Automatic Segmentation of Skin Cancer Images using Adaptive Color Clustering

Dana Elena Ilea and Paul F. Whelan
Vision Systems Group
School of Electronic Engineering
Dublin City University, Dublin 9, Ireland
danailea@eeng.dcu.ie

Abstract

This paper presents the development of an adaptive image segmentation algorithm designed for the identification of the skin cancer and pigmented lesions in dermoscopy images. The key component of the developed algorithm is the Adaptive Spatial K-Means (A-SKM) clustering technique that is applied to extract the color features from skin cancer images. Adaptive-SKM is a novel technique that includes the primary features that describe the color smoothness and texture complexity in the process of pixel assignment. The A-SKM has been included in the development of a flexible color-texture image segmentation scheme and the experimental data indicates that the developed algorithm is able to produce accurate segmentation when applied to a large number of skin cancer (melanoma) images.

Keywords: Skin cancer, Image segmentation, Adaptive-SKM, Diffusion-based filtering.

1 Introduction

Skin cancer is one of the most common types of cancer and it can affect people at any age. It is a malignant tumor that develops changes in the skin texture and color, but it can be cured in more than 90% of cases, if the skin tumor is detected and treated in the early stages. There are two types of skin cancer, namely malignant melanoma and non-melanoma (basal cell and squamous cell carcinoma) [1]. Melanoma is more dangerous and can be fatal if untreated and a number of commercially available systems are designed for the analysis of pigmented skin lesions. Two representative systems are the SolarScan [2] and the microDERM dermoscopy unit [3]. It is useful to note that these systems are designed primarily to accurately capture skin images and not for automated detection of skin cancer images which is the aim of the image segmentation technique detailed in this paper.

The first step in the identification of the skin cancer is to determine accurately the lesion boundaries using various image processing techniques that provide important information for accurate diagnosis. The clinical parameters of pigmented lesions are known as ABCD's of skin cancer (asymmetry, border irregularity, color variation and diameter greater than 6 mm [1]) and are calculated with respect to the lesion's border. Automated skin cancer detection in dermoscopy images is a challenging task due to several reasons: a) low contrast between the lesion and the surrounding skin, b) irregular and fuzzy lesion borders, c) artifacts such as skin texture, air bubbles and hair, and d) non-uniform coloring inside the lesion [4]. There has been a significant amount of research work dedicated to the development of automated methods for segmentation of skin cancer images and the most relevant publications include the work of Binder et al [5], Ganster et al [6] and Xu et al [7].

In this paper the color and texture features are included in a compound mathematical descriptor to optimally extract the melanoma in skin cancer images. In this regard, texture is analyzed using the Local Binary Patterns (LBP)[8], while color information is extracted from the input data using the Adaptive SKM clustering procedure.

2 Image Segmentation Algorithm

The main components of the developed image segmentation algorithm are illustrated in Fig. 1. The key component of the algorithm is the Adaptive Spatial K-Means clustering algorithm that is included in the development of a split and merge color-texture segmentation framework.

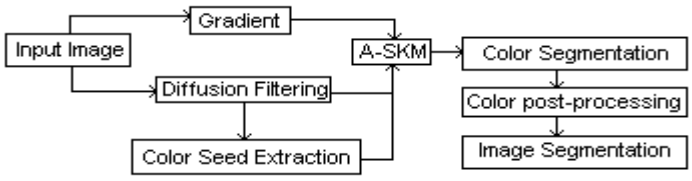

Fig. 1. Overview of the image segmentation algorithm.

The performance of the standard K-Means algorithm is modest since this space partitioning technique is sensitive to local variation in color and image noise. In addition the cluster initialization for standard K-Means algorithm is based on a random procedure and this may force the algorithm to converge to incorrect decisions (local minima). To alleviate the problems generated by the random initialization procedure we initialize the cluster centers with the dominant colors which are the peaks of the color histogram calculated from the image resulting after the application of color quantization (we applied linear quantization and the resulting image has 8 colors for each color axis – 3 bits resolution). We applied color quantization because the color histogram of the original image is sparsely populated and the peaks in the histogram are not statistically relevant.

