Automated quantification and reconstruction of collagen matrix from 3D confocal datasets

J. WU*, B. RAJWA†¶, D. L. FILMER*, C. M. HOFFMANN‡, B. YUAN‡, C. CHIANG‡, J. STURGIS† & J. P. ROBINSON†§

*Department of Biology, †Department of Basic Medical Sciences, ‡Department of Computer Science, and §Department of Biomedical Engineering, Purdue University, West Lafayette, IN 47907, U.S.A.
¶Department of Biophysics, The Jan Zurzycki Institute of Molecular Biology and Biotechnology, Jagiellonian University, ul. Gronostajowa 7, 30–387 Kraków, Poland

Key words. 3D reconstruction, collagen fibre, confocal microscopy, ECM, tracking algorithm.

Summary

The geometrical structure of fibrous extracellular matrix (ECM) impacts on its biological function. In this report, we demonstrate a new algorithm designed to extract quantitative structural information about individual collagen fibres (orientation, length and diameter) from 3D backscattered-light confocal images of collagen gels. The computed quantitative data allowed us to create surface-rendered 3D images of the investigated sample.

Introduction

Extracellular matrix (ECM) is an important structural entity composed of collagens, elastin, glycoproteins, proteoglycans (PGs) and glycosaminoglycans (GAGs). ECM provides not only the mechanical framework to maintain tissue integrity, but also the biophysical and biochemical environment to support and modulate cell growth, migration and differentiation (Gua et al., 1991; Adams & Fiona, 1993; Aumailley & Gayraud, 1998; Bruckner-Tuderman & Bruckner, 1998; Iozzo, 1998).

Recent research has shown that the properties of ECM are highly correlated with its structural characteristics (Berthiaume et al., 1996; Fitton et al., 1998; Ranucci et al., 2000). The dominant structural components of ECM are collagens. Orientation of collagen fibres induced by traction forces can guide cell migration and cell orientation (Guido & Tranquillo, 1993; Takakuda & Miyairi, 1996; Knapp et al., 1999). In addition, it has been shown that diameter and length of collagen fibres are significantly correlated with the mechanical properties of ECM (Parry et al., 1978; Flint et al., 1984; Craig et al., 1987; Trotter & Koob, 1989; Hsu et al., 1994; Derwin & Soslowsky, 1999). Because interactions between ECM and cells occur on a nanometric scale and are important for the development of cellular engineering and understanding tissue repair mechanisms (Curtis & Wilkinson, 1998), we and other researchers have become increasingly interested in determining orientations, lengths and diameters of fibres in non-processed and non-fixed collagen matrices (Brightman et al., 2000; Ottani et al., 2001; Voytik-Harbin et al., 2001; Roeder et al., 2002).

Microscopical observation of ECM has been a great challenge. The use of light microscopy for ECM study was pioneered by Virchow, who found in 1910 that orientation of collagen fibres in flat sections of cornea could be determined after interstitial injection of chromic acid (Virchow, 1910). In the 1960s it was shown that collagen fibres could be clearly distinguished by polarizing microscopy (Sweat et al., 1964). More recently, polarized light microscopy has been used to obtain quantitative information about the collagen organization in the cornea and sclera (Newton et al., 1995) and other tissues (Dickey et al., 1998). However, the traditional visualization techniques based on polarized light and/or histochemical staining and conventional thin-sectioned preparations of ECM structures do not provide a satisfactory means of visualizing ECM in three dimensions. Moreover, use of histochemical dyes in the latter method has two basic drawbacks: the staining procedure usually not only significantly changes the 3D geometry of ECM but also requires fixation, thus ruling out any live-cell studies. Additionally, owing to poor penetration, pH or osmotic imbalance of chemical fixatives, dislocation or agglutination of components of the proteoglycans and other non-collagenous proteins occurs during the fixation process (Kellenberger et al., 1992).

A relatively new, non-invasive method of ECM observation in three dimensions is based on a natural source of contrast...
provided by absorption and refractive-index changes in the microscopic structure of tissue (Gunzer et al., 1997; Strohmaier et al., 1997; Brightman et al., 2000). These phenomena give rise to reflection, and can be utilized by confocal microscopy, which forms images by collecting laser light backscattered from a small localized volume of ECM (Dunn et al., 1996).

