Introduction

With the growing demand of high-resolution subsurface characterization from 3D seismic surveying, the size of 3D seismic datasets has been dramatically increasing, and correspondingly, the process of interpreting a seismic dataset is becoming more time consuming and labor intensive. In recent years, there has been increased interest in computational seismic interpretation. Specifically, methods based on supervised machine learning have had great success in salt body delineation, fault and fracture detection, horizon extraction, and facies analysis (e.g. Barnes and Laughlin (2005); Coléou et al. (2003); Guillen et al. (2015); Zhao et al. (2015); Qi et al. (2016); Wang et al. (2015); Ramirez et al. (2016); Figueiredo et al. (2015)).

Additionally, in wide-ranging fields, supervised machine learning has proved to be the most successful machine learning paradigm by far. Supervised machine learning algorithms require labels to perform training. However obtaining labels for large volumes of seismic data is a very demanding task. Furthermore, while the amount of data is continuously growing, the ability of human experts to label data remains limited. Thus, the time and effort saved by techniques based on supervised machine learning can be over-shadowed by the time required to obtain accurate labels from the seismic data. This is why we believe there should be more work done on *weakly-supervised learning*, especially in the area of seismic interpretation where the amount of data is often in the hundreds of gigabytes.

Weakly-supervised machine learning is the case when the learning is done with missing or few labels or with partially labeled data. In this work, we propose a weakly-supervised framework for labeling seismic structures using Non-Negative Matrix Factorization (NMF) with additional sparsity and orthogonality constraints. Specifically, given some weakly-labeled seismic data (CeGP, 2015) extracted from the Netherlands North Sea dataset (dGB Earth Sciences, 1987) using the method proposed in (Alfarraj et al., 2016; Alaudah and AlRegib, 2016), we are given a single label for every image (e.g. an image with a fault label will indicate that it has a fault structure somewhere within it.) Our goal is then to use these image-level labels and map them to specific locations within the seismic image where the structure is most likely to be. Since an image $\mathbf{x}_i \in \mathbb{R}_+^{n\times m}$ has only a single label (as opposed to around $n \times m$ labels), this learning problem is a weakly-supervised one.

Orthogonal Non-Negative Matrix Factorization

Non-negative Matrix Factorization (NMF) (Lee and Seung, 1999, 2001) is a commonly used matrix factorization technique that is closely related to many unsupervised machine learning techniques such as k-means and spectral clustering (Ding et al., 2005). NMF is used for decomposing a non-negative matrix $\mathbf{X} \in \mathbb{R}_+^{N_p\times N_s}$ into the product of two lower-rank matrices $\mathbf{W} \in \mathbb{R}_+^{N_p\times N_f}$, and $\mathbf{H} \in \mathbb{R}_+^{N_f\times N_s}$ such that both $\mathbf{W}$ and $\mathbf{H}$ are non-negative. In other words,

$$\mathbf{X} \approx \mathbf{WH} \quad \text{s.t.} \mathbf{W}, \mathbf{H} \geq 0. \tag{1}$$

Here, N_f is the number of components (or the *rank*) in which $\mathbf{X}$ will be represented. In our work, $\mathbf{X}$ represents a data matrix where each column is a single seismic image in vector form. The data matrix $\mathbf{X}$ has N_s such images, each of which is a vector of length N_p. NMF factorizes this data matrix into two non-negative matrices, a basis matrix $\mathbf{W}$ and a coefficient matrix $\mathbf{H}$. In clustering terms, the columns of $\mathbf{W}$ represent N_f number of clusters in the data, whereas the columns of $\mathbf{H}$ represent the memberships of each of the images to the different clusters in the data. Here, the clusters can represent different seismic structures like salt domes, faults, or horizons.