To eliminate the problems associated with the classic K-Means algorithm we developed a new clustering scheme named the Adaptive Spatial K-Means. The novelty of this algorithm resides in the process of assigning the data points into K (for this implementation K=5) clusters by evaluating along with the pixel color information, two more distributions that sample the local color smoothness and the local texture complexity. In this regard to sample the local color smoothness, the image is filtered with an adaptive diffusion scheme [9], while the texture complexity is sampled by filtering the input image with a gradient operator. Thus, during the space partitioning process, the developed algorithm attempts to optimize the fitting of the diffusion and gradient distributions in a local neighborhood (5×5, 7×7, 9×9 and 11×11 window size-chosen adaptively, see Eq. 1) around the pixel under analysis, with the diffusion and gradient distributions for each cluster. The purpose of the A-SKM algorithm is to minimize the following global objective function:

$$J = \sum_{j=1}^{K}\sum_{i=1}^{n}\left[\left\| x_i^{(j)} - c_j \right\| + \underset{w\in[5\times5,11\times11]}{\arg\min} KSM(H_{Diff_i}^{(j)}(w), H_{Diff}^{(j)}) + \underset{w\in[5\times5,11\times11]}{\arg\min} KSM(H_{Grad_i}^{(j)}(w), H_{Grad}^{(j)}) \right] \quad (1)$$

where n is the total number of pixels, $H_{Diff_i}^{(j)}$ is the local color smoothness distribution (calculated from the diffusion filtered image) for the data point x_i, c_j is the center of the cluster j, $H_{Diff}^{(j)}$ is the local color smoothness distribution for cluster j, $H_{Grad_i}^{(j)}$ is the local texture complexity (calculated from the gradient data) for data point x_i, $H_{Grad}^{(j)}$ is the texture complexity for cluster j and KSM is the Kolmogorov-Smirnov metric bounded in the interval [0,2]. The local neighborhood is sampled by the local window size given by the parameter w. The developed algorithm is convergent and to improve the computational overhead we have applied the K-Means algorithm in a standard form for the first 5 iterations and then spatial and color continuity constraints are evaluated. The color segmentation result is further refined by joining the image regions resulting from the A-SKM algorithm, based on the evaluation of inter region variability. In this regard, we have developed a post-processing step that joins the image regions resulting from the A-SKM algorithm based on region adjacency graph and color similarity.

The image segmentation method used for this implementation is based on a split and merge framework that evaluates adaptively the color and texture information. The first step of the algorithm recursively splits the image hierarchically into four sub-blocks using the texture information extracted using the Local Binary Pattern (LBP) method [8]. The splitting decision evaluates the uniformity factor of the region under analysis that is sampled using the Kolmogorov Smirnov metric. In this regard, the pairwise similarity values of the four sub-blocks are calculated and the ratio between the highest and the lowest similarity values is compared with a split threshold value. The region is splitted, if the ratio is higher than the split threshold. During the splitting process two distributions are calculated, the LBP distribution that defines the texture and the distribution of the color labels computed using the A-SKM algorithm. Merging is the second step of the image segmentation scheme

and joins the adjacent regions that have similar color-texture characteristics by evaluating a merging importance (MI) value. The main novelty of this algorithm is the development of a segmentation scheme that adapts the importance of color and texture in the merging process based on the local color uniformity of the regions evaluated. This process is controlled by the weights w_1 and w_2 as follows:

$$MI(r_1,r_2) = w_1 * KSM(TD_1,TD_2) + w_2 * KSM(CD_1,CD_2) \quad (2)$$

where r_1 and r_2 are the adjacent regions under evaluation, w_1 and w_2 are the weights for texture and color distributions, KSM defines the Kolmogorov-Smirnov Metric, TD_i is the texture distribution for region i and CD_i is the color distribution for region i. The weights w_1 and w_2 are calculated as follows:

$$K_i = \frac{\arg\max_C(CD_i)}{N_i},\ K_i \in (0,1],\quad i = 1,2 \quad \text{and} \quad w_2 = \frac{\sum_{i=1}^{2} K_i}{2},\quad w_1 = 1 - w_2 \quad (3)$$

where N_i is the number of pixels in the distribution CD_i. To further refine the boundaries between the regions resulting after merging we applied a pixel-wise classification procedure that exchanges the pixels situated at the boundaries between various regions. For more details about the image segmentation algorithm the reader can refer to [10].