In backscattered-light confocal imaging (CLSM-BSL) the pixel brightness in an image is proportional to the intensity of the light reflected from the corresponding tissue element. Therefore, voxels with significantly different optical properties can be differentiated. CLSM-BSL is a difficult technique owing to a cumbersome optical set-up and the presence of numerous image artefacts. Fibres on the confocal reflectance images seem to have varying intensity or even small gaps. This does not indicate a lack of continuity but arises from the fact that the reflecting surface is tilted so steeply that part of the reflected light does not reach the objective. Artefacts on confocal images are intensified when volume data are visualized via surface rendering. As a result, a final 3D visualization often exhibits false positives (spurious branches) or false negatives (erroneous discontinuities of fibres) (Levoy, 1988).

The complexity of fibrous networks makes quantification of 3D information very difficult as well. Methods based on Fourier transform are limited to the analysis of overall fibre orientations (Pourdeyhimi et al., 1997; Redon et al., 1998). The Hough transform, which is generally used for line detection, is unreliable for irregular curves such as collagen fibres (Hough, 1962; Davidson & Clarke, 1999). Skeletonizing (thinning) with a direct tracing method can identify individual fibres, but is time-consuming and produces many artefacts owing to variations on the boundaries (Krucinska, 1999; Pourdeyhimi et al., 1996). Measurement of the elliptical parameters of a fibre’s image could determine its 3D orientation, but this approach almost always has 180° angle ambiguity and depends on many additional assumptions (Yurgartis, 1987; Hine et al., 1993, 1995). Pattern matching methods might measure 3D orientation of unidirectional or short fibres (Clarke et al., 1993, 1995; Davidson et al., 1997), but could not be applied to a 3D image dataset of collagen matrix in our research because collagen fibres are long and non-unidirectionally distributed in 3D space. Together, these techniques, which are usually used to quantify 2D information, cannot be applied straightforwardly to a 3D case.

In this report, we used confocal microscopy operating in a backscattered-light mode to acquire a z-series of optical sections from the model collagen matrix. The scanned images had low signal-to-noise ratio, and the radius of collagen fibres (below 600 nm) was close to the resolution of the imaging system (c. 270 nm). To quantify collagen fibre organization quickly and accurately, we developed novel middle-line tracing software based on our original algorithm, which was specifically tuned towards CLSM-BSL images of fibrous collagen. The software was able to compute fibre orientation, diameter and length on the basis of 3D confocal datasets. The extracted quantitative data allowed us to build 3D surface-rendered models.

Materials and methods

Sample preparation and image acquisition

Purified type I collagen from bovine dermis (Collagen Corporation, Fremont, CA, U.S.A.) was obtained as a sterile solution in 0.012 N HCl; 8 mL chilled collagen was mixed with 1 mL PBS, and 1 mL of 0.1 M NaOH was added and mixed and the pH of the solution was adjusted to 7.4 ± 0.2. Collagen gelation was initiated in Lab-Tek chambered cover glasses by warming the neutralized solution to 37 °C.

Samples were imaged without staining using the Bio-Rad MRC1024 (Bio-Rad Laboratories, Hemel Hempstead, U.K.) confocal microscope via a 60×, 1.4-NA oil-immersion lens. Illumination was provided by a 488-nm laser and backscattered light was detected by a photomultiplier tube. The x-y plane pixel size was set to 0.1–0.25 µm and z-steps of 0.2–0.6 µm were used to optically section the samples. Three-dimensional data sets consisted of between 50 and 80 image slices.

Tracing algorithm

Figure 1 shows the outline of our method of extraction of 3D quantitative data from confocal datasets. The 3D image dataset was converted to binary form after Gaussian filtering (Fig. 2 shows the kernel). The threshold value was set by a user. However, our algorithm can also use thresholds found by thresholdings procedures (Otsu, Entropy or Minimum Error). Let 3D image space in the discrete space $Z^3$ be $V = \{v = (x, y, z) \in Z^3 | 0 \leq x \leq L, 0 \leq y \leq W, 0 \leq z \leq H\}$, where $L$, $W$ and $H$ indicate the size of $V$. The element of $V$, $v$ is called a voxel or point, and is classified into an object point or a background point.

Definition 1. The box of $u$ with radius $r$ is the cubic set of points such that

$$B(u, r) = \{v \in V | ||v - u|| \leq r\}$$

for any point $u \in V$.

