MULTI-SCALE AND UPSCALING OF DIGITAL ROCK PHYSICS WITH A MACHINE THAT CAN LEARN ABOUT ROCKS

R. Sungkorn, A. Morcote, G. Carpio, G. Davalos, Y. Mu, A. Grader, N. Derzhi, J. Toelke
Ingrain Inc., 3733 Westheimer Rd., Houston, Texas, 77027, U.S.A.

This paper was prepared for presentation at the International Symposium of the Society of Core Analysts held in St. John's Newfoundland and Labrador, Canada, 16-21 August, 2015

ABSTRACT
Digital rock physics (DRP) is becoming a standard tool for rock characterization. DRP utilizes 2D and 3D digital images of rock samples to analyze petrophysical and geological properties. The ability to apply DRP to a large rock sample opens a way for economic exploration and recovery of hydrocarbon. Nevertheless, due to the well-known multi-scale nature of rocks and limitations in imaging technology, less than 1% by volume of a rock sample will be digitally acquired and analyzed. Undoubtedly, relevancy and representativeness of DRP remain hotly debated topics in oil and gas industry.

Machine learning (ML) has recently accelerated advances in many industries. ML brings together multiple disciplines such as computer science, statistics, and natural science to create algorithms that can learn from data. DRP can harness the power of ML to learn from its data, the digital image of rocks, to generate breakthroughs in the oil and gas industry.

In this paper, we present a framework that combines advances in DRP and ML to characterize rock samples at a large scale, not only a tiny part of it. The framework is based on an understanding that a rock consists of multi-scale rock fabrics intermixed spatially. These rock fabrics are captured as groups of patterns within a digital image when they are smaller than the image resolution being used. We developed ML algorithms that can automatically learn about rock fabrics and their patterns. This learning process can be iteratively repeated down to an image resolution that resolves the smallest or the most significant rock fabrics. Thus, the framework integrates DRP paradigm to achieve a truly multi-scale analysis. Also, DRP and ML analysis determine the optimum number and optimum locations for further acquisition and analysis of rock fabrics at a higher resolution.

INTRODUCTION
Rocks are well-known to inherit complex heterogeneous structures with a broad spectrum of scales. For example, pores within a rock sample can range in size from nanometers to millimeters. Ehrenberg [1] carried out laboratory measurements of porosity and permeability of the same rock at different sample scales; the smaller plug samples were

drilled from the larger cores. The study found that the larger core samples had generally lower porosity but higher measured permeability values than the plugs. It can be implied that small-scale rock samples do not adequately represent all features found in large-scale samples. In addition, it is rarely feasible to perform laboratory measurements of large-scale rock samples and, measurements of small-scale rock samples are typically limited to a small amount of samples due to extensive time and cost.

Digital rock physics (DRP) aims at providing qualitative and quantitative understanding of flow transport units as well as geometrical properties of rocks. Some of the rock properties are extremely difficult, if not impossible, to measure in the laboratory. Thus, DRP in conjunction with laboratory measurements, will compliment and complete well log analysis with not only detailed information but also with new kinds of insights. Such well logs enhance analysis of reservoirs and will open a way for economic exploration and better recovery of hydrocarbons.

The use of DRP involves three steps: (a) *digital imaging* to create a digital representation of a rock in 2D and 3D at a scale and resolution that will resolve rock features such as pores, organics, and grains; (b) *digital image processing* to categorize pixels/voxels in 2D and 3D respectively, with similar properties, and (c) *digital rock analysis* to digitally model desired rock properties using the digital image of the rock [2, 3, 4].

The following discussions are applicable for both 2D and 3D images. For the sake of simplicity, the term "image" refers to 2D and 3D images and "pixel" refers to both image pixel and voxel, unless otherwise stated.

Scientists and researchers have been analysing rock properties using multi-scale DRP [5] [6, 7, 8, 9]. Figure 1 shows on the left a typical multi-scale DRP paradigm and on the right the DRP paradigm introduced in this paper. The main differences are the use of rock fabrics instead of rock features and the recursive process to obtain information from the small-scale rock fabrics. A rock fabric is defined as a combination of rock features. Similar rock fabrics have similar properties or follow similar property trends.

