A SET OF DESCRIPTORS FOR AUTOMATIC CLASSIFICATION OF SCATTERERS IN SEISMIC SECTIONS

S. Maciel and R. Biloti

email: *susanne@unb.br*

keywords: *Machine Learning, SVM, diffractions, descriptors*

ABSTRACT

Seismic diffractions are mainly induced by edges, tips and small structures, so diffraction imaging can extract valuable information to identify subsurface scattering features. We investigate the possibility to image and characterize diffractions using pattern recognition methods. To this end, we look at kinematical and dynamical aspects of diffraction operators under a determined velocity model and we propose a set of attributes that better distinguish diffractions from reflections. These attributes are used as descriptors of imaging points on a seismic section to perform automatic classification using supervised and unsupervised algorithms. We evaluate the method using synthetic and GPR data. For synthetic data, we show results from amplitudes picking within a range of error on velocities to indicate the method sensivitness on velocity model. For real datasets, velocity analysis is performed. Results indicate that the method is robust even for low signal-to-noise ratio datasets.

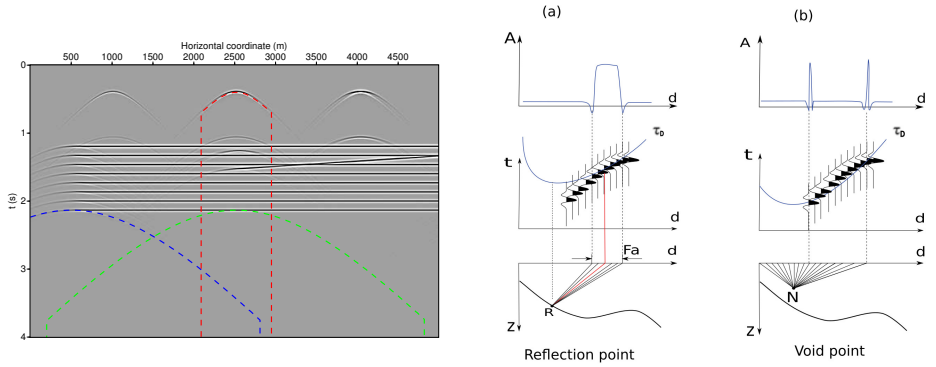
INTRODUCTION

The theory of seismic wave propagation in acoustic media is used to unriddle seismograms into realistic Earth models. Traditional processing of seismic data generally uses information from the reflected wavefield to obtain images of interfaces at the subsurface. An important aspect when interpreting subsurface images is the identification of small scale features, such as faults, channels and fractures. Instead of promoting reflections, seismic energy interacting with these structures results in diffractions, when their dimensions are smaller than the acoustic wavelength emitted during seismic acquisition (Trorey, 1970; Klem-Musatov, 1994). This information has been used for high-resolution imaging (Khaidukov et al., 2004; Fomel et al., 2007) or local velocity analysis (Sava et al., 2005; Reshef and Landa, 2009).

Even though theoretical studies on diffracted waves have been developed in detail since 1962, with the remarkable work of Keller (1962) on the geometrical theory of diffractions, diffraction imaging is being considered by oil and gas industry only for a few years until today, mainly within regions with high density of fractures, such as shales and carbonates. Recent works on this theme can be found on Sturzu et al. (2014); Kowalski et al. (2014); Burnett et al. (2015); Grasmueck et al. (2015)

Diffracted seismic waves are characterised by some peculiar attributes. Trorey (1970) showed analitically that for a single truncated plane reflector, the phase of the diffractions suffers a reversal of 180° on either side of the reflecting edge. Diffracted waves are recorded as significantly lower energy than reflected waves. Their amplitudes decay faster than it would be by simple geometrical spreading, what makes diffractions to be treated as noise in traditional seismic processing. These phenomena can also be observed on seismograms of controlled experiments of physical modeling of a simulated fault (Hilterman, 1970).

In isotropic media, diffraction traveltimes are approximated by a conventional double-square-root (DSR) equation. The amplitudes extracted along the elementary diffractions, known as the diffraction operator, forms a curve that was used by Tabti et al. (2004) to determine the so-called Fresnel aperture, used to enhance Kirchhoff-type depth-migration.



