

# SEGMENTATION OF NON-CONVEX REGIONS WITHIN UTERINE CERVIX IMAGES

Shiri Gordon and Hayit Greenspan

Biomedical Engineering Dept, Faculty of Engineering,  
Tel-Aviv University,  
Tel-Aviv 69978, Israel

## ABSTRACT

The National Cancer Institute has collected a large database of uterine cervix images, termed “cervigrams” for cervical cancer screening research. Tissues of interest within the cervigram, in particular the lesions, are of varying sizes and complex, non-convex shapes. The current work proposes a new methodology that enables the segmentation of non-convex regions, thus providing a major step forward towards cervigram tissue detection and lesion delineation. The framework transitions from pixels to a set of small coherent regions (superpixels), which are grouped bottom-up into larger, non-convex, perceptually similar regions, utilizing a new graph-cut criterion and agglomerative clustering. Superpixels similarity is computed via a combined region and boundary information measure. Results for a set of 120 cervigrams, manually marked by a medical expert, are shown.

**Index Terms**— cervicography images; segmentation; graph algorithms

## 1. INTRODUCTION

Cervicography is one of the methods for cervical cancer screening that uses visual testing based on the color change of abnormal cells that turn white when exposed to acetic acid (acetowhite). The cervicographic images (cervigrams) handled in this work were selected from a large database collected by the National Cancer Institute (NCI), National Institute of Health (NIH). NCI in collaboration with the National Library of Medicine (NLM), NIH, is developing a unique Web-based database of the digitized cervix images to study the evolution of lesions related to cervical cancer. The cervix region can be divided into three main tissues: The squamous epithelium (SE), which is smooth and pink; The columnar epithelium (CE) that appears red and irregular; and the acetowhite (AW) region which is a transient, white-appearing epithelium following the application of acetic acid (Figure 1(a))<sup>1</sup>. The separation between the AW and the SE tissues presents an extremely difficult image segmentation task. The AW tissue can be identified by its color, which is lighter than the color of the

surrounding SE tissue and by its boundary, but earlier studies have shown that a large overlap exists between the color distributions of these two tissues [1]. The boundary itself, may vary in quality and is not always clearly visible. The AW lesions are of varying shape and can be located in different places within the cervix region, thus no shape constraints can be used to identify them.

Recent works analyze high-resolution colposcopic images (e.g [2, 3]). Segmentation efforts for cervigrams, which resemble a low-magnification colposcopic image, were recently introduced in [1, 4]. In these works pixel-based segmentation based on color features is used. Sample segmentation maps generated by unsupervised clustering using a mixture of Gaussians model in the color feature space, as in [1], are shown in Figure 2(b). Due to the large tissue overlap in color space, only a small part of the AW tissue is detected. Furthermore, the pixel-based segmentation is very noisy and each segment is broken into many disjoint components.

In the current work a new segmentation framework is proposed, that shifts the pixel-based analysis to a region-based (superpixel) scheme, and augments the feature space to include both region color and edge information. A preprocessing stage [1] extracts the cervix region from the cervigram image and removes specular reflections (SR) and illumination field artifacts to provide the input image as in Figure 1(b). Utilizing the morphological watershed transform, the image is over-segmented into a large set of superpixels (Section 2). A similarity matrix between superpixels, that combines region and edge information (Section 3) is used in an agglomerative clustering framework utilizing a new graph-cut criterion (Section 4). A hierarchy of segmentation maps is generated, with the segments extracted enabling the representation of long, thin and *non convex* shapes within the cervix.

## 2. SUPERPIXELS GENERATION

Using the watershed transform [5], cervigrams are over segmented into superpixels: The image is initially preprocessed to smooth out noise and enhance tissue boundaries using the morphological toggle-contrast transform [5] with a small structuring element (SE) (a disk of radius  $r = 1$ ). The toggle-contrast sharpens the image edges without boosting the con-

<sup>1</sup>A colored version is available at <http://www.eng.tau.ac.il/~hayit/>.

trast of structures smaller than the considered SE. A color gradient image is generated next [6]. The magnitude of the initial gradients is linearly normalized to the range of (0 : 1). A smoothed version of the normalized gradients and a single threshold generates the regional minima for the watershed transform. The threshold is set lower than the energy of the image boundaries that need to be preserved (defined experimentally). The detected regional minima are post-processed to eliminate regions smaller than 3 pixels and are imposed on the normalized gradient image. The watershed transform is applied to this modified gradient image and over-segments the image into superpixels (Figure 1(c)).