The process begins with an image of a large rock sample acquired at a relatively coarse resolution to cover a large field of view. At this stage, rock fabrics larger than the image resolution are resolved while smaller ones are unresolved. A rock fabric is considered resolved when it is represented, in every direction, by at least two pixels. Then, the unresolved rock fabrics are segmented into groups. Information concerning the unresolved rock fabrics is analyzed from additional images acquired at a finer resolution and smaller field of view. The information from resolved and unresolved rock fabrics are fused and populated back into the large-scale image. DRP analysis, of desired properties, is carried out using the large-scale image.

Multi-scale DRP provides a promising method to characterize rocks. Nevertheless, due to limitations in imaging technology, currently available multi-scale DRP methods still suffer from shortcomings. They can be roughly summarized as follow:

- ***Image scale***: despite advances in image processing and imaging technology, the size of a "*large*" rock sample is still limited, at best, to several centimeters and often only to a few millimeters range. Properties derived from a large sample provide higher reliability and relevancy.
- ***Image resolution***: in rocks, pore size may span several orders of magnitudes. For complex rocks, such as shales small scale pores play an important role in transport properties as well as in the total porosity of the rock. Therefore, an image resolution adequately small to resolve these small pores should be used. However, such image will have a drastic reduction in the field of view.
  Lemmens & Richards [10] created an impressive high-resolution, 12 mm in length 2D-SEM image. 12800 images where stitched and tiled together. It resolved pores from millimeter to nanometer-scale. Their approach accurately provides detailed rock properties with extensive time and resources requirements. Set aside the practical aspect in 2D imaging, the approach is undoubtedly infeasible for 3D imaging.
- ***Representative elementary area/volume (REV)***: REV is defined as the size of sub-samples in which a measured property is approximately independent from location. The definition is arguably invalid when a rock has a mild level of heterogeneity.
- ***Unresolved rock fabrics sampling locations***: unresolved rock fabrics are divided into groups of similar properties, e.g. pixel intensity and CT number. Images of these groups are acquired at a higher-image resolution to determine their properties. The image locations are typically chosen manually and qualitatively. However, it is extremely difficult to make a reasonable and consistent selection in 3D due to complex process in human perception of volume from texture [11].

Recent advances in computing hardware and machine learning (ML) have accelerated innovations and breakthroughs in many industries. ML brings together multiple disciplines such as computer science, statistics, and natural science to create algorithms that can learn from data. These algorithms have the ability to build a model from data and/or training data without strict instructions. Detailed discussion regarding ML can be found in [12] and [13]. Examples of ML-based computer vision applications include autonomous vehicle technology [14], automatic tumor detection [15], and object recognition [16]. Digital images produced in DRP can be also considered as data. Based on this perspective, DRP can harness the power of ML to discover and learn from its data.

In the computer vision community, a texture is loosely defined by complex visual patterns formed by distinct features. Such features can be extracted using various mathematical models such as intensity histogram, co-occurrence matrix, and Gabor filter. Detailed discussion regarding texture analysis can be found in [17].

A paradigm for texture analysis usually involves four steps: (a) *keypoint detection* to limit the analysis only to meaningful areas, (b) *feature extraction* to quantitatively represent texture using appropriate models, (c) *feature classification* defines clusters of keypoints corresponding to a perceptually homogeneous texture, and (d) *texture segmentation* to construct area/volume based on the feature clusters. Texture in computer vision is similar to the rock fabrics in DRP. Therefore, the texture analysis paradigm can be used with some modifications to discover rock fabrics in DRP.

