2015 IEEE International Symposium on Multimedia

Interactive Image Segmentation Using Elastic Interpolation

Sachin Meena
Department of Computer Science
University of Missouri
Columbia, MO, 65211 USA
Email: sm3z5@mail.missouri.edu

Kannappan Palaniappan
Department of Computer Science
University of Missouri
Columbia, MO, 65211 USA
Email: Pal@missouri.edu

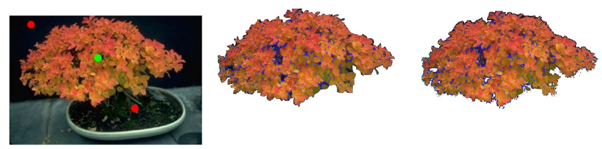
Guna Seetharaman
US Naval Research Laboratory
Washington, DC 20375 USA
Email: guna@ieee.org

***Abstract*—Elastic body splines (EBS) belong to a family of splines introduced for biomedical image registration. EBS models the elastic deformation of a homogeneous isotropic elastic body subjected to external forces. The task of interactive image segmentation is framed as a semi-supervised interpolation where the basis functions are learned using the user provided seed points to model and predict the labels for the unlabeled pixels. Seed points are sparse compared to other methods that may require scribbles and regions. The spline for interpolating labels is the EBS which we compare to our previous work using Gaussian Elastic Body splines (GEBS) [1] for the task of interactive image segmentation. Experimental results show that the EBS is about 14 percent better, in terms of accuracy, than GEBS and significantly better than random walk and graph cut based segmentation. EBS is also 2.5 times faster than GEBS.**

***Keywords*-Interactive image segmentation, semi supervised regression, elastic body splines.**

I. Introduction

Image segmentation is a fundamental problem in image processing and computer vision. The goal is to extract regions associated with objects of interest from the given image. Automatic segmentation methods are often employed for the task of segmentation however these methods can cause over segmentation and fail to produce satisfactory result. These methods also fail in the case of complex image scenes with weak boundary edges, foreground and background sharing similar color distributions or textured images. To segment a region of interest one can use manual annotations but human ground truth annotations can be both time consuming and expensive. Semi-automatic methods often work well in such cases as they lie some where in between the automatic methods and high quality manual annotations. Semi-automatic methods use the human knowledge to guide the segmentation process. As a result interactive image segmentation has received a lot of attention and popularity in recent years to segment images as it overcomes some of the inherent problems associated with very precise unsupervised image segmentation. Earlier work in this area are live-wire [2] and intelligent scissors [3] where users provide some information about the location of the object boundary that is then used to guide the image segmentation task. Both live-wire [2] and intelligent scissors [3], being an edge based methods are highly susceptible to noise and often miss weak object boundaries. Active contour models such as



(a) 3 marked pixels (b) True (c) Proposed method

Figure 1. Image from Berkeley data set [4]. From left to right Seed Points, Ground Truth, EBS (proposed method). The interactive user input points sparsely sampling the foreground (green) and background (red) are just single pixel but shown as thick dots for ease of visibility. No pre-processing or post processing is performed to generate the result.

[5] and [6] can be interactively initialized using a curve close to the desired object boundary from which the curve evolves under forces based on shape regularization and local edge information [5] or region information [6]. One of the main drawbacks of these methods is solutions are only locally optimum hence the final results are greatly affected by the initial input provided by the user. Graph based methods such as [7], [8] and [9] frame the interactive segmentation task as a graph cut problem, the goal here is to find a minimum cost cut that best separates the foreground from the background. Random walker [10] is another graph based algorithm that uses a different approach where a random walker jumps from one unlabeled pixel to another until it reaches a foreground or a background pixel.

The input to interactive image segmentation methods are primarily provided in two ways, scribble (brush strokes) or by drawing a rectangular box around an object. Recently, in [1] we have suggested sparse inputs in the form of points for the task of interactive image segmentation. Only eight to ten pixels are marked and the sparse input leads to a very fast image segmentation process. In this work we use the same frame work as suggested in [1] for the interactive image segmentation task but use a different basis function in the governing equation. The spline functions used are Elastic Body Splines (EBS) which are analytic solutions of the Navier equilibrium PDE that models the deformation of an elastic body. EBS have been applied to the task of

978-1-5090-0379-2/15 $31.00 © 2015 IEEE
DOI 10.1109/ISM.2015.78

biomedical image registration [11] and in this work we propose to use it for interactive image segmentation. In section 2 we provide a brief description of the original EBS based image registration problem. In section 3 we adapt the EBS framework for interactive image segmentation and include relevant mathematical details. In section 4 we provide quantitative and qualitative results of the proposed framework and finally in section 5 we provide a summary of our work along with possible directions for future work.

