Quantitative comparison of segmentation algorithms for FIB-SEM images of porous media

M. Salzer[a,*], T. Prill[b], A. Spettl[a], D. Jeulin[c], K. Schladitz[b], V. Schmidt[a]

–

martin.salzer@uni-ulm.de, torben.prill@itwm.fraunhofer.de,

aaron.spettl@uni-ulm.de, dominique.jeulin@ensmp.fr,

katja.schladitz@itwm.fraunhofer.de, volker.schmidt@uni-ulm.de

–

[a] Institute of Stochastics, Faculty of Mathematics and Economics, Ulm University, Helmholtzstraße 18, D-89069 Ulm, Germany.

[b] Department of Image Processing, Fraunhofer Institute for Industrial Mathematics, Fraunhofer-Platz 1, D-67663 Kaiserslautern, Germany.

[c] MINES ParisTech, PSL Research University, CMM Centre de morphologie mathématique, 35 rue Saint-Honoré, F-77300 Fontainebleau, France.

[*] Corresponding author. Phone: +49 (0)731/50-23555, Fax: +49 (0)731/50-23649.

–

FIB tomography has proven to be capable of imaging porous structures on a nano-scale. However, due to shine-through artefacts, common segmentation algorithms often lead to severe dislocation of individual structures in

z-direction. Recently, a number of approaches have been developed, which take into account the specific nature of FIB-SEM images for porous media. In the present study, we analyse three of these approaches by comparing their performance based on simulated FIB-SEM images. Performance is measured by determining the amount of misclassified voxels as well as the fidelity of structural characteristics. Based on this analysis we conclude that each algorithm has certain strengths and weaknesses and we determine the scenarios for which each approach might be the best choice.

1 Introduction

Porous materials are of current interest within a wide range of applications such as heat flow management and heat exchange (Clyne, Golosnoy, Tan & Markaki, 2006), drug delivery (Khanafer & Vafai, 2006), energy storage (Less et al., 2012) and catalysis (Kortunov et al., 2005) just to name a few. For all of these and many other applications the properties of the corresponding materials are not fully determined by simple measurements like the absolute porosity or average pore size. Instead, various complex structural properties of the material, such as size and shape of individual pores (or particles) and even more important their spatial relation to each other have to be considered.

One way to gain a better understanding of these underlying structural properties is to reconstruct the different phases (solid phase and pore phase) from a 3D image of the corresponding specimen. Once this is accomplished various mathematical tools (e.g. numerical simulations or statistical analyses) can be applied either directly to the resulting binary image, or a more advanced mathematical model can be matched to the contained structures. For porous structures on a nano-scale, FIB-SEM tomography (Holzer & Cantoni, 2011) has shown to be capable to adequately perform the first step of this workflow, which is to obtain a 3D image of the corresponding microstructure. However, due to the way FIB-SEM images are acquired, segmenting the solid phase from the pore phase becomes a non-trivial problem. FIB-SEM images are obtained by imaging the surface of the specimen with a scanning electron microscope (SEM) and subsequently removing this slice with the aid of a focused ion beam (FIB). This step is repeated until the desired amount of slices is obtained.

For solid materials methods developed for other imaging techniques often can be applied successfully. For porous materials it is often possible to use an epoxy resin to fill the pore space and thereby avoid the problems of imaging porous materials.

However, epoxy resin infiltration is not possible for closed pores and might be even problematic for open pores since the process of infiltration might alter the underlying structure. In these cases the problem of segmentation becomes more complex. Due to the principle of serial-sectioning, parts of the material are visible through the pore phase. Therefore, when applying segmentation algorithms the distance of solid material to the sensor is often significantly underestimated. As shown in (Gunda, Choi, Berson, Kenney, Karan, Pharoah & Mitra, 2011) errors made during the step of image processing have a direct influence on the structural properties we seek to understand.

Therefore several new methods have been developed. Although most of these methods have been validated against manual segmentations (Salzer, Thiele, Zengerle & Schmidt, 2014) or even simulated data (Prill, Schladitz, Jeulin, Faessel & Wieser, 2013) these results only have a limited value. Every microstructure produces a very different combination of artefacts that has to be considered. Therefore, one method might provide good results for one data set but perform very badly on another one.

In this paper, we present a study of three algorithms that are tested on artificially generated FIB-SEM images for multiple scenarios. This has advantages over the way results were validated before. First, using artificially generated images allows us to compare the results of every segmentation algorithm against a voxel-rendering of the original random structures used for the simulation. Contrary to manual segmentation,

which is always subjective, this provides us with objective criteria. Second, we compare different segmentation algorithms on the same datasets. This enables us to compare the measured performance of different algorithms which is not possible when they are applied to different datasets. Third, we study various configurations for the underlying stochastic models, namely different objects (cylinders and spheroids) and different degrees of porosity. This allows us to study the effect of the microstructure on the quality of the obtained segmentation. Finally, we are able to combine the last two steps and compare how well different approaches adapt to a wider range of datasets (e.g. to test the robustness of an algorithm). The paper is organised as follows. In Section 2, we describe the different geometric structures and how they are used to simulate FIB-SEM images. In Section 3, we give a quick overview of the three segmentation algorithms and how the parameters are chosen. In Section 4, we then present the analysis performed on the obtained binary images. Finally, Section 5 provides a summary of the study and a discussion of the results.

2 Data

In the following we give a brief overview on the different data sets used in this paper and how they were obtained.

2.1 Random Set Models

All geometric structures used in this study are realisations of a 3D germ-grain process (see Matheron, 1967; Schneider & Weil, 2008; Chiu, Stoyan, Kendall & Mecke, 2013). This is a stochastic process, where grains are placed at points of a point process. In the data sets designated by Sp_035, Sp_060 and Sp_080 a Boolean model was used, which means that spheres are placed on the points of the Poisson point process. The grains were drawn from a collection of spheres with radii uniformly distributed between 8 and 12 voxels. As indicated in the designation, realisations with volume fractions of 35%, 60% and 80% respectively, were generated.

