Summary Page

Quantification of Bone Remodeling in the Proximity of Implants

Author: Hamid Sarve, Carina B. Johansson, Joakim Lindblad, Gunilla Borgefors and Victoria Franke Stenport
Address: Centre for Image Analysis, Box 337 S-751 05 Uppsala, Sweden
E-mail: hamid@cb.uu.se
Phone: +46184713471
Fax: +4618553447

Keywords: *Color Segmentation, Bone Implants, Histomorphometry*

1. **What is the original contribution of this work?**
   The framework of the method, the specific implementation to solve the particular task, and the comparison of the automatic and the manual method.

2. **Why should this contribution be considered important?**
   This work shows the benefits and drawbacks of a color based automatic method. It is compared to the manual method; for some measurements, the difference is according to our experience lower than the variation in the material and hence within acceptable error limits. However, for other measurements, the automatic method requires further development to match the manual method.

3. **What is the most closely related work by others and how does this work differ?**
   L. Ballerini et al: Image Segmentation by a Genetic Fuzzy c-Means Algorithm Using Color and Spatial Information, LCNS Volume 3005/2004, pages 260-296. This work differs in type of classifier used. Moreover, there is no comparison of the proposed method and the manual method in Ballerini et al.

4. **How can other researchers make use of the results of this work?**
   This work gives a methodology for the automatic analysis of histological samples and is of high practical value for biomaterial researchers and others needing sophisticated color segmentation.

5. **Has this work been presented/submitted elsewhere?**
   No

6. **Which form of presentation is preferred: Oral or Poster?**
   Oral

Quantification of Bone Remodeling in the Proximity of Implants

Hamid Sarve[1], Carina B. Johansson[2], Joakim Lindblad[1], Gunilla Borgefors[1], and Victoria Franke Stenport[3]

[1] Centre for Image Analysis, Swedish University of Agricultural Sciences, Box 337, SE-751 05 Uppsala, Sweden
{hamid,joakim,gunilla}@cb.uu.se
[2] Department of Clinical Medicine, Örebro University, SE-701 85 Örebro,Sweden
carina.johansson@ikm.oru.se
[3] Biomaterials Research Center, Göteborg University, SE-405 30 Göteborg,Sweden
victoria.stenport@biomaterials.gu.se

Abstract. In histomorphometrical investigations of bone tissue modeling around screw-shaped implants, the manual measurements of bone area and bone-implant contact length around the implant are time consuming and subjective. In this paper we propose an automatic image analysis method for such measurements. We evaluate different discriminant analysis methods and compare the automatic method with the manual one. The results show that the principal difference between the two methods occurs in length estimation, whereas the area measurement does not differ significantly. A major factor behind the dissimilarities in the results is believed to be misclassification of staining artifacts by the automatic method.

1 Introduction

The aim of this work is to develop an automatic quantification method for evaluation of bone implants. The use of automatic methods enables the investigator to focus more on other interesting tissue reaction topics that occurs around the implant rather than solely the percentages of bony contact and bone areas. Moreover, an automatic method yields a more objective measurement, since a manual measuring process heavily depends on the operator.

The medical device market is rapidly growing and the research field related to medical devices is of high interest today. At our laboratories we focus on several *in vivo* test models with retrieval of samples and subsequent histomorphometrical investigations. However, there is no consensus on how to perform these tests and each laboratory may have their own routines and standards. At our laboratories, evaluating the remodeling of bone tissue around medical devices/implants is performed on histologically stained un-decalcified cut and ground sections (10μm) with the implant *in situ* (the so called *Exakt* technique [1]). This technique is time and money consuming and it does not allow serial sectioning. Most often, one section of 10μm is prepared from the mid portion of,

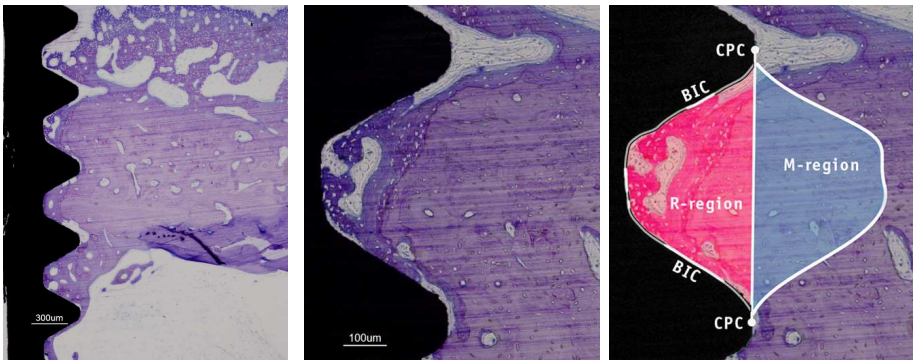

Fig. 1. Oculars of 4× (left) and 10× (middle) were used for capturing the above images. The screw-shaped implant (black), bone (purple with a number of hollow spaces) and soft tissue (light blue) are shown. Right: marked regions of interest.