(a) Synthetic dataset. **(b)** Fresnel aperture.

Figure 1: Left: Synthetic dataset. Right: Top - Diffraction operator associated with a reflection (a) and a void point (b). Vertical axis is the amplitude, and horizontal axis is the horizontal coordinate of mid point. Middle - Diffraction traveltime is tangent to the reflection traveltime at the specular reflection event (a). For void points , there is no tangency to any events, and the diffraction operator shows random peaks (b). Bottom - Location in depth of the imaging point associated with a reflection (a) and a void point (b).

Tabti et al. (2004) proposed a version of Kirchhoff migration with Fresnel aperture selection, which provides better resolution on reflection imaging. Diffraction imaging may be performed if part of the amplitude inside reflection Fresnel aperture is removed (Bona and Pevzner, 2015). According to Tabti et al. (2004), to every image point y it is associated a diffraction operator $d(y)$, which is a vector of dimension n, the number of traces in the section. Each element of $d(y)$ is defined as $d_k = W(\xi_k, y)\partial_t U(\xi_k, t)|_{t=\tau_D(\xi_k,y)}$ and ξ_k is the coordinate that parametrizes the k-th trace, is defined as the vector of all seismic amplitudes to be stacked by a Kirchhoff migration for imaging point y and midpoint coordinate ξ. $W(\xi, y)$ is a weight function, $U(\xi, t)$ is the seismic data measured at position ξ and time t, and $\tau_D(\xi, y)$ is the traveltime of the elementary diffraction of a point scatterer in y. In isotropic media with small lateral variations on the velocity model, for a common-offset configuration τ_D is computed by double-square root approximation

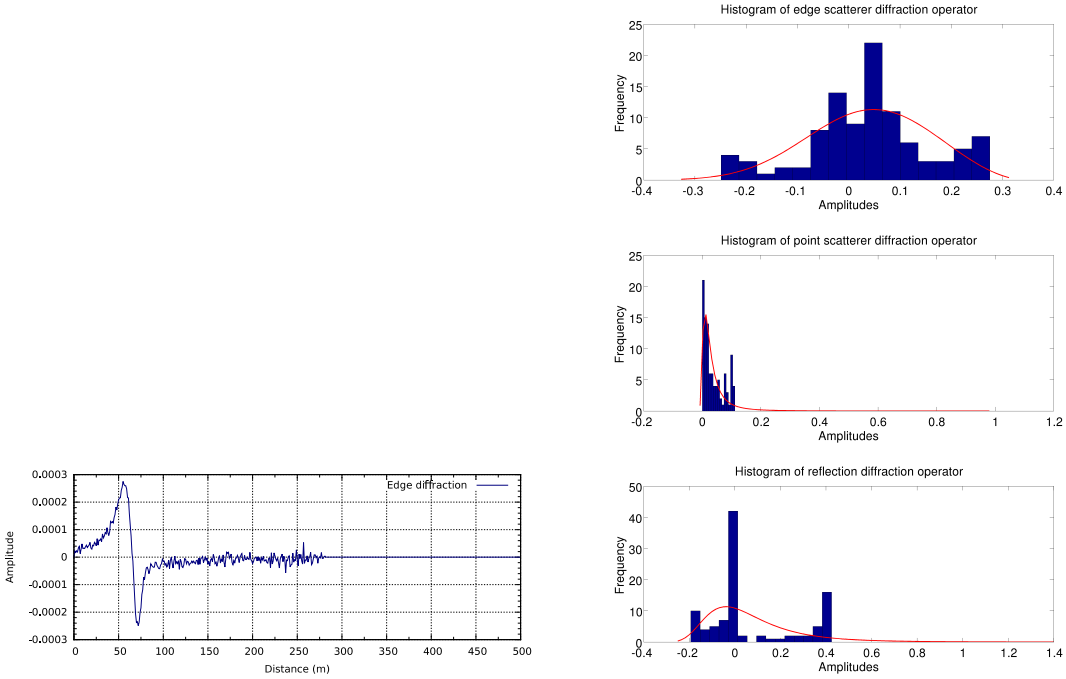
$$\tau_D(\xi; y) = \sqrt{\left(\frac{t_0}{2}\right)^2 + \left(\frac{\xi - h - \xi_y}{v_{rms}}\right)^2} + \sqrt{\left(\frac{t_0}{2}\right)^2 + \left(\frac{\xi + h - \xi_y}{v_{rms}}\right)^2}, \quad (1)$$