A modified version of the 'cherry-pit' model (see Torquato, 1984) was used in the data sets Cy_014, Cy_024 and Cy_057, with volume fractions of 14%, 24% and 57% respectively. Instead of spheres, the grains were cylinders with radii uniformly distributed between 8 and 12 and length of 1200 voxels in Cy_057 and 600 voxels for the other two. An impenetrability parameter d_{Hc} was used to control the overlapping of the cylinders. The cylinders were regarded as having an impenetrable core with radius $(1 - d_{\mathrm{Hc}})r$, where r is the radius of the grain. For volume fractions higher than 0.3 it is difficult to impossible to generate a process with non-intersecting cylinders. Thus, the parameters were chosen as $d_{\mathrm{Hc}} = 0$ in Cy_014, which means no intersection of the cylinders is allowed. For Cy_057 ($d_{\mathrm{Hc}} = 0.05$) and Cy_024 ($d_{\mathrm{Hc}} = 0.3$), more intersection was permitted, to achieve the higher volume fractions. The direction angles θ (altitude) and φ (longitude) of the cylinders follow a distribution, which is given by (see Schladitz, Peters, Reinel-Bitzer, Wiegmann & Ohser, 2006)

$$p(\theta, \varphi) = \frac{1}{4\pi} \frac{\beta \sin\theta}{(1 + (\beta^2 - 1)\cos^2\theta}, \quad \theta \in [0, \pi), \varphi \in [0, 2\pi).$$

The anisotropy of the cylinders is characterized by the parameter β. The parameter ranges from $\beta = 1$ for the data set Cy_014, which means that the cylinders are oriented isotropically, to $\beta = 0.1$ in Cy_057, which corresponds to an almost perfect alignment along the z-axis. In Cy_024 an intermediate value of $\beta = 0.25$ was chosen. All models have been generated with the MAVIlib software (Fraunhofer ITWM, Department of Image Processing, 2011).

2.2 FIB-SEM Forward Simulation

All synthetic image data sets were generated with the FIB-SEM simulation tool presented by Prill & Schladitz (2013). The parameters are chosen close to a typical experimental setup, used for imaging geometric structures. More precisely, the beam energy was 5 keV and pixel size, which equals the beam width, was 50 nm. Each data set consists of 100 slices. With an edge length of 400 pixels for every image the simulated volume is $20\mu\text{m} \times 20\mu\text{m} \times 5\mu\text{m}$. For every pixel 10,000 electrons were simulated with a structuring element multiplier $m = 3$. The images were simulated on a cluster, taking between 1 and 5 hours each. The simulation covered both the BSE and the SE signal, but only the SE signal was used for later steps. An example slice of a realisation of a Boolean model and the corresponding simulated FIB-SEM slice is given in Figure 1.

3 Segmentation Algorithms

In this section we give a quick introduction to the segmentation algorithms and how the corresponding parameters were chosen.

3.1 Thresholding [THRESH]

In order to obtain a reasonable and simple reference segmentation we used global thresholding with some additional postprocessing. For global thresholding a voxel (x, y, z) is assigned to the solid phase if and only if the corresponding grey value $I(x, y, z)$ is above a certain threshold T. To obtain a more competitive and thus meaningful reference segmentation we added two further steps. First, we used two thresholds (T_{lower} and T_{upper}), instead of one. Here, all voxels whith a grey value $T_{\text{lower}} \leq I(x, y, z) \leq T_{\text{upper}}$ were assigned to the solid phase, all other voxels were assigned to the pore phase. This significantly reduced artefacts caused by relatively high grey values that occur at the border of particles. After this step, we applied a median filter with a radius $r = 2$ to compensate for small isolated clusters of misclassified voxels. The parameters were determined by the following approach: For each $T_{\text{lower}} \in [0, 255]$ the second parameter T_{upper} was chosen so that the corresponding binary images after the median filter would match the volume fraction optimally. For all of these sets $(T_{\text{lower}}, T_{\text{upper}})$ the set with minimum error ratio was chosen as the final set of parameters.

Local thresholding as described by Sauvola & Pietikäinen (2000) has been tested but the result did not improve significantly compared to global thresholding (for details on the used implementation see Shafait, Keysers & Breuel, 2008). This was expected as the

images are widely homogenous (e.g. no illumination differences throughout the images), thus a local approach was unlikely to provide any benefits over a global one.

3.2 Local Threshold Backpropagation [LTB]

In (Salzer, Spettl, Stenzel, Smått, Lindén, Manke & Schmidt, 2012) a new approach to automatic segmentation of FIB-SEM images has been proposed. This approach is based on estimating a reasonable local threshold for the last slice right before the corresponding substructure was being cut off. These thresholds are then backpropagated to earlier slices, hence local threshold backpropagation (short: LTB). In the original description of the algorithm only one threshold was mentioned that served as a lower bound. For this paper we used a second theshold that serves as an additional upper bound. Again, this removed artefacts induced by high grey values at the borders of particles. Due to different properties of the material and the surface, these illumination effects did not occur in the data set studied by Salzer, Spettl, Stenzel, Smått, Lindén, Manke & Schmidt (2012). This can be considered as a generalisation where we derive the earlier version of the algorithm by setting the upper threshold to the maximum possible grey value (in this case 255). The postprocessing step was adapted to the datasets by manually choosing adequate parameters for various filters and cluster detections.

3.3 Valley Detection [VALLEY]

The second algorithm was first described in (Salzer, Thiele, Zengerle & Schmidt, 2014). In this approach we attempt to detect the last and first occurrences of structures by

selecting outstanding local minima and maxima, respectively. By this we estimate the "valley" in grey values in z-direction that represents the pore phase between two microstructures. Similar to the previously described algorithm minor modifications have been made. For the main step we ignored all local extrema within a certain range of grey values, determined by $[T_{\text{lower}}, T_{\text{upper}}]$. Further, we added a second (upper bound) threshold to all steps during postprocessing that included thresholding. Like above this change can be considered a generalisation of the original method where the upper threshold is set to the maximum grey value. Again these and all other parameters were adjusted manually to provide a good optical fit.

