

## Image integration with learned dictionaries and application to seismic monitoring

Youli Quan, Tiejuan Zhu and Jerry M. Harris, Stanford University; Roy M. Burnstad, Saudi Aramco; Sergio E. Zarantonello, Algorithmica LLC

### Summary

Sparse coding can be applied to train an overcomplete dictionary on time-lapse seismic data or images. The learned dictionary generally consists of sparse representations of one or more images. We then use such sparse representations, along with  $L_1$ -regularization techniques, to predict missing values in seismic images by solving an inverse problem. The practical outcome of the proposed methodology can be a significant reduction in field operational costs by requiring only sparse instead of dense surveys, and by integrating in the seismic images the information captured by the learned dictionary from previous time-lapse and baseline images. A synthetic example is presented to test the method.

### Introduction

Time-lapse seismic (4-D seismic) detects saturation and pressure changes in a reservoir, thereby helping optimize recovery and reducing the costs of field operations. In CO<sub>2</sub> enhanced oil recovery and storage, for example, injected CO<sub>2</sub> is used to increase oil production and also to sequester a portion of the injected CO<sub>2</sub>. A “real-time” flood performance diagnosis using 4-D seismic and other monitoring tools is important to process optimization and storage safety (Kuuskraa and Ferguson, 2008; Harris *et al.*, 2007).

Time-lapse seismic usually consists of one baseline and several time-lapse surveys. The image integration methodology in this setting is to adaptively train an overcomplete dictionary (or basis functions) on a set of baseline and previous time-lapse images, and then use the learned dictionary to predict missing values in the current seismic image. For certain cases, “missing data recovery” and “image integration” have similar meaning and are used interchangeably in this abstract.

Another potential application of image integration is quasi-continuous or real-time seismic monitoring. The concept was applied to missing data in an approach to quasi-continuous 4-D monitoring (Arogunmati and Harris, 2009, 2010). In this situation we need to collect and process data much more frequently than in conventional 4-D seismic. This, in practice, may signify sparse acquisition because of operational costs and/or technical acquisition issues. Sparse data usually results in low spatial resolution and/or incomplete subsurface images. However, quasi-continuous (and real-time) monitoring has high temporal resolution. If we use learned dictionaries to integrate time-lapse subsurface images, we can recover the missing values

and may also improve the spatial resolution. We demonstrate this application via a synthetic example. Arogunmati and Harris (2010) used a data-estimation-based method for quasi-continuous monitoring using sparse surface data.

The image integration with learned dictionaries is a sparse-representation-based approach to predicting missing values in an image (e.g., Selesnick, 2004; Elad *et al.*, 2005; Zhang, 2006; Herrmann *et al.*, 2008; Liu and Sacchi, 2004). The underlying principle used in this study is related to Elad *et al.* (2005) and Zhang (2006). Classic interpolation methods (such as polynomial-based approaches) are often not satisfactory on difficult cases. Nonlinear and adaptive approaches are expected to exhibit better performance. In our approach, we expect the signal (or image) to be sparse, that is, to be represented by few non-zero coefficients in the basis expansion. Then, an inversion with  $L_1$ -regularization can be solved to recover the missing values in the signal.

The use of sparsity and  $L_1$ -regularization for an inversion problem is closely related to compressed sensing (see, e.g., Donoho, 2006; Candès *et al.*, 2006; Candès and Wakin, 2008; Lustig *et al.*, 2007). In compressed sensing, we want to recover a sparse signal that is encoded by a “small” set of random projections (or observations). Because of the sparsity and randomness, the signal can be recovered with  $L_1$ -regularization, even when the observation data is highly incomplete.

Sparsity plays an important role in an  $L_1$ -regularized inversion problem. A signal exhibits different sparsity when using different dictionaries or basis functions. We need to select a proper sparse transform for a given signal. The wavelet transform and discrete cosine transform are examples of generic sparse representations. The use of overcomplete dictionaries may improve the sparsity and match more patterns in a signal (Mallat, 2008). An overcomplete dictionary is a set of redundant basis functions; namely, the number of basis functions is greater than dimensionality of the signal. There are a variety of algorithms for learning overcomplete dictionaries (e.g., Aharon *et al.*, 2006; Lee *et al.*, 2007). The sparse coding algorithms introduced by Lee *et al.* (2007) are used to train overcomplete dictionaries in this study.

