any facial surgical treatment to guarantee an optimal, functional and aesthetical outcome. Facial surgeries require accurate planning based on available preoperative diagnostic datasets. The current clinical practice is to segment the facial muscles manually which is a tedious and time consuming task and also increases the overall cost of the surgery.

Image segmentation is an ill-posed problem, therefore, additional constraints are needed in order to constrain the solution space and achieve the desired solution. Due to the presence of noise, weak edges, imaging artifacts and inhomogeneous interior, the use of image information alone often gives unsatisfactory segmentation results. In order to overcome this problem, prior knowledge about the shape and location of the target structure can be used to achieve desired segmentation results. The biggest hinderance, however, lies in unifying diverse prior information from different domains into a single framework.

A wide variety of methods has been proposed over the last two decades. Of all these methods MRF models provide a harmonious way to incorporate different types of constraints into the segmentation problem. Although MRF holds a lot of promise due to their flexibility in combining information from different domains elegantly, they did not receive much attention due to the lack of efficient discrete optimization techniques. With the introduction of graph-cut by Boykov and Jolly,[1] relatively dormant MRF based models came back strongly.

Further author information: (Send correspondence to Tahir Majeed)
Tahir Majeed: E-mail: tahir.majeed@unibas.ch, Telephone: +41 61 265 9658
Ketut Fundana: E-mail: ketut.fundana@unibas.ch, Telephone: +41 61 265 9658
Philippe Cattin: E-mail: philippe.cattin@unibas.ch, Telephone: +41 61 265 9655



(a) Side View (b) Frontal View (c) Axial View

Figure 1. (a,b) 3D and (c) 2D views of the masseter muscle (color online).

Traditional graph-cut approaches are very successful at finding the global optimum solution in cases where the object of interest can be distinguished from its adjacent structures. However, they fail where the adjacent structures are very similar in appearance to the target structure since the edges are weak.

Many researchers have tried to incorporate prior shape information into MRF based energy functionals, such as in.[2–11] It is, however, very difficult to incorporate prior shape knowledge into graph-cut based approaches. The approaches of Slabaugh and Unal[6] and Zhu-Jacquot,[8] tried to spatially constrain the graph-cut segmentation by incorporating parametric shape information. The parametric techniques, however, have the drawback that only shapes that can be roughly represented geometrically can be segmented, which is often not the case. Freedman and Zhang[2] used unsigned distance maps of the shape template where the 0-level set defines the shape prior contour. They encoded this prior shape knowledge in the second order clique potential of the MRF. Xu *et al.*[12] used a combination of active contours and graph-cut optimization to iteratively deform the contour. Statistically learned shape knowledge can be incorporated in the energy function as shown by El-Zehiry and Elmaghraby[7] and Malcolm *et al.*[5] Malcolm *et al.*[5] incorporate prior shape knowledge through nonlinear shape priors. They learn the prior shape knowledge through statistical learning and use an iterative graph-cut approach to segment objects.

In this paper we propose a novel segmentation approach integrating shape priors from 3D CT datasets into graph-cut. We show the applicability of the approach on the difficult problem of segmenting the masseter muscle Fig. 1.

2. METHOD

2.1 Segmentation Model

We formulate our muscle segmentation problem by minimizing the energy functional

$$E(\mathbf{z}|\mathbf{I},\mathbf{S}) = E^I(\mathbf{z}|\mathbf{I}) + E^S(\mathbf{z}|\mathbf{S}), \tag{1}$$

where $\mathbf{z}$ is a binary variable which defines the segmentation, $\mathbf{I}$ is the observed image data and $\mathbf{S}$ is the shape prior. Equation (1) consists of an intensity term $E^I(\mathbf{z}|\mathbf{I})$ and a shape prior term $E^S(\mathbf{z}|\mathbf{S})$ which are based on the intensity and shape information respectively. Both the energy terms $E^I(\mathbf{z}|\mathbf{I})$ and $E^S(\mathbf{z}|\mathbf{S})$ consists of the first and the second order clique potential functions of the Gibbs energy function.

