Multi Dimensional Color Histograms for Segmentation of Wounds in Images

Marina Kolesnik[1], Ales Fexa

Fraunhofer Institute for Media Communication, Schloss Birlinghoven, D-53754 Sankt-Augustin, Germany

Abstract. The work investigates the use of multi dimensional histograms for segmentation of images of chronic wounds. We employ a Support Vector Machine (SVM) classifier for automatic extraction of wound region from an image. We show that the SVM classifier can generalize well on the difficult wound segmentation problem using only 3-D dimensional color histograms. We also show that color histograms of higher dimensions provide a better cue for robust separation of classes in the feature space. A key condition for the successful segmentation is an efficient sampling of multi-dimensional histograms. We propose a multi-dimensional histogram sampling technique for generation of input feature vectors for the SVM classifier. We compare the performance of the multi-dimensional histogram sampling with several existing techniques for quantization of 3-D color space. Our experimental results indicate that different sampling techniques used for the generation of input feature vectors may increase the performance of wound segmentation by about 25%.

1. Introduction

Chronic skin wounds affect many people and take a long time to heal. Systematic measurement of the physical dimensions of a chronic wound is an excellent way to record the progress of healing. Normal practice of wound care includes weekly check-up of a patient at which an image of wound is acquired. A clinician draws a contour around the wound and assesses its size by comparing contours in subsequent images. This is a time consuming and subjective process. The work here attempts at developing an automatic procedure for *automatic segmentation* of wound region in wound images.

Even for restricted instances of wound image segmentation, the use of simple features is not sufficient for reliable differentiation of image pixels onto different classes. An efficient separation of classes can be achieved if features are derived from various histograms counted in a local neighborhood of image pixels [1]. Further improvement is obtained if multiple local histograms are linked together thus resulting into a single multi-dimensional histogram. Feature space generated by a sampling of such multi-dimensional histogram provides most efficient local description of image pixels.

Several methods for the histogram sampling have been suggested in the literature. Chapelle et al. [2] downsize the original color range with 255 bins down to 16 bins. This reduces the size of any 3-dimensional color histogram to 16^3=4096. Experiments with a smaller number of bins have produced worse image classification, whereas a larger number of bins have not been tested on the ground of limited computational resources. Pietikäinen et al. [3]

[1] To whom correspondence should be addressed: marina.kolesnik@imk.fraunhofer.de

apply a more advanced approach by dividing each color dimension into *N* bins with an equal number of entrees. They confirm that best classification accuracy is obtained using the sampling into 16x16x16 bins. This kind of sampling was found to be efficient for wound segmentation [4]. However, both methods sampled each color dimension independently, which cannot be fully justified because it does not take into account the inherent dependency of different color dimensions in natural objects.

In this work we employ 3-D color histograms to generate a set of features, which are then used as input to the Support Vector Machine Classifier (SVM) [5], [6]. We show that a single histogram of higher dimensions provides a better description of pixels of one class than a collection of several 1-D histograms. Also, the role of different sampling techniques cannot be underestimated. Our experiments indicate that different sampling techniques make a profound impact on the quality of wound segmentation.

The paper is organized in six sections. We start by introducing the multi-dimensional *Histogram Sampling* (Section 2). We proceed by describing SVM^{light} implementation classifier and generation of input feature vectors. (Section 3). Next, we look at performance of the SVM classifier by conducting experiments when using a single 3-D color histogram versus three 1-D histograms, and different sampling techniques (Section 4). Discussion of segmentation results concludes the paper (Section 5).

2. Histogram Sampling

We distinguish between the 3-D color space and a higher dimensional feature space, which consists of feature vectors attributed to image pixels. The *Histogram Sampling* as introduced [3] uses a normalized 1-D histogram *M* of an image, or an averaged histogram of a set of images, and samples it into a number of *L* bins, each one constituting an equal fraction of pixels *1/L:*

$$\sum_{k=l-1}^{l} M(k) = \frac{1}{L}$$

for all bins *l=1,…,L*. Note that such a sampling automatically gives a denser bin distribution for those histogram parts with larger number of elements thus providing an optimal sampling of histogram entries into bins.

Let *H* be a local histogram *H* computed in a neighborhood of pixel *(i,j).* Next, *H* is sampled into the same set of bins, *L*. A number of histogram elements falling into each bin defines one feature per bin. Thus, elements falling into first *L-1* bins (the last bin is excluded as dependent on the previous ones) define *L-1* features, which, when taken together, form a point in the (*L-1)*-dimensional feature space. The point coordinates define a feature vector associated with the pixel *(i,j).* Fig. 1a shows an example of the *Histogram Sampling*.