3.3.1 Median Filter [VAL-MED]

During the present study we included one additional segmentation that we denote by VAL-MED. This segmentation was derived from VALLEY by applying a median filter with radius $r = 2$ to the segmentation. This additional step of postprocessing was not part of the original algorithm, but it removes a huge amount of small clusters of misclassified voxels that occurred for images with spheroids and a high volume fraction. For more details see Section 4.2.

3.4 Morphological Segmentation [MORPH]

The morphological segmentation method was described in (Prill, Schladitz, Jeulin, Faessel & Wieser, 2013). For experimental data the segmentation starts with a preprocessing step, but due to the perfect alignment and the low signal to noise ratio of the simulated

data, this was omitted here. The segmentation is based on the detection of morphological features on the z-profiles of the images. First a thresholding excludes pixels that can be identified as void phase from their grey value. Then a morphological half-gradient image identifies pixels which are considered void phase due to the specific shape of the artifacts in the FIB-SEM data set. Finally, minima of a certain minimal depth are identified by a morphological reconstruction. Then the features are combined to generate a preliminary segmentation which is then refined using a constrained watershed transformation on a morphological gradient image. The detailed description and a validation with synthetic data can be found in (Prill, Schladitz, Jeulin, Faessel & Wieser, 2013).

The segmentation algorithm has 5 free parameters which where chosen according to the guidelines given in (Prill, Schladitz, Jeulin, Faessel & Wieser, 2013). The parameters are threshold on the grey value (T_GV in Figure 1), a threshold on the morphological gradient (T_MG), a minimal dynamic for the minima (D_MIN), a length of the structuring element in the morphological half-gradient (L_HG) and a threshold on the half-gradient image (T_HG). The values chosen for these parameters are shown in Table 1.

4 Analysis and Results

In this section we provide a brief description of the measurements used to characterise the various segmentations, present the corresponding results and give some interpretation.

Prior to the analysis all segmentation images were cropped and rescaled – alignment, which is usually necessary for experimental data, was not performed as the simulated

images did not contain any shift. The cropping aperture in z-direction was determined by analysing the error ratio on a slice-to-slice basis. There, two general trends have been observed. On the one hand LTB performs significantly worse for the last 5–10 slices, on the other hand VALLEY and MORPH show a higher error ratio for the first 10–15 slices. Both effects are expected as LTB relies on the last occurrence and VALLEY/MORPH on the first occurrence of a microstructure to detect it. Thus microstructures that end or begin (respectively) outside of the operation window cannot be detected properly. These problems have been known before, but this is the first time we present a quantification. Based on this analysis, we excluded the first 15 and last 10 slices from later steps. For an example of both described trends see Figure 2.

4.1 Basic Measurements

For a first overview we considered the error for the estimated volume fraction and the amount of misclassified voxels – the latter for each phase individually and for both together. For the volume fractions MORPH seems to provide the best fit with the lowest number for both the mean and the maximum difference to the original data. However with the exception of LTB and THRESH numbers are relatively similar, for more details see Table 2.

The error ratios for misclassified voxels as shown in Table 3 are more informative (to see how these errors are distributed amongst both phases see Table 4). When comparing different scenarios for individual segmentation algorithms, we can observe two different trends: For THRESH the error ratio decreases significantly with higher volume fractions.

This effect is expected as an increasing volume fraction leads to decreasing porosity and thereby to less porosity-related artefacts during segmentation. This effect of decreasing error ratios for decreasing porosity can also be observed for MORPH, which produces the lowest error ratios of all algorithms for the two previously mentioned scenarios.

For LTB and VALLEY (VAL-MED) on the other hand this does not seem to be true. Contrary to the less-porosity-less-artefacts-trend both perform surprisingly well for the cylinder scenario with the lowest porosity. We suspect that the corresponding mechanisms to detect microstructures based in the z-profile perform better when the distance between two subsequent microstructures (in z-direction) is larger, which is the case for higher porosities. This effect however cannot be observed for the other scenarios (i.e. comparing Cy_024 to Cy_057 or two of Sp_035, Sp_060 and Sp_080), which might be related to the introduction of additional thresholds as described in Section 3. Presumably for higher volume fractions more voxels are classified based on the thresholding criteria instead of the corresponding z-profile.

For all tested images MORPH and VAL-MED perform better than THRESH. VALLEY outperforms THRESH for all images but one, i.e. Sp_080, which has the highest volume fraction. LTB one the other hand outperforms THRESH for three images with the lowest volume fraction. Although for each image one of the other approaches performs better, LTB still might have its use for cases where huge differences in illumination occur. All other approaches at some point rely on a global threshold. For images where it is impossible to choose this parameter adequately, LTB still might be the method of choice.

4.2 Structural Analysis

Although measurements like the error ratio already give a good first impression of the quality of segmentation these measurements are far from sufficient to judge a segmentation algorithm. As mentioned in Section 1, for most applications the performance of a certain material is determined by various structural properties of both the pore phase and the solid phase. Thus, in order to give a meaningful judgment we have to consider how the errors measured above influence the preservation of important structural properties. In the following we provide the results of some of the most common structural characteristics, namely the covariance function and granulometry (Matheron, 1967). The covariance function provides information about the typical length scale in the image and was computed in all three spatial dimensions. In order to reduce the number of plots x and y were combined into one curve. See Figures 4 and 5 and for results of covariance function in xy- and z-direction, respectively. For granulometry a series of morphological openings is performed before measuring the volume fraction. This yields the distribution of particle sizes in the binary image. Results for granulometry for all images and all algorithms (with the original Boolean model as a reference) can be found in Figure 6.