where t_0 is the zero-offset two-way time for image point y, ξ_y is the coordinate of y, h is the half-offset and v_{rms} is the RMS velocity. In this study, for diffration imaging in time we use the Double Square Root equation in the limits of time migration. For diffraction imaging in depth, we use paraxial traveltime calculation provided by Seismic Unix package (Cohen and Stockwell, 2013).

Attributing to point y the sum of elements of $d(y)$ over the migration aperture leads to Kirchhoff migration.

A diffraction traveltime curve associated to an image point located on a reflector is tangent to the reflection traveltime at the specular reflection event (see Fig. 1(b))(Schleicher et al., 1997). This point becomes a tangential region when the source has limited bandwidth, which is defined by Tabti et al. (2004) as Fresnel aperture, and by Schleicher et al. (1997) as the minimum aperture for true-amplitude depth migration. As illustrated on Fig. 2(a), the Fresnel aperture turns the diffraction operator associated with a reflector to have a *plateau* shape. In the case of a tip or edge scatterer, the associated elementary diffraction traveltime corresponds to the diffracted seismic event. Thus, its Fresnel aperture extends theoretically to infinity, and the diffraction operator shape will vary according to the nature of the scatterer: with a $180°$ phase shift if an edge diffraction, or an approximated gaussian shape if a point scatterer (see Fig. 2(a)).

This fact is used by Figueiredo et al. (2013) to classify imaging points. They apply a two-class k nearest neighbours (kNN) pattern recognition technique to amplitudes along diffraction operators to distinguish between diffractions, reflections or absence of scattering energy.



(a) Amplitudes. **(b)** Histograms.

Figure 2: Left: Amplitudes collected along indicated curves on Figure 1(a). Point scatterer diffraction operator (top); Edge diffraction operator (middle); Diffraction operator associated with a reflection (bottom). Right: Diffraction operators' associated PDFs are similar to their histograms.

We present a set of routines to perform automatic detection of diffractions on unmigrated data using pattern recognition techniques. We propose an extension of Figueiredo et al. (2013) approach to diffraction imaging, using a set of features to better distinguish diffraction operators. Classification is performed with Machine Learning algorithms. For the diffraction imaging task, we have used Support Vector Machines (SVM), after a study by Kotsiantis et al. (2006) pointing that SVM approach can present the best accuracy for classification of a waveform dataset. The next session includes a brief Machine Learning background, explaining how SVM works.

MACHINE LEARNING BACKGROUND

Machine learning is an intersection field between computer science and statistics, that explores the construction of algorithms that can learn from and make predictions on data. Such algorithms operate by building a model from a set of input examples in order to accomplish a given task, rather than following strictly static program instructions to make data-driven predictions.

Input examples are numerically described by an ensemble of quantities that characterizes an object, denominated *descriptor*. The object may not be entirely described by the descriptor, but descriptors for different classes of objects should be different enough to allow the discrimination of the objects.

A descriptor is composed by experimental measures, or theoretical calculations that describe the structure of the object. The major hypothesis is that descriptors capture some important characteristic of the object, and then a mathematical function can generate a mapping between the descriptor space and a property space, where classes are defined. Another hypothesis is that objects with similar descriptors must have similar properties. In many cases, the task of building descriptors is equivalent to find the best classifier for a problem.

