MORE THAN PRETTY IMAGES – TOWARDS CONFIDENCE BOUNDS ON SEGMENTATION THRESHOLDS

Moonen P.[*1,2] and Sénéchal P.[†1]

[1] CNRS/ Univ Pau & Pays Adour/ E2S UPPA, DMEX-IPRA, UMS3360, France
[2] CNRS/ Total/ Univ Pau & Pays Adour/ E2S UPPA, LFCR-IPRA, UMR5150, France

Keywords: Gaussian mixture model, segmentation, classification, uncertainty

Summary: We present an approach to assess the uncertainty associated to classifying voxels into different material phases. The approach consists in a spectral deconvolution of the grey-level histogram using a Gaussian mixture approach, followed by an iterative classification procedure based on the occurrence frequency. As phase attributions become increasingly more uncertain as iterations proceed, confidence bounds on the final segmentation are naturally obtained.

1. INTRODUCTION

X-ray tomography yields a three-dimensional dataset, in which each voxel has a grey-value. Ideally, distinct material phases are mapped onto different grey values. Yet, image noise, partial volume effects and phase-contrast artifacts all cause grey-level variations. Consequently, segmentation techniques based on grey-level intensity are prone to some degree of uncertainty, and this uncertainty propagates to all derived quantities such as volume fractions and interfacial areas.
Assessing the degree of uncertainty is not straightforward. A brute force approach could consist in varying the segmentation thresholds and evaluating their impact on a quantity of interest. But over which range should one very the threshold?
In this work, an alternative approach is proposed. The central idea is to deconvolve the global grey-level histogram into Gaussian kernels, each supposedly corresponding to a material phase. Subsequently, an iterative procedure enables attributing each voxel to a phase. The attribution is done by comparing the relative occurrence frequency of a voxel's grey level to an iteration-specific probability threshold. The threshold is lowered with each subsequent iteration until all voxels are classified. The algorithm thus yields the most probable material class for each voxel, and the probability associated to this attribution. The latter enables defining confidence bounds on derived quantities.

2. CONFIDENCE-BASED CLASSIFICATION METHOD

The proposed method is based on the global grey-level histogram of a tomographic dataset. Ideally, a histogram would be composed of distinct Dirac-type distributions, corresponding to each of the n material phases in the considered sample volume. Artefacts (e.g. noise) and partial volume effects tend to spread out these distributions, yielding a distribution resembling a sum of Gaussian kernels.
A Gaussian mixture model (GMM) is a probabilistic model that assumes that the grey level values of each voxel in the dataset are generated from a mixture of a finite number of Gaussian distributions, one per material phase. The characteristic parameters (weight, mean and variance) of these Gaussian distributions are unknown. Several techniques exist to determine them such as expectation maximization and variational inference [1], but they generally require the number of material phases to be known *a priori*. To overcome this limitation, one can infer the number of phases using a parameter-free Markov Chain that relies entirely on Gibbs sampling [2]. This is the approach used in the current paper.
In a second step, the probability that a voxel belongs to a given material is assessed by means of an iterative procedure. Hereby each iteration corresponds to a confidence level P_{conf}. During the first iteration only the most probable phase attributions are carried out, and the confidence level is gradually lowered with each subsequent iteration until all voxels are attributed. The attribution itself is based on the grey level of the voxel under study. One calculates the relative share of each of the n material phases to the total number of voxels with this color. Material phases for which the relative share is smaller than $1 - P_{conf}$ are considered unlikely and are eliminated

* e-mail: peter.moonen@univ-pau.fr
† e-mail: pascale.senechal@univ-pau.fr

as possible options. The voxels for which only one option remains are attributed to the corresponding phase with confidence level P_{conf}. Optionally, one can repeat this step, but instead of considering grey values of individual voxels, one can use the average grey-value of the immediate neighborhood of a voxel. This tends to smoothen the interior of material phases, while preserving the edges. After this optional step, the confidence threshold is lowered and the procedure is repeated. In this way, one iteratively classifies all voxels and simultaneously constructs a confidence map.