An extension of *Histogram Sampling* for the case of multi dimensional histogram is built upon recursive sampling. Consider 3-D color space in which *M* being a 3-D image histogram. Let M_1 be a selected image histogram of one color dimension. Next, M_1 is being sampled into *L* equal sized bins. Elements E_l of M_1, falling into each bin, *l*, form a set of entries for the computation of 1-D histogram of next color dimension. Let histogram M_{11} be generated. Next, M_{11} is sampled into *L* equal sized bins, too. The process of sampling is repeated recursively for all bins and three dimensions of the color space. It generates a total number of L^3 bins which gives rise to a set of L^3*-1* features. These features form a feature

vector attributed to a central pixel of window used for the computation of local histogram. The multi-dimensional *Histogram Sampling* can be easily extended to a general case of *N*-dimensional histogram in which case a set of L^N*-1* features would be generated. Fig. 1b illustrates the *Histogram Sampling* in case of two dimensions.

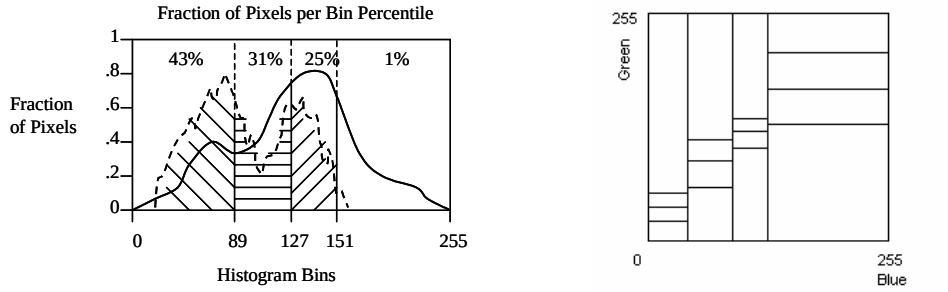

Figure 1. **a)** Histogram Sampling into four bins. The image histogram (solid line) is sampled into four equal sized bins. The local histogram (dashed line) is sampled into same four bins. The Histogram Sampling defines three features indicating a fraction of entries falling into three first bins: 43%; 31% and 25% in our example. **b)** Schematic illustration of the multi-dimensional Histogram Sampling in case of two dimensions. Blue and Green channels are used for the generation of 15 color-based features.

3. SVM-based wound segmentation in the color feature space

SVM is an approach for supervised classification of data into two classes [6]. In this work we use SVM^{light} - implementation of the SVM classifier available for research application at http://svmlight.joachims.org/ [7]. SVM classification is performed in two stages. The aim of the first, *training stage*, is to find an *optimal separating hyperplane* which divides the set of test examples into two classes. Note that each test example has to bear a label of either class. During the second, c*lassification stage*, each input point is attributed a label according to the side this point appears with respect to the hyperplane. A more detailed account of SVM's is out of the scope of this paper and here we will only discuss those aspects of the Training and Classification Stage, which are specific for wound segmentation.

Input to the SVM^{light} is a set of feature vectors attributed to selected image pixels. We use manually segmented images of wounds to compose a *training set* of feature vectors attributed to pixels of wound and non-wound class. Our experiments suggest that a balanced contribution of feature vectors from two classes improves the quality of classification. We therefore select an approximately equal number of evenly distributed pixels from across the wound and non-wound regions for the generation of input for the training stage.

SVM^{light} offers three optional kernels such as linear, polynomial and radial one. In our experiments the radial kernel performed best followed by the polynomial and linear one. These results are consistent with earlier experiments on image classification [8].

The choice of feature space is crucial for the performance of the SVM classifier. In a "good" feature space, input elements originating from either wound or skin class, would

form two volume clusters, which are widely separated from each other and easy to classify. In spite of the fact, that the color histogram technique is a very simple method, it has shown good results for image indexing and segmentation [1]. Below we investigate the impact of two factors on the performance of the SVM classifier: 1) the use of three 1-D color histograms versus a single 3-D color histogram for the generation of input feature vectors and 2) different quantization techniques employed in the histogram sampling.