In general, we find the trends described in Section 4.1 supported. However, it is difficult to find a consistent interpretation, as individual artefacts have a different impact on each measurement. For example, for Cy_014 and Cy_024 the covariance function supports the observation that LTB performs quite well on these images. However, the granulometry shows significant differences for Cy_014. Visual inspection shows that

some of the misclassified voxels appear outside of partially detected cylinders. For Morph, which has a similar error rate but a better fit for the original granulometry function, less partially detected cylinders can be observed. Instead the errors derive from completely missing cylinders (see Figure 7 for a visual example of both cases). While partially detected cylinders disturb the estimation of particle size, an appropriate object extraction algorithm might be able to reconstruct the missing part. For completely missing cylinders this seems less likely.

Valley also performs relatively poor under granulometry for some of the images. Similar to LTB, visual inspection shows small misclassified clusters within correctly classified regions of foreground voxels (see Figure 8 for a visual example). Although of relatively small size, these clusters become quite big under a morphological opening and thus strongly influence the volume fraction measured by granulometry. In this case, however, we were able to remove these artefacts by adding one step of postprocessing. As these misclassified clusters mostly consist of individual voxels they are easily removed by a small-ranged median filter. This is reflected by Val-Med performing significantly better in all granulometries than Valley. Combined with consistent decrease for global error rates (as shown in Figure 3) and no significant negative impact on the covariance function this suggests that the additional postprocessing resolves the issue of misclassified clusters without lowering other aspects of the segmentation.

Furthermore, results of granulometry help with interpretation of error ratios for Thresh. The maximal particle size of each binary image is related to the smallest radius for which granulometry is zero. With exception of Thresh (and Valley) this radius is roughly

the same for all methods (including the original binary image). For THRESH, however, we observe a significant amount of particles with a greater size and the maximal particle size is significantly overestimated. This becomes most obvious for Cy_057 and Sp_080. For these data sets error ratios of THRESH are close to other approaches, however, granulometry reveals that it fails to cover important structural features.

Results for covariance functions in z-direction generally seem to be more widespread than for x and y-direction which is fitted quite well for most cases. Due to the principle of serial-sectioning, it is more difficult to estimate the position and elongation of an object in z-direction than in x and y direction. Therefore, it is expected that differences in quality of segmentation primarily become visible in z-direction.

5 Summary and Discussion

In this paper we present a study comparing various segmentation algorithms for porous media. Using the algorithm introduced in (Prill & Schladitz, 2013), FIB-SEM images were simulated for multiple scenarios, which are based on Boolean models and cover different geometric structures and porosities. By using such artifical images the "ground truth" is known and we have a common base of very different data sets to be used as input for all algorithms. Therefore, after a short explanation of the investigated segmentation algorithms THRESH (based on global thresholding), MORPH (Prill, Schladitz, Jeulin, Faessel & Wieser, 2013), LTB (Salzer, Spettl, Stenzel, Smått, Lindén, Manke & Schmidt, 2012) and VALLEY/VAL-MED (Salzer, Thiele, Zengerle & Schmidt, 2014),

these algorithms were applied to all simulated FIB-SEM images. The obtained binary images were compared to the known corresponding ground truth. For a first evaluation the amount of misclassified voxels and the volume fraction for each obtained image was computed. Because a voxel-based comparison does not reflect the relevance or severity of misclassifications, we also analysed the amount of structural information that was preserved. This is achieved by comparing structural characteristics, i.e., the covariance function and granulometry.

We conclude that different segmentation algorithms can be recommended for different scenarios. For samples with a very low porosity Thresh might already provide sufficient results, because the shine-through artefacts are negligible. Valley/Val-Med on the other hand provides the best results of all approaches for highly porous samples. For intermediate and low porosities Morph seems to be the best choice. Finally, LTB is suitable for images where huge illumination differences prevent other approaches from being used, because they require global thresholds at some point. In practice, such additional artefacts depending on the experimental setup or the samples themselves are likely to play an essential role deciding which algorithm provides the best results, therefore our recommendations are to be seen as a general guideline.

References

Chiu, S.N., Stoyan, D., Kendall, W.S., and Mecke, J. (2013) *Stochastic Geometry and Its Applications*, Wiley, Chichester, 3nd edition.

Clyne, T., Golosnoy, I., Tan, J., and Markaki, A. (2006) Porous materials for thermal management under extreme conditions. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 364(1838), 125–146.

Fraunhofer ITWM, Department of Image Processing (2011) MAVIlib – modular algorithms for volume images, `http://www.itwm.fraunhofer.de/en/departments/image-processing/microstructure-analysis/mavi.html`.

Gunda, N.S.K., Choi, H.W., Berson, A., Kenney, B., Karan, K., Pharoah, J.G., and Mitra, S.K. (2011) Focused ion beam-scanning electron microscopy on solid-oxide fuel-cell electrode: Image analysis and computing effective transport properties. *Journal of Power Sources*, 196(7), 3592–3603.

Holzer, L. and Cantoni, M. (2011) Review of FIB-tomography, in *Nanofabrication Using Focused Ion and Electron Beams: Principles and Applications* (edited by I. Utke, S.A. Moshkalev, P. Russell), pp. 410–435, Oxford University Press, New York.

Khanafer, K. and Vafai, K. (2006) The role of porous media in biomedical engineering as related to magnetic resonance imaging and drug delivery. *Heat and Mass Transfer*, 42(10), 939–953.

Kortunov, P., Vasenkov, S., Kärger, J., Fé Elía, M., Perez, M., Stöcker, M., Papadopoulus, G.K., Theodorou, D., Drescher, B., McElhiney, G., Bernauer, B., Krystl, V., Kočiřík, M., Zikánová, A., Jirglová, H., Berger, C., Gläser, R., Weitkamp, J., and Hansen, E.W. (2005) Diffusion in fluid catalytic cracking catalysts on various dis-

placement scales and its role in catalytic performance. *Chemistry of Materials*, 17(9), 2466–2474.