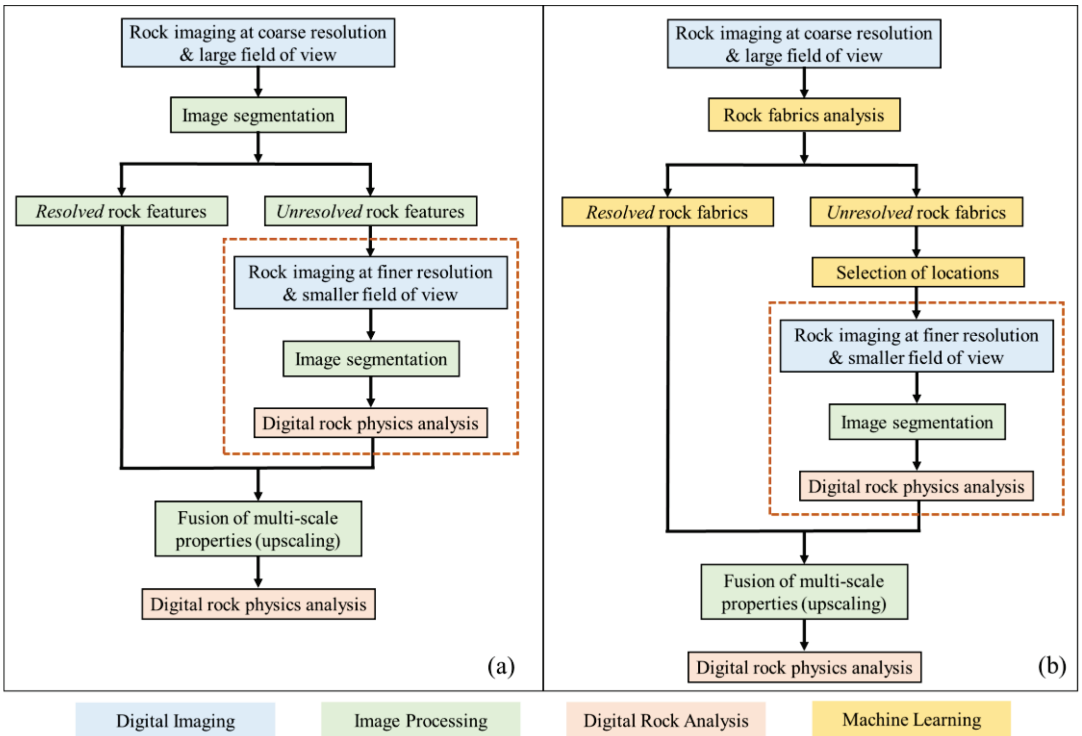

Figure 1: Left: generalized paradigm for multi-scale digital rock physics. Right: generalized paradigm for multi-scale digital rock physics using machine learning. Steps are highlighted according to the tasks shown in the bottom. Dotted block indicates a nested digital rock physics paradigm for unresolved rock features.

ROCK IMAGING

Rocks inherit complex multi-scale heterogeneous structures. A variety of imaging and detection techniques have been used to gain insights into rocks. Ideally, the image resolution being used should resolve all significant rock features and provide a reasonably large field of view (i.e. image scale). Due to limitations in imaging technology both image resolution and image scale are overly compromised.

Figure 2a shows a schematic image with multi-scaled objects. Overlaying the objects, we have grid cells. Large objects encompass significant amount of grid cells and will be resolved. In contrast, small objects are significantly smaller than the grid cells and will be unresolved. Figure 2b shows a digitized representation of Figure 2a image. The gray scale (i.e. intensity) of the grid cells directly relates to the object area covered by the grid cell. A grid cell has low intensity (dark gray) when completely covering a feature while high intensity (light gray) is presented in a grid cell that covers only a tiny part of a feature. It can be seen that larger objects are fairly well represented by digitized grid cells. On the other hand, small objects are smeared out in digitized grid cells and are unrecognizable.

Nevertheless, digitized grid cells are generated based on the interaction between an object and the physics of the imaging technique being used. A group of unresolved grid cells contains information regarding patterns of unresolved objects. A sophisticated mathematical model which quantifies patterns of grid cells intensity, e.g. co-occurrence matrix, local binary patterns and Gabor filter, must be used [15] [17]. Since averaged intensity of these patterns is approximately similar, the widely used averaging approach cannot distinguish them. Figure 3 shows an example of multi-scale rock image. A large-scale rock image, in the center, contains unresolved rock fabrics which have different image intensity patterns. On the sides, we have high-resolution images

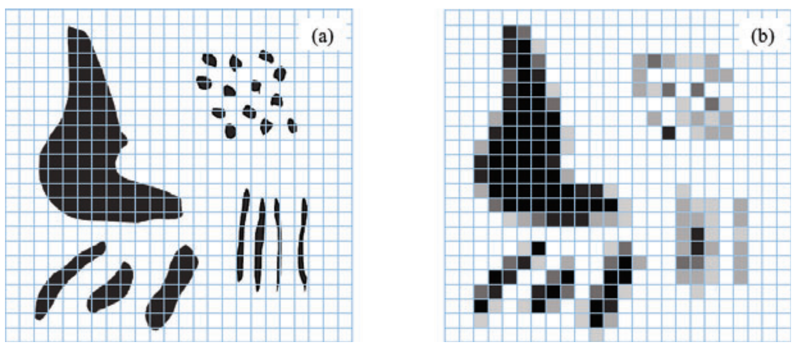

Figure 2: Illustration of a comparison between image resolution (grid size) with resolved rock features on the left of (a) and unresolved rock features on the right of (a). Illustration similar rock features acquired and digitized (b).