3. RESULTS

The proposed classification strategy was applied to a tomographic acquisition of a sphere pack clogged with fine sand. Figure 1 illustrates the different steps of the method. A slice through the sphere pack is shown in (a) and the corresponding histogram in (b). GMM enables decomposing the histogram into three overlapping Gaussian kernels (see b), respectively corresponding to air, sand and ceramic spheres. Application of the method outlined above yields the segmentation shown in (c) and its corresponding confidence map in (d). The confidence map reveals that the bulk of the fine sand particles can reliably be segmented, yet that the exact location of the border of this phase is error-prone. The 50%-value indicates that the bounding voxel has a 50-50 chance to belong to either phase touching the border. This impacts the estimated volume fraction of the sand phase. If one only considers the attributions with over 95% confidence, this volume fraction is estimated at 27%. Lowering the confidence level to 50% yields a volume fraction of 28%. While it is clear that the obtained "confidence levels" are not by any means a rigorous statistical quantity, they do reflect the inherent uncertainty of intensity-based segmentation methods and they do avoid the subjective action of manually selecting optimal thresholds. As such, we believe that they are a valuable contribution to render imaging-based results more quantitative.

References

[1] Christopher M Bishop. Pattern recognition and machine learning. 2006.

[2] C.E. Rasmussen. The Infinite Gaussian Mixture Model. in: S.A. Solla, T.K. Leen and K.-R. Müller (eds.), Advances in Neural Information Processing Systems 12, 554–560, MIT Press, 2000.

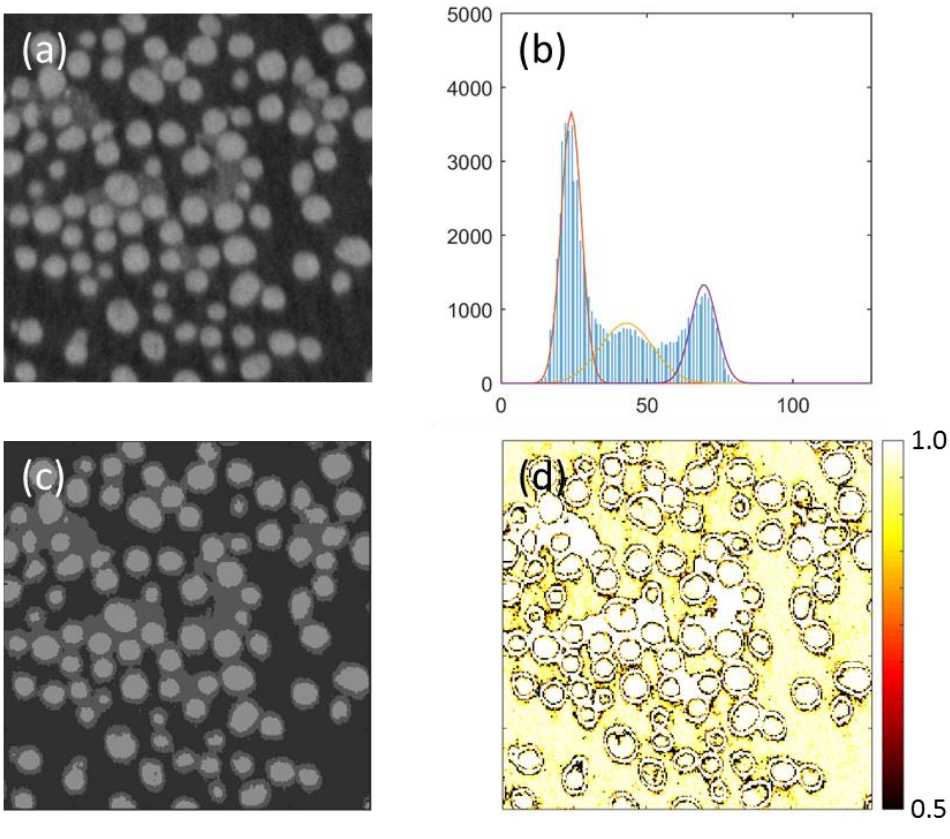

Figure 1: (a) slice through a tomographic dataset, (b) corresponding histogram and inferred Gaussian kernels, (c) segmented slice, and (d) corresponding uncertainty map.