How robust is the SVM wound segmentation?

Marina Kolesnik[1] *and Aleš Fexa*[2]

[1]Fraunhofer Institute for
Applied Information Technology
Schloss Birlinghoven, 53604 Sankt Augustin
GERMANY
Tel. +49-2241 14 3421, Fax: +49-2241 14 1506
E-mail: marina.kolesnik@fit.fraunhofer.de
URL: http://viswiz.imk.fraunhofer.de/~marina

Institute for Algorithms and Cognitive Systems
Department of Computer Science
Universität Karlsruhe
76128 Karlsruhe
GERMANY
E-mail: fexa@iaks.uni-karlsruhe.de

ABSTRACT

This paper investigates the robustness of automatic wound segmentation. The work builds upon an automatic segmentation procedure by the Support Vector Machine (SVM) - classifier presented in [8],[9]. Here we extend the procedure by incorporating textural features and the deformable snake adjustment to refine SVM-generated wound boundary. The robustness of SVM-based segmentation is tested against different feature spaces using a long sample of training images featuring a broad variety of wounds' appearance. Recommendations drawn from these experiments provide a useful guideline for the development of a software support system for the visual monitoring of chronic wounds in wound care units.

1. INTRODUCTION

Automatic detection of wound surface area in images is a desirable tool for clinicians involved in chronic wound care. Although simple in its formulation, the problem proved to be challenging for a computer. Neither algorithmic approaches, employing the existing color image processing techniques, nor attempts to develop a dedicated image acquisition system, .could not yet produce robust solution.

The list of image processing techniques probed for wound segmentation comprises a broad bunch of algorithms starting with relatively simple edge-based grey-scale segmentation and going the whole way up to sophisticated statistical analysis of color and multi-spectral images. Edge-based algorithms build upon the existing edge detection algorithms, c.f. Canny, [2] and use the detected wound boarders for segmentation of wound region [1]. Despite of optimistically low magnitudes of erroneous segmentation of up to 7% reported in [2], it is doubtful that the edge-based methods could be reliable in general as images featuring wound and skin typically exhibit numerous edge segments in the vicinity of the true wound boundary. This high amount of false edges may also be confusing for the deformable snake algorithm, applied for wound segmentation in [5], [6]. With no a priory input helping to discriminate between the false edge segments and the wound boarder, it is difficult to draw an initial contour required by the snake algorithm. Another family of wound segmentation approaches exploits regional features based on either color or texture, or their combination. Final decision on the segmentation is then taken using various classification techniques such as conditional thresholding [3], [4], neural networks [5], case-based reasoning [7], and support vector machines [8], [11].

Another outstanding aspect related to the development of new wound segmentation techniques is the lack of an accepted method for objective assessment of how efficient any new segmentation technique is. Accuracy of automatic segmentation is typically compared against manual one in terms of bias (offset of the mean of multiple measurements) and precision (repeatability or variation) [6] or by computing percentile rate of misclassified pixels [9]. Different segmentation techniques have never been a subject to a direct comparison among each other in terms of acquisition conditions, image quality, tolerance against different wound and skin appearances, processing time, or a level of required human intervention.

This work continues the development of the SVM-based wound segmentation algorithm introduced in [8]. Input features to the SVM in this algorithm are generated by a recursive sampling of multi-dimensional color histograms. The efficiency of different sampling techniques for wound segmentation was investigated in [9], where it was shown that the recursive sampling of 3-D color histograms employed in feature generation produces better segmentation results as compared to the sampling of one-dimensional histograms. In the work here we extend the multi-dimensional sampling approach used for the generation of color features towards incorporation of two textural features (Section 2). We evaluate the efficiency of wound segmentation in a series of experiments with a rather large training sample of wound images (Section 3). We furthermore refine the SVM-generated wound contour through a multiscale snake adjustment algorithm (Section 4). Conclusions summarize our experience in automatic wound segmentation and give recommendations for a semi-automatic support system for wound surface evaluation from images.

2. FEATURE GENERATION

2.1 Color features

The right choice of feature space is crucial to the SVM-classifier for seamless separation of two classes. The reader is referred to [9] for a detailed account on the sampling of 3-D color histograms used for generation of input feature vectors for the SVM-based segmentation.