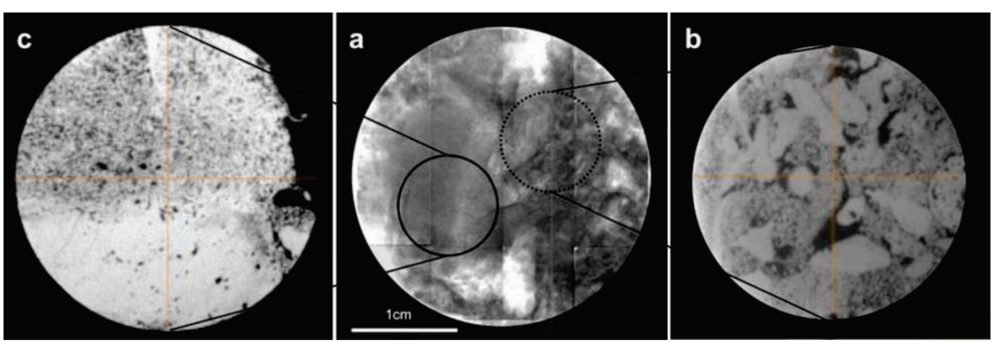

Figure 3: An example of unresolved rock fabrics in a large-scale image (a). Image of the unresolved rock fabrics are acquire at higher resolution (b) and (c). Image courtesy of ADCO Ltd.

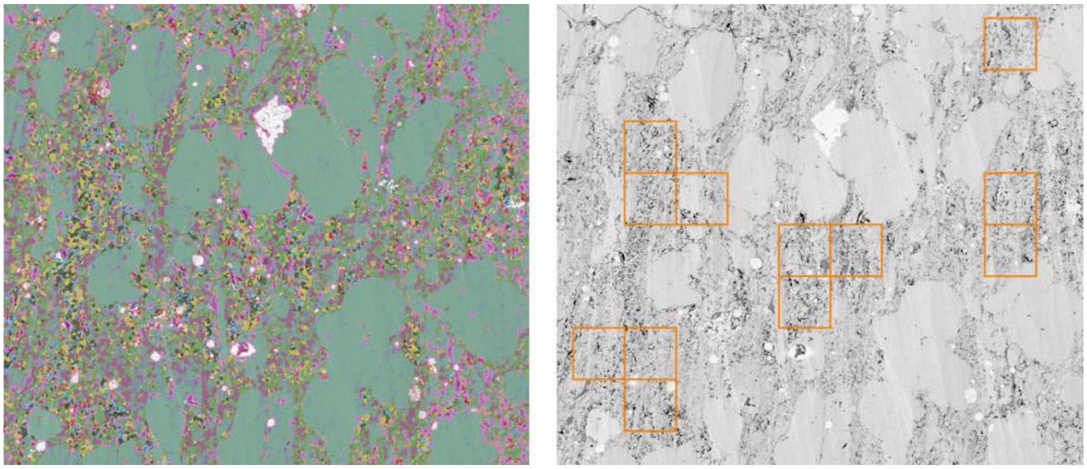

Figure 4: Eagle Ford sample. Segmented rock fabric overview (left) with optimum size and number of areas to capture all rock fabrics (right).

MULTI-SCALE DRP WITH ML

DRP has been improved significantly in recent years. Variations of multi-scale DRP (Figure 1a) are widely adopted [5] [6] [7] [8] [9]. However, applications of DRP for reservoir characterization are still limited due to shortcomings discussed earlier. In this section, we discuss a scalable multi-scale DRP paradigm with ML techniques for rocks. The differences from previous multi-scale DRP are the use of ML to learn about resolved and unresolved rock fabrics presented in a rock sample. We also use ML to identify optimum size and location for further analysis of the unresolved rock fabrics.