for example, a screw shaped implant [2, 3]. The section is analyzed both qualitatively and quantitatively in a light microscope. The final analysis step in the light microscope involves manual measurements. This is time consuming and subjective and therefore it would be of great value if this manual measuring process could be an objective automatic method.

1.1 Background

The most important step in the automatic analysis is, as is almost always the case, the segmentation. Any reliable measurement has to be preceded by a proper and accurate segmentation of the different regions in the images. In many segmentation techniques presented for medical images, the emphasis is placed on the gray-scale pixel value [4], rather than the color information (if such information exists). However, histological staining provides significant color information that should be taken into account in the segmentation.

Color segmentations methods can be divided into four groups: pixel-based, area-based, edge-based and physics-based segmentation [5]. For this work, the focus has been on pixel-based methods. Examples of such methods are supervised or unsupervised pixel-based classifiers with several color channels as features, proposed in [6], and a method for bone segmentation which uses genetic fuzzy c-means clustering based on the color content presented in [7].

2 Material and Methods

Screw-shaped implants of commercially pure titanium were retrieved from rabbit bone after six weeks of integration. This study was approved by the local animal committee at Göteborg University, Sweden. The screws with surrounding bone

were processed in the laboratories according to internal standards and guide-lines [2], resulting in 10μm un-decalcified cut and ground sections . The sections were histologically stained prior to light microscopical investigations. The histological staining method used on these sections, i.e. Toluidine blue mixed with pyronin G, results in various shades of purple stained bone tissue: old bone light purple and young bone dark purple. The soft tissue gets a light blue stain.

For the suggested method, 1024x1280 24-bit RGB TIFF images were acquired by a camera connected to a Nikon Eclipse 80i light microscope. Two different oculars were used, 4× and 10×, giving pixel sizes 2.3μm and 0.9μm, respectively. The regions of interest (ROIs) are marked in Fig. 1: the gulf between two center points of the thread crests (CPC) denoted R (reference area); the R flipped about the line connecting the two CPCs, denoted M (mirrored area) and regions where the bone is in contact with the screw, denoted BIC. Desired quantifications involve BIC length estimation and area of different tissues in R and M; they are calculated for each thread (gulf between two CPCs) expressed as percentage of total length or area [8].

2.1 Method

2.1.1 Segmentation The segmentation was done by pixel-wise classification [9], using the low intensity of the implant and the color staining of the bone-region. The classification requires prior training which was done in a supervised manner; an expert marked the different regions using a mouse based interface and the RGB values of the corresponding region were stored in a file.

In addition to the three RGB channels, the HSV transformed channels were also included in the feature space. The RGB to HSV transformation (described in [10]) is a non-linear function, hence avoiding a linear relation between the features. It was noticed that for the H channel, the classes are not normally distributed and the distributions of the classes are overlapping to a large extent. Furthermore, this channel contains a considerable amount of noise.

The classes are assumed to be multivariate normally distributed and can be grouped in two clusters: implant and bone/soft-tissue. This is shown for intensity values in Fig. 2. At the bone/implant interfacial region there is an overlap of the implant and the bone or soft tissue. Due to the BIC estimation, these regions are of special interest. This overlap is not visible in Fig. 2, since the number of pixels in the interface is rather small.

A number of classifiers were tested on three sample images. Initially, two classifiers with the assumption of multivariate normally distributed classes were set up. The covariance matrices were assumed to be equal and unequal, generating a linear (LDA) and a quadratic (QDA) classifier [11], respectively.

The described distinct grouping can be used so that rather than segmenting all three classes at once, the segmentation is divided into two sub-segmentations with two partial classifiers. Firstly, the implant is discriminated against the bone/soft tissue, which is seen as one class. Secondly, the bone and the soft tissue are classified within the group. Three such classifiers with all combinations of LDA and QDA were tested, see Table 1 in Sec. 3.