3 Experiments and Results

To evaluate the performance of the proposed algorithm, we use six representative skin lesion images depicted in Figs. 2 and 3. It can be noticed that the boundaries of some lesions are not well defined since parts of melanoma have characteristics of healthy skin tissue. To be able to determine the accuracy of the developed algorithm we constructed the ground truth by tracing manually the outline of the melanoma (see Fig. 4).

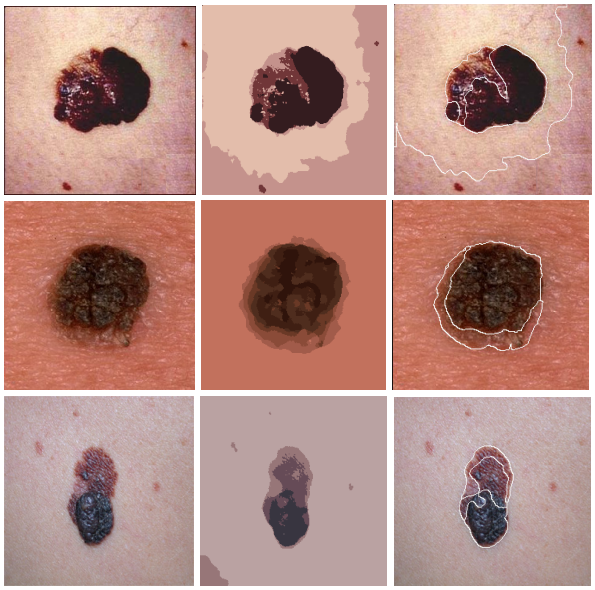

Fig. 2. Skin cancer segmentation results: First column: Original images (from top to bottom M1, M2, M3). Second column: Color segmentation results. Third column: Image segmentation results.

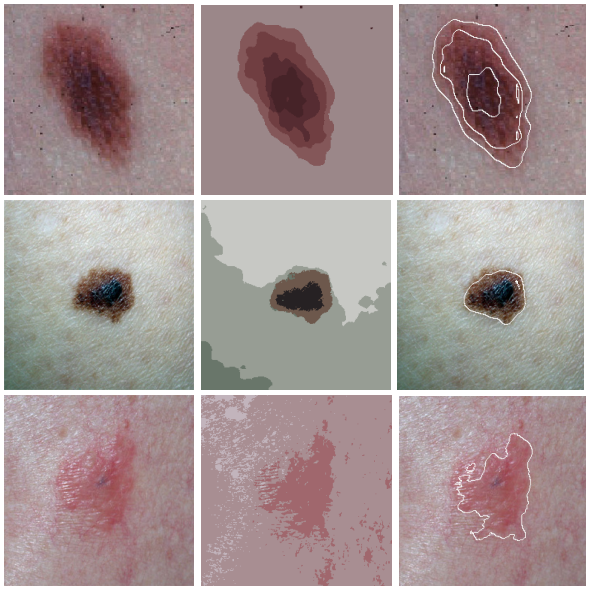

Fig. 3. Skin cancer segmentation results: First column: Original images (from top to bottom M4, M5, M6). Second column: Color segmentation results. Third column: Image segmentation results.