The present multi-scale DRP paradigm (Figure 1b) begins with a digital imaging of a rock sample at a large scale. This image will be called overview throughout this section. The overview (Figure 4, right) might contain resolved and/or unresolved rock fabrics.

In the second step, rock fabrics (Figure 4, left) in the overview are detected and segmented using the texture analysis method discussed earlier. However, rock fabrics are different from image texture commonly encountered in computer vision. The main difference is that rock fabrics tend to have pattern at individual pixel level not at edge or blob level [13] [14] [15] [16] [17].

We developed a novel *rock fabric analysis* method, based on the texture analysis paradigm (Figure 1b). Rock fabrics key-points are detected using the method discussed in Appendix A. They are mostly located within an area with a rock fabric.

Then, *rock fabric features* of the area around the keypoints are computed using the method discussed in Appendix B. In this method, four rock fabric features or attributes are used: contrast, homogeneity, entropy, and variance.

Consequently, keypoints are clustered using the four rock fabric features. We developed an unsupervised ML method for clustering high-dimensional data. It automatically learns data and finds an appropriate number of clusters. Based on the understanding of rock images, it is reasonable to postulate that keypoints within similar clusters have similar rock fabrics. For example, in Figure 5, green represents areas with medium to high intensities (unresolved organics and pores) while magenta represents areas only with medium intensities (might be unresolved organics or very small pores). These clusters can be used as a model for segmentation of the whole overview. During segmentation, rock fabric attributes for every pixel/voxel are computed. Distance from clusters is computed using, for example, Euclidean and Mahalanobis distance. A general discussion regarding data clustering can be found in [12].

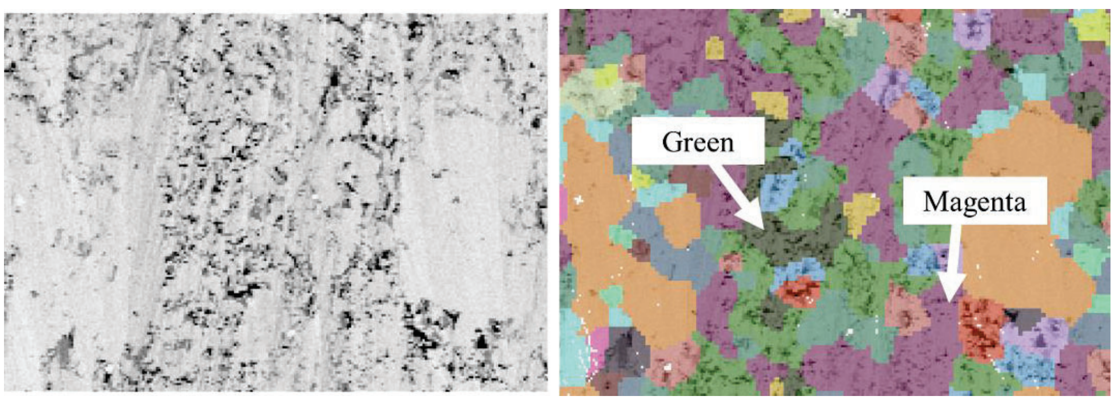

Figure 5: Close-up view of overview image (left) and its corresponding segmented fabrics (right).

The goal of rock fabric segmentation is to gain knowledge about unresolved rock fabrics. Detailed information of the unresolved fabrics is needed in order to characterize the large-scale image. This can be done by sub-sampling unresolved rock fabrics. An optimum amount of sub-sampling is desired to minimize expenses while keeping high level of accuracy and reliability. We developed an optimization algorithm for spatial data analysis which it determines the most suitable locations for further analysis (Figure 4). The algorithm finds a combination of areas that contains all fabrics with broadest variety within the fabric. Information from sub-sampling areas can be used in fusion of multi-scale information later.

An illustration of an extraction of multi-scale correlation extraction is shown in Figure 7. We give an example of 2D porosity upscaling in this paper. The method can be used directly to upscale properties in 2D and 3D. The suggested area within an overview image is acquired at a resolution that adequately resolved rock features and segmented into phases (e.g. pore and organic matters). A multi-scale correlation is obtained by correlating, for each fabric, the intensity of the overview image pixels to the porosity obtained from the area cover by the pixel in the high-resolution image. Examples of the extraction are shown as plots in Figure 7. A multi-scale correlation for each fabric is derived by fitting a function to the data (shown as solid lines in the plots). Then, the porosity is upscaled by populating information from the multi-scale correlations for each fabric to all pixels in the overview image (Figure 7).