Less, G., Seo, J., Han, S., Sastry, A., Zausch, J., Latz, A., Schmidt, S., Wieser, C., Kehrwald, D., and Fell, S. (2012) Micro-scale modeling of li-ion batteries: Parameterization and validation. *Journal of the Electrochemical Society*, 159(6), A697–A704.

Matheron, G. (1967) *Éléments pour une théorie des milieux poreux*, Masson et Cie.

Prill, T. and Schladitz, K. (2013) Simulation of FIB-SEM images for analysis of porous microstructures. *Scanning*, 35(3), 189–195.

Prill, T., Schladitz, K., Jeulin, D., Faessel, M., and Wieser, C. (2013) Morphological segmentation of FIB-SEM data of highly porous media. *Journal of Microscopy*, 250(2), 77–87.

Salzer, M., Spettl, A., Stenzel, O., Smått, J.H., Lindén, M., Manke, I., and Schmidt, V. (2012) A two-stage approach to the segmentation of FIB-SEM images of highly porous materials. *Materials Characterization*, 69, 115–126.

Salzer, M., Thiele, S., Zengerle, R., and Schmidt, V. (2014) On the importance of FIB-SEM specific segmentation algorithms for porous media. *Materials Characterization*, 95, 36–43.

Sauvola, J. and Pietikäinen, M. (2000) Adaptive document image binarization. *Pattern Recognition*, 33, 225–236.

Schladitz, K., Peters, S., Reinel-Bitzer, D., Wiegmann, A., and Ohser, J. (2006) Design of acoustic trim based on geometric modeling and flow simulation for non-woven. *Computational Materials Science*, 38(1), 56–66.

Schneider, R. and Weil, W. (2008) *Stochastic and Integral Geometry*, Springer, Heidelberg.

Shafait, F., Keysers, D., and Breuel, T.M. (2008) Efficient implementation of local adaptive thresholding techniques using integral images. *Document Recognition and Retrieval XV*, 6815(1), 681510.

Torquato, S. (1984) Bulk properties of two-phase disordered media. i. cluster expansion for the effective dielectric constant of dispersions of penetrable spheres. *The Journal of Chemical Physics*, 81(11), 5079–5088.

Captions (including Tables/Figures)

	T_GV	T_MG	D_MIN	L_HG	T_HG
Cy_014	39	22	7	9	1
Cy_024	37	24	7	9	1
Cy_057	26	45	25	5	1
Sp_035	43	43	12	5	1
Sp_060	37	42	10	5	1
Sp_080	42	42	13	6	1

Table 1: Choice of parameters for MORPH for each data set.

algorithm	mean	min–max
Thresh	9.6%	4.8%–23.3%
LTB	10.5%	4.2%–19.8%
Morph	2.3%	0.4-4.6%
Valley	3.5%	0.3%–6.7%
Val-Med	3.2%	0.4%–5.5%

Table 2: Table of the absolute errors made for the volume fractions. For each segmentation algorithm the absolute difference to the correct volume fraction is computed for each image individually and then accumulated by mean, mininum and maximum values over all analysed images.

Im. / Algo.	Thresh	LTB	Morph	Valley	Val-Med
Cy_014	15.4%	10.3%	10.7%	7.0%	6.6%
Cy_024	14.1%	12.9%	8.7%	8.5%	7.4%
Cy_057	6.6%	13.2%	4.3%	7.1%	4.5%
Sp_035	22.1%	23.0%	12.3%	15.9%	13.9%
Sp_060	18.0%	22.3%	11.7%	13.5%	11.3%
Sp_080	10.1%	19.9%	7.8%	11.7%	8.9%

Table 3: Table of error ratios for misclassified voxels.

Im. / Algo.	Thresh (fg/bg)	LTB (fg/bg)	Morph (fg/bg)	Valley (fg/bg)	Val-Med (fg/bg)
Cy_014	58%/9%	28%/7%	38%/6%	29%/3%	28%/3%
Cy_024	40%/6%	21%/10%	31%/2%	25%/3%	23%/3%
Cy_057	5%/8%	14%/10%	4%/3%	8%/4%	6%/1%
Sp_035	28%/17%	35%/14%	17%/8%	24%/10%	22%/8%
Sp_060	14%/26%	26%/14%	7%/20%	14%/11%	12%/8%
Sp_080	6%/32%	20%/18%	2%/39%	10%/17%	7%/14%

Table 4: Table of error ratios for misclassified voxels that originally belong to the solid phase (fg) or pore phase (bg).

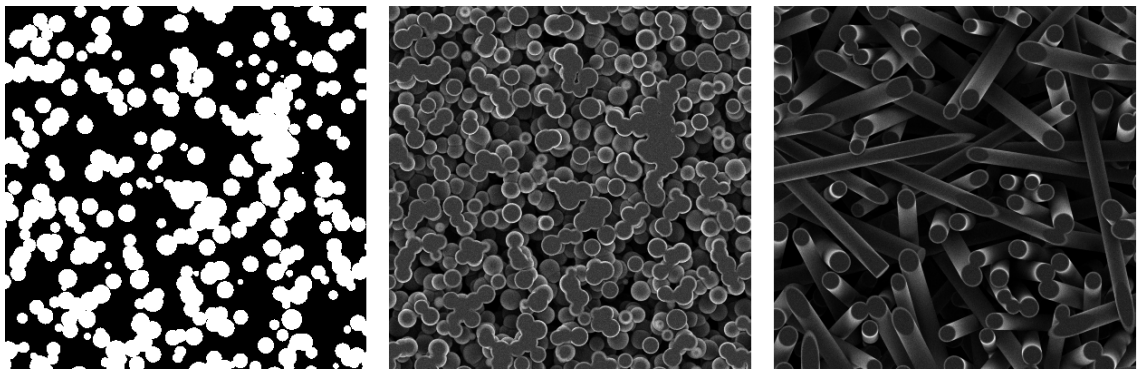

Figure 1: Slice of the data set for Sp_035 (left) and its corresponding simulated SEM slice (middle). On the right: a simulated SEM slice for Cy_024. Both SEM slices show the SE signal of the corresponding image.