2.2 Intensity Energy Function

The intensity information of the target anatomical structure is modeled by the energy functional $E^I(\mathbf{z}|\mathbf{I})$ which is based on the traditional graph-cut intensity based energy function of Boykov and Jolly[1] and is given by

$$E^I(\mathbf{z}|\mathbf{I}) = (1-\lambda)(1-\mu)\sum_{p\in\mathcal{P}} V_p^I(z_p) + \lambda(1-\gamma)\sum_{p\in\mathcal{P}}\sum_{q\in\mathcal{N}_p} V_{pq}^I(z_p, z_q), \tag{2}$$

where λ is the smoothness parameter, μ and γ are the shape prior parameters for data and smoothness terms respectively. $\mathcal{P}$ is the set of voxels, and $\mathcal{N}_p$ is a 6-neighborhood system for $\mathcal{P}$. The data term $V_p^I(z_p)$ encodes regional properties of the foreground and background. It defines the individual penalties for assigning voxel p to foreground and background which are calculated from the respective intensity histograms. The smoothness term $V_{pq}^I(z_p, z_q)$ encodes the boundary properties of the foreground. It defines a discontinuity penalty between adjacent voxels p and q when they are assigned different labels.

2.3 Shape Prior Energy Function

The prior shape information of the masseter muscle is encoded into the data and the smoothness terms of the energy function $E^S(\mathbf{z}|\mathbf{S})$ given by Eq. (3). Such encoding of the global shape information transforms it into the local constraints:

$$E^S(\mathbf{z}|\mathbf{S}) = (1-\lambda)\mu\sum_{p\in\mathcal{P}} V_p^S(z_p) + \lambda\gamma\sum_{p\in\mathcal{P}}\sum_{q\in\mathcal{N}_p} V_{pq}^S(z_p, z_q). \tag{3}$$

The data term $V_p^S(z_p)$ encodes how likely a particular voxel p is to belong to the foreground "1" and background "0" given the shape prior $\mathbf{S}$. An unsigned distance map of the aligned shape prior is calculated. Based on the unsigned distance map, a probability map for the voxels is calculated.

$$V_p^S(z_p = \text{“1”}) = -\ln Pr(S_p|\text{“1”}). \tag{4}$$

$$V_p^S(z_p = \text{“0”}) = -\ln Pr(S_p|\text{“0”}). \tag{5}$$

Equation (4) and (5) define the penalty voxel p incurs given that it belongs to foreground or background. These penalties are calculated by creating a probability map for the voxels. Two different probability maps, one for foreground and other for background are created. The foreground probability map considers only the voxels enclosed by the shape prior's contour. The voxels that are farthest away from the contour lie at the center of the shape and hence are likely to belong to foreground as compared to the voxels close to the contour, which is computed by Eq. (4) as the negative log-likelihood of belonging to foreground. Thus the voxels at the center of the shape prior incurs higher penalty to belong to the background as compared to voxels close to the contour. Voxels outside the contour incur 0 penalty for belonging to background. The opposite is true for the background probability map, where only the voxels outside the contour are considered. The voxels close to the contour are more likely to belong to the foreground as compared to the voxels farther away. Thus voxels farther away from the contour incur a higher penalty to belong to foreground as compared to voxels close to the contour. Voxels inside the contour incur 0 penalty to belong to the foreground. Freiman *et al.*[4] also creates a probability map by registering binary image segmentations where as our probability map is created from the unsigned distance map of the template shape's contour in 3D. The smootness term $V_{pq}^S(z_p, z_q)$ is similar to that of Freedman and Zhang.[2]

For a facial muscle, a mesh with a representative shape is used as the shape template. The shape template is calculated by averaging manually segmented muscles. It is then semi-automatically initialized in the dataset using landmark registration. Algorithm 1 summarizes the segmentation method.