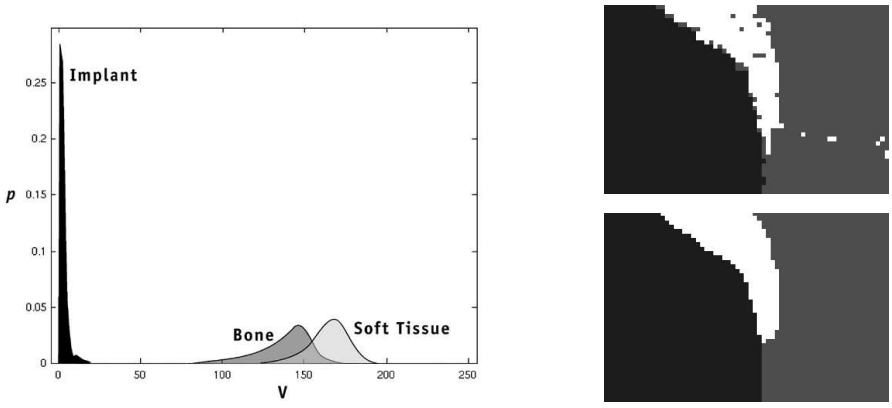

Fig. 2. Left: the distribution of the pixel in the V-channel for a sample image. Right: segmentation before (top) and after (bottom) the enhancement step. Black is the region classified as implant, gray the bone and white the soft tissue.

Some processing in order to enhance the segmentation output was performed: the implant is assumed to consist of one single piece, i.e., only the largest 8-connected implant region was kept. Moreover, it is assumed that the bone does not consist of small isolated pieces; separated bone regions with a minimum size of 7 pixels were removed. Finally, a 3x3 majority filter was applied.

2.1.2 Measurements The area measurement was done by summing up the pixels classified as bone in the R- and M-regions. In order to mask the regions, the CPCs had to be located. This was done by extracting a number of line profiles from halfway between the the thread peak (not necessary the CPC) and the lowest point in the thread gulf and up to 20 pixels below the thread peak; the centers of the line profiles were averaged and determined as the CPCs.

For the BIC-length estimation, *Koplowitz and Bruckstein*'s method for perimeter estimation of digitized planar shapes (the first method of the two presented in [12]) was used. This method requires a well defined contour, i.e. each contour-pixel shall have two neighbors only. The implant contour was derived by dilation with a 3×3 '+'-shaped structural element on the implant region in the segmentation map. The overlap between the dilated implant and the bone region was defined as the bone-implant contact. The dilation process may also yield some 'L' shapes, giving pixels with three neighbors. These pixels can be removed without affecting the length estimation. This was done by an erosion of the interface line with a 2×2 'L'-shaped structural element (in each direction). The dilation may also yield some loops. These were located by finding the pixels, that if removed, do not break the connectivity of the contour.

Table 1. Segmentations of three different images and their misclassification rates, ϵ and κ, Cohen's kappa [13]. Different classifiers were applied, quadratic discriminant analysis (QDA) and linear discriminant analysis (LDA).

	Segmentation procedure	ϵ	$1-\kappa$
(f)	1. Implant - Bone - Soft tissue: LDA	7.1%	12.2%
(e)	1. Implant - Bone - Soft tissue: QDA	6.1%	10.8%
(a)	1. Implant - Bone/Soft tissue: LDA 2. Bone - Soft tissue: LDA	5.8%	10.2%
(b)	1. Implant - Bone/Soft tissue: QDA 2. Bone - Soft tissue: LDA	5.8%	10.1%
(c)	1. Implant - Bone/Soft tissue: LDA 2. Bone - Soft tissue: QDA	6.0%	10.8%
(d)	1. Implant - Bone/Soft tissue: QDA 2. Bone - Soft tissue: QDA	6.0%	10.7%

3 Results

The different approaches for the segmentation described in Sec. 2.1 were tested on a set of three images. The averaged misclassification rates, ϵ, as well as *Cohen's kappa*, κ [13], are listed in Table 1. The misclassification rate, ϵ was computed as $\epsilon = 1 - \frac{\text{number of correct classified pixels}}{\text{total number of pixels}}$. As Table 1 shows, the average misclassification rates for these three images is lower for the classifier which used partial segmentations of the implant and bone/soft tissue; dividing the segmentation into parts yields a somewhat better segmentation result. Among the two best approaches (QDA-LDA and LDA-LDA), the QDA-LDA approach was adopted for the experiments as the first segmentation involves presumably unequal covariance matrices. It was also noticed in a test that the misclassification rate does not decrease when the H channel is included, even when the Cartesian components of the H channel, $cos(H)$ and $sin(H)$, were used as features. Therefore the H channel was excluded from the feature vector.