Learning algorithms are employed for classification or regression tasks. In classification, inputs are divided into two or more classes, and the algorithm must produce a model that assigns unseen inputs to one or more of these classes. The algorithms usually has two main phases: training and testing. On the

training stage, a training set of objects is used by the learner to build a general model about the space given that enables it to make predictions in new samples. The training set usually comes from some unknown probability distribution, but must be considered representative of the space of occurrences. Learning can be supervised or unsupervised. Pattern recognition methods that exploit *a priori* known information about the training data set are known as *supervised pattern recognition*, or in the more general context of machine learning, as *supervised learning*.

For instance, consider a seismic common-offset gather as input dataset for learning. The set of imaging points of the section, denoted here by Y, is the set of objects to be classified. A descriptor $\gamma : Y \to F$ maps an object $y \in Y$ onto a feature $f \in F$ on the feature space F. The elements in F are then mapped onto an instance space $L = \{\omega_1, \ldots, \omega_l\}$, by the classifier $\theta : F \to L$. The set of classes L might be, for example, composed by the classes point diffraction (ω_1), edge diffraction (ω_2), reflection (ω_3) or void point (ω_4). With a good representation of the input space in the training phase, further imaging points are automatically classified by the algorithm.

There are some well known Machine Learning techniques for clustering and classification of data. Further references are found for example on Theodoridis and Koutroumbas (2009).

METHOD

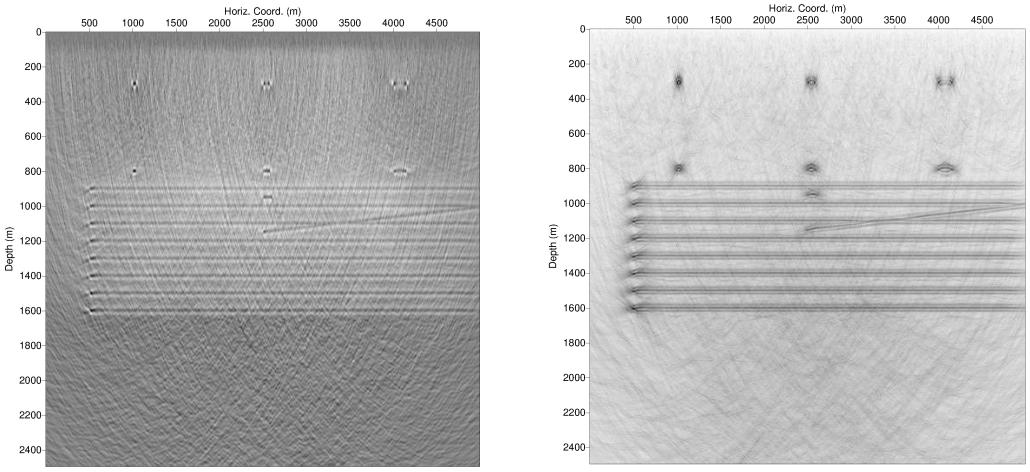
Building a descriptor for classification

Let the seismic section be represented by Y, a common-offset gather in our case. The objects to be classified are imaging points $y \in Y$. Figueiredo et al. (2013) used the diffraction operators presented by Tabti et al. (2004) as descriptor. In other words, Figueiredo et al. set their descriptor as $d(y)$ and the feature space as $\mathbb{R}^n$ (recall that n is the number of traces in the input common-offset section). This implies that the dimension of the feature space is dependent on the geometry of the common-offset gather, which is dataset dependent. Classification is performed by kNN algoritm, where training data is composed by diffraction operators associated with scatterer imaging points, labeled as *diffractions*, and diffraction operators associated with void image points, labeled as *noise*. In other words, their classifier θ_F maps vectors of $\mathbb{R}^n$ to ω_1 (meaning diffraction) or ω_2 (meaning *noise*). The training dataset $T_f \subset \mathbb{R}^n$ is built manually by the user, from a synthetic dataset with the same geometry of the common-offset to be analyzed, where the exact position and nature of the scatterers are known. They used Euclidean distance to measure the distance between neighbours.

