Computer Vision — CVWW’03

Proceedings of the 8th Computer Vision Winter Workshop

February 3–6, 2003
Valtice, Czech Republic

Ondřej Drbohlav (ed.)

Czech Pattern Recognition Society
Prague 2003
Robust Segmentation of Homogeneously Oriented Fibrils in Microscopic Biological Soft Tissue Images *

Pierre J. Elbischger†, Horst Bischof‡, Gerhard A. Holzapfel‡
+ University of Technology Graz, Institute for Computer Graphics and Vision, Austria
† University of Technology Graz, Institute for Structural Analysis – Computational Biomechanics, Austria

Abstract Soft tissue such as tendons, arteries, veins or skins are biological materials of particular interest. A deeper understanding of the foundations and interactions of structure and function of soft tissues, and, in particular, the associated mechanobiology is of fundamental importance. A thorough understanding of the complex interrelations between mechanical factors and the associated biological responses may help to improve diagnostics, which allow disease and injury to be treated earlier. One aim is to describe the concentration and structural arrangements of collagen fibers in biological soft tissues.

A fully automatic method for the structural analysis of light microscopic images from the outer most layer of blood vessels is proposed. Clustering in the RGB space and morphological operations are used to select fiber regions and mask non-fiber regions. Based on ridge and valley detection the fiber orientation of small fiber patches is robustly calculated. Finally a token-based region growing is used to combine the fiber patches to regions of a homogeneous fiber orientation. The extracted data is intended to be used in biomechanical models of blood vessels.

1 Introduction

Soft tissues are wide-ranging biological materials in which the cells are separated by extracellular material. They may be distinguished from hard (mineralized) tissues such as bones for their high flexibility and their soft mechanical properties. Examples of soft tissues are tendons, ligaments, blood vessels such as arteries and veins, skins or articular cartilages among many others.

A thorough understanding of the complex interrelation between mechanical factors and the associated biological responses may help to improve diagnostics which allow disease and injury to be treated earlier. It requires quantification of the mechanical environment of the involved tissues, i.e. geometrical and constitutive models of all tissue components involved. A greater understanding of the structure-function relationships of native tissues is also a prerequisite for appropriate repair and replacement tissues.

Hence, in particular, the quantitative knowledge of preferred orientations in biological soft tissues enhances the understanding of their general mechanical characteristics significantly. It is important to note that realistic structural models rely strongly on this knowledge. Collagen fibers are those components of many soft tissues that render the material properties anisotropic. In order to describe the anisotropic feature, appropriate geometrical data is required. The focus of this paper is the structural arrangement of collagen in biological soft tissues, in particular, in the adventitia of arteries.

The common morphologic structure, interesting for the biomechanical modelling of the adventitia appears in the images as red/magenta bundles of collagen fibrils (Figure 1 and 2). The fibrils within the bundles roughly share a common fiber orientation. The tissue preparation process causes several artifacts, which have to be located and properly treated by the analyzing process. The staining procedure, used to make the collagen fibers visible, is not stable and typically stains tissue samples differently. The algorithm has to be able to cope with the staining variabilities.

Section 2 gives a brief overview of the concept and addresses the tissue preparation process and the image acquisition setup. Next the image processing tasks, i.e. artifact segmentation, orientational data extraction and segmentation of homogeneously oriented regions are described. The study concludes with a comparison of the proposed algorithm to manual segmentations, and a discussion of the work.

2 Concept Overview

To relieve medical assistants from the tedious work of measuring structural data per hand, it is reasonable to strive for analyzing the tissue with the help of computer vision methods. The use of computer vision methods for the extraction of relevant structural data guarantees an objective (robust against human subjective impressions) way for data acquisition, and furthermore allows an exhaustive, automatic analysis of a tissue sample set, considerably large to gain statistically meaningful results. The aim of this study is the extraction of the following tissue features.

- The discrimination between fiber and non-fiber areas
- The extraction of regions with a homogeneous fiber orientation

---

*Financial support was provided by the "Amt der Stmk. Landesregierung – Abteilung für Wissenschaft und Forschung, Österreich" and a "DOC-Stipendium" of the "Österreichische Akademie der Wissenschaften"
Statistics about the fiber orientations (the size of a homogenous region and its mean orientation).