II. Elastic Body Splines

The Elastic Body Splines (EBS) belong to the class of 3D splines suggested by Davis in [11] and mainly used for elastic registration of bio-medical images. Given the displacement of landmark points in a source and a target image a transformation is determined that maps the corresponding landmarks in the source and the target image and interpolates the displacement for all the remaining points in the transformed image. The deformable displacement transformation is modeled as

$$\vec{d}(\vec{x}) = A\vec{x} + \vec{b} + \sum_{i=0}^{N} G(\vec{x} - \vec{p_i})\vec{c_i} \tag{1}$$

where $d(\vec{x})$ is the displacement vector of the landmark points. A, b and C are the coefficients for affine and non-linear elastic terms. Here the first sum represents a non-linear elastic transformation while the second expression represents a linear affine transformation. EBS splines are an analytic solution of the Navier equilibrium PDE.

$$\mu\nabla^2\vec{u}(\vec{x}) + (\mu + \lambda)\nabla[\nabla \cdot \vec{u}(\vec{x})] = \vec{f}(\vec{x}) \tag{2}$$

Here $\vec{u}(\vec{x})$ is the field at position $\vec{x}$, $\vec{f}(\vec{x})$ is the external force vector, μ, λ are the Lame coefficients, ∇^2, ∇ denote the Laplacian and gradient respectively. Eq. (2) models the deformation of a homogeneous isotropic elastic body subjected to loads. The solution of (2) is obtained by Galerkin vector method, that results in three coupled PDEs for the displacement $u(\vec{x})$ which are then transformed into three independent radially symmetric biharmonic PDEs. In [11] the authors have suggested to use rational force of type (3)

$$\vec{f}(\vec{x}) = \vec{c}/r(\vec{x}) \tag{3}$$

where $\vec{c}$ is a multiplicative constant vector. The elastic body spline function for the force $\vec{f}(\vec{x}) = \vec{c}/r(\vec{x})$ is given as

$$\mathbf{G}(\vec{x}) = \beta r(\vec{x})\mathbf{I} - \vec{x}\vec{x}^T/r(\vec{x}) \tag{4}$$

where $\beta = 8(1-\nu)-1$, $\nu = \lambda/2(\lambda+\mu)$ and $\mathbf{I}$ is the identity matrix. PDE (2) is solved using the force function (3). We refer to [13] and the references there-in for additional details regarding EBS.

Gaussian elastic body splines (GEBS) belongs to class of elastic splines suggested by Kohlrausch in [12] for bio-medical image registration. The spline function for GEBS is given as

$$\mathbf{G}(\vec{x}) = \Bigg[\Bigg(\left(4(1-\nu)-1\right)\frac{\mathrm{Erf}(\hat{r})}{r} - \sqrt{\frac{2}{\pi}}\sigma\frac{e^{-\hat{r^2}}}{r^2} + \sigma^2\frac{\mathrm{Erf}(\hat{r})}{r^3}\Bigg)\mathbf{I} + \Bigg(\frac{\mathrm{Erf}(\hat{r})}{r^3} + 3\sqrt{\frac{2}{\pi}}\sigma\frac{e^{-\hat{r^2}}}{r^4} - 3\sigma^2\frac{\mathrm{Erf}(\hat{r})}{r^5}\Bigg)\vec{x}\vec{x}^T\Bigg]. \tag{5}$$

were $\mathrm{Erf}(\cdot)$ is the standard error function, $\hat{r} = |\vec{x}|/\sqrt{2}\sigma$, $\mathbf{I}$ is the identity matrix, $\vec{x}\vec{x}^T$ is an outer product. The GEBS function is derived by also solving PDE (2) but using a Gaussian external force term as described in [12].