The regular NMF problem does not have a closed-form solution, and is typically solved by minimizing the following optimization problem:

$$\underset{\mathbf{W},\mathbf{H}}{\arg\min} ||\mathbf{X} - \mathbf{WH}||_F^2 \quad \text{s.t.} \mathbf{WH} \geq 0, \tag{2}$$

where, $||\cdot||_F$ is the Frobenius norm. Lee and Seung (1999) showed that NMF can be used to learn a "parts-based" representation of the data, where each feature would represent a localized "part" of the data. In practice, this is rarely achieved using the formulation in equation 2. In order to remedy this, we initialize matrix $\mathbf{W}$ using k-means on the data matrix $\mathbf{X}$. Then we impose a sparsity constraint on these initial features using the following sparsity measure:

$$\rho(\mathbf{w}) = \frac{\sqrt{N_p} - ||\mathbf{w}||_1/||\mathbf{w}||_2}{\sqrt{N_p} - 1}. \tag{3}$$

In order to enforce this constraint, we follow the algorithm proposed by Hoyer (2004). Additionally, to enforce that each feature $\mathbf{w}_i$ represents a single class, we impose an orthogonality constraint on the coefficient matrix $\mathbf{H}$. The problem then becomes:

$$\arg\min_{\mathbf{W},\mathbf{H}} ||\mathbf{X} - \mathbf{W}\mathbf{H}||_F^2 + \lambda_1 ||\mathbf{W}||_F^2 + \lambda_2 ||\mathbf{H}||_F^2 + \gamma_1 ||\mathbf{H}\mathbf{H}^T - \mathbf{I}||_F^2 \quad \text{s.t.} \mathbf{W},\mathbf{H} \geq 0 \ \text{ and } \ \rho(\mathbf{w}_i) = \rho_w, \tag{4}$$

where $\mathbf{I}$ is an identity matrix, and λ_1, λ_2, γ_1, and ρ_w are constants. In order to solve the problem in equation 4, we derive the following multiplicative update rules for $\mathbf{W}$, and $\mathbf{H}$ where

$$\mathbf{W}^t = \mathbf{W}^{t-1} \odot \frac{(\mathbf{X}\mathbf{H}^{t-1^T} + \varepsilon)_{ij}}{\mathbf{W}^{t-1}\mathbf{H}^{t-1}\mathbf{H}^{t-1^T} + \lambda_1 \mathbf{W}^{t-1} + \varepsilon)_{ij}}, \tag{5}$$

$$\text{and } \ \mathbf{H}^t = \mathbf{H}^{t-1} \odot \frac{(\mathbf{W}^{t^T}\mathbf{X} + \gamma_1 \mathbf{H}^{t-1} + \varepsilon)_{ij}}{\mathbf{W}^{t^T}\mathbf{W}^t\mathbf{H}^{t-1} + \gamma_1(\mathbf{H}^{t-1}\mathbf{H}^{t-1^T}\mathbf{H}^{t-1}) + \lambda_2 \mathbf{H}^{t-1} + \varepsilon)_{ij}}. \tag{6}$$

Here, $\odot$ represents element-wise multiplication, and the superscript indicates the iteration number. These update rules for $\mathbf{W}$ and $\mathbf{H}$ are applied successively until both $\mathbf{W}$ and $\mathbf{H}$ converge.

Weakly-Supervised Seismic Structure Labeling

Once $\mathbf{W}$ and $\mathbf{H}$ have converged, each column of $\mathbf{H}$, $\mathbf{h}_n$ indicates the features used to construct image n. Since every feature in $\mathbf{W}$ should correspond to a single seismic structure, we can then map the coefficients of each image into the seismic structures that make up the image

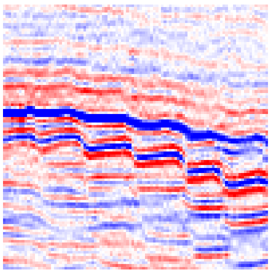
$$\mathbf{Y}_n = \mathbf{W}(\mathbf{Q} \odot (\mathbf{h}_n \mathbf{1}_{1\times N_l})) \qquad \forall n \in [1, N_s]. \tag{7}$$

Here we define $\mathbf{Q} \in \{0,1\}^{N_f \times N_l}$ as a *cluster membership* matrix such that the element $Q_{ij} = 1$ if the feature $\mathbf{w}_i$ belongs to structure j, and $\mathbf{1}_{1\times N_l}$ is a vector of ones of size $1 \times N_l$. Then, each location i in image n corresponds to the seismic structure given by

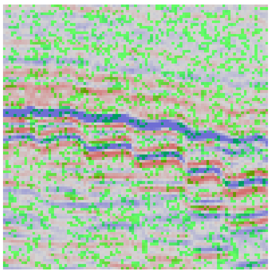
$$\text{structure}_i = \underset{j}{\text{argmax}} \ \mathbf{Y}_{nj}. \tag{8}$$