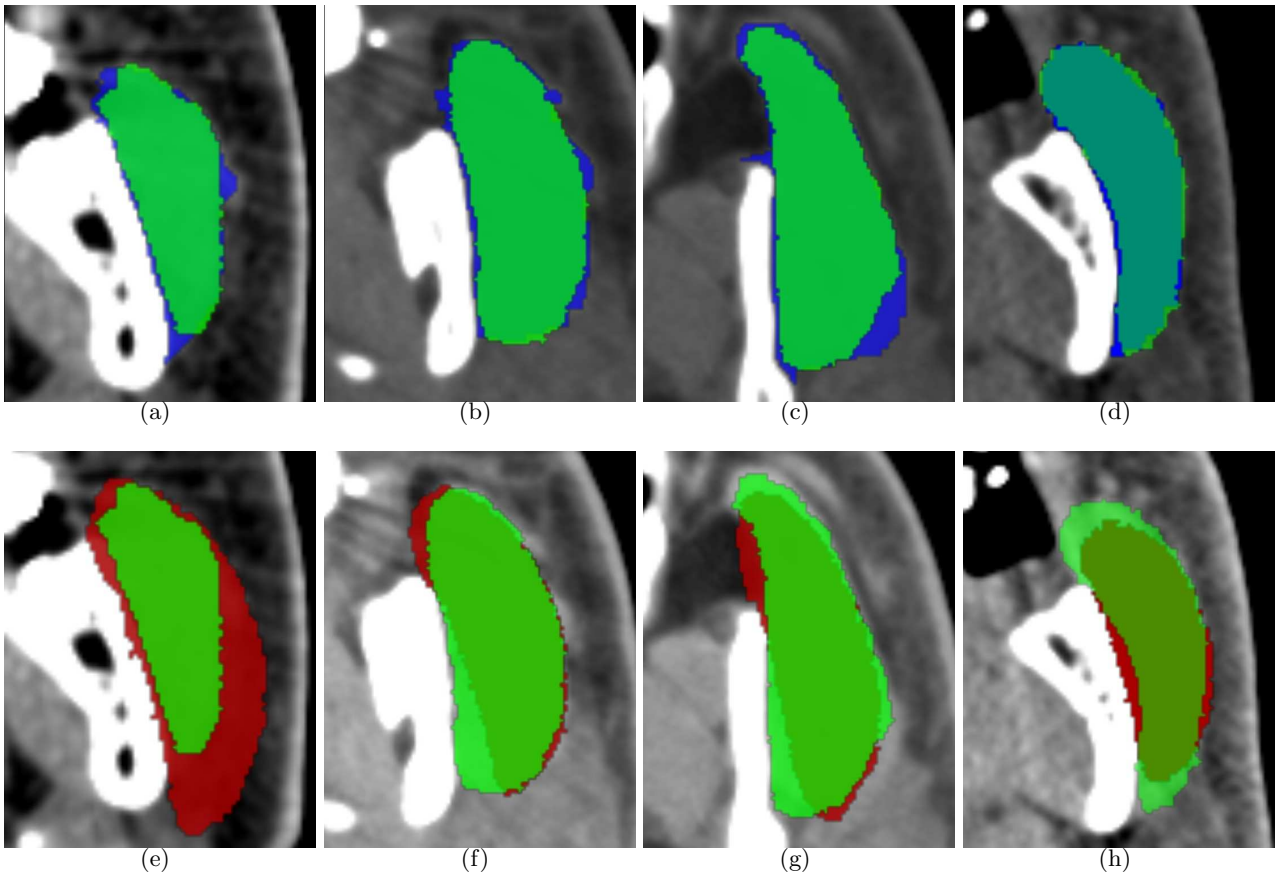


Figure 2. Qualitative segmentation results in 2D. The top row (a..d) shows the segmentation result in blue, while the bottom row (e..h) shows the shape prior in red. The overlayed green in both rows show the ground truth (color online).

