EXTRACTION OF CURVED FIBERS FROM 3D DATA

Gerd Gaiselmann[✉,1], Ingo Manke[2], Werner Lehnert[3,4] and Volker Schmidt[1]

[1]Institute of Stochastics, Ulm University, 89081 Ulm, Germany; [2]Helmholtz Centre Berlin for Materials and Energy, Institute of Applied Materials, 14109 Berlin, Germany; [3]Forschungszentrum Jülich, Institute of Energy and Climate Research (IEK-3: Electrochemical Process Engineering), 52428 Jülich, Germany; [4]Modeling in Electrochemical Process Engineering, RWTH Aachen University, Germany
e-mail: gerd.gaiselmann@uni-ulm.de, manke@helmholtz-berlin.de, w.lehnert@fz-juelich.de, volker.schmidt@uni-ulm.de

ABSTRACT

A segmentation algorithm is proposed which automatically extracts single fibers from tomographic 3D data of fiber-based materials. As an example, the algorithm is applied to a non-woven material used in the gas diffusion layer of polymer electrolyte membrane fuel cells. This porous material consists of a densely packed system of strongly curved carbon fibers. Our algorithm works as follows. In a first step, we focus on the extraction of skeletons, *i.e.*, center lines of fibers. Due to irregularities like noise or other data artefacts, it is only possible to extract fragments of center lines. Thus, in a second step, we consider a stochastic algorithm to adequately connect these parts of center lines to each other, with the general aim to reconstruct the complete fibers such that the curvature properties of real fibers are reflected correctly. The quality of the segmentation algorithm is validated by applying it to simulated test data.

Keywords: fiber-based material, fiber extraction, non-woven, stochastic model, synchrotron tomography.

INTRODUCTION

The fuel cell technology provides an efficient way to convert hydrogen into electricity. This allows the replacement of oil products as energy carrier, *e.g.*, in automobiles or submarines. In the present paper we concentrate on proton exchange membrane fuel cells (PEMFC). A key component of PEMFC is their gas-diffusion layer (GDL), which has to provide many functions for an efficient operation of fuel cells like to supply the catalyst layer with gas, water storage and evacuation, mechanical support of membrane, etc., (Mathias *et al.*, 2003; Hartnig *et al.*, 2008).

It is generally accepted that the microstructure of GDL has a significant impact on their functionality. For example, the length and curvature of single fibers influence their mechanical strength, stiffness, and electrical resistance (Teßmann *et al.*, 2010), where the arrangement and the form of fibers can be changed by the production process or as a result of degradation phenomena. To better understand the coherence of the microstructure of the considered material and the physical processes therein, the development of stochastic models describing the geometry of GDL is of great importance. Typically, these models are built in two steps (Thiedmann *et al.*, 2008; Gaiselmann *et al.*, 2012). First, a stochastic model for single fibers is developed. Then, in a second step, the single-fiber model is used to construct a stochastic microstructure model for the entire 3D system of fibers. Therefore,

in order to build an adequate single-fiber model, we first need to develop a segmentation algorithm which automatically extracts single fibers from tomographic 3D data. In the literature, some works on this topic has been done so far (Altendorf and Jeulin, 2009; Dokládal and Jeulin, 2009; Teßmann *et al.*, 2010). In this paper, we propose another algorithm for single-fiber extraction which is developed in particular for densely packed systems of strongly curved fibers. It extends and modifies similar algorithms which have recently been developed for straight (Thiedmann *et al.*, 2008) and curved (Gaiselmann *et al.*, 2012) planar fibers on the basis of 2D SEM images.

Our algorithm can be described in the following way. In a first step the 3D image is binarized by global thresholding. Then, we focus on the extraction of skeletons, *i.e.*, center lines of single fibers from the binarized image. Therefore, using tools from image analysis, *i.e.*, the Euclidean distance transformation, global thresholding and skeletonization, the center lines are extracted. Due to irregularities like noise or binarization artefacts, it is only possible to extract relatively short fragments of center lines. Thus, in a second step, we consider a stochastic algorithm to adequately connect these parts of the center lines to each other, with the general aim to reconstruct the complete fibers such that the curvature properties of real fibers are reflected correctly. The quality of the segmentation algorithm is validated by applying it to simulated test data.

The paper is organized as follows. First, we describe the algorithm for automatic extraction of single fibers from tomographic 3D data. Then, a technique to fit the parameters of the segmentation algorithm and to validate the quality of segmentation is briefly explained. A final section summarizes the results.