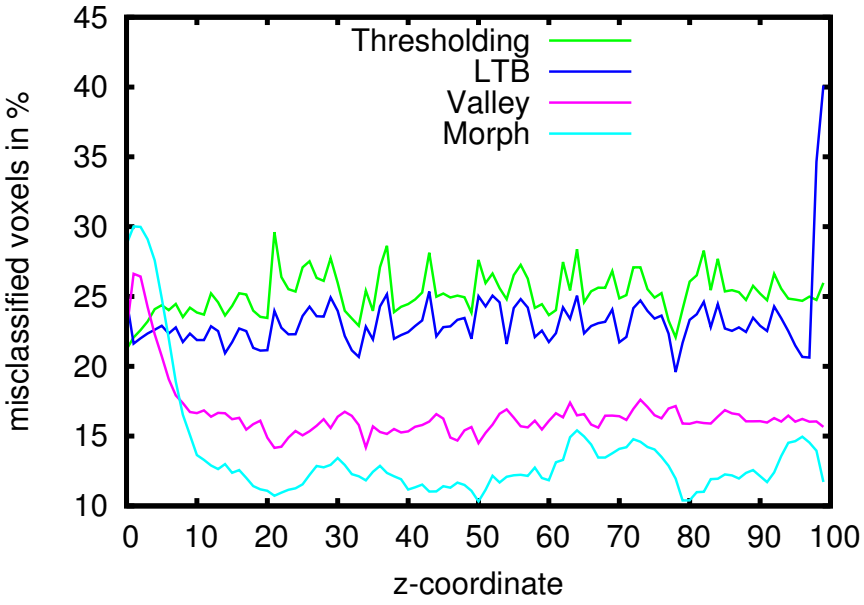

Figure 2: Error rates in the reconstructed binary image of Sp_035 for given z-coordinates. Peaks can be observed in the beginning (MORPH and VALLEY) and the end (LTB).

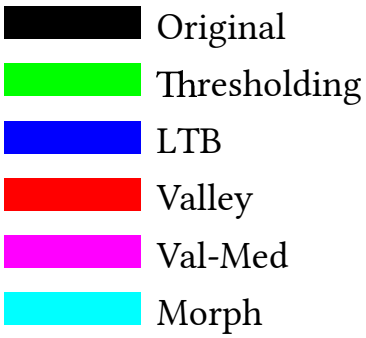

Figure 3: Color table. In all of the following plots the same color is used for the corresponding algorithm.

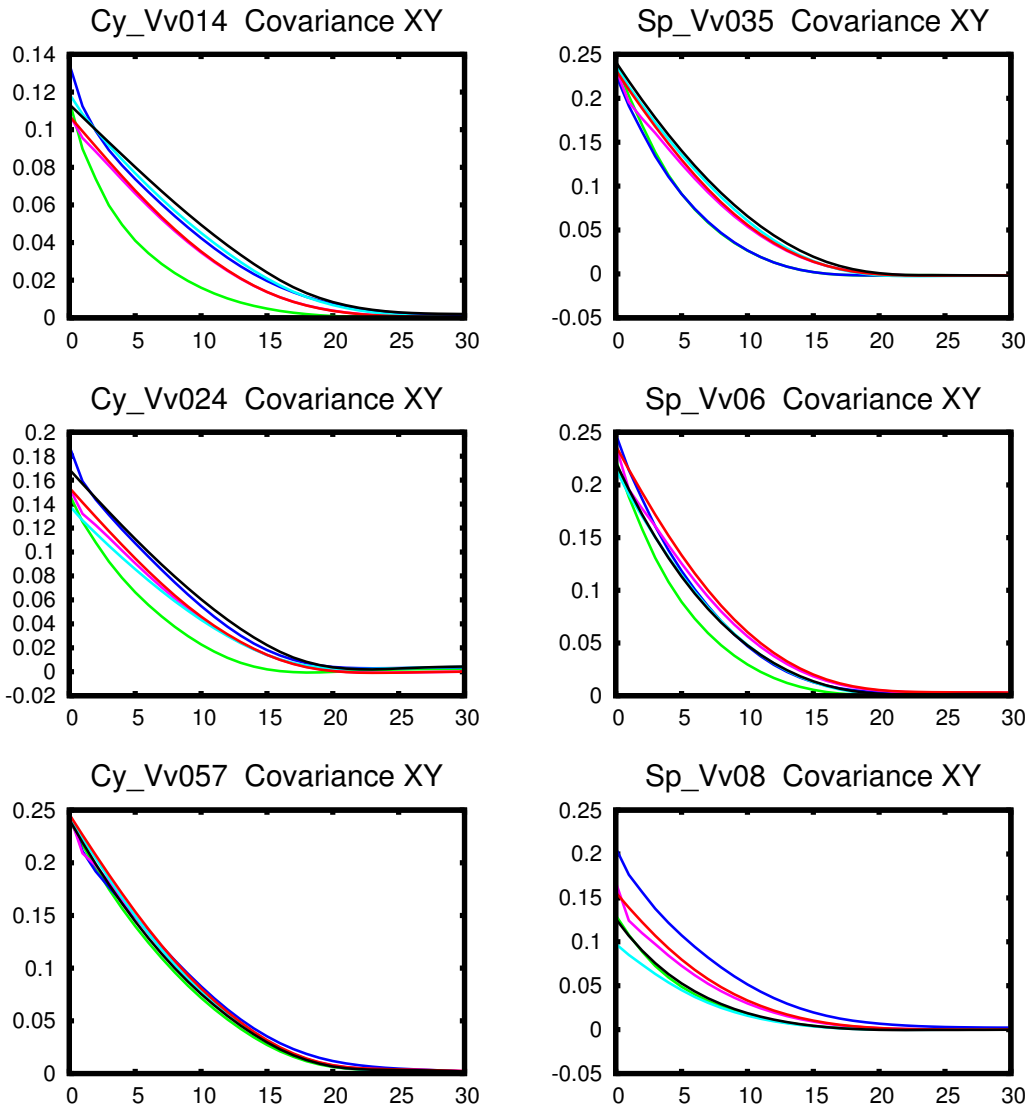

Figure 4: Plots for covariance in x- and y-direction combined (cylinders on the left, spheres on the right, volume fractions increasing downwards).