The automatic method was applied on a set of six images (three screw sections and the two magnifications per section). As training set for each section, manually segmented images of one thread from each of the two other sections were used. The differences between the two methods are summarized in Table 2 and the results for every individual thread of one section is illustrated in Fig. 3.

As Fig. 3 shows, the two estimators differ significantly for the BIC length estimation; this measure is consistently overestimated by the automatic method compared to the manual method. However, for the area measurement they are more similar. Moreover, the magnification does affect the result; the deviation in BIC is somewhat larger for 4× magnification.

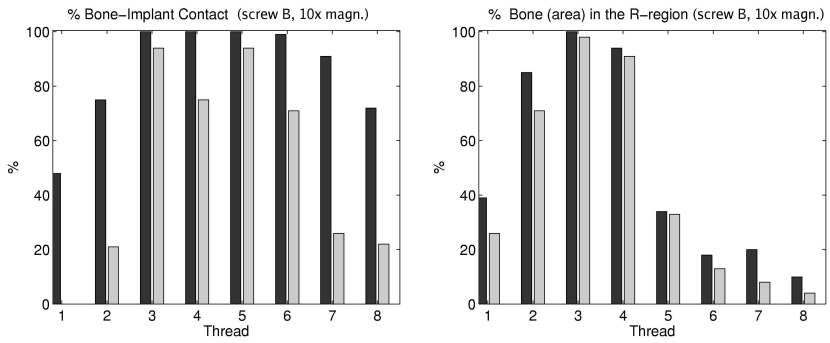

Fig. 3. The result for the two methods for one image; the percentage of BIC between two CPCs (left), and the percentage of the bone in the R-region (right) respectively. The dark bars represent the automatic method and the bright bars the manual one.

Table 2. The averaged absolute difference of the results between the methods and the averaged result of the two methods for three screw sections (eight threads/section) at two levels of magnifications. BIC_{diff} is calculated as $\frac{1}{n}\sum_{i=1}^{n}|BIC_{Ai} - BIC_{Mi}|, n = 8$, where BIC_A is the percentage of BIC between two CPCs for the automatic method and BIC_M the corresponding measure for the manual method. R_{diff} and M_{diff} are calculated in the same manner.

Section		BIC_{diff}	R_{diff}	M_{diff}	$\overline{BIC_A}$	$\overline{R_A}$	$\overline{M_A}$	$\overline{BIC_M}$	$\overline{R_M}$	$\overline{M_M}$
A	10×	17.5	4.1	1.5	61.5	47.3	41.8	44.0	43.1	40.5
A	4×	21.1	4.3	1.5	65.1	47.4	41.8	44.0	43.1	40.5
B	10×	35.3	7.0	1.9	85.7	50.0	34.6	50.4	43.0	32.8
B	4×	37.3	4.9	6.4	87.6	47.1	27.0	50.4	43.0	32.8
C	10×	11.8	3.3	2.9	59.6	55.4	58.9	48.6	52.4	56.5
C	4×	15.1	2.6	3.5	63.8	54.3	56.3	48.6	52.4	56.5

4 Discussion

As the results of the segmentation of the histological images show (see Sec. 3), the BIC length was overestimated by the automatic method. In some interfacial parts of the sections, where there was a tendency of implant loosening (rather often occurring as an artifact due to the so called "water-planning effect") the tissue was very darkly stained. Such regions may be judged as non-contacts rather than contact and with the aid of the naked-eye-interpretations in the microscope, in combination with the option of zooming in closer, it is possible to determine if it is contact or not. The automatic method, however, failed to distinguish such regions from actual bone since the intensity and the color content are very similar. This results in a non-actual BIC-line (see Table 2 and Fig. 4).

Moreover, although the manual and automatic method correspond well for the area measurement, there are diversions to a small extent in the result for this measure, such as the area of the M-region of thread 4, screw section B (see Fig. 4). As discussed in Sec. 2.1, the bone and soft tissue regions are more difficult to separate from each other than the implant and the other regions. The classification is complicated further if the training set is from another screw with slightly differently stained tissue. This results in some misclassification. However, this can possibly be avoided by including more images from different screws in the training set or by a preprocessing step.

5 Summary

An automatic (after an initial training step) method for measurement of bone area and estimation of bone-implant contact length in images of screw-shaped implants was evaluated. The method was compared with a manual method.