The analysis of the collagen fibers is based on digital light microscopic images of thin histological tissue samples. The samples stem from human subclavian arteries and have been extracted from the adventitia by cutting out axial and radial tissue patches. The preparation process consists of three steps that are necessary to obtain samples suitable for the exploration within a transmitted light microscope:

1. To make the interesting structures visible the histological samples have to be very thinly sliced. To avoid large slicing artifacts (mainly distortions) the tissues are embedded into Acrylat, which is a relatively stiff material and resistant against shearing forces.

2. Next, with a microton the embedded tissue samples are thinly sliced (3μm).

3. To make the fibers visible the sliced tissue samples are stained with an Elastica-Van-Gieson staining. The collagen brils appear red/magenta, the elastic brils appear brown. Unfortunately, the coloring process is not very stable and depends strongly on each of the many sub steps necessary to obtain a colored tissue.

A 3-CCD camera (Sony DXC-930P), that is build into the optical system of the microscope, is used to take images of size 760 × 570. The optical magnification (objective ×20, adapter ×0.6) and the dimensions of the CCD chip (6.6 mm × 8.8 mm) lead to a total resolution of 1μm per pixel. The optical magnification has been chosen in accordance to obtain a good tradeoff between the field of view and the necessary amount of detailed structural information of collagen brils. Intensity inhomogeneities of the light microscope are corrected in the capturing process.

The analysis process can be divided into three major tasks: (i) detection and segmentation of artifacts (ii) determination of fiber orientation (iii) segmentation of homogenously oriented fiber regions.

2.1 Detection and Segmentation of Artifacts

Artifacts that arise in the tissue preparation have to be detected and excluded from analysis. These may be summarized as:

- Tissue distortion caused by slicing (e.g. only back light is visible).
- Elastic fibers that are also stained.
- Sliced tissue which are too thick. No reliable information about the brils can be extracted.
- Pieces of other tissues are included – e.g., muscle tissue.
- Unstable coloring – different tissue sets are differently colored.
- Staining impurities.

Some typical artifacts that occur in the histological images are shown in Fig. 2.

The artifacts do not contain any information about the fiber structure and disturb the analysis of the real fiber properties. Generally, the entire process is driven by the motivation to robustly detect actual fiber regions. To classify detected fibers as non-fibers is acceptable, whereas the reversed is not allowed. If an image satisfies the following conditions, it contains enough inherent information about the collagen fibers to allow an unsupervised fiber segmentation:

- Fibers within a single image are colored equally.
- The main part of the image is covered with fibers.
- Fibers emerge in bundles that feature a common fiber direction rather than as single fibers.

As long as these constraints are satisfied, the fibers form a compact cluster in the RGB space. A robust clustering algorithm [1] [4] is used to detected the cluster. It uses the Mahalanobis distance [8] and an α-trimmed estimator for robust covariance estimation [5] to iteratively refine the estimate of the ellipsoidal cluster, and stops if a stable clustering is found. The result of the algorithm is a binary mask that discriminates fibrous and non-fibrous regions. Morphological operators [11] are used to generate a more compact mask (see Fig. 3)
In order to obtain a single gray-level image for further analysis, a Hotelling transform [3] also known as Principal Component Analysis (PCA) can be applied to the feature vectors associated with fibers. The first principal component is then used to transform the entire RGB image. With this approach the gray level image preserves the maximal information and is adapted to the variations of staining. Fig. 4 shows the resulting transformed image.