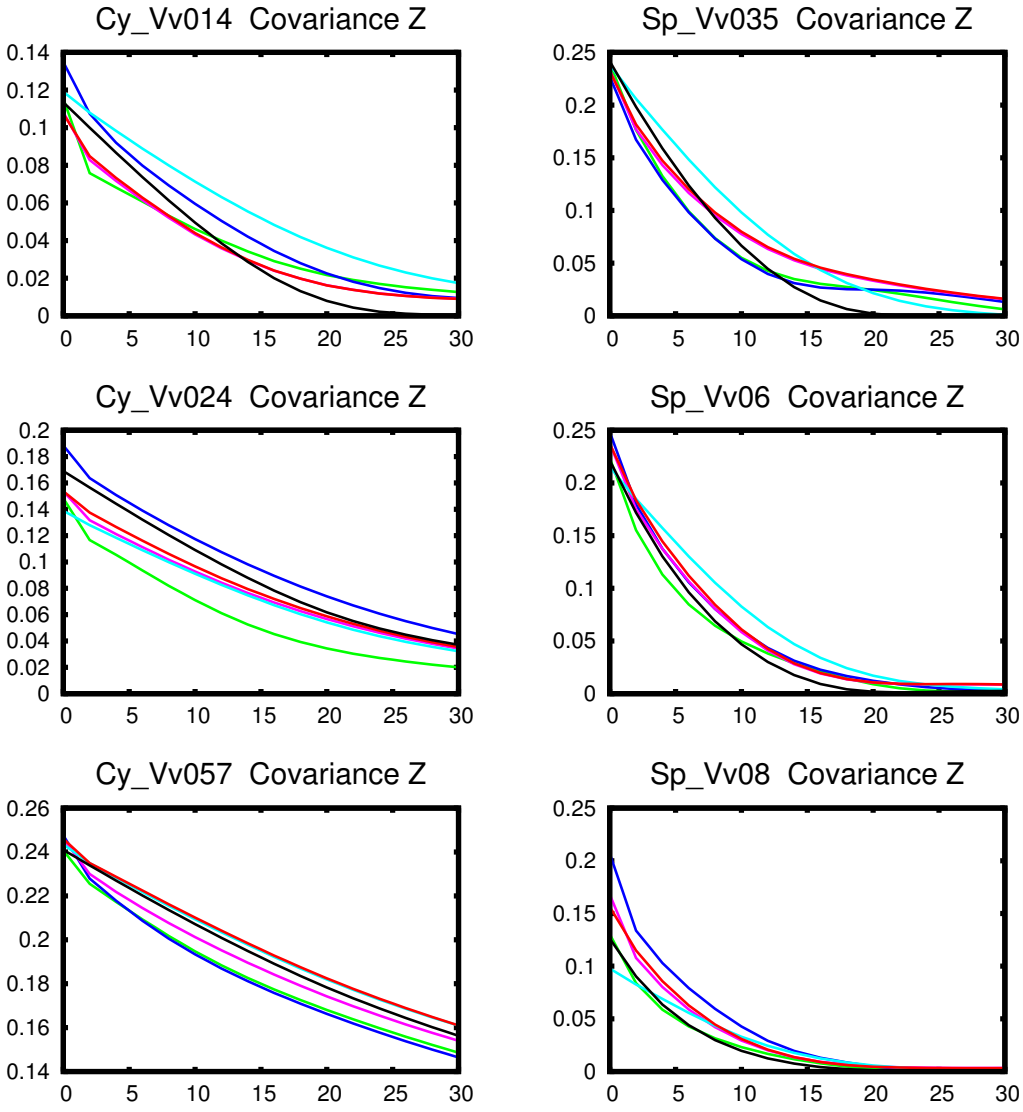

Figure 5: Plots for covariance in z-direction (cylinders on the left, spheres on the right, volume fraction increasing downwards).