### 3. REGION & EDGE SIMILARITY MATRIX

In a typical graph-based image representation,  $G(V, E)$ , each vertex is a point in the feature space used, associated with a single image pixel. Each edge,  $(i, j) \in E$ , is weighted by the pairwise similarity,  $w_{ij}$ , between nodes  $i$  and  $j$  [7]. The weights define the symmetric  $n \times n$  similarity matrix  $W$ , where  $n$  is the number of vertexes (image pixels). In the current work the graph representation is shifted from a pixel based representation to a region based representation. Each vertex within the graph is a superpixel within the over-segmented image. Each edge is weighted by the similarity between corresponding superpixels. The size of the similarity matrix is considerably reduced, which is crucial in terms of the computational load of agglomerative clustering. The similarity between superpixels is composed from a region similarity and from an edge cost.

**Region similarity:** Each superpixel is modeled by a Gaussian distribution in  $Lb$  feature space (from the  $Lab$  color space)<sup>2</sup>. Region similarity between two neighboring superpixels,  $w_r(i, j)$ , is measured using the symmetric version of the Kullback-Leibler divergence between two Gaussians,  $SKL_{ij}$  [8], locally scaled by a self-tuning kernel [9]:

$$w_r(i, j) = \exp \left( -\frac{SKL_{ij}}{\sqrt{\sigma_i \sigma_j}} \right), \quad (1)$$

where the scalars  $\sigma_i, \sigma_j$  are the local scaling factors. The  $\sigma_i$  factor is calculated as the median  $SKL$  value between superpixel  $i$  and its  $K$  nearest neighbors ( $K \leq 5$ , depending on the actual number of neighboring superpixels). Using this kernel, region similarities are normalized to the (0 : 1) range, where higher values indicate stronger similarity.

**Edge cost:** Two superpixels are considered neighbors if they possess a common watershed boundary. The edge cost,  $w_e(i, j)$ , is defined using the average value (energy) of the normalized gradients along this common boundary:  $e_{ij}$ . An edge separating between two tissues within the cervix, is composed of many such boundary portions with varying energy

<sup>2</sup>The  $Lb$  combination was found best for separating AW lesions from their SE surrounding. Details are omitted here due to lack of space.

$\ell_{MinMax}(C_p, C_q)$	$\ell_{WMCut}(C_p, C_q)$	$\ell_{NMCut}(C_p, C_q)$
$=$	$=$	$=$
$\frac{S(C_p, C_q)}{S(C_p, C_p)S(C_q, C_q)}$	$\frac{\sum_{i \in C_p, j \in C_q} a_{ij} w_{ij}}{\sum_{i \in C_p, j \in C_q} a_{ij}}$	$\frac{\ell_{WMCut}(C_p, C_q)}{S(C_p, C_p)S(C_q, C_q)}$
(1)	(2)	(3)

**Table 1.** Different linkage measures between two clusters.

values (Figure 1(c)). It is often the case that boundary portions within the tissue's region possess similar energy values to those of the boundary portions on its edge (especially when textured regions are involved). The edge cost is therefore locally scaled to emphasize the energies of between-tissues boundary portions and to smooth out the energies of within-tissue boundary portions. The scaled energy,  $\hat{e}_{ij}$ , is calculated as:

$$\hat{e}_{ij} = \left( 1 - \exp \left( - \left| \frac{e_{ij} - \min_{(k,l) \in N}(e_{kl})}{e_{ij} - \max_{(k,l) \in N}(e_{kl}) + \epsilon} \right| \right) \right) e_{ij}, \quad (2)$$

where  $(k, l) \in N$  are the neighboring boundary portions (connected on both ends of boundary  $(i, j)$ ). Note that this enhancement is performed on the superpixel level rather than the pixel level, thus encompassing a more global information. The edge cost is calculated by applying the negative transform [6] to the  $\hat{e}_{ij}$  values, thus edges with low energy attain high values (to be consistent with the region similarity).

Using the region similarity and the edge cost between two superpixels, the overall similarity is defined by the square root of their multiplication:  $w_{ij} = \sqrt{w_r(i, j)w_e(i, j)}$ .