DESCRIPTION OF ALGORITHM

We present an algorithm to automatically detect single fibers from tomographic 3D data of fiber-based materials. As an example of application, we use this algorithm for 3D image data of non-woven GDL, gained by synchrotron tomography.

IMAGING AND MATERIAL

The segmentation algorithm introduced in this paper is applied to tomographic image data of non-woven GDL. More precisely, we consider 3D image data of H2315 GDL produced by the company Freudenberg FCCT KG. The GDL consists of strongly curved carbon fibers which run mainly in horizontal direction (Figs. 1 and 2). Moreover, there occur fiber bundles within the GDL, *i.e.*, several fibers which run parallel to each other. By means of high resolution synchrotron tomography, 3D image data of this material was gained. The conditions of acquisition were the following. A W-Si monochromator with an energy resolution of $\Delta E/E = 10^{-2}$ was used to obtain a monochromatic X-ray beam. The beam energy was adjusted to 15 keV in order to achieve optimal contrast for fibers. A radiographic set of 1500 projections and 500 flatfields were taken and, subsequently, reconstructed to a 3D volume. The image data is given in the dimension $[0, 3.3] \times [0, 2.2] \times [0, 0.2]$ mm^3, where the voxel size is equal to 0.833 μm. In order to obtain reasonable computational efforts for the single-fiber extraction algorithm which is introduced in the next sections, we consider only a cut-out of the image data defined on the volume $[0, 0.625] \times [0, 0.625] \times [0, 0.2]$ mm^3. Note that this volume corresponds to $[0, 750] \times [0, 750] \times [0, 240]$ voxels, where the size of each voxel is 0.833 μm.

PREPROCESSING OF IMAGE DATA

In this section, we explain several steps of image preprocessing which are necessary for the detection of fibers from 3D image data. To begin with, the synchrotron image is median filtered by a $3 \times 3 \times 3$ kernel in order to reduce noise. Then, the image data is binarized by global thresholding with some threshold value t, which will be specified later on (Fig. 1).
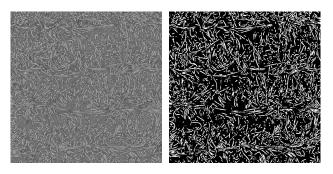

Fig. 1. *2D slice of synchrotron data (left) and binarized image (right).*

The binarized image consists of one large foreground cluster representing the fiber system and small foreground clusters which obviously do not contribute to the fibers. To remove them, an algorithm proposed in Hoshen and Kopelmann (1976) is used for the detection of isolated small clusters. Thus, small clusters with a size below a certain limit are removed where this limit has been set to 1000 voxels since all clusters (excluding the largest one) have a volume of less than 1000 voxels. The binarization threshold t is chosen such that the foreground phase of the binary image, after removing the dispensable small clusters, has a volume fraction of 23.5 %. This volume fraction of the fiber system is known from manufacturers of non-woven GDL.

In the next step we focus on the extraction of skeletons, *i.e.* center lines of single fibers, from the binarized image. Since there exists many fibers which touch each other and run parallel (Fig. 2), the extraction of center lines can not be accomplished just by a standard skeletonization procedure for binary images, as described *e.g.* in Soille (2003); Burger and Burge (2007).
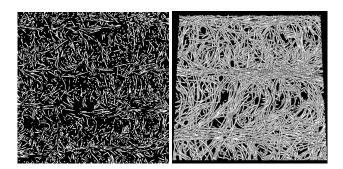

Fig. 2. *2D slice (left) and small cut-out (right) of binarized 3D image.*

Standard skeletonization algorithms would recognize touching fibers as one single fiber (Fig. 3). Thus, these skeletonization algorithms are not applicable in our case.

Fig. 3. *Fibers running in parallel (left) and standard skeleton (right).*

The basic idea of our approach to the extraction of center lines is to benefit from the fact that all fibers have the same radius $r = 4.75$ μm known from manufacturers. Thus, the center lines of perfect fibers should have a minimal distance r to the background phase. We therefore determine the Euclidean distance transformation from foreground to background, *i.e.*, we associate each voxel of the binarized image B with its shortest Euclidean distance (given in μm) to the background phase and denote the resulting image by D_B. Taking *e.g.* binarization or discretization artefacts into account, we consider all voxels $v \in D_B$ with a shortest distance $D_B(v) \in [r - \sqrt{3} \cdot 0.833, r + \sqrt{3} \cdot 0.833]$ to the background as candidates for voxels representing the center lines and set those voxels to foreground. The new image is denoted by V. Subsequently, the foreground phase of V is skeletonized. This means that foreground voxels, *i.e.*, those belonging to the objects we are interested in, are changed to background voxels in a way that the remaining (still voxel-given) paths have a thickness of one voxel, where the connectivity properties of B have to be preserved (Soille, 2003; Burger and Burge, 2007).
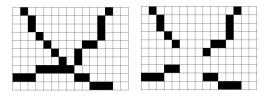