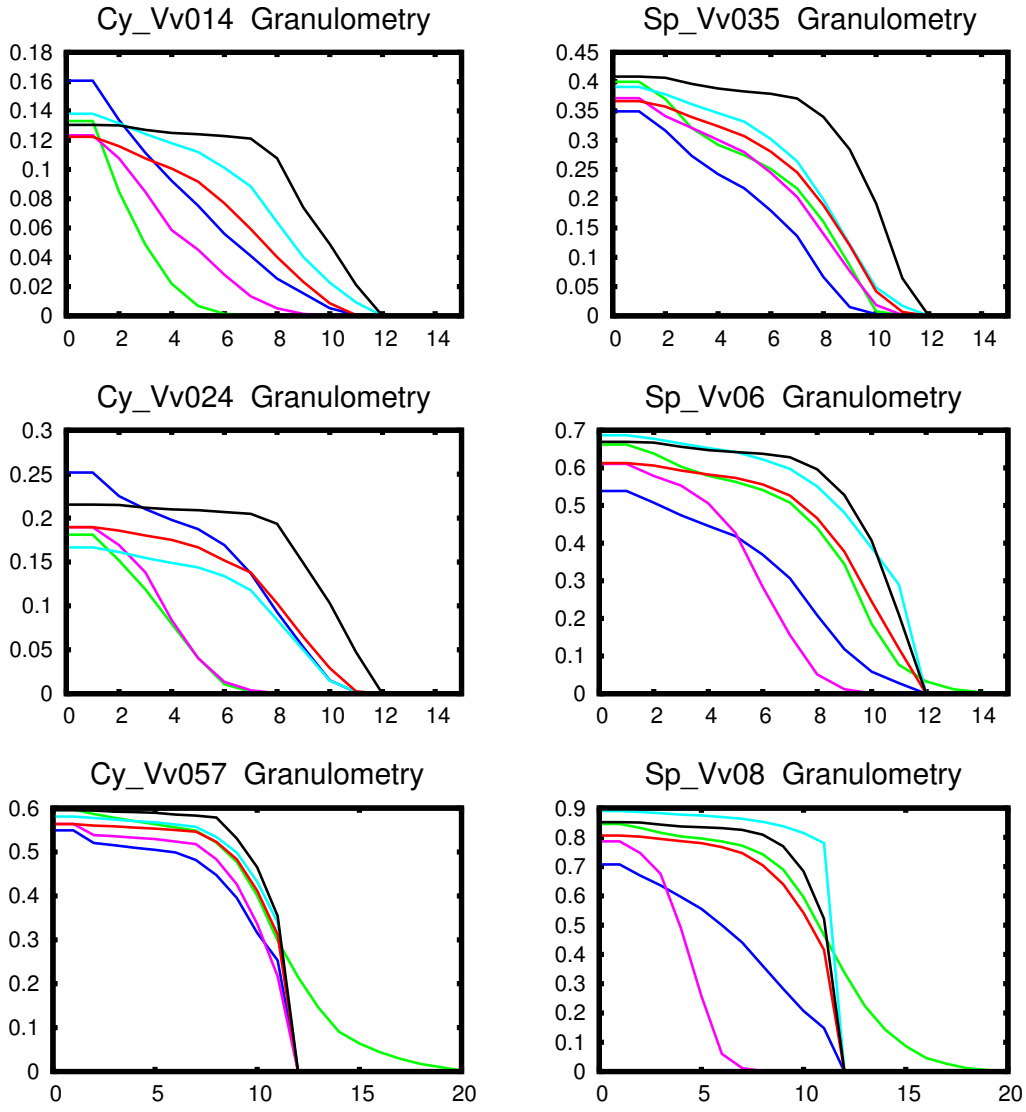

Figure 6: Plots for granulometry (cylinders on the left, spheres on the right, volume fraction increasing downwards).

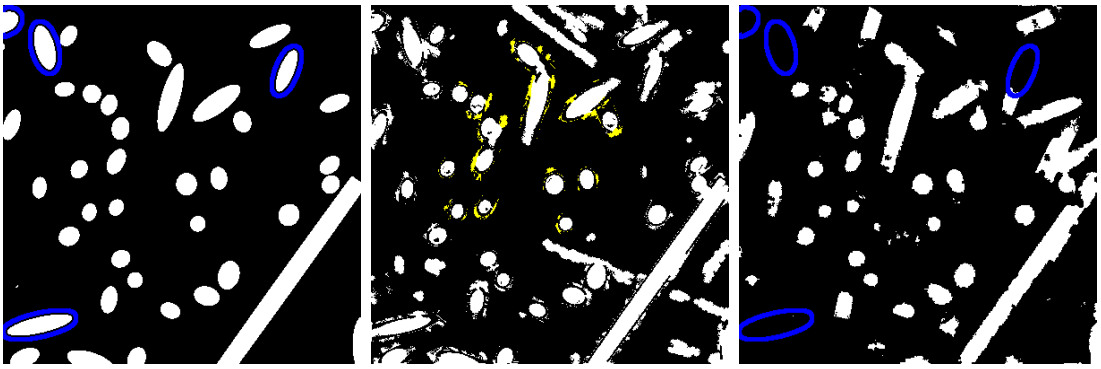

Figure 7: Visual example for different types of segmentation errors. Left: the original binary image as a reference. Middle: LTB with small clusters of misclassified voxels (partly highlighted with yellow). Right: MORPH with some missing cylinders (blue). For particle size distribution (e.g. granulometry) these missing structures have little negative impact, since the distribution is estimated based on present structures. The smaller clusters however have a significant impact. However, they are presumably easier to remove in later processing, including particle-extraction.

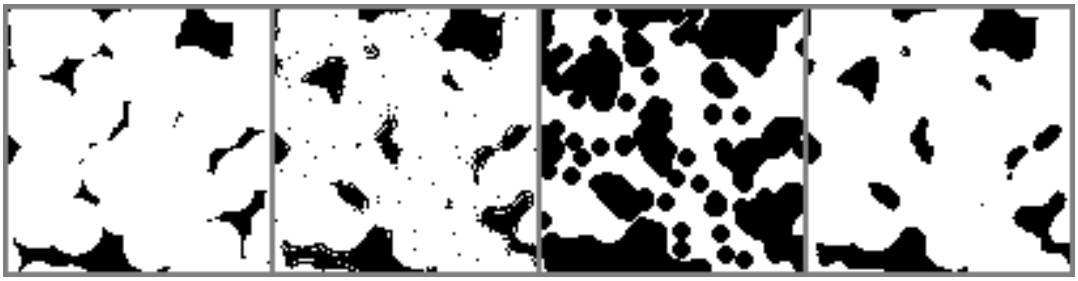

Figure 8: Visual example for the artifacts produced by VALLEY. From left to right: The original binary as a reference. Then, the binary as obtained by VALLEY with small clusters of misclassified voxels. The third image shows how these small clusters get enlarged by a morphological erosion, which is performed during granulometry. On the right: the result of VAL-MED with most of the clusters removed.