### Motivation

This study is motivated by the recovery of missing data in a crosswell seismic monitoring survey. Figure 1 shows the survey geometry. Crosswell seismic was used to monitor

## Image integration with learned dictionaries

CO<sub>2</sub> EOR at McElroy Field in West Texas (Harris *et al.*, 1995). Baseline surveys were acquired in 1993 and post-injection surveys were collected in 1995. One problem in this project was that a large amount of post-injection data could not be acquired near the pay zone in Well 1068 and Well 1202 (Lazaratos and Marion, 1997). Figure 2 shows the *P*-wave traveltimes data from the baseline survey (1993) and the time-lapse survey (1995). It was difficult to perform reliable time-lapse tomography between these two wells because the 1995 survey had very limited aperture. We wonder if we can recover part of the missing data, assuming only a small part of the data in the 1995 survey were affected by the CO<sub>2</sub> flooding. We consider using the complete 1993 baseline data as training data to find a dictionary and then use it to estimate some of the missing data in the 1995 survey.

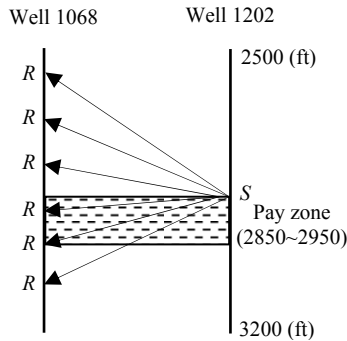


Figure 1: A common source gather in the crosswell geometry. A crosswell survey consists of many source gathers at different depths. Traveltimes from source (*S*) to receivers (*R*) are used for velocity tomography.

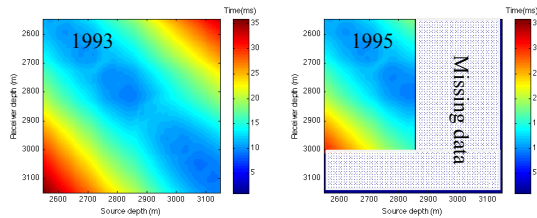


Figure 2: *P*-wave traveltimes data. Each column is traveltimes picks from a source gather. Survey 1993 is baseline that has complete data coverage. However, the 1995 time-lapse survey is incomplete due to an acquisition technical problem.

### Method

Let  $y$  be a signal vector with a length of  $N$ . Assume that we have an  $N \times K$  (usually  $K > N$ ) matrix  $A$  that decomposes the signal  $y$  as

$$Ax = y, \quad (1)$$

where the vector  $x$  of length  $K$  consists of the expansion coefficients of a sparse representation of the signal. We hope  $x$  has the fewest non-zero elements as possible, so that  $y$  can be sufficiently represented by a small number of columns of  $A$ ; namely,  $A$  is a sparse representation of  $y$ . The inverse discrete cosine transform (IDCT) is an example of the generic transforms commonly used for constructing the matrix  $A$ . For IDCT, the size of matrix  $A$  is  $N \times N$ , and its elements can be expressed as

$$A_{ij} = \frac{\min(j, \sqrt{2})}{\sqrt{N}} \cos\left(\frac{(2i-1)(j-1)\pi}{2N}\right).$$

Suppose there are  $M$  missing data values in  $y$ . We rearrange data vector order as

$$y = \begin{pmatrix} y_a \\ y_m \end{pmatrix}, \quad (2)$$

where  $y_a$  is the available data vector (set of “observations”) with a length of  $N-M$ , and  $y_m$  is missing data vector with a length of  $M$ . Because each element in  $y$  corresponds to a row in matrix  $A$ , we also need rearrange the rows in matrix  $A$ . According to the ordering used in equation 2, we reorganize matrix  $A$  as

$$A = \begin{pmatrix} A_a \\ A_m \end{pmatrix}. \quad (3)$$

Here,  $A_a$  and  $A_m$  correspond to available data  $y_a$  and missing data  $y_m$ , respectively. Using equations 2 and 3, we rewrite equation 1 as

$$A_a x = y_a \quad (4)$$

$$A_m x = y_m. \quad (5)$$

We cannot make any use of equation 5, because  $y_m$  is unknown. However, if we are able to solve equation 4 for  $x$ , then we use the solution  $x^*$  and equation 1 to recover the complete signal  $y$  as

$$y = Ax^*.$$

The number of equations (rows) in equation 4 is  $N-M$ , which is less than the number of unknowns  $K$ . The theory of compressed sensing (Donoho, 2006; Candès *et al.*, 2006) provides an approach to solving this under-determined problem. To recover  $x$ , we consider solving an  $L_1$ -regularization problem