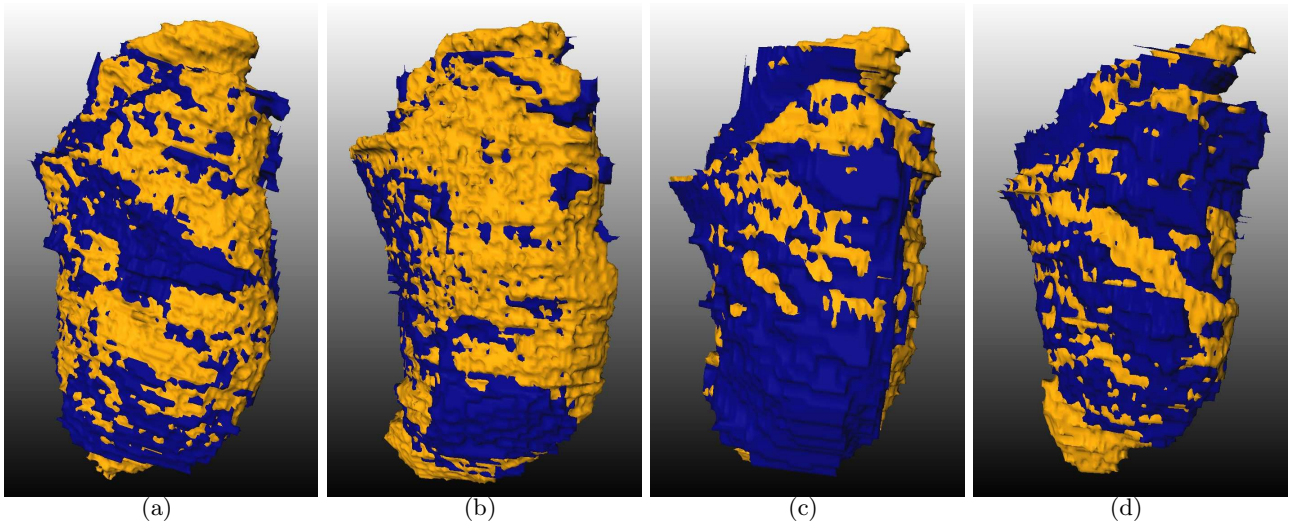


Figure 3. Qualitative segmentation results in 3D of left muscle (a-d) of our method shown in blue and overlayed in orange is the ground truth (color online).

Algorithm 1

Require: An initialized shape template.
Ensure: Graph-Cut segmentation of target muscle.

1. Compute: $V_p^S(z_p)$ as explained in Sec. 2.3.
   (a) Compute: $V_p^S(z_p =$ "1").
   (b) Compute: $V_p^S(z_p =$ "0").
2. Compute: $V_{pq}^S(z_p, z_q)$.
3. Create graph according to the energy function $E(\mathbf{z}|\mathbf{I},\mathbf{S})$.
4. Compute the optimal segmentation using graph-cut.

		Similarity Measure					
Data	Artifact	Dice Coefficient		Sensitivity%		Specificity%	
		Our	Freedman	Our	Freedman	Our	Freedman
1	Mild	0.903	0.872	94.34	96.59	99.01	98.30
2	Mild	0.804	0.841	91.81	91.92	98.02	98.56
3	Mild	0.909	0.897	92.13	96.55	99.46	99.03
4	Mild	0.912	0.865	93.20	93.62	99.53	99.05
5	Mild	0.895	0.909	87.52	92.83	99.59	99.43
6	Mild	0.892	0.864	84.38	92.77	99.58	98.08
7	Mild	0.887	0.838	94.33	93.37	99.41	99.06
8	Mild	0.903	0.922	90.66	96.33	99.37	99.22
9	Mild	0.904	0.875	96.86	97.41	99.32	99.02
10	Mild	0.882	0.916	84.69	93.56	99.65	99.51
11	Severe	0.829	0.283	92.40	94.55	96.71	48.76
12	Severe	0.819	0.799	93.45	92.76	98.68	98.50
13	Severe	0.843	0.804	90.91	78.59	98.65	99.08
14	Severe	0.879	0.569	96.37	97.66	98.50	90.51
15	Severe	0.821	0.781	97.38	96.37	97.11	96.35
16	Severe	0.845	0.823	77.25	81.20	99.55	98.74
17	Severe	0.850	0.831	95.86	88.81	98.91	99.08
18	Severe	0.844	0.771	94.23	85.23	98.40	98.03
19	Severe	0.894	0.375	86.35	96.97	99.38	70.39
20	Severe	0.857	0.509	97.31	96.89	98.05	87.94
21	Severe	0.755	0.595	89.20	96.03	97.73	93.89