### 4. AGGLOMERATIVE CLUSTERING OF SUPERPIXELS USING NORMALIZED-MEAN CUT

Given a graph model represented by the pairwise similarity matrix,  $W = (w_{ij})$ , perceptually similar superpixels can be clustered using a graph-cut clustering framework [7]. The classical cut between two clusters  $C_1, C_2$  is defined as:  $S(C_1, C_2) \equiv \sum_{i \in C_1} \sum_{j \in C_2} w_{ij}$  and the similarity within a cluster  $C_1$  is defined as  $S(C_1, C_1)$ . Additional cut criterions exist [10], as well as a variety of cluster merging and splitting algorithms. The current work focuses on an agglomerative clustering algorithm, where clusters are built bottom up to optimize the cut criterion and a full hierarchy of segmentations is enabled. This framework requires a linkage measure,  $\ell$ , corresponding to the selected cut criterion. During each iteration of the agglomerative process, the two clusters  $C_p$  and  $C_q$ , that have the largest pairwise linkage between them are merged.

The *MinMax cut* criterion and the corresponding linkage measure  $\ell_{MinMax}$  (Table 1-(1)), are shown in [10] to provide the best clustering results. This criterion is defined as minimizing the classical cut, while maximizing the within-clusters similarities. It is known for its tendency to generate balanced

clusters. The *weighted-mean cut* (WMCut), minimizes the weighted-mean value of the similarities along the cut. It was recently introduced for clustering of superpixels [11]. In the  $\ell_{WMCut}$  linkage measure (Table 1-(2)) the weights,  $a_{ij}$ , are defined as the length of the corresponding boundary portions. The WMCut doesn't favor clusters with a small number of boundary portions [11], but in agglomerative clustering large clusters tend to form even larger clusters, while smaller clusters tend to be left alone [10].

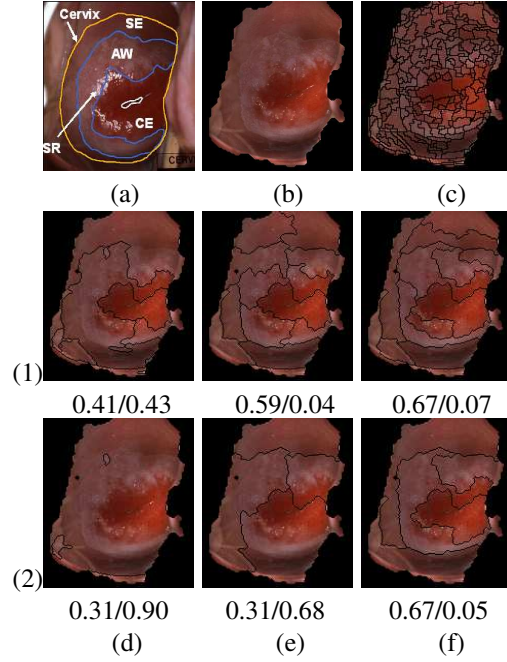
In the current work a new cut criterion, termed the *normalized - mean cut* (NMCut), combines the benefits of the MinMax-cut and the WMCut. This criterion (and the corresponding  $\ell_{NMCut}$  linkage measure, Table 1-(3)) minimizes the WMCut while maximizing the within-clusters similarities. Using the NMCut more balanced clusters are generated, but when several neighboring clusters have similar within-cluster similarities (denominator), the clusters are merged according to the actual boundary energy and not the number of boundary portions being considered (numerator). This observation is of major importance in the case of cervigrams, where neighboring tissues may possess very similar within-cluster similarities due to their color resemblance and where tissues are elongated and non convex, thus a small number of boundary portions may exist between their sub-parts.

## 5. RESULTS

Experiments were performed on a test set of 120 preprocessed cervigrams for which manual segmentations by an NCI expert are available. A visualization of the segmentation process is presented in Figure 1. The original cropped image, with the AW manual markings in blue, is shown in (a). The watershed superpixel boundaries are shown in (c). Visual inspection indicates that the superpixels do not cross any important tissue boundaries, in particular, an overlap with the expert marked AW boundaries can be detected, as desired. Thus, the superpixels can be used as an initialization for the follow-up merging process. Segmentation maps generated following the agglomerative clustering are shown for levels  $K = 15$  (1) and  $K = 6$  (2) of the hierarchy. Results of the proposed NM-Cut criterion are shown in column (f). Overall, we note that the segments generated capture well the non-convex and elongated shape of the tissues within the cervix.