We propose different choices for the feature space and descriptor. Instead of using the diffraction operator itself as the descriptor, we compute some quantities from the diffraction operator, which are less dependent of the geometry of the dataset. In other words, our descriptor γ is defined as $\gamma(y) = \phi(d(y))$, where ϕ is a set of measures of $d(y)$. Thus, our feature space is $\mathbb{R}^m$, where m is the number of characteristics quantities computed from $d(y)$. Usually, $m \ll n$. This has the side effect to make the classification problem cheaper when compared to the strategy of Figueiredo et al. (2013).

What makes diffractions different from reflections on the diffraction operator space $\mathbb{R}^n$ is essentially their shapes. For each imaging point, the associated diffraction operator is seen as a random variable. In this way, $\phi(d(y))$ is a set of shape parameters of the probability distributions associated to the amplitudes of diffraction operators. Figure 2(b) shows the histograms of diffraction operators associated with an edge diffraction, a point diffraction and a reflection. The histogram is a tool to show the frequency function of a distribution, or the number of sample values falling into a certain specified range. To build a histogram, one must take every class interval as the basis of a rectangle with height $\frac{v}{ph}$, where v denotes the number of sample values in the class, h is the length of the class interval and p is the number of classes. The area of any rectangle in the histogram is equal to the corresponding class frequency $\frac{v}{p}$. For large p this may be expected to be approximately equal to the probability that an observed value of the variable will belong to the corresponding class interval, which is equal to the integral of the frequency function over the interval. This means that the histograms of diffraction operators are similar to their respectives Probability Density Functions, if diffraction operators are seen as random variables. (For references on histograms see, for example, Cramer (1946)).

Statistical measures, such as *skewness* and *kurtosis* are commonly as used shape parameters of probability densities (Theodoridis and Koutroumbas, 2009). In Statistics, skewness is a measure of the asymmetry



(a) Sum

(b) Standard Deviation

Figure 3: Lower moments.

of a probability density function. It is zero if the distribution is symmetric, like the normal distribution. If the mass of the distribution is concentrated on the left side, skewness is positive, and it is negative if it is concentrated on the right side. Kurtosis is a measure of the peakdness or flattening of a probability density function. Normal distribution has kurtosis equal to 3. Distributions that are more outlier-prone than the normal distribution have kurtosis greater than 3; distributions that are flatter have kurtosis less than 3. It is possible to see in Figure 2(b) that diffraction operators associated with point diffractions are highly skewed and present higher kurtosis, while reflections are less skewed, and present lower kurtosis. Edge diffractions presents diffraction operators closer to gaussian distribution, due to their phase inversion, that distributes almost equally negative and positive amplitudes.

Besides skewness and kurtosis, in Probability Theory, there is a number of statistical attributes that are used to describe the shape of a Probability Density Function (PDF). We propose using central moments. Central moment is the expected value of a specified integer power of the deviation of the random variable from the mean. The third and fourth central moments are generally used for shape parameter description. They are related to the skewness and kurtosis of a function, which may be seen as standardized central moments. The various moments form a set of values by which the properties of a probability distribution are usefully characterised. A distribution can be characterized by location, scale and shape parameters. The location parameter shifts the entire distribution left or right, the scale parameter compresses or stretches the entire distribution and the shape parameter changes the shape of the distribution in some other way. Higher-order central moments relate only to the spread and shape of the distribution. Lower-order moments, such as mean, are related to the location of high frequencies on a PDF.