2.2 Determination of Fiber Orientation

The determination of the fiber orientation is based on a ridge and valley analysis. The gray value profile of a single fiber looks simple. If the fibers appear as bright lines in the image, their profiles in transverse direction show a dark-bright-dark transition that forms a ridge. If the fibers appear dark or they are located near by each other and cause a small gap in between (like in our case), the profile features a bright-dark-bright transition or a valley. Ridges and Valleys feature two significant properties in their profiles. First, a high negative (positive) second derivative. Second, they should be a local maximum (minimum). The directional second derivative of a 2D image in the direction \( u \) can be calculated with the Hessian according to

\[
(\nabla_u^2 I) = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix}.
\]

(1)

The directional second derivative can be approximated by the second-order term of its Taylor series expansion eq.(2). Thus,

\[
\frac{\partial^2 I}{\partial u^2} \approx (u) \cdot (\nabla_u^2 I)(u)^T.
\]

(2)

The eigenvalues of \( \nabla_u^2 I \) are equal to the minimal and maximal values of the second directional derivative. Their associated eigenvectors point into the corresponding directions. These two directions are orthogonal to each other.

Compared to the algorithm described in [2], which is designed to work on images, where fibers appear as continues curvy structures, in our case fibers do not appear as single continues structures. Usually fibers are grouped into bundles of fibers where one fiber may suddenly end/start, or will be occluded from other fibers. This fact makes it difficult to pursue a single fiber in a bundle. However, to gather an overall statistical data of the tissue it is not necessary to know the exact shape of a single fiber within the tissue. It is sufficient to know where bundles of nearly parallel fibers are located, and what their mean fiber orientations are. For this case it is sufficient to determine partial fiber segments and group similar patches. There are several benefits that make the detection by ridge and valley analyses attractive:

- The detected fiber pixels form small fiber strips in a natural way.
- The detection of fiber segments that collect a couple of pixels results in a more robust algorithm with respect to noisy data.
- The small fiber segments can be used in an additional task to pursue the trace of a single fiber within a fiber bundle in order to gain information about its mean elongation.

During the process two images are generated, one for ridges and the other for valleys. On the positions of the classified pixels the directional information (transverse direction to the eigenvector) is recorded. All possible directions in the 2D space are mapped into the angular space ranging from \(-\frac{\pi}{2}\) to \(+\frac{\pi}{2}\), whereby the value 0 corresponds to the vertical axis. To guarantee fiber strips with one pixel thickness a morphological thinning operation [11] is performed. Strips that consist of fewer than 4 pixels do not have an acceptable angular resolution and are removed. The resulting valley and ridge images are shown in Figures 5(b) and
Figure 5: Masked region of the input image (a), thinned ridges (b), thinned valleys (c), valid ridges (d), valid valleys (e) and valid ridges and valleys (f).

Figure 6: A possible fiber strip and the distances between the various end points. The black line represents the largest distance of two end points (a). Basic masks used to search for end points, where gray pixels have the same label as the center pixel (b).

5(c). To produce consistent fiber strips the ridges and valleys are cross-checked with each other. For each ridge (valley) pixel a cross-check operation with its nearest neighbor valley (ridge) pixel is performed. If their directional differences are lower than a predefined threshold, the pixel is assumed to be a real ridge (valley) pixel, otherwise it is removed. Figures 5(d) and 5(e) show the valid fiber strips. Due to the fact that the intersection of detected ridge and valley pixels are zero, it is possible to combine the strips into one label image (see Fig. 5(f)), where each fiber is associated with a distinct label and the background (no detected fiber) is represented by the label 0.

The directions of the fiber strips are derived from the end points of a strip. This way of calculating the direction is more robust than using the directions of the single pixels, which are obtained by the ridge and valley analysis. The end points that yield the largest Euclidean distance are chosen to represent the real fiber strip end points and determine their angles (see Fig. 6).

2.3 Segmentation of Homogenously Oriented Fiber Regions

A token-based region growing is used to segment regions of a homogenous fiber orientations. In order to enable the region growing it is necessary to establish neighborhood relations between the fiber strips. A distance map [6] that can be derived from the label image is used to create small patches around each fiber strip. A threshold that determines the maximal distance between the points within a patch and the corresponding fiber strip, restricts the patch size to a certain value. Figure 7 shows the resulting label image of this operation.

The next task is the selection of proper seed points (patches) for the region growing process. For each fiber patch the following data is calculated:

- The Euclidian distance \( d_p \) between the determined end points
- The orientation of the fiber strip \( \alpha_p \)
- Statistical parameters (mean \( \bar{s}_p \) and standard deviation \( s_p \)) derived from the orientation data of the distinct fiber strip pixels. When calculating the statistics, the circular character of the angular distribution has to be taken into account [10].