Fig. 4. *Skeleton (left) and elimination of crosspoints (right).*

In order to transform the skeletonized image into vector data, we first remove the crosspoints of the voxel-paths with respect to the 26-neighborhood (Fig. 4) and subsequently represent the remaining voxel paths by polygonal tracks. For further information about this transformation the reader is referred to Gaiselmann *et al.* (2012). Thus, we end up with a family of polygonal tracks, which can be interpreted as a graph structure, representing parts of the center lines of fibers.

RECONSTRUCTION OF FIBERS

In this section, we discuss a stochastic algorithm to adequately connect the polygonal tracks, whose extraction from the 3D synchrotron image has been explained in the previous section. Recall that these polygonal tracks, denoted by $p_1, \ldots, p_n$ for some $n \geq 1$, represent relatively short parts of the center lines of the fibers to be detected. Our algorithm working on 3D data is an extension of the 2D extraction algorithm proposed in Gaiselmann *et al.* (2012). Furthermore, a similar approach to connect polygonal tracks was considered in Jeulin and Kurdy (1992) in order to reconstruct incomplete grain boundaries. The algorithm introduced in this manuscript puts us in a position to construct complete fibers such that the curvature properties of real fibers are reflected correctly. Hence, for an adequate reconstruction of complete center lines, the polygonal tracks considered in the previous section have to be appropriately connected, *i.e.*, we seek for sequences of polygonal tracks (parts of center lines) representing the courses of complete fibers. The flow-chart of the algorithm shown in Fig. 5 gives an overview of the consecutive steps of the algorithm.
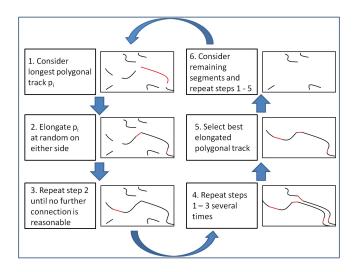

Fig. 5. *Flow-chart of the random fiber-reconstruction.*

More precisely, we first consider the polygonal track p_i with the largest Euclidean length. Then, we connect p_i on either side to other polygonal tracks. Since, for a given polygonal track, there can be several possibilities of connecting other polygonal tracks with it, a decision rule has to be established which chooses the next polygonal track for connection. For connecting the end-segment ℓ_i of a polygonal track p_i to the end-segment ℓ_j of another polygonal track p_j, say, we are looking for all end-segments $\{\ell_1, \ldots, \ell_k\} \setminus \{\ell_i\}$ and $\{\ell_1, \ldots, \ell_k\} \setminus \{\ell_j\}$ with an endpoint belonging to a certain cone around the considered endpoint of

ℓ_i and ℓ_j, respectively. If there is no line segment in these cones, we stop the connection of polygonal tracks at this stage, where the cone is given as follows. Note that the task of the cone is to ban connections which are not suitable. Let $|\ell_{ij}|$ be the length of the line segment ℓ_{ij} connecting ℓ_i and ℓ_j and $\sphericalangle(\ell_i, \ell_{ij})$ denotes the angle between the line segments ℓ_i, ℓ_{ij} (Fig. 6). Then, an arbitrary line segment ℓ_j is in the cone of ℓ_i, if $|\ell_{ij}| < \ell_{max}$ and $\sphericalangle(\ell_i, \ell_{ij}) < \alpha_{max}(|\ell_{ij}|)$ as well as $\sphericalangle(\ell_j, \ell_{ij}) < \alpha_{max}(|\ell_{ij}|)$, where $\alpha_{max}(|\ell_{ij}|) = \left(\frac{\ell_{max} - |\ell_{ij}|}{\ell_{max}}\right)\left(\frac{\pi}{2} + \alpha\right)$ for some parameters $\ell_{max} > 0$ and $\alpha \in \mathbb{R}$. In other words, the maximum angle $\alpha_{max}(|\ell_{ip}|)$ is large if $|\ell_{ip}|$ is small and vice versa.
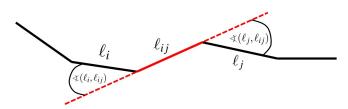