A truly multi-scale DRP is achieved by repeating the procedure above recursively on the unresolved rock fabrics until all rock fabrics are resolved. A method designed for upscaling and fusion of multi-scale rock properties is discussed in [9]. It is noted that, the multi-scale DRP with ML technology discussed here can be directly applied to multi-dimensional rock images. We use 2D images in this paper only for the sake of simplicity. It is also worth to note that, the methods present in this paper are implemented using graphics processing unit (GPU), which results in a computational time of approximately 100 seconds for a 2D image with a dimension of 2000 x 3000 pixels.

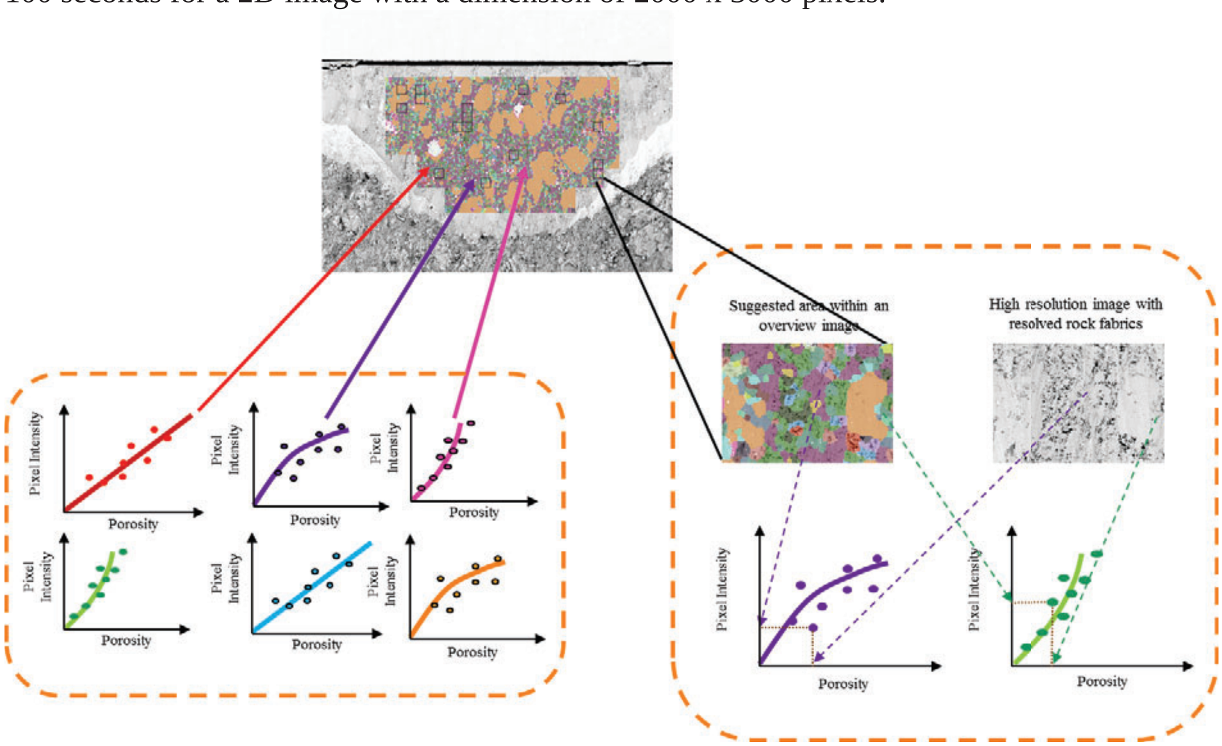

Figure 7: Illustration of multi-scale correlation extraction based on fabrics (bottom right) and population of the multi-scale correlation back on the overview image (bottom left).