Results

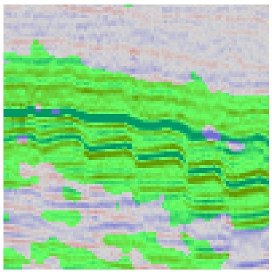
In order to test this method, we apply it on a subset of a 1000 seismic images from the LANDMASS dataset CeGP (2015) that includes weakly-labeled images extracted from the Netherlands North Sea F3 block (dGB Earth Sciences, 1987). We put these images in vector form, and use them to construct



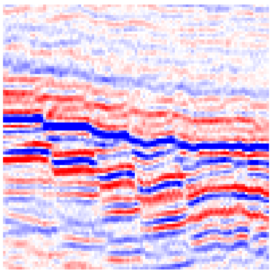
(a) The seismic amplitude



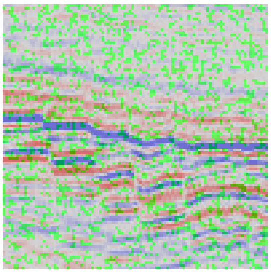
(b) Random initialization of faults



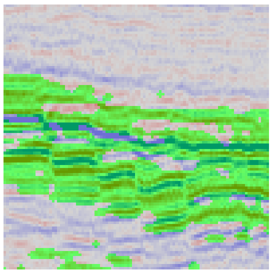
(c) The labeled faults overlaying the seismic image



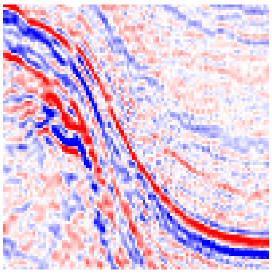
(d) The seismic amplitude



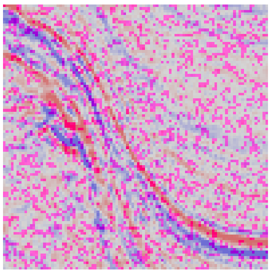
(e) Random initialization of faults



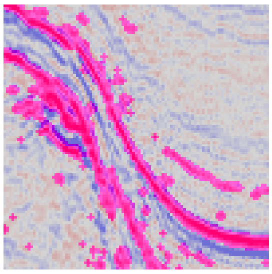
(f) The labeled faults overlaying the seismic image.



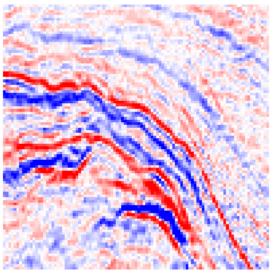
(g) The seismic amplitude



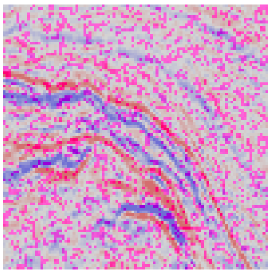
(h) Random initialization of salt dome boundary



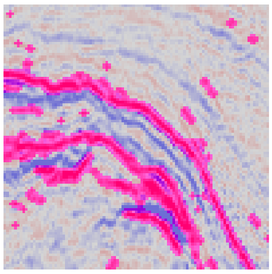
(i) The final labeled salt dome boundaries overlayed on the seismic image



(j) The seismic amplitude



(k) Random initialization of salt dome boundary



(l) The labeled salt boundaries overlaying the seismic image

Figure 1: The results of labeling faults (shown in green) and salt dome boundaries (shown in magenta) in the vertical sections of inlines 410 and 450 from the F3 seismic dataset by the proposed Orthogonal NMF scheme. Note the good match between the labeling and the original seismic image.

matrix $\mathbf{X}$. We then apply k-means on these images along with the sparsity constraint, and use the results to initialize the feature matrix $\mathbf{W}$. The coefficient matrix $\mathbf{H}$ is initialized with random positive real numbers. We empirically select the values of λ_1, λ_2 and γ_1 as: 0.1, 0.5, and 5 respectively. The sparsity of the initial features ρ_w is set to 0.85. The larger this term the more localized the results would become. We then apply the multiplicative update rules shown in equation 5 and 6 successively until they both converge. Finally, we apply equations 7 and 8 to find the final location of the seismic structures for each image in **X**. Figure 1 show a subset of the results for two fault and two salt dome images from vertical inlines 410 and 450. We observe that there is a good match between the labeled seismic structure and the structure in the original seismic image and that further processing can easily yield the specific location of each structure. Furthermore, it is important to note that this method is not limited to faults and salt dome boundaries, and can be applied to any other seismic structure as long as sufficient weakly-labeled training data is provided.