The parameters \( d_p \) and \( s_p \) are used to generate a scalar value \( w_p \) for each patch that describes its suitability to be a seed patch. It is given by the weights

\[
    w_p = p_{d_p(n)} \cdot \left[ 1 - p_{s_p(n)} \right],
\]

where \( w_p \) ranges between 0 and 1, and

\[
    p_{d_p(n)} = \frac{d_p(n)}{\max(d_p)},
\]

\[
    p_{s_p(n)} = \frac{s_p(n)}{\max(s_p)}
\]

are the normalized angle values and distance values, respectively.

To reduce the computational costs of the growing process, a region adjacency graph [9], [7] is used. A region adjacency graph features some essential benefits compared with the pixel-oriented description:

- Only a few thousand patches instead of hundreds of thousands pixels have to be concerned. This speeds up the growing process by approximately 60% to 70% (the additional costs for the generation of the map are included!)
- If one region is merged into another all regions that are adjacent to the new one are known immediately. The boundary does not have to be scanned pixelwise.
Figure 8: Grown regions (false color representation of the angle values)

Figure 9: Manually segmented image

- Subsequent processing steps can make use of the region adjacency graph that, in general, has a still lower complexity after the region growing.

Region growing that merges neighboring regions with similar characteristics is used to segment regions of a homogeneous fiber orientation. The circular character of the angle distribution has to be considered [10]. Figure 8 shows the finally grown regions. The angle values are depicted in false color representation.

3 Comparison to Manual Segmentation

To compare the proposed algorithm to a manual segmentation procedure, eight images where manually segmented. Polygonal regions of homogeneous fiber orientation are manually segmented, and a straight line is used to indicate the mean fiber orientations (see Fig. 9).

On average the algorithm segmented approximately 60% of an image into fiber areas. Because of the large efforts, necessary for a manual segmentation, on average only 14% of the image areas have been manually segmented, i.e. only for those regions it was easy to identify a "meaningful" fiber orientation. Table 1 shows the area statistics of the grown and manually segmented regions.

The overlap between the manually segmented regions and the automatically segmented regions is 92%. These regions are used for our comparison (see Fig. 10).

Table 1: Area statistics of the grown regions and manually segmented regions (the percentages are related to the entire image area (760 x 570)). GR: grown regions, MS: manually segmented regions. Nr. GR: number of GR regions, Nr. MS: number of MS regions, Intersec.: intersecting area of GR and MS regions, Over Seg.: GR regions that are not contained in the MS regions, Under Seg.: MS regions that are not contained in the GR regions.

Table 2: Statistics of the angular differences of the manually segmented and the automatically segmented regions (related to the manually segmented regions).

- Subsequent processing steps can make use of the region adjacency graph that, in general, has a still lower complexity after the region growing.

Region growing that merges neighboring regions with similar characteristics is used to segment regions of a homogeneous fiber orientation. The circular character of the angle distribution has to be considered [10]. Figure 8 shows the finally grown regions. The angle values are depicted in false color representation.

3 Comparison to Manual Segmentation

To compare the proposed algorithm to a manual segmentation procedure, eight images where manually segmented. Polygonal regions of homogeneous fiber orientation are manually segmented, and a straight line is used to indicate the mean fiber orientations (see Fig. 9).

On average the algorithm segmented approximately 60% of an image into fiber areas. Because of the large efforts, necessary for a manual segmentation, on average only 14% of the image areas have been manually segmented, i.e. only for those regions it was easy to identify a "meaningful" fiber orientation. Table 1 shows the area statistics of the grown and manually segmented regions.

The overlap between the manually segmented regions and the automatically segmented regions is 92%. These regions are used for our comparison (see Fig. 10).

The distribution of the angular differences over the entire image is zero-mean, with a standard deviation of 10.75 degrees. The angular difference (standard deviation) within a single manually segmented region is about 14 degrees. The zero-mean distribution confirms the high reliability of the algorithm.