CONCLUSION

We present a multi-scale digital rock physics (DRP) method using machine learning (ML) for rock fabrics characterization and scaling of rock properties. The method integrates knowledge in geology, physics, and computer science. We developed new algorithms based on the ones used in computer vision and pattern recognition communities. The rock fabric analysis discovers rock fabrics both resolved and unresolved by the image resolution being used. Also, it has the capability to characterize large-scale rock images by iteratively learning about fabrics.

Additionally, it can be applied for selecting optimum number of meaningful areas/volumes for laboratory measurements using whole core and cuttings.

It is also important to note that, despite the use of unsupervised ML methods, we designed the method based on the concept of intelligence augmentation (IA) [21]. Thus, experts can integrate their knowledge into the analysis to maximize benefits.

REFERENCES

1. Ehrenberg, S.N., "Whole core versus plugs: Scale dependence of porosity and permeability measurements in platform carbonates," *AAPG Bulletin*, (2007) **91**, pp. 835-846.
2. Dvorkin, J., Derzhi, N., Diaz, E., Fang, Q., "Relevance of computational rock physics," *Geophysics*, (2011) **76**, pp. E141-E153.
3. Andrä *et al*, "Digital rock physics benchmarks – Part I: Imaging and segmentation," *Computers & Geosciences*, (2013) **50**, pp. 25-32.
4. Andrä *et al.*, "Digital rock physics benchmarks – Part II: Computing effective properties," *Computers & Geosciences*, (2013) **50**, pp. 33-43.
5. Øren P.E., Bakke, S., "Process based reconstruction of sandstones and prediction of transport properties," *Transport in Porous Media*, (2002) **46**, pp. 311-343.
6. Grader, A.S., Clark, A.B.S., Al-Dayyani, T., Nur, A., "Computations of porosity and permeability of sparic carbonate using multi-scale CT images," *Society of Core Analysis International Symposium 2009*, SCA2009-10.
7. Sok, R.M., Vaslot, T., Ghous, A., Latham, S., Sheppard, A.P., Knackstedt, M.A., "Pore scale characterization of carbonates at multiple scales: Integration of micro-CT, BSEM and FIBSEM," *Society of Core Analysis International Symposium 2009*, SCA2009-18.
8. Khalili, A.D., Arns, C.H., Arns, J.-Y., Hussain, F., Cinar, Y., Pinczewski, W.V., "Permeability upscaling for carbonates from the pore-scale using multi-scale Xray-CT images," *SPE/EAGE European Unconventional Resources Conference and Exhibition 2012*, pp. 606-622.
9. Sungkorn, R., Gundepalli, V.A., Mu, Y., Grader, A., Sisk, C., Bhakta, S., Dvorkin, J., Toelke, J., "Multiscale permeability trends using digital rock physics," *Society of Core Analysis International Symposium 2014*, SCA2014-43.
10. Lemmens, H., Richards, D., "Multiscale imaging of shale samples in the scanning electron microscope," in Camp, W., Diaz, E., Wawak, B., *Electron microscopy of shale hydrocarbon reservoirs: AAPG Memoir*, (2013) **102**, pp. 27-35.
11. Tory, M., Möller, T., "Human factors in visualization research," *IEEE Transactions on Visualization and Computer Graphics*, (2004) **10**, pp. 1-13.
12. Bishop, C.M., *Pattern Recognition and Machine Learning*, Springer, Cambridge U.K., (2006).
13. Bengio, Y., "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, (2009) **1**, pp. 1-127.

14. Li, Q., Chen, L., Li, M., Shaw, S.-L., Nüchter, A., "A sensor-fusion drivable-region and lane-detection system for autonomous vehicle navigation in challenging road scenarios," *IEEE Transactions on Vehicular Technology*, (2014) **63**, pp. 540-555.
15. Menze *et al*., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, (2014), pp. 33.
16. He, K., Zhang, X., Ren, S., Sun, J., "Delving deep into Rectifiers: Surpassing human-level performance on ImageNet Classification," *arXiv:1406.4729*, (2015).
17. Materka, A., Strzelecki, M., *Texture Analysis Methods – A Review*, Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels, (1998).
18. Lowe, D.G., "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, (2004) **60**, pp. 91-110.
19. Yang, J., Yu, K., Gong, Y., Huang, T., "Linear spatial pyramid matching using sparse coding for image classification," *Proceedings of CVPR*, (2009).
20. Haralick, R.M., "Statistical and structural approaches to texture," *Proc. IEEE*, (1979) **67**, pp. 1786-804.
21. Licklider, J.C.R., "Man-computer symbiosis," *IRE Transactions on Human Factors in Electronics,* (1960) **1**, pp. 4-11.