Conclusions

In conclusion, we have proposed a method based on Orthogonal Non-Negative Matrix Factorization for weakly-supervised labeling of seismic structures. A sparsity constraint on the initial features used allows the seismic structures to be easily localized. We have showed that "rough" image-level labels of specific seismic structures can be mapped into finer more localized location within the seismic volume. Results obtained by labeling fault regions and salt dome boundaries from the Netherlands F3 block are very promising.

Acknowledgements

The authors would like to acknowledge the support of the Center for Energy and Geo Processing (CeGP) at the Georgia Institute of Technology and King Fahd University of Petroleum and Minerals (KFUPM).

References

Alaudah, Y. and AlRegib, G. [2016] Weakly-supervised labeling of seismic volumes using reference exemplars. In: *2016 IEEE International Conference on Image Processing (ICIP)*. 4373–4377.

Alfarraj, M., , Alaudah, Y. and AlRegib, G. [2016] Content-adaptive Non-parametric Texture Similarity Measure. *2016 IEEE Workshop on Multimedia Signal Processing (MMSP)*.

Barnes, A.E. and Laughlin, K.J. [2005] *Investigation of methods for unsupervised classification of seismic data*. 2221–2224.

CeGP [2015] LANDMASS: Large North-Sea Dataset of Migrated Aggregated Seismic Structures. *http://cegp.ece.gatech.edu/codedata/LANDMASS/default.htm*.

Coléou, T., Poupon, M. and Azbel, K. [2003] Unsupervised seismic facies classification: A review and comparison of techniques and implementation. *The Leading Edge*, **22**(10), 942–953.

Ding, C., He, X. and Simon, H.D. [2005] On the Equivalence of Nonnegative Matrix Factorization and Spectral Clustering. *Proceedings of the fifth SIAM International Conference on Data Mining (SDM)*, (4), 606–610.

dGB Earth Sciences, B. [1987] The Netherlands Offshore, The North Sea, F3 Block - Complete. *https://opendtect.org/osr/pmwiki.php/Main/Netherlands/OffshoreF3BlockComplete4GB*.

Figueiredo, A., Silva, F.B., Silva, P., de O. Martins, L., Milidiú, R.L. and Gattass, M. [2015] *A Clustering-based Approach to Map 3D Seismic Horizons*. 1166–1170.

Guillen, P., Larrazabal, G., González, G., Boumber, D. and Vilalta, R. [2015] Supervised Learning to Detect Salt Body. *Society of Exploration Geophysicists*.

Hoyer, P.O. [2004] Non-negative Matrix Factorization with Sparseness Constraints. *The Journal of Machine Learning Research*, **5**, 1457–1469.

Lee, D. and Seung, H. [2001] Algorithms for non-negative matrix factorization. *Advances in neural information processing systems*, (1), 556–562.

Lee, D.D. and Seung, H.S. [1999] Learning the parts of objects by non-negative matrix factorization. *Nature*, **401**(6755), 788–91.

Qi, J., Lin, T., Zhao, T., Li, F. and Marfurt, K. [2016] Semisupervised multiattribute seismic facies analysis. *Interpretation*, **4**(1), SB91–SB106.

Ramirez, C., Larrazabal, G. and Gonzalez, G. [2016] Salt body detection from seismic data via sparse representation. *Geophysical Prospecting*, **64**(2), 335–347.

Wang, Z., Hegazy, T., Long, Z. and AlRegib, G. [2015] Noise-robust detection and tracking of salt domes in postmigrated volumes using texture, tensors, and subspace learning. *Geophysics*, **80**(6), WD101–WD116.

Zhao, T., Jayaram, V., Roy, A. and Marfurt, K.J. [2015] A comparison of classification techniques for seismic facies recognition. *Interpretation*, **3**(4), SAE29–SAE58.