Figure 1 presents an example comparison between the different linkage measures, using the similarity matrix of Section 3. Segment boundaries generated using the  $\ell_{WMCut}$ ,  $\ell_{MinMax}$  and  $\ell_{NMCut}$  are presented in columns (d), (e) and (f), respectively. Inspecting the  $K = 6$  level: the  $\ell_{WMCut}$  provides large segments that cover most of the image plain and some small isolated segments. Segments generated by the  $\ell_{MinMax}$  criterion are balanced but more convex. The  $\ell_{NMCut}$  criterion provides more accurate overlap between the generated segments and the AW region. The level of  $K = 15$  clusters shows an improved overlap quality in all three mea-

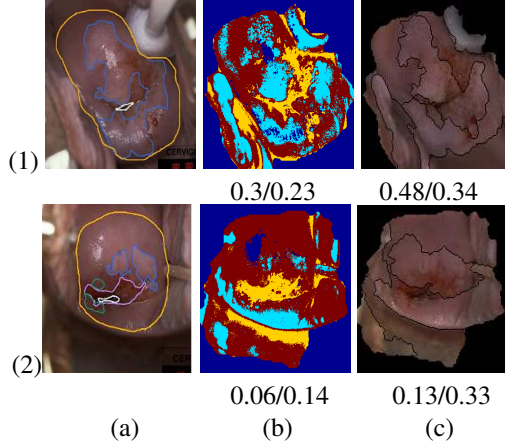
sures, with  $\ell_{NMCut}$  still more accurate.



**Fig. 1.** (a) Original image, marked are the different tissues of interest; (b) preprocessed cervigram; (c) Superpixels; Segmentation maps using different linkage measures: (d)  $\ell_{WMCut}$ , (e)  $\ell_{MinMax}$ , (f)  $\ell_{NMCut}$ , for  $K = 15$  (1) and  $K = 6$  (2). Segments boundaries imposed in black. Corresponding Dice/FP measures are listed.

In order to quantitatively evaluate the results, the generated segments quality per AW tissue is assessed. A mask is generated (per level) as an ensemble of the segments that have any overlap with the actual (expert-marked) AW tissue. The Dice metric,  $\frac{S \cap R}{S \cup R}$ , and the false-positives (FP) metric,  $\frac{S \cap \hat{R}}{\hat{R}}$ , are then used to quantify the overlap between the automated and expert-based AW delineation. Here,  $S$  is the AW mask area as extracted by the algorithm,  $R$ , is the expert's segmentation area and  $\hat{R}$  the complement of that area. Corresponding Dice and FP measures are listed below each segmentation result of Figure 1 (note that a more accurate segmentation overlap is correlated with high Dice and low FP). A Dice of 0.67 is achieved with the  $\ell_{NMCut}$  measure, with a low FP result. This is a very strong result which correlates well with the visual appearance. As a comparison, segmentation results using the unsupervised pixel segmentation framework that use color-only information [1], provides the result of  $Dice = 0.2$  and  $FP = 0.1$  for this image.

Figure 2 shows two additional cervigram segmentation examples. Existing pixel-based color-only segmentation (b), is compared with the proposed region-based segmentation (c). The images are very complex and this is reflected in the low Dice values. A clear advantage can be seen with the proposed algorithm, both via visual inspection as well as quantitatively.



**Fig. 2.** (a) Input image. Manual markings of AW imposed in blue; (b) Pixel-based segmentation: AW- light blue, SE-dark red, CE-yellow; (c) Proposed region-based ( $K = 6$ ) segmentation. Corresponding Dice/FP measures are listed.