4.1 Computation of feature vectors

A training set of input feature vectors for the case of N wound images and 3-D color histogram is obtained as follows. For every image, about 2000 evenly distributed pixels are selected from the wound region and approximately the same number from outside of the wound region. This gives rise to a set of about 4000 pixels for each image. Pixels from image background are counted as belonging to "not a wound class". Pixels from a boundary region around the wound of about 12 pixels wide are not selected so as to exclude a confusing mixed wound/skin region. A 3-D color histogram is computed for every image using the selected pixels. A *3-D average histogram* is generated by summing up all the 3-D color histograms and dividing the sum by the number of images, N. This *average histogram* is sampled into 64 bins using the multi-dimensional *Histogram Sampling.*

The ordering of color dimensions in the RGB space used for the recursive sampling depends on the level of differentiation between wound and skin provided by these dimensions: color dimensions with higher differentiation are sampled first. This resulted into the blue/green/ red-ordering of dimensions for the recursive sampling.

Computation of feature vectors for SVM segmentation utilizes the bins resulted from the *Histogram Sampling* of the *average histogram*. Computation of feature vectors for each pixel of an image to be segmented takes the following steps: 1) Generation of 3-D *local histogram* in a local window of about 75x75 pixels; 2) Sampling the *local histogram* into 64 bins resulted from the sampling of the *average histogram*; 3) Composing a 63-element feature vector out of entries of the *local histogram* falling into the first 63 bins.

4. Experiments and results

We present several experimental trials each one testing a specific way of computation of input feature vectors. Six images of different wound types were used for the training of the SVM in each experimental trial. Each of these images was then segmented using the trained SVM. As required by the training, the images were manually segmented onto a wound and "non wound" region.

The quality of segmentation in each trial was measured by counting an average rate of erroneously classified pixels as follows. Let W_m be a number of wound pixels and let S_m be a number of "non-wound" pixels in a manually classified image. Similarly, let W_c and S_c be a number of pixels classified as wound and "non-wound" in a computer-segmented image, respectively. Error rate for misclassified wound pixels is given by the normalized intersection of the manual wound segment and the computer-generated non-wound segment:

$$E_{wound} = (W_m \bigcap S_c) / W_m$$

where $\bigcap$ denotes the intersection of two sets. Similarly, the error rate for misclassified "non-wound" pixels is given by the normalized intersection of the manually defined non-wound segment and the computer-generated wound segment:

$$E_{skin} = (S_m \bigcap W_c) / S_m$$

Full classification error is then given by the sum of the above error rates:

$$Error = (E_{wound} + E_{skin}) / 2$$

5.1 Single multi-dimensional histogram versus multiple 1-D histograms

A series of experiments tested the quality of segmentation when sampling 1) three 1-D histograms and 2) a single 3-D color histogram. In the first experimental trial each 1-D histogram was sampled into 22 bins generating 63 input features per pixel. In the second trial each dimension of the 3-D color histogram was sampled into 4 bins giving rise to 63 features. The generated features were firstly used for the training of the SVM classifier on six wound images and, after that, for the segmentation of each one of these six images, independently. All other segmentation parameters used by the SVM classifier were kept identical in all experiments.

Image	3 x 1-D	3-D
1	0.63	0.36
2	1.74	1.46
3	1.25	1.38
4	2.41	1.93
5	1.86	1.56
6	0.61	0.47
Average	1.42	1.19

Table 1. The error percentile, *Error*, of erroneously classified pixels resulted from the segmentation of six images. Input feature vectors are generated by the independent sampling of 1-D histograms (column 3 x 1-D), and the sampling of 3-D color histograms (column 3-D).

Segmentation results in Table 1 display a convincing advantage in the performance of the SVM classifier for the case of 3-D color histogram over a corresponding collection of three 1-D histograms. We extent this conclusion to a more general statement: a better performance of multi-dimensional histogram can be explained by the fact that the sampling in the 3-D color space indeed exploits the inherent dependency of color dimensions usually shown by complex natural objects. This is especially true for the human skin. Light remitted from skin is a complete spectrum. Consequently, the 3-D color histogram is a gross approximation of the true remitted light. The composition of the spectrum for the skin depends on the mixture of scatters and absorbers in the skin, each one of these affecting each primary dimension of the spectrum. This results into correlation of R, G and B parts of the RGB histogram, which is exploited by the multi-dimensional Histogram Sampling.