2.2 Textural features

Although drawn from image contrast variations, textural features have the advantage of being entirely independent from color features. Given a collection of textural features, which separates well on wound and non-wound region, would greatly increase the robustness of SVM.-based segmentation.

Investigation across a wide variety of different wound images has identified two textural features that exhibit robust separation of the wound surface area from background. These are a *Local Contrast* variation (LC) and a *Local Binary Pattern* (LBP). Both textural features have been computed for grayscale images and in a local window as follows.

Local contrast variation records the difference between an average value of pixels above the central value and an average value of pixels below the central value:

$$LC = (P - N)/(P + N)$$

where P is a local average value of all pixels whose gray values is greater or equal than that of the gray value of central pixel, and N is a complementary average gray value of remaining pixels whose brightness is less than the brightness in the center. In case N is equal to zero, 0.8 (the average contrast) has been assigned as the contrast value LC. The size of the local window employed is about 9x9 pixels.

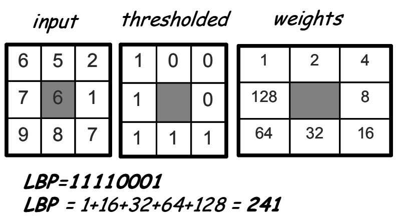

Fig.1. Computation of the LBP value.

The principle of the *Local Binary Pattern* is depicted in the Fig. 1. The LBP-operator thresholds a 3x3 neighborhood by the value of the central pixel followed by the summation of the resulting binary pattern multiplied by corresponding binary weights. The advantage of the LBP is its invariance against monotonic brightness variations across image spatial domain. We use a generalized version of the LBP-operator called LBP_{16}^{riu2} [10], which is computed in a 5x5 neighborhood using interpolated brightness values and is rotation invariant.

Given the fact that the sampling of 3-D color histograms generates robust separation of color features, we extent the histogram sampling approach by incorporating above textural features. First, a 2D texture histogram using the LC and LBP_{16}^{riu2} textural features is generated. Next, the multidimensional histogram sampling is applied for the sampling of the LC-dimension into 4 equal-size bins and the LBP-dimension into eighteen bins. Due to uneven distribution of 18 discreet LBP values, - about 30% of pixels typically exhibit the largest possible value of 17, the LBP sampling was carried out manually into the following set of 4 bins: {{0, 1, 2, 3}, {4, 5, 6, 7, 8, 9, 10, 11}, {12, 13, 14, 15, 16}, {17}}.

3. SVM-BASED WOUND SEGMENTATION

3.1 SVM training

A total of 50 wound images representing different wound types have been employed for the training of the SVM classifier. A wound region in each of these images has been manually segmented. Three training runs have been carried out, each one employing one of the following feature set: 1) color features only, 2) textural features only, and 3) the combination of color and textural features. The training images were captured by a dedicated image acquisition device consisting of the Olympus C-2500L camera surrounded by a ring of LED lights, which provided homogeneous illumination of a wound surface. The training wound images were of different sizes cropped from original color images of 1280x960 pixels in size.

3.2 Quality evaluation

In our experiments, we use the trained SVM for the segmentation of new wound images, i.e. the ones not employed in the training. Each SVM generated wound segment has been compared against the corresponding manually drawn wound segment by reckoning the number of misclassified pixels. The quality of SVM segmentation is measured as an average number of erroneously classified pixels from wound region and that ones from background (see [9] for details).

3.3 Segmentation results

Three segmentation trials have been carried out using the three SVM classifiers trained in the different feature spaces as indicated in 3.1. Each SVM classifier has been employed for the segmentation of 23 new wound images. These images were captured by the same acquisition device as the training images.

Table 1 illustrates the quality of segmentation in the three trials. The table gives the average magnitude of misclassified pixels in the segmented sequence of the 23 images as well as the encountered minimal and maximal error magnitudes.

	Color	**Texture**	**Color&Text.**
Average	6,56	22,16	5,80
Max.	23,77	52,12	30,45
Min.	0,66	1,38	0,47

Table 1. Percentile error magnitudes resulting from the SVM segmentation in the color, texture and the combined color & texture feature space.

Fig. 2. Wound image overlaid with the SVM-generated contour. Error magnitude is 0,47%.