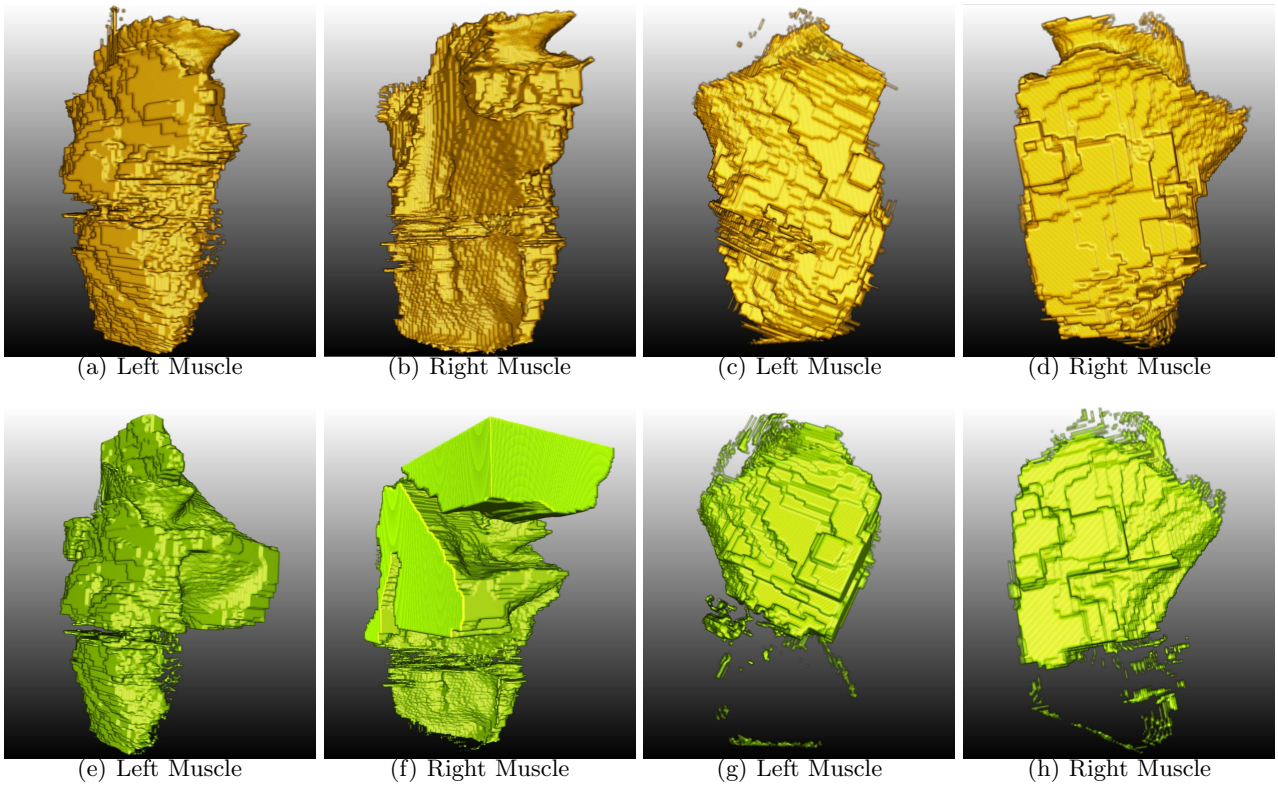
Table 1. Quantitative comparison of our method with that of Freedman and Zhang.[2]

3. EXPERIMENTAL RESULTS

The method was tested on 21 CT datasets of the masseter muscle with dataset dimensions 79-118 × 148-214 × 125-329 voxels and spacing 0.3-0.5 × 0.3-0.5 × 0.3-1 mm^3. The datasets were noisy with almost all possessing mild to severe imaging artifacts such as high-density artifacts caused by *e.g.* the very common dental fillings and dental implants. The datasets were chosen randomly from the hospital repository so that they could represent anatomical variations. The ground truth segmentations were done by a medical expert. Due to the limited number of datasets and ground truth the method was evaluated using a Leave-One-Out approach. The ground truths were used to calculate the mean shape which was then used as the shape template. The parameters were optimized on three datasets and then the same parameters were used on the rest of the datasets. The

parameter values used were $\lambda = 0.3$, $\mu = 0.02$ and $\gamma = 0.004$. The dice coefficient, sensitivity and specificity of the segmentation were calculated as similarity measures to ascertain the accuracy of the method.

Table 1 gives a quantative evaluation of our results and the results of applying Freedman and Zhang[2] to the same datasets for comparison, to which our technique is closely related. Figure 2 shows qualitative results of our technique in 2D while Fig. 3 shows the results in 3D for the datasets without noise and Fig. 4 (a..d) shows our results in 3D over noisy datasets. Figure 4 also shows a qualitative comparison of our method with that of Freedman and Zhang.[2]



(a) Left Muscle (b) Right Muscle (c) Left Muscle (d) Right Muscle

(e) Left Muscle (f) Right Muscle (g) Left Muscle (h) Right Muscle

Figure 4. Qualitative comparison of the segmentation results in 3D of our method with that of Freedman and Zhang[2] over noisy datasets. The top row (a..d) shows our segmentation results, while the bottom row (e..h) shows the results of Freedman and Zhang[2] for the same noisy dataset (color online).

We have also conducted a statistical comparison of our approach with Freedman and Zhang[2] and performed statistical analysis (t-test with signed interval of 0.01). For this we split-up the available datasets depending on the severity of imaging artifacts. Statistical analysis showed, that for the datasets without imaging artifacts both methods, ours and Freedman and Zhang,[2] performed statistically equally well ($p = 0.51$), but when comparing datasets with imaging artifacts our method performed statistically significantly better ($p = 0.009$) as shown in Fig. 4. The figure shows our results from (a..d) over noisy datasets while (e..h) show the results of Freedman and Zhang[2] over the same noisy datasets. This clearly shows the increased robustness gained by including a shape prior in the data term as proposed.