5.2 Comparison of different sampling techniques

In order to perform an independent testing of the multi-dimensional *Histogram Sampling* technique, we have conducted a series of experiments in which other quantization methods were involved. Five sampling techniques have been tested:

1. *Independent Sampling* (IS) performs independent sampling of each color histogram into *N* bins with an equal number of histogram entrees [4].
2. *Learning Vector Quantization* (LVQ) performs the quantization of unlabeled data vectors into a smaller set of codebook vectors. Each data vector is then represented by its nearest codebook vector. An initial set of random codebook vectors is trained so as to minimize the error of misclassification of data vectors. We use an optimized LVQ1 training algorithm [10] for the quantization of 3-D histogram.
3. *Vector Quantizer Design* (LBG-VQ) [9] is a lossy data compression method based on the principle of block coding. The reason of applying the LBG-VQ for the sampling of 3-D histogram is similar to the motivation of any image compression algorithm, namely, the need to downsize an original dataset by extracting most important information while leaving out the rest. We use the LBG-VQ algorithm for the coding of wound images. Code vectors characterizing image pixels are used for generation of input features: each image pixel is attributed a feature vector that gives a fraction of occurrences of coding vectors in a local window. 63 coding vectors used by the LVG-VQ compression give rise to a same number of 63 features composing the elements of a feature vector.
4. *Random Density Estimation* (RDE) employs the *Voronoi Diagram* which, in our case, is a partition of color space into Voronoi cells, each of which consists of elements closer to one particular object than to any others. The advantage of RDE is that the shape of Voronoi cells varies with the density of elements of the 3-D color histogram. Because most of the histogram elements are concentrated within an ellipsoid of revolution around the axis R=G=B, one would expect that flexibly shaped Voronoi cells could “better” partition the area within the ellipsoid than the square-shaped bins of the *Histogram Sampling*. The following iterative procedure was used: 1) Select *N (N=250)* random color vectors out of the elements of 3-D color histogram; 2) Construct the Voronoi Diagram using the selected color vectors; 3) Compute a number of histogram elements falling into each Voronoi cell; 4) Delete a color vector with the smallest number of histogram elements contained in its Voronoi cell; 5) Update the Voronoi Diagram down to *N-1* cells; Step 4 and Step 5 were repeated until a required number of cells *N (*in our case *N=64)* is obtained.
5. *Histogram Sampling (HS)* is the multi-dimensional *Histogram Sampling* technique (Section 3) applied to the 3-D color histogram.

IS	LVQ	LBQ-VQ	RDE	HS
1.42	7.86	1.35	1.12	1.19

Table 2. Average segmentation error for the different sampling techniques. The average error is computed over six wound images.

Table 2 shows how average error of segmentation of six wound images is affected by the use of the above sampling techniques. As evidenced by the error values, the RDE sampling provides the lowest rate of misclassified pixels. This can be explained by the fact

that the Histogram Sampling based on the Voronoi Diagram provides an optimal partitioning of elements of the 3-D color histogram.

5.3 Examples of wound segmentation

The capability of SVM classifier to segment a wound was tested with numerous images. Here we show the result of segmentation of three test images from the sequence of six ones used in the previous experiments. Input feature vectors were obtained by the sampling of 3-D color histograms. Examples of segmentation in Fig. 2 show that the SVM-classifier produces a fairly reliable segmentation of wound tissue despite of large variations in brightness of skin and quite a different appearance of wounds.

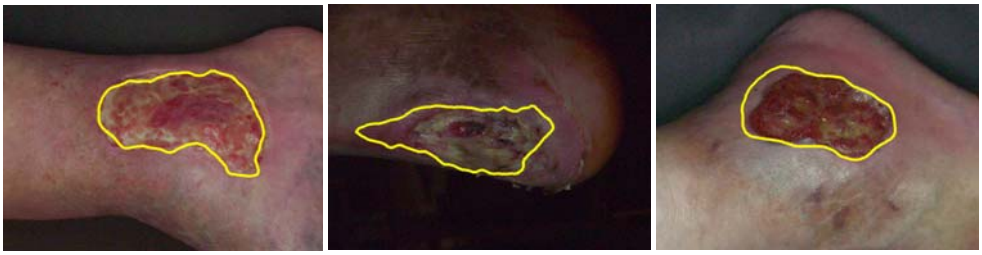

Figure 2. Three examples of wound segmentation. 63 input features were computed locally in the window of 95 pixels. The SVM classifier employed the polynomial kernel. Corresponding error rates for misclassified pixels are given in Table 1, images 1, 4 and 6, column 3-D.