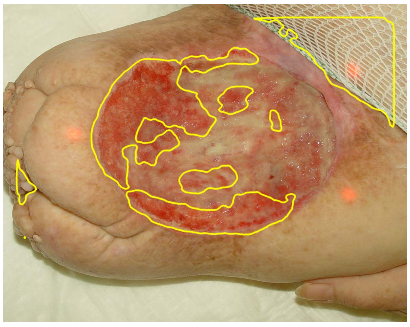

Fig. 3. Wound image overlaid with the SVM-generated contour. Error magnitude is 30,45%.

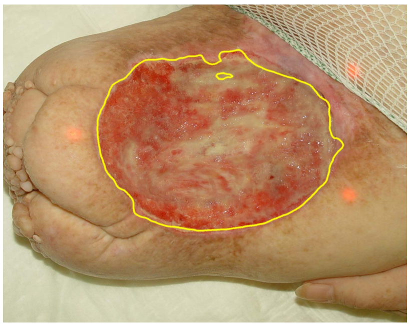

Fig. 4. Wound image overlaid with the SVM-generated contour. Error magnitude is 1,25%.

As one can see from the error magnitudes, the average robustness of segmentation increases in case of the combined color and texture feature space. This is despite of the fact that high error magnitudes are observed if only the textural features are employed. One notes, however, that the quality of segmentation varies dramatically with the error magnitudes ranging from as low as 0,47% and up to 30,45%. These two extreme segmentations are illustrated in Fig. 2 and Fig.3, respectively. The figures show wound images together with the computer-generated wound contours. The SVM segmentation in Fig.3 basically fails. This is most likely due to misleadingly yellow color of wound's matter, which also obscures the textural features. Wound bandage ads to the confusion of textural features too. As the wound image in Fig. 3 proved to be particularly difficult for the SVM segmentation, we have conducted an additional investigation of this case. We have trained the SVM using this single wound image and segmented this image by the trained SVM classifier. The result in Fig. 4 shows a considerable improvement in the quality of segmentation.

4. WOUND CONTOUR REFINEMENT

One certain conclusion can be drawn based on the above experiments: the SVM-based segmentation cannot generate as fine wound contour, as a human would do. Even in case of robust segmentation with low error rate, the SVM-generated contour does not follow a wound boundary exactly. To compensate this drawback, we have tested the capability of snake adjustment approach [12] for refining the SVM-generated contour.

The snake algorithm adjusts a given initial contour to the underlying intensity gradient of the wound image. It is straightforward to use the SVM-generated contour as the initial one in the snake adjustment algorithm. The problem however is that a typical wound image exhibits numerous edges with sufficiently large intensity gradient near the true wound boundary. In the presence of these "false" edges, the snake contour adjustment attracted towards these edges becomes unstable. One possible solution stabilizing the contour adjustment is a multiscale modification of the snake adjustment algorithm using image pyramid.

The multiscale snake adjustment begins the processing on a high level of the Gaussian pyramid with coarse resolution and gradually proceeds towards the pyramid level of original size. First, the Gaussian pyramid of image gradient is generated. Next, the initial contour is computed by projecting the SVM-generated wound contour onto a starting pyramid level of smallest size. The snake adjustment is then applied on the starting pyramid level generating a coarsely adjusted contour. Finally, the adjusted contour is projected onto the next pyramid level and used as the initial contour in the next round of hierarchical snake adjustment. This "coarse-to-fine" strategy is much more likely to converge on a true wound contour. Fig. 5 illustrates the refined wound contour using the multiscale snake adjustment.

Fig. 5. An example of wound image with the overlaid contour resulted from the SVM segmentation (left) and with the refined contour using the multiscale snake adjustment (right).

5. CONCLUSIONS

We have presented the SVM wound segmentation procedure followed by the refinement of the SVM-generated wound contour using the deformable snake adjustment. The robustness of the SVM segmentation has been tested with the sample of 23 wound images not employed in the training of the SVM classifier. All images have been captured by the same acquisition device. The magnitude of segmentation error is 5.8% on average, although the maximum error recorded is 30,45% for the texture & color feature space and 23,77% for color features. It is proper to ask which magnitude of the segmentation error is perceived as acceptable by a human observer. Since the segmented wound area is not uniquely associated with the error, i.e., many different segmentations can produce the same error magnitude, this very same error magnitude invokes different visual assessment of the quality of segmentation. This outstanding issue requires an additional investigation before a robust method of comparing multiple segmentations can be suggested. Our observations indicate that segmentation errors ranging within 5% and 7% are perceived as being of acceptable quality, whereas error magnitudes below 3% count as being very good.