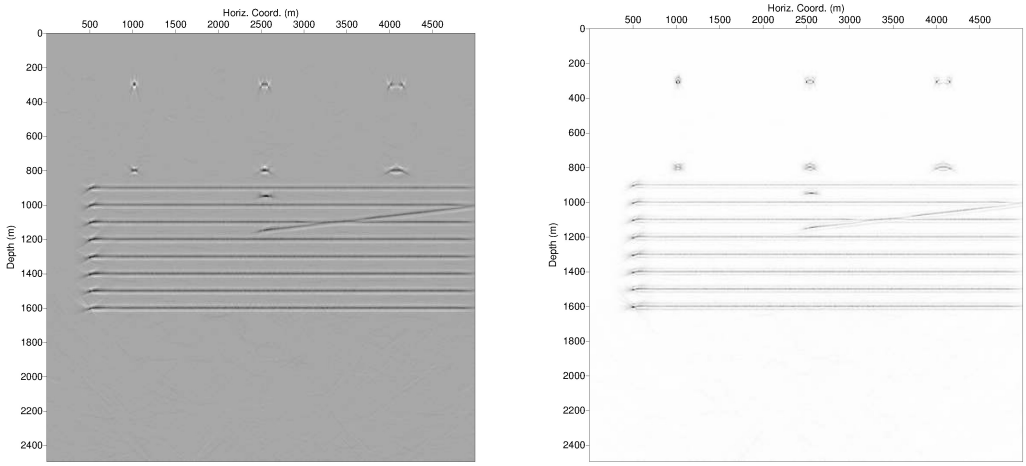
For a discrete scalar function $f(x)$, its k-th central moment about the mean is defined as

$$m_k = \sum_{x=1}^{N} (x - \mu)^k f(x), \tag{2}$$

where $\mu = \frac{1}{N} \sum_{x=1}^{N} f(x)$. If we have an infinite number of central moments, we can completely describe the function (see Cramer (1946) for details).

For each imaging point of a common-offset section, we compute a vector composed by the six first central moments of the diffraction operator. This vector will play the role of the descriptor in the classification problem. On Figures 3(a) and 3(b), it is shown in a gray scale the first and second central moments of the section in Figure 1(a). Note that the first moment corresponds to the mean of the diffraction operator amplitudes, which may generate an scaled version of Kirchhoff imaging. Figures 4(a) and 4(b) show the scaled third and fourth central moments, where noise is significantly suppressed.

Figure 5 shows how imaging points are distributed on cuts of the feature space $\mathbb{R}^6$. In order to make



(a) Skewness **(b)** Kurtosis

Figure 4: Higher moments.

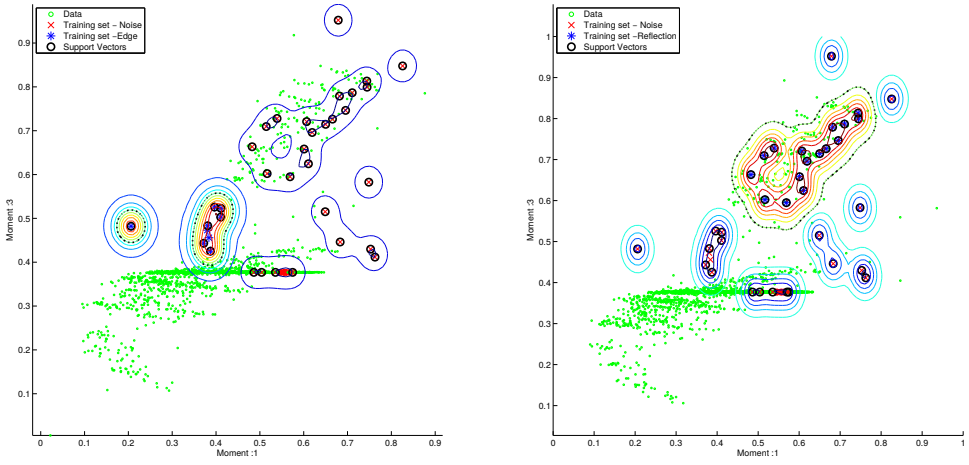
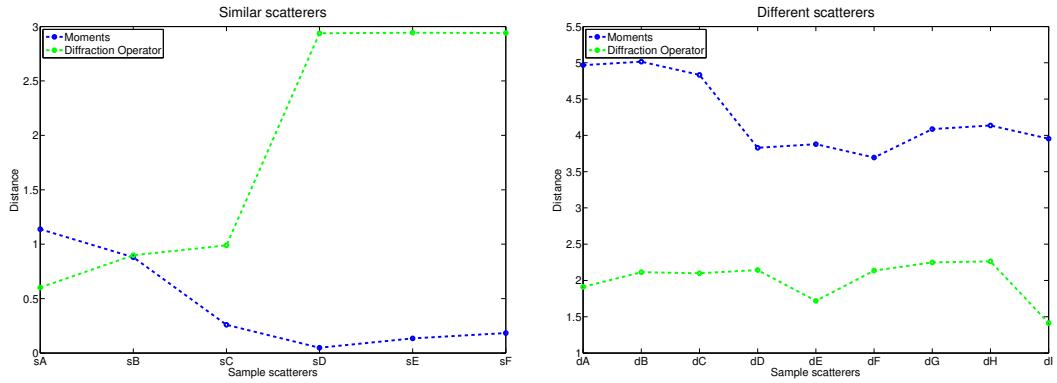