Definition 2. The boundary of the box $B(u, r)$ is the set of points such that

$$\partial B(u, r) = \{v \in B(u, r) | ||v - u|| = r\}$$

Definition 3. The box $B(u, r)$ is a function that maps 3D image set from $V$ to $\{0, 1\}$ on a basis of intensity values of the points. Given the point $u \in X$, the minimal Euclidean distance from $u$ to the background point set $X$ is denoted as $d(u, X)$, and the squared minimal Euclidean distance is denoted as $D(u, X)$.
Fig. 1. The outline of the tracing procedures. (a) The 3D grey-level image set is converted to the binary image set by filtering and thresholding. The brightness of each object voxel is encoded as the distance to the nearest voxel in the background by Euclidean distance mapping. The tracing procedure starts marching from the unmarked global maximal point (GMP) and forms branches consisting of GMP and local maximal points through recursive marching and marking. The tracing step size is automatically regulated based on local structure. (b) Subsequent processing removes the spurious side branches and connects the branches with similar orientations, recovering the continuity. Finally, the quantitative data, including fibre length, orientation and diameter, are computed, and 3D structure is reconstructed on the basis of these data. The tracing procedure is time consuming if the matrix structure is complex. The processing step might also be computationally intensive if many branches are found.

Fig. 2. Three-dimensional Gaussian smoothing kernel.

Fig. 3. (a) The illustration of definitions of $X$, $\bar{X}$, $u$, $B(u,r)$, $\partial B(u,r)$ and $M(x)$. Detailed description in text. (b) Euclidean distance mapping (EDM) converted a binary image to a grey-scale image in which pixel value codes the Euclidean distance from each pixel within the object $X$ to the nearest background $\bar{X}$. Description of GMP and LMP in text.
Definition 3. The medial axis (MA) or skeleton of X is denoted as \( M(X) \). The MA can be described as a subset of \( X \), consisting of those points \( u \) that there is such a \( B(u,r) \) of the maximal radius that is completely contained within \( X \) and intersects with the boundary of \( X \) in at least two distinct points: 
\[
M(X) = \{ u \in X | B(u,r) \subset X \text{ and } d(u,X) \geq d(v,X), v \in B(u,r) \text{ and } \text{card}(B(u,r) \cap \partial X) \geq 2 \},
\]
where card denotes cardinality of sets.

Definition 4. The global maximal point (GMP) is the point \( u_0 \in X \) whose \( D(u_0,\bar{X}) \geq D(w,\bar{X}) \) for any point \( w \in X \), and \( d(u_0,\bar{X}) > f_{\text{max}} \), where \( f_{\text{max}} \) is the minimal radius of a fibre in the binary image set.

Definition 5. The local maximal point (LMP) is the point \( v \in X \cap \partial B(u,r), \) where \( u \in M(X) \) and \( r = d(u,\bar{X}) \), such that \( D(v,\bar{X}) \geq D(w,\bar{X}) \) for all \( w \in Q(v,\delta) \), and there is no other local maximal point in \( Q(v,\delta) \). Here \( Q(v,\delta) = X \cap \partial B(u,r) \cap B(v,\delta) \), with radius \( \delta = 1 \). Theoretically, the local maximal point \( v \) should satisfy \( D(v,\bar{X}) > D(w,\bar{X}) \) for all \( w \in Q(v,\delta), w \neq v \). Owing to discretization errors and noise in the image set, there may exist \( D(v,\bar{X}) > D(w,\bar{X}) \). To resolve this situation, the point set \( B(v,r) \) with \( r = d(v,\bar{X}) \) is marked as the just-found point \( v \).

Tracing and marking

The first step to extract MA from confocal images is to calculate the minimal squared Euclidean distance \( D(u,\bar{X}) \) for every point \( u \in X \). Our algorithm uses the well-known distance-mapping method to compute \( D(u,\bar{X}) \) (Danielsson, 1980; Borgefors, 1984, 1996; Ragnemalm, 1993). Subsequently, the tracing procedure starts marching from the GMP. From point \( u \) the current branch is searched to find the LMPs \( u_1, u_2, \ldots, u_n \) on \( \partial B(u,r) \) with \( r = d(u,\bar{X}) \). Then, the set of points \( B(u,r) \) is marked. If \( n = 1, u_1 \) is added into the current branch and serves as the next searching point. Otherwise, the current branch is put into a branch list, and \( n \) new branches, consisting of \( u \) and \( u_n \), are created. Point \( u_1 \) is subsequently used to find the LMPs on \( \partial B(u,r) \) with \( r = d(u,\bar{X}) \). By recursive marching and marking, a series of branches consisting of GMPs and LMPs are created.

Processing branches

Owing to noise and discretization errors, the coarse boundary of fibres creates many artefactual side branches. To extract medial axes of fibres, the spurious side branches are removed, and the branches along the medial axis of the same collagen fibre are connected according to the following algorithm.