Our experiments suggest that neither choice of the feature space is capable of making the SVM segmentation stable for any new wound image. The combination of color and texture features reduces the average magnitude of the segmentation error as compared to the use of only color features. Still, it does not ensure the generated segment to be of acceptable quality for any incoming wound image. The followed snake refinement step is only successful if the initial SVM contour lies in the vicinity of the true wound boundary. Consequently, the current segmentation procedure may fail and it therefore will not be useful to the clinician in its fully automatic form.

Of course, there still remains a question whether a much larger training sample than that one used in our experiment would increase the robustness of the SVM segmentation so as it will guarantee that maximum segmentation error is below the indicated quality margin of about 6%. Currently however, we believe that semiautomatic procedures (c.f. [3], [13]) have an advantage over any fully automatic segmentation because the user would be offered an option to correct the automatic wound contour if this fails at some locations. In a semiautomatic system, the user would control the final segmentation decision thus excluding the possibility of total segmentation failure. For instance, a semiautomatic system based on the SVM segmentation may have an option to retrain the SVM classifier interactively based on user's input. It is sensible to train the SVM classifier with an image sample of a single patient featuring the development of particular wound. Our experiments show that the SVM segmentation of any new image of the same wound is highly reliable with the error magnitude below 2%.

REFERENCES

[1] Z. Zhang, W. V. Stoecker, andR. H. Moss, Boarder detection on digitized skin tumor images", in *IEEE Trans. Med. Imaging,* 19(11), 2000, pp.1128-1143.

[2] S. Lankton, "Wound Measurement: Edge Based Segmentation for Medical Instrumentation", Report # ECE4580, Georgia Institute of Technology. Available: *http://shawn.lankton.com/projects-wound.htm.*

[3] A.A. Perez, A. Gonzaga, and J.M. Alves, "Segmentation and Analysis of Leg Ulcers Color Images", in *Proc. of the Int. Workshop on Medical Imaging and Augmented Reality (MIAR'01),* 0-7695-1113-9/01 © 2001 IEEE.

[4] B. Acha, C. Serrano, and J.I..Acha, "Segmentation and classification of burn images by color and texture information", *Journal of Biomedical Optics,* 10(3), May/June 2005, pp. 034014-1-034014-11.

[5] T.D. Jones, "Improving the precision of leg ulcer area measurement with active contour models", *PhD thesis,* University of Glamorgan, 2000.

[6] P. Plassmann and T. D. Jones, "Improved active contour models with application to measurement of leg ulcers", *Journal of Electronic Imaging 12(2),* April 2003, pp. 317-326.

[7] M. Galushka, H. Zheng, D. Patterson, and L. Bradley, "Case-Based Tissue Classification for Monitoring Leg Ulcer Healing", *Proc. of the 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05),* 1063-7125/05 © 2001 IEEE.

[8] M. Kolesnik and A. Fexa, "Segmentation of wounds in the combined color-texture feature space", in *Proc. of SPIE Medical Imaging 2004: Image Processing,* 16-19 Februar, 2004, San Diego, CA, pp. 549-556.

[9] M. Kolesnik and A. Fexa, "Multi-dimensional Color Histograms for Segmentation of Wounds in Images"**,** *in Proc. of the Int. Conference on Image Analysis and Recognition (ICIAR'2005),* Toronto, Canada, September 28-30, 2005, *LNCS* Vol. 3656/2005, pp.1014-1022.

[10] Timo Ojala, Matti Pietikäinen and Topi Mäenpää. Grey Scale and Rotation Invariant Texture Classification with Local Binary Patterns, *Proc. Sixth European Conference on Computer Vision, Dublin, Ireland,* 1:404-420, 2000.

[11] B. Belem, "Non-Invasive Wound Assessment by Image Analysis", *PhD thesis,* University of Glamorgan, 2004.

[12] E. Trucco and A. Verri, "Introduction Techniques for 3-D Computer Vision", 1998 © Printice-Hall.

[13] H. Odunchu, A. Hoppe, M. Clark, R. J. Williams, and K.G. Harding, "Analysis of Skin Wound Images Using Digital Color Image Processing: A Preliminary Communication", *Lower Extremity Wounds,* 3(3), 2004, pp. 151-156.