As stated earlier that our contributions is in adding an additional shape prior in the data term while Freedman and Zhang[2] had the shape prior only in the smoothness term. This new shape prior in the energy functional makes it more robust against noisy datasets. The shape prior in Freedman and Zhang[2] has the effect of smoothing out regions which results in a portion of the muscle completely smoothed out. This can be handled by our method that incorporates the shape prior in the data term. This also gives us flexibility to encode shape knowledge in areas where the intensity alone does not provide enough knowledge because of the artifacts. In such regions,

through our shape prior in the data term, we can offset intensity likelihood and say that based on the spatial location of the voxel with respect to the shape prior, how likely the voxel is to belong to the muscle. This is the main motivation of adding a shape prior in the data term.

While the segmentations obtained through our method were generally good as can be seen in Fig. 2 and Fig. 3, there were some areas where it leaked because the shape prior was not accurate enough. The average running time of the segmentation was within 1-5 seconds. Although the method was not as accurate as a human expert, the results are clinically acceptable. All results are computed using the freely available implementation of Max-Flow Min-Cut algorithm of Boykov and Kolmogorov.[13]

4. CONCLUSION

This paper presents a new way to incorporate shape priors in the MRF framework and utilize graph-cut to obtain the global solution to segment masseter muscle from CT datasets. Segmenting the muscle is a challenging task because of the presence of soft tissues in close proximity, but the method shows that by incorporating prior shape knowledge, clinically acceptable results can be achieved. The contribution of this paper is in transforming global shape knowledge into the regional energy term.

In future work we plan to replace the fixed shape template, employed in the research, by a statistical model.

ACKNOWLEDGMENTS

This work has been supported by the NCCR/CO-ME research network of the Swiss National Science Foundation.

REFERENCES

[1] Boykov, Y. Y. and Jolly, M.-P., "Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images," in [*ICCV*], **1**, 105–112 (2001).

[2] Freedman, D. and Zhang, T., "Interactive Graph Cut Based Segmentation with Shape Priors," in [*CVPR*], 755–762 (2005).

[3] Chittajallu, D. R., Shah, S. K., and Kakadiaris, I. A., "A Shape-Driven MRF Model for the Segmentation of Organs in Medical Images," in [*CVPR*], 3233–3240 (2010).

[4] Freiman, M., Kronman, A., Esses, S. J., Joskowicz, L., and Sosna, J., "Non-parametric Iterative Model Constraint Graph Min-Cut for Automatic Kidney Segmentation," in [*MICCAI*], **13**, 73–80 (2010).

[5] Malcolm, J., Rathi, Y., and Tannenbaum, A., "Graph Cut Segmentation with Nonlinear Shape Priors," in [*ICIP*], 365–368 (2007).

[6] Slabaugh, G. G. and Unal, G., "Graph Cuts Segmentation Using an Elliptical Shape Prior," in [*ICIP*], 1222–1225 (2005).

[7] El-Zehiry, N. and Elmaghraby, A., "Graph Cut Based Deformable Model with Statistical Shape Priors," in [*ICPR*], 1–4 (2008).

[8] Zhu-Jacquot, J., "Graph Cuts Segmentation with Geometric Shape Priors for Medical Images," in [*Southwest Symposium on Image Analysis and Interpretation*], 109–112 (2008).

[9] Das, P., Veksler, O., Zavadsky, V., and Boykov, Y., "Semi-Automatic Segmentation with Compact Shape Prior," *Image and Vision Computing* **27**(1), 206–219 (2009).

[10] Ali, A. M., Farag, A. A., and El-Baz, A. S., "Graph Cuts Framework for Kidney Segmentation with Prior Shape Constraints," in [*MICCAI*], **10**(1), 384–392 (2007).

[11] Song, Q., Liu, Y., Liu, Y., Saha, P. K., Sonka, M., and Wu, X., "Graph Search with Appearance and Shape Information for 3-D Prostate and Bladder Segmentation," *MICCAI* **13**, 172–180 (2010).

[12] Xu, N., Ahuja, N., and Bansal, R., "Object Segmentation Using Graph Cuts Based Active Contours," *Computer Vision and Image Understanding* **107**(3), 210–224 (2007).

[13] Boykov, Y. and Kolmogorov, V., "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision," *PAMI* **26**(9), 1124–1137 (2004).