APPENDIX

A. Rock keypoint detection

Natural rock images inherit highly heterogeneous arrangements of rock features. Dividing images into tiles leads to an overestimation and classification of rock fabrics. In order to (1) limit further analysis only to meaningful areas, (2) keep computational expense tractable and (3) obtain reasonable number and classification of rock fabrics, keypoints are detected. Concepts of keypoint detection are widely used in computer vision community [18] [19]. Since rock fabrics have different characteristics from commonly encountered image texture, previous methods for keypoint detection are not applicable. Our rock keypoint detection algorithm (step b in Figure A1) begins with (a) *discrete wavelet decomposition* of the image up to desired level. This step ensures that rock fabrics at multiple scales will be captured. Then, (b) pixel/voxel *gradient of images* obtained from previous step is computed. (c) *Laplacian of Gaussian* (LoG) is computed on the gradient images to locate points of variation of pixel/voxel intensity. (d) *Keypoints detection* within rock fabrics are detected by locating maxima in LoG images. Keypoints at the edges of rock features are eliminated by limiting keypoints within certain value of maxima (e.g. 80% of maxima).

B. Rock fabric features

There are models for image texture available in literature [17]. It is known that features of rock fabrics, especially the unresolved rock fabrics, are in pixel/voxel level. Additionally, similar rock fabric may have different orientations in an image. Therefore, a model that quantitatively describes rock fabric features and is rotation invariance must be selected. Note that, for the sake of naming consistency, rock fabric features (in DRP) are used interchangeably with texture features (in computer vision). We use Haralick texture features [20] which based on gray-level co-occurrence matrix (GLCM). They can be used to quantify spatial distribution and auto-correlation of pixel/voxel pairs. The GLCM, $P_{i,j}$, is constructed from probability of intensity j next to intensity i in defined directions and distance. In this paper, we select an appropriate set of Haralick texture features to obtain maximum separation between rock fabrics (for following features classification) namely, (1) contrast, (2) homogeneity, (3) entropy, and (4) variance. Their mathematical description can be found in [20].

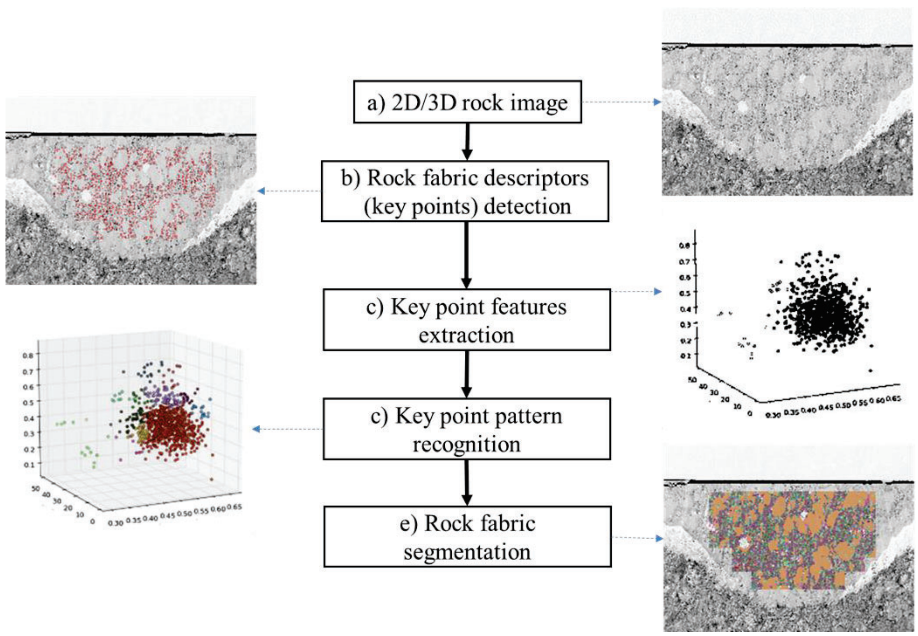

Figure A1: Flow diagram for rock fabric recognition and segmentation.