Mean Dice and FP measures calculated over the entire, 120 image test set, are presented in Table 2. A quantitative comparison across the possible linkage criterions for the entire archive is presented in the first three rows. The  $\ell_{NMCut}$  is shown to provide the best results for both levels of the agglomerative hierarchy. The last four rows of the Table compare across possible region similarity measures between superpixels. In the proposed framework, each superpixel is represented by a Gaussian in  $Lb$  feature space and similarity between superpixels is calculated using a local scaling process of the symmetric KL distance. We term this the kl-Lb method. A similar scheme can be applied to the  $VS$  feature space (value and saturation channels of the  $HSV$  color feature space), hereon termed the kl-VS method. Different schemes may be used to represent the superpixel content and to measure superpixel similarities. For example, a superpixel can be represented by its color histogram (with 256 bins per channel) and the EMD [11] can be used for similarity. We name this method EMD-Lb and EMD-VS, using the  $Lb$  and  $VS$  color features, respectively. Table 2 presents the average Dice and FP measures calculated over the test set for each of these methods, in two levels of the agglomerative hierarchy. The best results are those achieved by using the kl-Lb method, as presented in this paper.

## 6. DISCUSSION

This work presents a new framework appropriate for segmentation of elongated, non-convex regions within the cervix. It includes clustering of superpixels using both region and edge information. A new graph-cut criterion, suitable for the special shape and color characteristics of regions within the cervix is introduced. The various components of the frame-

	Dice		FP	
	$K = 6$	$K = 15$	$K = 6$	$K = 15$
$\ell_{WMCut}$	0.09(0.03)	0.17(0.06)	0.53(0.26)	0.27(0.14)
$\ell_{MinMax}$	0.13(0.04)	20(0.06)	0.23(0.07)	0.09(0.03)
$\ell_{NMCut}$				
kl-Lb	<b>0.16(0.05)</b>	<b>0.22(0.07)</b>	<b>0.15(0.04)</b>	<b>0.07(0.02)</b>
kl-VS	0.13(0.03)	0.19(0.06)	0.20(0.06)	0.09(0.03)
EMD-Lb	0.13(0.03)	0.18(0.05)	0.16(0.04)	0.05(0.01)
EMD-VS	0.11(0.02)	0.18(0.05)	0.19(0.06)	0.05(0.01)

**Table 2.** Average segmentation results ( $mean(var)$ ) using Dice and FP measures on two levels of the agglomerative hierarchy (120 cervigrams).

work, including the superpixel representation, the superpixel similarity measure, and the linkage measure in the agglomerative clustering, are validated quantitatively. The framework is shown to outperform unsupervised, pixel-based schemes proposed to date. The current work provides a hierarchy of segmentations maps. We are currently investigating means for utilizing the collection of maps towards tissue extraction and lesion delineation. Extensive validation, using manual markings from several medical expert, is currently underway.

**Acknowledgement:** We would like to thank the Communications Engineering Branch, NLM, NIH and the Hormonal and Reproductive Epidemiology Branch, NCI, NIH, for the data and support of the work.

## 7. REFERENCES

- [1] H. Dvir, S. Gordon, and H. Greenspan, "Illumination correction for content analysis in uterine cervix images," in *proc. of IEEE WS MM-BIA, CVPR*, 2006, p. 95.
- [2] V. Van Raad, Z. Xue, and H. Lange, "Lesion margin analysis for automated classification of cervical cancer lesions," in *Proc. of SPIE Medical Imaging*, 2006, vol. 6144.
- [3] H. Lange, "Automatic detection of multi-level acetowhite regions in rgb color images of the uterine cervix," in *Proc. of SPIE Medical Imaging*, 2005, vol. 5747, pp. 1004–1017.
- [4] Y. Srinivasan et al., "A probabilistic approach to segmentation and classification of neoplasia in uterine cervix images using color and geometric features," in *Proc. of SPIE Medical Imaging*, 2005, vol. 5747, pp. 995–1003.
- [5] P. Soille, *Morphological image analysis. Principles and applications*, Springer-Verlag, Berlin, 2003.
- [6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prantice-Hall, Inc, 2002.
- [7] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. on PAMI*, vol. 22, no. 8, pp. 888–905, 2000.
- [8] S. Kullback, *Information theory and statistics*, Dover, NY, 1968.
- [9] L. Zelnik-Manor and P. Perona, "Self-tuning spectral clustering," in *Proc of Neural Information Processing Systems*, 2004.
- [10] C. Ding and X. He, "Cluster merging and splitting in hierarchical clustering algorithms," in *Proc. of IEEE ICDM*, 2002, pp. 139–146.
- [11] R.J. O'callaghan and D.R. Bull, "Combined morphological-spectral unsupervised image segmentation," *IEEE Trans. on Image Processing*, vol. 14, no. 1, pp. 49–63, 2005.