Three types of connections exist between identified branches, depending on what their terminal points connect to: (1) only one other branch, (2) more than one other branch, or (3) no other branch. Therefore, the branches are classified into five categories:

- Normal branch: both terminals are type 1.
- End branch: one terminal is type 1, and the other is type 3.
- Side branch: one terminal is type 2, and the other is type 3.
- Joint branch: one terminal is type 2, and the other is either type 1 or type 2.
- Free branch: both terminals are type 3.

The algorithm classifies the branches and removes side branches identified as below a preset of length. It then reclassifies the branches and connects the normal and end-branches. The joint branches having similar orientations are connected to maintain the cross-links between fibres. The connection procedure works as follows: the joint branch has more than one connected branch at one of its terminals such as OA in Fig. 4. Since the orientation of each branch has already been measured, to select which branch should be connected at point O (a cross-link), the orientation difference between joint branch OA and branches OB, OC and OD is calculated. If the resulting value is less than a preset value, two branches are considered to have similar orientations and are connected to form one branch. For example, OA is connected to OC to form branch AOC. Similarly, OB and OD are connected to form branch BOD. In order to recover the continuity of the fibres, the algorithm also connects branches that have similar orientations and only a small gap of a predefined value between them. Eventually, each processed branch represents the medial axis of an individual fibre in 3D image space. This procedure is non-iterative: the removal of side-branches and the connection of branches is performed only once per dataset. Figure 5 illustrates the steps for extracting the medial axes of fibres.

Quantification and 3D reconstruction

Each processed branch consists of a series of points \( u_i, i = 1, \ldots, n \), whose position \( (x_i,y_i,z_i) \) and radius \( r = d(u,\bar{X}) \) are recorded. Thus the fibre length is computed as \( \sum_{i=1}^{n} d(u_{i-1},u_i) \), and the mean fibre radius is the average of \( r \). The 3D orientation of the fibre is defined by the \( xy \) and \( xz \) angles, which are computed by linear regression of the points \( u_i, i = 1, \ldots, n \), projected to the \( xy \) and \( xz \).
planes, respectively. To reconstruct the 3D structure of the collagen matrix, the algorithm uses cylinders connecting $u_{n+1}$ and $u_n$.

**Results and discussion**

Gels of bovine collagen type I were synthesized and imaged with a confocal microscope operating in backscattered-light mode. The presented algorithm was able to process collected 3D image sets and measure orientations, lengths and diameters of individual fibres in 3D space. Figure 6 presents the image slices scanned at different depths within a sample, and Fig. 7 shows their reconstructed 3D structure.

The new image-processing technique described in this paper automatically tracked collagen fibres through a 3D image dataset. Unlike other published methods, our tracing procedures did not create the fibre midline as a series of
connected voxels (or pixels in two dimensions), but as a series of tracing elements (boxes). The tracing step size (size of box) was automatically regulated with respect to variations in local fibre structure, and therefore this technique not only reduced computational complexity but also provided detailed structural information without compromising the quality of 3D reconstructions. Our algorithm identified and removed unwanted points forming side branches. Thus, the resulting midlines were accurate, robust and not dependent on structural variations. Furthermore, the ability to maintain the cross-links between fibres and recover the fibre continuity gave rise to accurate measurement of fibre length and orientation.

Three-dimensional information about individual fibres is critical in investigating fibre formation, distribution and interaction, as well as the overall mechanical properties of collagen matrix. It is also very important for research on cell–ECM interactions. Integration of confocal microscopy using

Fig. 6. Three-dimensional image dataset consisting of 50 optical slices of collagen type I matrix. Images were recorded using confocal microscopy in backscattered-light mode.
a backscattered-light mode with a novel tracing technique to derive 3D quantitative data from collagen matrix images is a promising new approach to our long-term goal of understanding the 3D structure of biomaterials in order to design and manipulate fully tissue-engineered complexes.

Acknowledgements

Part of this research was supported by a reinvestment grant from Purdue University. B.R. was supported in part by a grant from the Polish Committee for Scientific Research (KBN 6PO4A07820).

References


Fig. 7. Three-dimensional reconstruction of type I collagen matrix based on quantitative data extracted from the image dataset presented in Fig. 3. Not all fibres are presented because of discretization errors and thresholding. The ‘nodules’ (intercross between fibres) are an artefactual result of imaging fibres located very close to each other. They could be removed by additional processing to make a more accurate representation. However, they do not interfere with the identification of fibre orientation.