Fig. 6. *Illustration of the notion ℓ_{ij} and $\sphericalangle(\ell_i, \ell_{ij})$.*

Suppose that the line segments $\ell_{i_1}, \ldots, \ell_{i_m}$ are in the cone of ℓ_i and ℓ_i is in the cones of the segments $\ell_{i_1}, \ldots, \ell_{i_m}$, i.e., a connection of the line segment ℓ_i to one of the line segments $\ell_{i_1}, \ldots, \ell_{i_m}$ is possible. Then the probability that a polygonal track p_i is connected to a polygonal track $p_j \in \{p_{i_1}, \ldots, p_{i_m}\}$ via a line segment ℓ_{ij} depends on

1) the values of $\sphericalangle(\ell_i, \ell_{ij})$ and $\sphericalangle(\ell_j, \ell_{ji})$, where the connection probability increases if the differences between the directions of ℓ_i, ℓ_j and ℓ_{ij} get smaller, and

2) the length $|\ell_{ij}|$ of the line segment ℓ_{ij}, where the connection probability decreases with increasing length of the line segment ℓ_{ij}.

In particular, a weight ω_{ij} is considered for each line segment ℓ_{ij} possibly connecting p_i and p_j, which is given by

$$\omega_{ij} = \exp\left(-\frac{1}{2}\left(\frac{|\ell_{ij}|}{\ell_{max}}\right)^2\right) \cdot$$
$$\exp\left(-\frac{1}{2}\left(\frac{\sphericalangle(\ell_i, \ell_{ij}) + \sphericalangle(\ell_j, \ell_{ji})}{2\alpha_{max}(|\ell_{ij}|)}\right)^2\right).$$

The weights ω_{ij} are normalized such that the sum of the normalized weights $\widetilde{\omega}_{ij}$ equals 1. The polygonal track p_i is then connected with p_j with probability $\widetilde{\omega}_{ij}$.

After having connected p_i on either side with other polygonal tracks according to the rule described above, we continue connecting the elongated polygonal track with other tracks until this is not possible anymore, i.e., we can not find any polygonal track which end-segments belong to the cones of the elongated polygonal track, and vice versa. We then end up with a completely elongated polygonal track.

Since polygonal tracks are connected according to a random selection rule, it is obvious that the result of one single run cannot be assured to be optimal. To find nearly optimal solutions, the simulation of connections, i.e., the random construction of elongated polygonal tracks, is repeated 1000 times for the longest (initial) polygonal track. Among these 1000 random elongated polygonal tracks we choose the best one according to a certain evaluation criterion. This evaluation criterion is constructed such that false connections within the randomly elongated fibers are penalized. In the present paper, the evaluation of an elongated polygonal track p is based on the gray-values of those voxels of the underlying tomographic image which are hit by p. More precisely, we discretize the polygonal track p dilated by the unit sphere $b(o, 1) \subset \mathbb{R}^3$. By $I(v_1), \ldots, I(v_{k(p)})$ we denote the gray-values of the 3D tomographic image at the voxels $v_1, \ldots, v_{k(p)}$ hit by the dilated polygonal track $p \oplus b(o, 1)$. As evaluation criterion $e(p)$, we choose the mean of these gray-values, i.e., $e(p) = \sum_{i=1}^{k(p)} I(v_i)/k(p)$. Note that the larger $e(p)$ is the better the polygonal track p is evaluated. The idea of this evaluation criterion is that the gray-values of voxels located on the fibers are larger than the remaining ones. Also note that the extraction algorithm is not restricted to the mean gray-value of fibers. Alternative evaluation criteria could be e.g. the smallest gray-value, the variance of gray-values or a function introduced in Teßmann *et al.* (2010) which takes advantage of the relation of cylindrical structures to the eigenvalues of its Hessian matrices and emphasize those voxel which exhibit the greatest likelihood of belonging to the center lines of fibers.