$$\min \|x\|_1 \quad \text{subject to } A_a x = y_a. \quad (6)$$

Here,  $\|x\|_1$  is  $L_1$  norm defined as

## Image integration with learned dictionaries

$$\|x\|_1 = \sum_i |x_i|.$$

The missing data recovery problem appears more difficult than the compressed sensing (CS) one, because the randomness required in CS is not guaranteed in the case of missing data recovery. Zhang (2006) had a brief theoretical discussion on the conditions for a successful recovery of missing data  $y_m$ , and remarked that a proper selection of the matrix  $A$  (basis functions or dictionaries) is the most important issue related to the recovery conditions. Prior knowledge of the data is helpful for selecting a proper generic sparse transform. Figure 3 is a 1-D example to demonstrate this method using IDCT. There are two gaps with different lengths in the signal, which are the missing data  $y_m$  in equation 5; rest is the available data  $y_a$  in equation 4. The recovery is reasonable, though gap  $m_2$  has a relatively large recovery error.

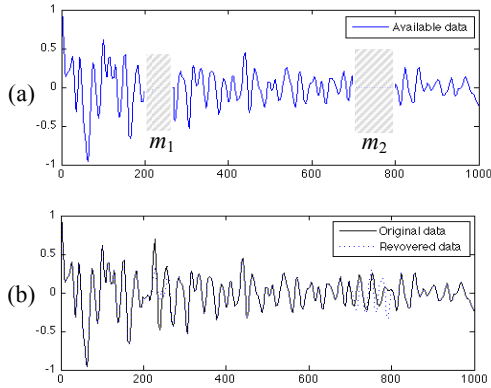


Figure 3: (a) 1-D signal with two segments blocked out ( $m_1$  and  $m_2$ ).  $L_1$ -regularized inversion is used to recover the missing data. (b) Comparison of the recovered and known signals.

In addition to generic sparse representations, we propose using learned dictionaries that may have better sparsity for certain types of data. Sparse coding provides a class of algorithms for finding overcomplete dictionaries that capture high-level features in the data. We use the sparse coding algorithms introduced by Lee *et al.* (2007) in this study. The goal of sparse coding is to represent an input signal  $y$  by a small number of basis functions in matrix  $D$  as

$$Dx = y. \quad (7)$$

Here,  $x$  is a vector of expansion coefficients. It is desired to have the fewest non-zeros values as possible in  $x$ . Matrix  $D$  has a dimension of  $N \times K$ . The basis functions can be overcomplete ( $K > N$ ), so that more patterns in the data can

be matched. The dictionary  $D$  in equation 7 is constructed by sparse coding using a large training data set  $Z$ . Assuming the size of training data set is  $L$ , then

$$Z = (y_1, \dots, y_L),$$

$$S = (x_1, \dots, x_L),$$

and equation 7 becomes

$$DS = Z. \quad (8)$$

Sparse coding algorithms can be applied to find dictionary  $D$  by solving an optimization problem

$$\min \sum_{i=1}^L \left( \frac{1}{2} \|y_i - Dx_i\|_2^2 + \lambda \|x_i\|_1 \right),$$

where  $\lambda$  is a weighting factor. Details on learning dictionaries can be found in Lee *et al.* (2007), Aharon *et al.* (2006), and in the abundant literature on the subject.

Learned dictionary  $D$  is just a special case of the matrix  $A$  in equation 1. After having dictionary  $D$ , we can use the same procedure as described in equations 2-6 for missing data recovery.

For the application to time-lapse monitoring, training data  $Z$  contains a collection of available time-lapse images. At reservoir-time (or slow-time)  $t^{(k)}$ , for example, we may use some or all data from  $y^{(0)}, y^{(1)}, \dots, y^{(k-1)}$  as training data  $Z^{(k-1)}$  to find dictionary  $D^{(k-1)}$ . Then the missing values in  $y^{(k)}$  can be recovered through

$$D^{(k-1)}x = y^{(k)}.$$

When new time-lapse images are available, we can adaptively train a new overcomplete dictionary and predict missing values in the current image.

### Example

We present a time-lapse synthetic example to demonstrate and test the image integration with learned dictionaries. The field example described in the previous section (see Figures 1-2) is under processing.