III. Framework for Interactive Image Segmentation

For an interactive image segmentation task we learn the classification function $\vec{d}(\vec{x})$ of the following spline form,

$$\vec{d}(\vec{x}) = \mathbf{A}\vec{x} + \vec{b} + \sum_{i=0}^{N} \mathbf{G}(\vec{x} - \vec{p_i})\vec{c_i}. \tag{6}$$

such that

$$\sum_{i=0}^{N}\vec{c_i} = \vec{0} \quad , \sum_{i=0}^{N}\vec{c_i}p_{ij} \quad = \vec{0} \quad j = 1, \ldots, 5 \tag{7}$$

where $\vec{x}$ is a feature vector, $\mathbf{G}$ is a spline matrix, $\vec{c_i}$ are the spline coefficients, $\vec{p_i}$ are the labeled pixels and the matrix $\mathbf{A}$ and $\vec{b}$ are coefficients for the linear mapping. The constraints in (7) ensure that the system of equations (6) has a unique solution. In [1] we have used GEBS (8) as the spline function for interactive image segmentation. In this paper we change the basis function to the EBS form, so the classification function is given as,

$$\vec{d}(\vec{x}) = \mathbf{A}\vec{x} + \vec{b} + \sum_{i=0}^{N} \mathbf{G}_{EBS}(\vec{x} - \vec{p_i})\vec{c_i}. \tag{8}$$

where $\mathbf{G}_{EBS}$ is given by (4).

For the task of interactive segmentation $\vec{d}(\vec{x})$ is learned from the user supplied seed points such that $\vec{d}(\vec{x})) = +\vec{1}^T$ for the foreground pixels and $\vec{d}(\vec{x}) = -\vec{1}^T$ for the background pixels. Let $\vec{w}$ be the vector of all the EBS coefficients given as,

$$\vec{w} = \begin{bmatrix}\mathbf{C}_F & \mathbf{C}_B & \mathbf{A} & \vec{b}^T\end{bmatrix}^T \tag{9}$$

where $\mathbf{C}_F$ and $\mathbf{C}_B$ are the elastic coefficients corresponding to the foreground and the background seed pixels.

$$\mathbf{C}_F = \left[\vec{c}_1^T \ldots \vec{c}_f^T\right], \quad \mathbf{C}_B = \left[\vec{c}_1^T \ldots \vec{c}_b^T\right].$$

The EBS mapping determined by the weights $\vec{w}$ that we are solving for, is given by the relationship, $\mathbf{Y} = \mathbf{L}\vec{w}$, where $\mathbf{Y}$ is the set of displacement vectors for the user defined seed points. When matrix $\mathbf{L}$ is singular or close to being singular, computational techniques such as SVD can be used to obtain the EBS coefficients. The EBS coefficients are estimated by solving the matrix equation,

$$\vec{w} = \mathbf{L}^{-1}\vec{Y} \tag{10}$$

where,

$$\mathbf{L} = \begin{bmatrix} \mathbf{K} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{O} \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} \mathbf{G}_{FF} & \mathbf{G}_{FB} \\ \mathbf{G}_{BF} & \mathbf{G}_{BB} \end{bmatrix} \tag{11}$$

$$\mathbf{G}_{FF}(\vec{r}) = \begin{bmatrix} \mathbf{G}_{11}(\vec{r}_{11}) & \ldots & \mathbf{G}_{1f}(\vec{r}_{1f}) \\ \vdots & & \vdots \\ \mathbf{G}_{f1}(\vec{r}_{11}) & \ldots & \mathbf{G}_{ff}(\vec{r}_{ff}) \end{bmatrix} \tag{12}$$

with $\vec{r}_{ij}$ is the distance between feature vectors $\vec{p}_i$ and $\vec{p}_j$ of the foreground seed points. Hence, $\mathbf{G}_{FF}$ is the matrix of EBS (4) functions defined only over the foreground pixels. $\mathbf{G}_{FB}$, $\mathbf{G}_{BB}$ and $\mathbf{G}_{BF}$ are similarly defined,

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_F & \mathbf{I}_F \\ \mathbf{P}_B & \mathbf{I}_B \end{bmatrix}, \tag{13}$$

where the identity matrices are given by,

$$\mathbf{I}_F = \left[\mathbf{I}_1 \ldots \mathbf{I}_f\right], \quad \mathbf{I}_B = \left[\mathbf{I}_1 \ldots \mathbf{I}_b\right], \tag{14}$$

and,

$$\mathbf{P}_F = \begin{bmatrix} x_{11}\mathbf{I} & \ldots & x_{15}\mathbf{I} \\ \vdots & & \vdots \\ x_{f1}\mathbf{I} & \ldots & x_{f5}\mathbf{I} \end{bmatrix} \tag{15}$$

where x_{ij} is the j^{th} feature for i^{th} pixel. $\mathbf{P}_B$ is similarly defined. The vector $\vec{Y}$ consists of $\vec{Y}_F$ with values +1 for the foreground seed points and $\vec{Y}_B$ with values -1 for the background seed points,