Subsequently, each of the (short) polygonal tracks $\{p_{i_1}, \ldots, p_{i_k}\}$ belonging to the best elongated polygonal track p is removed from the original set $\{p_1, \ldots, p_n\}$ of polygonal tracks, i.e., the fiber reconstruction procedure described above is now applied to the reduced set $\{p_1, \ldots, p_n\} \setminus \{p_{i_1}, \ldots, p_{i_k}\}$ of polygonal tracks. This procedure is repeated until the reduced set of (short) polygonal tracks is empty. The computational time of the algorithm strongly depends on the number of existing parts of center lines to be connected. In our case, we observe 6053 segments which leads to a computational time of 83 minutes on a desktop computer with a quad-core processor and 12 GB ram.

VALIDATION OF THE ALGORITHM

FITTING OF PARAMETERS

To find appropriate values for the parameters α and ℓ_{max} of the fiber-extraction algorithm described in the previous sections, we use a pragmatic fitting technique. In a first step, we manually choose values for α and ℓ_{max} such that the segmentation algorithm generates fibers which are realistic in the following sense. The fibers extracted for given values of α and ℓ_{max} from the experimental data set of non-woven GDL are used to fit a stochastic single-fiber model, where synthetic fibers drawn from this model are composed to a random 3D system of non-overlapping fibers, see *e.g.* Gaiselmann *et al.* (2012). Then, in order to obtain realistic gray-value images, we add random noise to the synthetic fiber system and choose values of α and ℓ_{max} such that the morphology of the resulting simulated 3D image is in a good visual accordance with that of the tomographic 3D image for non-woven GDL gained by synchrotron tomography (Fig. 7).
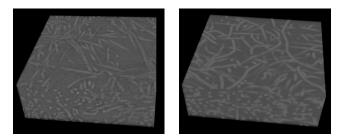

Fig. 7. *3D synchrotron data (left) and synthetic fiber system with random noise (right).*

CHECKING THE QUALITY OF SEGMENTATION

The quality of the segmentation algorithm proposed in this paper is validated by applying it to simulated test data. The starting point for this is the simulated 3D system of non-overlapping fibers considered in the previous sections. The major advantage of this virtual data set is that the morphological properties, *e.g.* the curvature, of its fibers are known. This allows us to perform an exact validation of the fiber-extraction algorithm described in the previous sections, where we applied our algorithm to a noisy version of the synthetic 3D image described above for different values of the parameters α and ℓ_{max}.
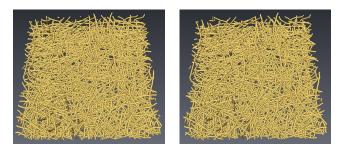

Fig. 8. *Simulated (left) and corresponding extracted (right) fibers.*

It turned out that $\alpha = 0.4$ and $\ell_{max} = 60$ is an optimal choice of parameters which yields the best results with respect to the reconstruction of connections between the extracted (short polygonal) fragments of center lines of fibers, *i.e.*, for this choice of α and ℓ_{max}, the most correct and the fewest wrong connections between the extracted fragments of center lines of fibers have been obtained. More precisely, in this case the algorithm correctly sets 89% of connections, where 7% of connections are not set, and 4% are set wrongly. For a visual comparison of the simulated 3D system of non-overlapping fibers and the system of fibers extracted from a noisy version of this 3D image (Fig. 8). In morphological analysis of fiber-based materials, it is of great importance that from a statistical point of view the extracted fiber tracks have the same structural characteristics as the original fibers. As first characteristic to validate the single-fiber extraction algorithm, we consider the number of fibers in the simulated and extracted fiber systems. Since 7% of connections are not set, we extract a larger number of fibers (903) by the introduced algorithm as existing simulated fibers (781). But when considering only the fibers with an Euclidean length larger than 50 μm (which is the minimum length of the 95% largest simulated fibers), then the number of extracted fibers is equal to 767 and the number of simulated fibers is 742. The small fibers can be seen as artefacts of the proposed segmentation algorithm. In order to check if curvature properties are adequately represented by the extracted fibers, we only compute a curvature characteristic for simulated and extracted fibers larger than 50 μm. More precisely, in order to validate our extraction algorithm, we considered the tortuosity $\tau(p)$ of a polygonal track $p = (a_0, \ldots, a_k)$, given by the starting and end points $a_0, \ldots, a_k$ of its line segments $\ell_1, \ldots, \ell_k$, which is defined by $\tau(p) = \sum_{i=1}^{n-1} d(a_i, a_{i+1})/d(a_1, a_n)$, where $d(a, b)$ denotes the Euclidean distance between $a, b \in \mathbb{R}^3$. Fig. 9 and Table 1 show that the tortuosity is very well reflected by the polygonal tracks extracted from the noisy version of simulated tracks.
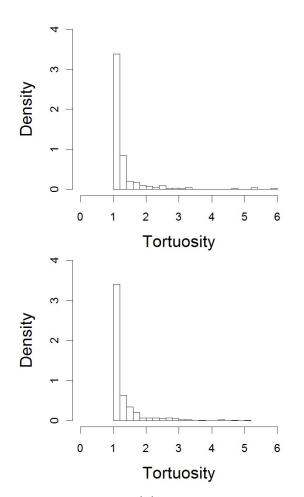
Fig. 9. *Histogram of $\tau(p)$ for simulated (top) and corresponding extracted (bottom) fibers.*