5. Conclusion and future work

The *Histogram Sampling* technique generates the efficient set of feature vectors, which, when inputted into the SVM-classifier, enable the reliable segmentation of wound region in images. The generalized multi-dimensional *Histogram Sampling* of 3-D color histograms further improves the discrimination of feature vectors.

Processing time needed for the SVM training depends linearly on the number of input feature vectors (i.e. of the number of training images), but also, on their "quality" in terms of how well these can be separated into two classes. For about 2000 feature vectors from the wound class and the same amount from outside the wound region, the observed training time is of the order of 2 minutes (Pentium, 1000 MHz). If, however, the training feature vectors are not widely separated in the feature space, the convergence of searching for the support vector may become problematic.

Our experimental results indicate that the sampling of 3-D color histogram generates input features with a better discrimination than those ones obtained by the independent sampling of 1-D histograms: the quality of wound segmentation in our experiments was improved by as much as 20%-30%. It is therefore always advantageous to employ the single 3-D color histogram for the generation of input feature vectors used by the SVM for wound segmentation. Experiments with different quantization techniques have lead to an unexpected result. Although the Learning Vector Quantization technique provide a "better"

partitioning of multi-dimensional feature space in a sense that cell distribution is related to a density of space elements, the quality of wound segmentation is significantly worse. Also surprisingly, the rate of misclassified pixels resulted from the Voronoi Diagram sampling is comparable with the error rate when the multi-dimensional *Histogram Sampling* is used.

More words are to be said with regard to 3-D color histograms. Despite of their simplicity, 3-D color histograms provide an efficient cue for the description of different image objects, which are in our case, of course, wounds. Color histograms are invariant to translation and rotation and change only slowly under change of angle of view and scale. As a result, *Histogram Sampling* generates image features, which are fairly invariant to small variations in brightness and scale. The multi-dimensional Histogram Sampling is therefore provides best cumulative measure characterizing image objects locally. Note that the 3-D color histogram can be easily extended to higher dimensions by adding other discreet distributions related, for instance, to texture. Applying the *Histogram Sampling* to the extended multi-dimensional histogram would certainly generate highly efficient local description of image pixels.

A less optimistic conclusion of this work is this one: however robust and good the SVM segmentation is, it cannot produce a wound contour which is as fine as the manual one drawn by a clinician. It seems that there should be an additional and independent mechanism that complements region segmentation on a final stage of contour generation. The aim of our future research will be aimed at the fusion of two processing methods -the SVM segmentation and wound contour detection.

6. References

1. Swain, M.J., Ballard, D.H.: Color Indexing. International Journal of Computer Vision, 7:1 (1991) 11-32.
2. Chapelle, O., Haffner, P., and Vapnik, V.: SVMs for histogram-based image classification ", IEEE Trans. on Neural Networks, 9 (1999)
3. Pietikäinen, O., Nieminen, S., Marszalec, E., and Ojala T: Accurate Color Discrimination with Classification based on Feature Distributions. *Proc. of the Inter. Conference on Pattern Recognition (ICPR'96),* (1996) 833-838.
4. Kolesnik, M., Fexa, A.: Segmentation of wounds in the combined color-texture feature space. *Proc. of SPIE Medical Imaging 2004: Image Processing,* 16-19 Feb. 2004, San Diego, CA, 549-556.
5. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, New York (1995)
6. Vapnik, V.: Statistical Learning Theory. John Wiley, New York (1998)
7. Joachims, T.: Making large-Scale SVM Learning Practical. In Advances in Kernel Methods - Support Vector Learning, B. Schölkopf and C. Burges and A. Smola (eds.), MIT Press, (1999)
8. Schölkopf, B., Sung, K., Burges, C., Girosi, F., Niogi, P., Poggio, T., Vapnik, V.: Comparing support vector machines with gaussian kernels to radial basis function classifiers. A.I. Memo. No. 1599, MIT (1996)
9. Linde, Y., Buzo, A., Gray, R., M.: An Algorithm for Vector Quantizer Design, *IEEE Transaction on Communications*, pp. 702-710, January (1980)
10. Kohonen, T., Kangas, J., Laaksonen, J., Torkola, K.: LVQ_PAK: A program package for the correct application of Learning Vector Quantization algorithms, *Proc. of the Inter. Conference on Neural Networks,* Baltimore, (1992) 1725-1730.