The comparison points out the following benefits of the automatic segmentation procedure:

- It is possible to segment a much larger area.
- A finer angular resolution is achievable.
- The automatic analysis is more time efficient.

The comparison proves the reliability of the fully automatic algorithm, and shows that the algorithm produces results that are comparable with a manual segmentation. Because of its finer resolution, the algorithm provides even

![Figure 8: Grown regions (false color representation of the angle values)](image1)

![Figure 9: Manually segmented image](image2)

<table>
<thead>
<tr>
<th>Nr. GR</th>
<th>Nr. MS</th>
<th>GR Area</th>
<th>MS Area</th>
<th>Intersec.</th>
<th>Over Seg.</th>
<th>Under Seg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>164</td>
<td>164</td>
<td>97.5%</td>
<td>27.9%</td>
<td>39.1%</td>
<td>30.7%</td>
<td>1.9%</td>
</tr>
<tr>
<td>1824</td>
<td>1824</td>
<td>94.2%</td>
<td>27.9%</td>
<td>39.1%</td>
<td>30.7%</td>
<td>1.9%</td>
</tr>
<tr>
<td>1152</td>
<td>2000</td>
<td>68.6%</td>
<td>20.5%</td>
<td>10.4%</td>
<td>40.2%</td>
<td>0.8%</td>
</tr>
<tr>
<td>1277</td>
<td>2300</td>
<td>69.0%</td>
<td>17.5%</td>
<td>8.0%</td>
<td>87.0%</td>
<td>0.8%</td>
</tr>
<tr>
<td>1311</td>
<td>3400</td>
<td>62.2%</td>
<td>12.6%</td>
<td>11.7%</td>
<td>56.4%</td>
<td>0.8%</td>
</tr>
<tr>
<td>1352</td>
<td>7200</td>
<td>67.2%</td>
<td>7.1%</td>
<td>6.8%</td>
<td>61.0%</td>
<td>0.3%</td>
</tr>
<tr>
<td>1182</td>
<td>17000</td>
<td>70.0%</td>
<td>8.0%</td>
<td>8.2%</td>
<td>42.5%</td>
<td>0.7%</td>
</tr>
<tr>
<td>974</td>
<td>43000</td>
<td>68.5%</td>
<td>9.0%</td>
<td>4.5%</td>
<td>34.3%</td>
<td>1.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>Means</th>
<th>Standard Deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ</td>
<td>σ</td>
<td>μ</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>164</td>
<td>-2.0</td>
<td>11.0</td>
</tr>
<tr>
<td>29</td>
<td>2.8</td>
<td>11.5</td>
</tr>
<tr>
<td>43</td>
<td>-1.2</td>
<td>7.8</td>
</tr>
<tr>
<td>29</td>
<td>6.8</td>
<td>11.2</td>
</tr>
<tr>
<td>34</td>
<td>-0.8</td>
<td>10.3</td>
</tr>
<tr>
<td>27</td>
<td>0.6</td>
<td>11.2</td>
</tr>
<tr>
<td>51</td>
<td>3.8</td>
<td>10.9</td>
</tr>
<tr>
<td>39</td>
<td>-4.3</td>
<td>8.4</td>
</tr>
</tbody>
</table>

The automatic analysis is more time efficient.

The comparison proves the reliability of the fully automatic algorithm, and shows that the algorithm produces results that are comparable with a manual segmentation. Because of its finer resolution, the algorithm provides even...
more information, and is even more accurate than the results obtained from a manual segmentation.

4 Conclusion

An approach for the determination of fiber orientations in microscopic biological soft tissue images was proposed. The method determines regions of a homogenous fiber orientation. A few parameters such as the area of the regions and the statistical fiber orientation have been calculated for each region. These parameters may be used for mechanical models and serve as a basis for further analyses.

Besides the statistical information about the regions of homogenous orientation, the mechanical models require the global structural characteristics of the entire fibrous bundles. To obtain representative bundle-oriented statistics, it is necessary to combine homogenous fiber patches to larger but consistent bundles. The bundle construction procedure and also the tracing of single fibrils that gives information about the middle fiber elongation will be addressed in future work.

References