The accuracy of the algorithm is determined by computing the minimum Euclidian distance between the white pixels situated on the border of the lesion in the ground truth image and the white pixels from the border of the image segmented by our algorithm. To evaluate the errors we calculated the standard deviation, mean and RMS errors to measure the border displacement between the ground truth and the segmented skin images. The experimental data is depicted in Table 1 and it can be noticed that accurate results were obtained for the first five-melanoma images (M1 to M5), while the errors obtained for M6 image are larger. This is motivated by the fact that the interface between the melanoma and healthy skin is very fuzzy and this made difficult to trace precisely the outline of the skin cancer.

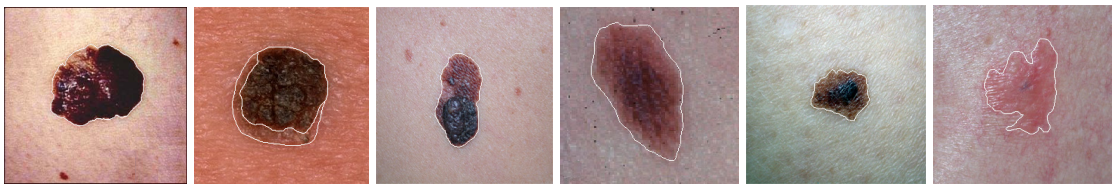

Fig. 4. Ground truth (manual annotation) for skin cancer images M1 to M6 depicted in Figs. 2 and 3.

Melanoma Image	Mean	Standard Deviation	RMS
M1	1.015	0.683	1.226
M2	1.561	1.202	1.971
M3	0.861	0.722	1.124
M4	1.684	1.405	2.194
M5	1.856	1.708	2.522
M6	3.981	3.525	5.318

Table 1. Point to curve errors between the ground truth and the automatic image segmentation algorithm. The mean and RMS errors are given in pixels.

Conclusions

The aim of this paper is to present a novel algorithm for segmentation of skin cancer images by evaluating adaptively the color and texture information. The main novelty of this approach is the development of an adaptive spatially coherent color-clustering scheme (A-SKM) that is included in the implementation of a color texture segmentation algorithm. The resulting color-texture algorithm proved to produce accurate segmentation of low-resolution skin cancer images that are defined by large color and texture non-uniformities. This research is on-going and we plan to investigate ways of improving accuracy and to evaluate the performance of our algorithm when applied to large collections of skin cancer images.

Acknowledgments

We would like to express our gratitude to Science Foundation Ireland (SFI) for supporting this research.

References

[1] NIH Consensus Conference (1992). Diagnosis and treatment of early melanoma, *JAMA* 268:1314-1319.

[2] http://www.polartechnics.com.au/Products/SolarScan/SolarScan.htm

[3] Barzegari M, Ghaninezhad H, Mansoori P, Taheri A, Naraghi Z and Asgari M (2005). Computer-aided dermoscopy for diagnosis of melanoma. *BMC Dermatology,* doi:10.1186/1471-5945-5-8.

[4] Celebi M, Aslandogan Y, Stoecker W, Iyatomi H, Oka H, and Chen X (2006). Unsupervised border detection in dermoscopy images. *Skin Research and Technology* (to appear).

[5] Binder M, Steiner A, et al (1994). Application of an artificial neural network in epiluminescence microscopy pattern analysis of pigmented skin lesions: a pilot study. *Br J Dermatol,* 130:460-465.

[6] Ganster H, Axel P, Roehrer R, Wildling E, Binder M and Kittler H (2001). Automated melanoma recognition. *IEEE Trans. Medical Imaging*, 20(3):233-239.

[7] Xu L, Jackowski M, Goshtasby A, Roseman D, Bines S, Yu C, Dhawan A, Huntley A (1999). Segmentation of skin cancer images. *Image and Vision Computing*, 17:65-74.

[8] Ojala T, Pietikainen M and Maenpaa T (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell*. 24(7): 971-987.

[9] Perona P and Malik J (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Machine Intell*., 12:629-639.

[10] Nammalwar P, Ghita O and Whelan PF (2004). Integration of feature distributions for colour texture segmentation. *17th International Conference on Pattern Recognition*.1:716-719