The automatic method uses discriminant analysis for segmentation of the bone, soft tissue and implant. Several approaches were tested. As Table 1 shows, two partial segmentations, firstly discriminating the bone against none-bone and secondly the soft tissue against the bone, yielded a lower misclassification rate. The first classifier uses quadratic discriminant analysis and the latter linear discriminant analysis. The RGB channels as well as the SV channels were used as features.

The difference in the area measurement is, according to our experience, lower than the variation in the material and hence within acceptable error limits. The BIC-estimation of the automatic method, however, requires further development to match the manual method.

6 Future Work

Further development of the segmentation process would be to include a classification of the artifacts. Furthermore, the bone classification can be divided into two subclasses, old bone and young bone. However, this is not a trivial task; the indicator for new bone is partly somewhat darker purple stains and partly the shape and number of the hollow spaces in the bone.

In addition to the 2D data, the samples have been imaged in a microtomography, hence 3D density images of the samples are recently available. A combined investigation of the 2D and 3D data, allowing a more extensive analysis, will follow.

Acknowledgments. Research technicians Petra Hammarström-Johansson, Ann Albrektsson and Maria Hoffman are greatly acknowledged for skillful sample preparations. This work was supported by grants from The Swedish Research Council, 621-2005-3402.

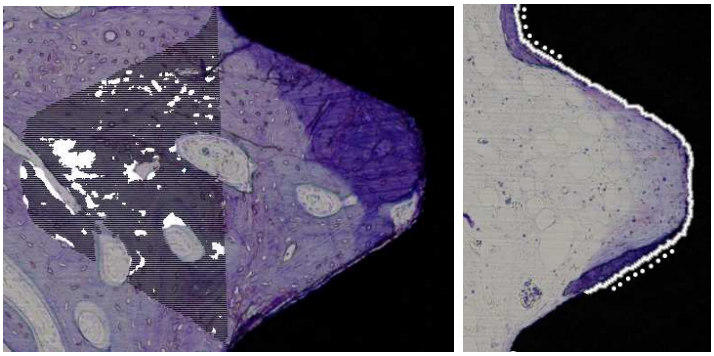

Fig. 4. Left: the identified bone regions by the automatic method in the M-region (marked as stripes). The bone-soft tissue classification misclassifies some lighter bone regions (marked as white regions), resulting in a minor underestimation of bone area. Right: identified BIC by the manual (dotted) and the automatic method (line). Histological staining artifacts result in overestimation by the automatic method.

References

1. Donath, K.: Die trenn-dunnschliffe-technik zur herstellung hist. präparate von nicht schneidbaren geweben und materialien. Der Präparator **34** (1988) 197–206
2. Johansson, C., Morberg, P.: Importance of ground section thickness for reliable histomorphometrical results. Biomaterials **16** (1995) 91–95
3. Johansson, C., Morberg, P.: Cutting directions of bone with biomaterials in situ does influence the outcome of histomorphometrical quantification. Biomaterials **16** (1995) 1037–1039
4. Milan, S., Michael, F.: Handbook of Medical Imaging. SPIE Press, Bellingham, USA (2000)
5. Skarbek, W., Koschan, A.: Colour image segmentation - a survey. Technical report, Technical University Berlin (1994)
6. Ranefall, P.: Towards automatic quantification of immunohistochemistry using colour image analysis. PhD thesis, Uppsala University, Uppsala (1998)
7. Ballerini, L., Bocchi, L., Johansson, C.B.: Image segmentation by a genetic fuzzy c-means algorithm using color and spatial information. In: LNCS. Volume 3005. (2004) 260–269
8. Johansson, C.: On tissue reactions to metal implants. PhD thesis, Department of Biomaterials / Handicap Research, Göteborg University, Sweden (1991)
9. O.Duda, R., et al.: Pattern Classification. John Wiley and Sons (2001)
10. Mathworks: Matlab Documentation R2006b, Function: rgb2hsv. (2006)
11. Johnson, R.A., Wichern, D.W.: Applied Multivariate Statistical Analysis. Prentice-Hall (1998)
12. Koplowitz, J., Bruckstein, A.M.: Design of perimeter estimators for digized planar shapes. In: Trans. on Pattern Analysis and Machine Intelligence. Volume 11. (1989) 611–622
13. Cohen, J.: A coefficient of agreement for nominal scales. Educational and Psychological Measurement **11** (1960) 37–46