We first created a reservoir model and ran flow simulations. After the flow simulations were completed, the resulting phase saturations and pressure data were converted to time-lapse seismic velocities and densities. Then the seismic models were used to compute synthetic seismic data. This simulation was done earlier as part of a CO<sub>2</sub> sequestration project. The model parameters were taken from an unmineable coalbed. In this study we only used the time-lapse velocity models ( $V_p^{(k)}, k=0,1,\dots,4$ ) as images for integration tests.

## Image integration with learned dictionaries

Figure 4 shows images selected at four different reservoir-times ( $t^{(1)}, \dots, t^{(4)}$ ). These images correspond to  $P$ -wave velocity changes ( $\Delta V_p^{(k)} = V_p^{(k)} - V_p^{(0)}$ ) caused by  $\text{CO}_2$  injection. At a given reservoir-time, we should use all available images to train a dictionary. In this example, we use baseline image  $V_p^{(0)}$  and two time-lapse images  $V_p^{(1)}$  and  $V_p^{(2)}$  to train a dictionary, and apply the dictionary to recover the missing values in a later image, e.g.,  $V_p^{(3)}$ . A basis function has the same dimension as the training image patch. For simplicity, we picked one row as the image patch in this study. Figure 5 is the learned dictionary that consists of 28 basis functions. Each row in  $V_p^{(k)}$  can be approximately represented by a small number of these basis functions.

To test the approach, we cut a strip (missing data) from the velocity model  $V_p^{(3)}$  (Figure 6). Then we use the learned dictionary and also the available data in  $V_p^{(3)}$  to recover the missing part (Figure 7). Compared to the original image, the relative recovery error is less than 1%.

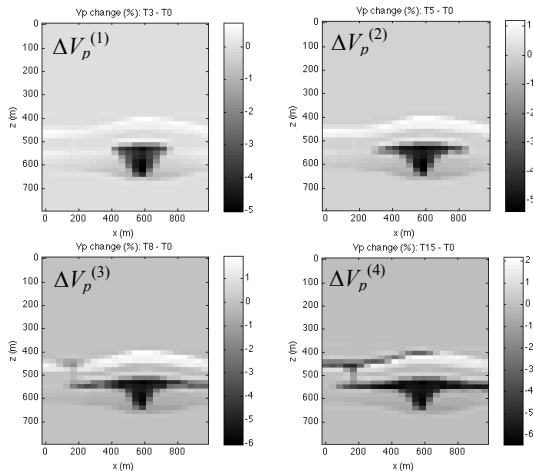


Figure 4:  $P$ -wave velocity changes (%) at four reservoir-times ( $t^{(1)}, \dots, t^{(4)}$ ). The  $\text{CO}_2$  plume grows with time. A leak can be seen after  $t^{(2)}$ .

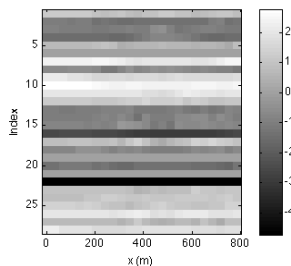


Figure 5: Learned dictionary. Each row is a basis function.

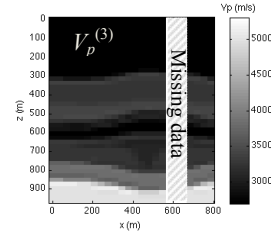


Figure 6: Velocity image at  $T_3$ . A strip is cut from the image as missing data.

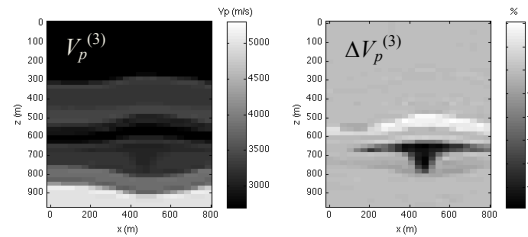


Figure 7: Recovered image. Left is absolute velocity and right is velocity change. Relative recovery error is less than 1%.

It should be pointed out that the basis functions used in this example are just one-dimensional functions. This is reasonable for this particular example because the geologic structure is mostly flat. In general, we would need to use 2-D image patches for the dictionaries. If we treat the reservoir-time as another dimension, 3-D image patches may be needed for the dictionaries.

## Conclusions

Learned dictionaries are an alternative for sparse representations using predefined bases or dictionaries. They offer the potential for better sparseness than the generic representations. The improved sparsity and  $L_1$ -minimization can be used to advantage for missing data recovery. Time-lapse seismic involves multiple surveys. 4-D multiple seismic surveys, therefore, provide good training data for finding overcomplete dictionaries adaptively.

## EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2011 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

## REFERENCES

- Aharon, M., M. Elad, and A. Bruckstein, 2006, K-SVD: An algorithm for designing over-complete dictionaries for sparse representation: *IEEE Transactions on Signal Processing*, **54**, no. 11, 4311–4322, [doi:10.1109/TSP.2006.881199](https://doi.org/10.1109/TSP.2006.881199).
- Arogunmati, A., and M. J. Harris, 2009, An approach for quasi-continuous time-lapse seismic monitoring with sparse data: Presented at the 79th Annual Meeting and International Exposition, SEG.
- Arogunmati, A., and J. Harris, 2010, A data-estimation based approach for quasi-continuous reservoir monitoring using sparse surface seismic data: Presented at the 72nd EAGE Conference and Exhibition.
- Candès, E., J. Romberg, and T. Tao, 2006, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information: *IEEE Transactions on Information Theory*, **52**, no. 2, 489–509, [doi:10.1109/TIT.2005.862083](https://doi.org/10.1109/TIT.2005.862083).
- Candès, E., and M. Wakin, 2008, An introduction to compressive sampling: *IEEE Signal Processing Magazine*, **25**, no. 2, 21–30, [doi:10.1109/MSP.2007.914731](https://doi.org/10.1109/MSP.2007.914731).
- Donoho, D., 2006, Compressed sensing: *IEEE Transactions on Information Theory*, **52**, no. 4, 1289–1306, [doi:10.1109/TIT.2006.871582](https://doi.org/10.1109/TIT.2006.871582).
- Elad, M., J. Starck, P. Querre, and D. Donoho, 2005, Simultaneous cartoon and texture image in painting using morphological component analysis (MCA): *Journal on Applied and Computational Harmonic Analysis*, **19**, no. 3, 340–358, [doi:10.1016/j.acha.2005.03.005](https://doi.org/10.1016/j.acha.2005.03.005).
- Harris, J., R. Langan, M. Van Schaack, S. Lazaratos, and J. Rector, 1995, High-resolution crosswell imaging of a west Texas carbonate reservoir: Part I — Project summary and interpretation: *Geophysics*, **60**, 667–681, [doi:10.1190/1.1443806](https://doi.org/10.1190/1.1443806).
- Harris, J. M., M. D. Zoback, A. R. Kovscek, and F. M. Orr Jr., 2007, Geologic Storage of CO<sub>2</sub>, in Global climate and energy project: SEP Report, [http://gcep.stanford.edu/research/technical\\_report.html](http://gcep.stanford.edu/research/technical_report.html).
- Herrmann, F., D. Wang, G. Hennenfent, and P. Moghaddam, 2008, Curvelet-based seismic data processing: A multiscale and nonlinear approach: *Geophysics*, **73**, no. 1, A1–A5, [doi:10.1190/1.2799517](https://doi.org/10.1190/1.2799517).
- Kuuskraa, V., and R. Ferguson, 2008, Storing CO<sub>2</sub> with enhanced oil recovery: DOE/NETL-402/1312/02-07-08.
- Lazaratos, S., and B. Marion, 1997, Crosswell seismic imaging of reservoir changes caused by CO<sub>2</sub> injection: *The Leading Edge*, **16**, 1300–1306, [doi:10.1190/1.1437788](https://doi.org/10.1190/1.1437788).
- Lee, H., A. Battle, R. Raina, and A. Ng, 2007, Efficient sparse coding algorithms: NIPS.
- Liu, B., and M. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: *Geophysics*, **69**, 1560–1568, [doi:10.1190/1.1836829](https://doi.org/10.1190/1.1836829).
- Lustig, M., D. Donoho, and J. Pauly, 2007, Sparse MRI: The application of compressed sensing for rapid MR imaging: *Magnetic Resonance in Medicine*, **58**, no. 6, 1182–1195, [doi:10.1002/mrm.21391](https://doi.org/10.1002/mrm.21391).

- Mallat, S., 2008, A wavelet tour of signal processing: The sparse way: Academic Press, 3rd edition.
- Selesnick, I., R. Van Slyke, and O. Guleryuz, 2004, Pixel recovery via L1 minimization in the wavelet domain: International Conference on Image Processing (ICIP), 3, 1819–1822.
- Zhang, Y., 2006, When is missing data recoverable: CAAM Technical Report TR06-15 Department of Computational and Applied Mathematics, Rice University.