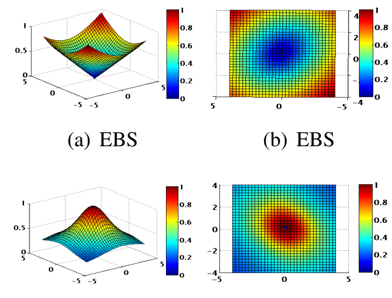
$$\vec{Y} = \left[\vec{Y}_F \quad \vec{Y}_B \quad \vec{O}\right]^T \tag{16}$$

where $\vec{O}$ is a vector of zeros.

Once the classification function (8) is learned for classification purposes we use zero (midpoint value between -1 and $+1$) to threshold the vector $\vec{d}(\vec{x})$ and assign label $\ell(\vec{x})$ to a pixel by taking consensus among vector elements of $\vec{d}(\vec{x})$,

$$\ell(\vec{x}) = \begin{cases} foreground, & \text{if majority } \vec{d}(\vec{x}) \text{ elements} \geq 0 \\ background, & \text{if majority } \vec{d}(\vec{x}) \text{ elements} < 0. \end{cases} \tag{17}$$

The classification functions for EBS and GEBS are shown in Fig. 3. As we can see from plots both the classification functions are symmetric in nature however the effect of seed points in case of EBS is global in nature while for GEBS the effect of seed point is more local in nature.



(a) EBS (b) EBS

(c) GEBS (d) GEBS

Figure 2. Classification functions for EBS with $\nu = 0.25$ and GEBS with $\nu = 0.25$ and $\sigma = 4.0$.

No.	Methods	F-measure
1	EBS	**0.8770±0.0993**
2	GEBS	0.7370±0.2273
3	RW	0.6928±0.2403
4	GC	0.3869±0.3485

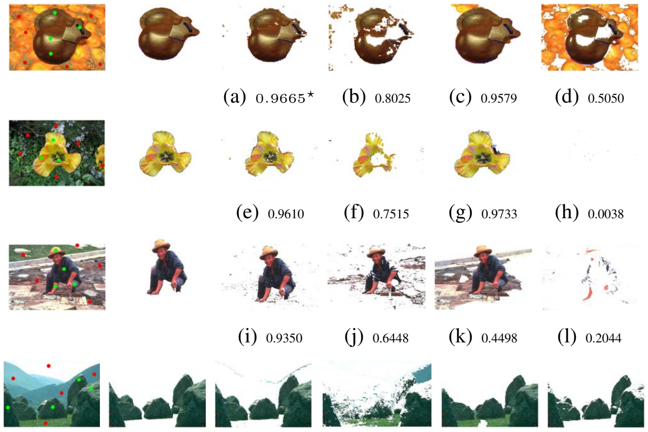
Table I
QUANTITATIVE EVALUATION : AVERAGE F-MEASURE ON 50 IMAGES FROM GRAB-CUT DATA SET [8] USING AN AVERAGE OF 8.46 MARKED PIXELS PER IMAGE.

IV. EXPERIMENTS

A. Quantitative Evaluation

We use images from the Grab-cut datasets [8] to provide the quantitative evaluation. Instead of using tri-maps seed points are provided manually by the user and then the segmentation is performed over entire image. In our experiment, we mark ten or less pixels as the initial seed points and compare the performances of EBS and GEBS spline functions and two popular methods Random walk (RW) [10] and Graph Cut (GC) [7] in our comparison. No pre or post processing operations are performed on the input image or the segmentation results obtained from different methods.

EBS and GEBS functions have parameters μ and σ that need to tuned for the task of interactive image segmentation. The optimum value of parameters is determined empirically. For experiments we set $\mu = .25$ for spline EBS and $\mu = .25$ and $\sigma = 1.0$ for spline EBS. As we can see from the Table I, of the four methods evaluated EBS-based interactive segmentation gives the most accurate results compared to manual ground truth. GEBS has significantly lower performance than the EBS which indicates the choice of force function is critical in the PDE (2) as it determines the behavior of the interpolating function (6). The performance of graph cut is quite poor mainly because it fails to learn the foreground-background color model accurately when only a sparse set (8 to 10) of pixels are marked as seed points.