Figure 5: Data descriptor: each imaging point is represented by a vector with 6 dimensions, where the i^{th} component is the i^{th} central moment of the corresponding diffraction operator. Here, we show how data is distributed on the plan composed by first (vertical axis) and third (horizontal axis) central moments. Curve levels indicates how SVM classifies data. Boundary layer is indicated in dashed line.



(a) **(b)**

Figure 6: Distances from similar and distinct scatterers in different metrics. On the left, it is shown the distances for six pairs of similar scatters, and on the right, nine pairs of different scatterers.

visualization feasible, we displayed only two dimensions of $\mathbb{R}^6$, for instance, the first and third dimensions, corresponding to the first and third central moments respectively. All imaging points of the common-offset section presented on Figure 1(a) are displayed in green. We have manually selected some imaging points corresponding to reflections, edge diffractions and void points, and displayed them in black, blue and red, respectively. For classification purposes, we have labeled reflections and void points as *noise*, and edge diffractions as *diffractions*. On Figure 5 right, it is shown how SVM performed classification. The support vectors are shown in black circles, and curve levels indicate how the Gaussian Radial Basis Function behaves on the nonlinear separable training set. Boundary layer is indicated in black.

Note that for this example, edge diffraction imaging points are grouped in a different region from reflection imaging points on the feature space, using Euclidean distance. This fact indicates that the descriptor γ is suitable for the purpose of identifying diffractions. A good descriptor should follow the golden rule: distances between feature vectors from similar scatterers should be small, and distances between scatterers of different type should be big. Guided by this principle, let us analyse the descriptors we have so far.

Consider two image points y_A and y_B, associated to scatterers located apart in the seismic section, and two other image point $y_{A'}$, $y_{B'}$ in the same horizontal coordinate of y_A and y_B, respectively, but vertically shifted to be over reflectors. A good descriptor should map y_A and y_B close together in the feature space, since both of them belong to the *diffractor class*. The same should hold for the mapping of $y_{A'}$ and $y_{B'}$, as they belong to the *non-diffractor class*. But most important of all, y_A and y_B must be mapped as far as possible from $y_{A'}$ and $y_{B'}$, since they belong to different classes.

However this is not what happens when the diffraction operator itself is used as descriptor. Since the horizontal coordinates of y_A and $y_{A'}$ are the same, high amplitudes on the diffraction operator associated with y_A will be placed on the same location of the high amplitudes of the diffraction operator associated with $y_{A'}$. This means that $\|d(y_A) - d(y_{A'})\|$ is small, where $\|\cdot\|$ states for the Euclidean distance. The same rationale holds for scatterers positioned far away from each other. Since the amplitude peak of $d(y_A)$ and $d(y_B)$ are in different positions, $\|d(y_A) - d(y_B)\|$ is big, giving no clue that both image points are of the same type.

On the other hand, the descriptor based on central moments is sensible to the shape of the distribution of amplitudes, but not to their spatial location. This means that, since $d(y_A)$ has a completely different shape from $d(y_{A'})$, $\|\gamma(y_A) - \gamma(y_{A'})\|$ is big. Furthermore, $\|\gamma(y_A) - \gamma(y_B)\|$ is small, since both image points have the same amplitude pattern over the diffraction operator.