Table 1. *Mean and variance of tortuosities computed for simulated and extracted fibers.*

	mean	variance
simulated fibers	1.28	0.31
extracted fibers	1.28	0.29

APPLICATION TO EXPERIMENTAL DATA

Finally, for the optimal parameter values $\alpha = 0.4$ and $\ell_{max} = 60$, we applied our fiber-extraction algorithm to the experimental 3D data of non-woven GDL, gained by synchrotron tomography. By visual inspection (Fig. 10), we found that the experimental data and the fiber system extracted from it are in good accordance. Notice that the extracted (polygonal) fibers are an important basis in order to develop and fit stochastic 3D models to the microstructure of fiber-based materials (Gaiselmann *et al.*, 2012).
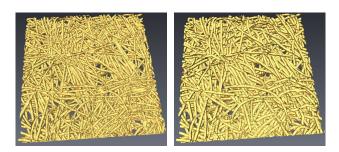

Fig. 10. *Binarized real data (left) and extracted fiber system (right).*

CONCLUSIONS

We introduced a method of structural image segmentation for fiber-based materials which automatically extracts single fibers from tomographic 3D data. To provide an example, we applied the algorithm to 3D synchrotron data of non-woven gas-diffusion layers which are a key component of polymer electrolyte membrane fuel cells. Additionally, the quality of the structural segmentation algorithm has been validated by applying it to simulated test data. The proposed segmentation algorithm has shown to be robust even in the presence of strongly curved fibers and image noise. Moreover, the algorithm is also able to handle tightly packed fiber systems.

ACKNOWLEDGEMENT

REFERENCES

Altendorf H, Jeulin, D (2009). Fiber separation from local orientation and probability maps. In: Wilkinson M, Roerdink J, eds. Proc 9th Int Symp Math Morph, 3336. University of Groningen, The Netherlands.

Burger W, Burge, MJ (2007). Digital image processing: An algorithmic introduction using java. Berlin: Springer.

Dokládal P, Jeulin D (2009). 3-D extraction of fibres from microtomographic images of fibre-reinforced composite materials. In: Wilkinson MHF, Roerdink JBTM, eds. Mathematical morphology and its application to signal and image processing. Lect Not Comput Sci 5720:126–36.

Gaiselmann G, Thiedmann R, Manke I, Lehnert W, Schmidt V (2012). Stochastic 3D modeling of fiber-based materials. Comp Mater Sci 59:75-86.

Hartnig C, Jörissen L, Kerres J, Lehnert W, Scholta J (2008). Polymer electrolyte membrane fuel cells (PEMFC). In: Gasik M, ed. Materials for fuel cells. Cambridge: Woodhead Publishing, 101–84.

Hoshen J, Kopelman R (1976). Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm. Phys Rev B 14:3438–45.

Jeulin D, Kurdy M B (1992). Directional mathematical morphology for oriented image restoration and segmentation. Proc. 8th ISS Congress. Irvine, CA 1991. Acta Stereol 11(Suppl I):545–50.

Mathias MF, Roth J, Fleming J, Lehnert, W (2003). Diffusion Media materials and characterisation. In: Vielstich W, Lamm A, Gasteiger H, eds. Handbook of fuel cells. London: J Wiley & Sons, 517–37.

Soille P (2003). Morphological image analysis: Principles and applications. 2nd ed. Berlin: Springer.

Teßmann M, Mohr S, Gayetsky S, Hassler U, Hanke R, Greiner G (2010). Automatic determination of fiber-length distribution in composite material using 3D CT data. EURASIP J Adv Signal Process 2010:545030.

Thiedmann R, Fleischer F, Lehnert W, Schmidt V (2008). Stochastic 3D modelling of the GDL structure in PEM fuel cells, based on thin section detection. J Electrochem Soc 155:B391–9.