Solid Earth Discuss., 7, C1994–C1998, 2016
www.solid-earth-discuss.net/7/C1994/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Solid Earth
Discussions

Interactive comment on "Beam-hardening correction by a surface fitting and phase classification by a least square support vector machine approach for tomography images of geological samples" *by* F. Khan et al.

F. Khan et al.

kersten@uni-mainz.de

Received and published: 9 March 2016

The critical remarks of referee #1 contained three main comments, and two additional remarks. Here we have addressed them as follows.

Comment #1: "Demonstrating the suitability of the workflow with only one individual sample is problematic. I think it is clear that the LS-SVM segmentation of an uncorrected image has to fail and one illustrative sample is definitively enough to show this. But the paper would be more sound if the robustness of each method (BH correction

and LS-SVM segmentation - not its combination) would be demonstrated with more test images, potentially more complicated ones."

Author's answer: Figure 3 and 5: We have agreed with the reviewer suggestion to provide further test images for the robustness of our proposed method. Therefore, Fig. 3 are modified and two more test images are added. The previous tested image of a rock core sample was already a complex structure example. However, we agree that the way BH correction was previously illustrated in Fig. 4 was not sufficient to observe the position of the fitted 2D-polynomial surface corresponding to the range of grey values of different phases. Therefore, a new Fig. 5 is added, and this also helps to respond and clarify this reviewer comment #3 (see below).

Comment #2: "The methodology is explained with mathematical rigor. However, especially the description of the LS-SVM method is hard to digest and it would be helpful if the authors explained the method also with their own words instead of referring to the standard literature. For instance, what constitutes the dimensions in the higher-dimensional feature space (does each material class represent one dimension?)? Later in the description it follows that the algorithm operates in dual space and Fig. 1b only shows two axes (which have identical labels!?) so why refer to a higher dimensional feature space in the first place? If I understood the method correctly, then each pixel in the remaining data set is assigned a class label according to the similarity with the class statistics of each material in the training data plus some internal regularization with a Gaussian kernel. Perhaps a high similarity also entails a higher weighting factor w? What is the job of the Gaussian kernel, or in other words, what happens if sigma and gamma is set too high or too low? I like the idea of providing a schematic like in Figure 1b. However, it is not self-explanatory, even after having read the main text. What are the properties x1 and x2 and what do the properties z1 and z2 stand for in the context of image classification?"

Author's answer: The reviewer comments are mainly about Section 2.3 which is basically about the mathematical background of the LS-SVM classifier, which was rearranged and some mathematical terms are further explained where needed. They are as follows.

p7, L7-19: A more detailed explanation to the diagram of non-linear SVM (Fig. 1) is added, where x1, x2 are now the features in x-space (input vector space), and z1, z2 are the features in z-space representing the higher-dimensional feature space.

Fig. 1b: The axes in z-space are now represented by z1, z2.

p7, L15-17, p10, L1-2 & L5-10: An additional explanation is added for the scheme in Fig. 1 to explain why it is important to transform data into higher-dimensional feature space, and also what makes LS-SVM so popular in dealing complex data sets.

p10, L21-26: The Gaussian radial basis kernel function is explained in more detail. The role of RBF in LS-SVM data classification, in particular, the significance of ïĄşïĂăvalue in Eq. 17, is now highlighted.

p12, L20-28: The regularization parameter and RBF kernel parameter values were obtained by applying integrated LS-SVM “tuning” function. The tuning parameter set is based on CSA along with cost function using cross-validation. This was the reason to apply the tuning to get the optimal values for the LS-SVM model in a more reliable and accurate manner. A manual selection of these parameters, and its effect on the data classification performance, was out of scope of the manuscript.

Comment #3: "The brief discussion of the BH-correction should be moved from the conclusions to the discussion section and extended substantially. I didn’t understand exactly how a strong material contrast leads to an over- or underestimation of gray values in each individual phase. Is it because the polynomial surface is a compromise between the spatial variability of intensities of all materials at once? What if the volume fraction of halite, clay and anhydrite would be more balanced (instead of mostly clay). Would the BH-correction then work at all? This is why I’d like to see at least a second sample for a completely different rock, where this issue is addressed. I don’t see why

a 3D correction algorithm would solve these issues. The true solution to the problem would be to have a separate BH model for each material. I also didn’t understand if the model surface for one individual 2D slice is applied to all other z-slices or if the fitting parameters are optimized individually for all slices."

Author’s answer: p13, L8-14: The conclusion part about BH correction is now moved to the discussion section.

p14, L20-25: Yes, it is true that 2D polynomial fitting is a compromise between the spatial variability of grey-values of all material at once (see new Fig. 5). The fitting to the density of a cloud of data (clay minerals in Fig. 5c) can miss the fit to the wide range of grey values in a complex multi-component material. As a consequence, in the residual grey-values in case of low contrast between phases, this can lead to over – or underestimation of the grey-values of each individual phase and can hamper the correct segmentation process. Therefore, a solution to this problem is to fit the range of each phase grey-values separately. We agree with the reviewer remarks that volume fitting may not be helpful to overcome this problem.

Figures 3a,b and 5: The reviewer asked for providing more images, specially on equally distributed phases in sample to see the trend of 2D-polynomail fitting. This comment has led to many additions in manuscript:

- p4, L26-31: ïĄ■XCT scanning properties of two new samples were described. - p5, L24-30: The reconstruction properties (image dimensions) were added. - p11, L18-27: The BH Performance description in Fig. 5a,b was added. - p12, L13-16, L25-26 & L31-32: The LS-SVM classification for a new sample B (Fig. 3b) was added. - p14, L1-4 & Fig. 7a,b: The classifier performance (ROC) explanation is extended and Fig. 7 was modified.

Minor comments:

1. "LS-SVM was trained with 1755 pixels, but the remaining 1,570,149 pixels is

nowhere near 1417x1417 or 1417x1417x450. Please check again."

Authors Response: p12, L 16-17: In fact, the remaining pixels number 1,570,149 was correct. Actually, only the grey-values of an object (rock core) inside a 2D-slice were tested for validation. In order words, any pixels outside of sample in 2D-slice did not count for any calculations in order to avoid computational cost, in particular, when applying LS-SVM method.

2. "Conclusions: "Without ... any requirement for prior knowledge" - Doesn't the definition of training data represent your prior knowledge of the materials in the image?"

Authors Response: p14, conclusion part: In fact, the classification on any test data set is based on the knowledge of input trained data points. “Requirement for prior knowledge” was out of context at this place and, therefore, was removed.

Interactive comment on Solid Earth Discuss., 7, 3383, 2015.