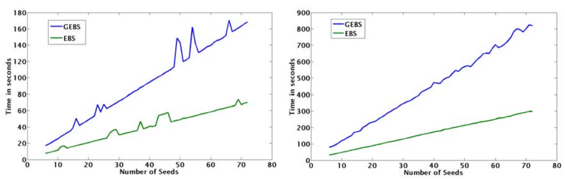
(a) 0.9665* (b) 0.8025 (c) 0.9579 (d) 0.5050

(e) 0.9610 (f) 0.7515 (g) 0.9733 (h) 0.0038

(i) 0.9350 (j) 0.6448 (k) 0.4498 (l) 0.2044

(m) TRUE (n) EBS 0.9739 (o) GEBS 0.5514 (p) RW 0.9013 (q) GC 0.9202

Figure 3. Qualitative results using images from Grab-cut [8] and Berkeley database [4]. From left to right input seeds, ground truth mask, EBS, GEBS [1], Random walk [10], Graph-cut [7]. * F-scores for the segmentation results. 9 pixels are marked as foreground and background. No pre-processing or post processing is performed to generate results.



(a) Image from Grab-cut dataset (b) Image from Berkeley dataset

Figure 4. Timing analysis for EBS and GEBS using images from Grab-cut and Berkeley datasets.

B. Timing Performance

We have also compared the speed of EBS and GEBS splines with different number of marked seed points. As shown in Fig. 4 the speed of EBS is 2.5 times faster that of GEBS. GEBS makes use of the error function (8) which is a computationally expensive operation and makes the GEBS performance slower compared to EBS. The experiments were performed on a Windows laptop with 2.4GHz CPU, 3GB RAM and MATLAB R2012a.

V. CONCLUSIONS

We have presented a novel adaptation of Elastic Body Splines (EBS) for interactive image segmentation. It uses the strength of elastic body splines to learn an interpolating pixel classifier function using sparse inputs, flexibly segmenting complex image for the task of interactive image segmentation. Quantitative and qualitative results suggest that the performance of EBS is significantly better than other methods like GEBS indicating that the choice of basis function is critical in the interpolation based interactive image segmentation framework. Experimental results also show that the EBS outperforms other widely used methods like random walk and graph cut.

ACKNOWLEDGMENT

This research was partially supported by U.S. Air Force Research Laboratory (AFRL) under agreement AFRL FA875014-2-0072. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of AFRL or the U.S. Government.

REFERENCES

[1] S. Meena, V. B. S. Prasath, K. Palaniappan, and G. Seetharaman, "Elastic body spline based image segmentation," in *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, 2014.

[2] A. X. Falcao and J. K. Udupa, "A 3D generalization of user-steered live-wire segmentation," *Medical Image Analysis*, vol. 4, no. 4, pp. 389–402, 2000.

[3] E. N. Mortensen and W. A. Barrett, "Intelligent scissors for image composition," in *Proc. Conf. on Computer Graphics and Interactive Techniques*, 1995, pp. 191–198.

[4] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *IEEE Int. Conf. Computer Vision*, 2001, pp. 416–423.

[5] F. Bunyak, K. Palaniappan, S. K. Nath, and G. Seetharaman, "Geodesic active contour based fusion of visible and infrared video for persistent object tracking," in *8th IEEE Workshop Applications of Computer Vision (WACV 2007)*, 2007.

[6] S. K. Nath, K. Palaniappan, and F. Bunyak, "Cell segmentation using coupled level sets and graph-vertex coloring," *Lecture Notes in Computer Science (MICCAI)*, vol. 4190, pp. 101–108, 2006.

[7] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in nd images," in *IEEE Int. Conf. Computer Vision*, 2001, pp. 105–112.

[8] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: interactive foreground extraction using iterated graph cuts," *ACM Transactions on Graphics*, pp. 309–314, 2004.

[9] A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr, "Interactive image segmentation using an adaptive gmmrf model," in *European Conf. on Computer Vision*, 2004, pp. 428–441.

[10] L. Grady, "Random walks for image segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, no. 11, pp. 1768–1783, 2006.

[11] M. H. Davis, A. Khotanzad, D. P. Flamig, and S. E. Harms, "A physics-based coordinate transformation for 3-d image matching," *IEEE Trans. on Medical Imaging*, pp. 317–328, 1997.

[12] J. Kohlrausch, K. Rohr, and H. Siegfried Stiehl, "A new class of elastic body splines for nonrigid registration of medical images," *Journal of Mathematical Imaging and Vision*, pp. 253–280, 2005.