Figure 6 shows the Euclidean distances for some pairs of sample scatterers on both feature spaces. We have chosen nine specific imaging points from the synthetic dataset presented on Figure 1(a), namely the tip diffractions located on coordinates $(1000, 300)$, $(2500, 300)$ and $(2500, 800)$, the edge diffractions located on $(500, 1100)$, $(500, 1200)$ and $500, 1600)$ and the reflections located on $(1000, 1100)$, $(2500, 1200)$ and $(2500, 1600)$. We formed randomly six pairs of similar scatterers, denoted by sK, $K = A, B, C, ...$, where sK is a pair of imaging points associated both to a scatterer from the same nature (a reflection, an edge

diffraction or a tip diffraction), and nine pairs of different scatterers, denoted by dK, $K = A, B, C...$, where dK is a pair of imaging points associated to different scatterers each. Note that similar scatterers are closer in γ, and different scatterers are more distant in γ, showing that the descriptor γ formed by moments is more representative in this sense than d, the diffraction operator itself used as the descriptor.

In supervised classifiers, the training set is a subset of feature space. This means that to use the whole diffraction operator as input, for each common-offset section analyzed by the algorithm, the training set must be generated according to the section geometry. This is a big disadvantage in the sense of classification task, since it would be more interesting if one could use the same training set for several data sets to be classified, regardless of their number particular number of traces. Associated with the fact that diffraction operator can achieve very big dimensions, it makes the task of building a good training set more challenging in approach of Figueiredo et al. (2013).

Workflow

Our workflow consists mainly on four steps. Pre-processing includes seismic treatments, which may vary depending on the nature of the dataset; velocity model building and ray tracing; and amplitudes picking along traveltimes trajectories, to build a cube of amplitudes. The next step consists of filtering of the cube sections, in order to remove undesirable peaks. Depending on your imaging purpose, filtering can be performed in many ways. With the dataset properly prepared, the next step is the feature space building, that consists on extracting statistical attributes from the cube of amplitudes, and building a vector space where each dimension represents an attribute. Finally, classification is performed, and the last step consists on selecting the best product for interpretation.

Pre-processing

Pre-processing of data must be taken with special care when dealing with seismic diffractions, due to their small amplitudes. Since signal strength of a diffracted wave has rapid decayment with the distance from the tangential portion with the reflected wave, it is common practice to apply a special type of gain, obtained by dividing traces by their envelopes, using a regularization parameter (Landa et al., 1987; Figueiredo et al., 2013). This is the same as obtaining the cosine of the phase of the signal.

The next step consists on building a velocity model to compute traveltimes. This can be done by several methods (see Jones, 2010), in time or depth domain. It is typically an iterative process and requires many runs of computationally intensive prestack depth migration. To produce accurate images of small structures using the method proposed here, it is necessary to build a sufficiently good velocity model, in order to correctly predict the kinematics of the diffractions.

Our method is presented for a single common-offset section, although the idea can be applied for other geometries, with appropriate adjustments. Each imaging point is seen as a point scatterer. Once traveltime table is available, amplitudes along diffraction curves for all imaging points are collected, leading to the construction of a cube of amplitudes, with dimensions $NX \times NZ \times NTR$, where NX and NZ are horizontal and vertical dimensions of velocity model, respectively, and NTR is the number of traces of input data.

For each vertical profile on a common-offset section, collecting amplitudes along the diffraction curve $d(y)$, and displaying it on the corresponding depth, generates a gather where diffraction operators are analyzed. It might be seen as a slice of the cube of amplitudes normal to NX direction. Tabti et al. (2004) presented some examples of those sections. Correlating with NMO corrected CMP gathers idea, we propose to call these sections as Diffraction Corrected Common-Offset gathers, or DC-CO gathers.

Filtering Note that when a diffraction hyperbola crosses reflection events or other scatterer events in the DC-CO gather, the corresponding diffraction operator has peaks of amplitudes which are not related to the imaging point. These peaks are removed by applying dip-filtering on Diffraction Corrected panel. Tabti et al. (2004) proposes the application of low-pass filtering, which can remove undesirable peaks. On the other hand, it might not be suitable for our purposes since low pass filtering may also destruct the original pattern of a diffraction operator shape, leading to further misclassification of the imaging point.

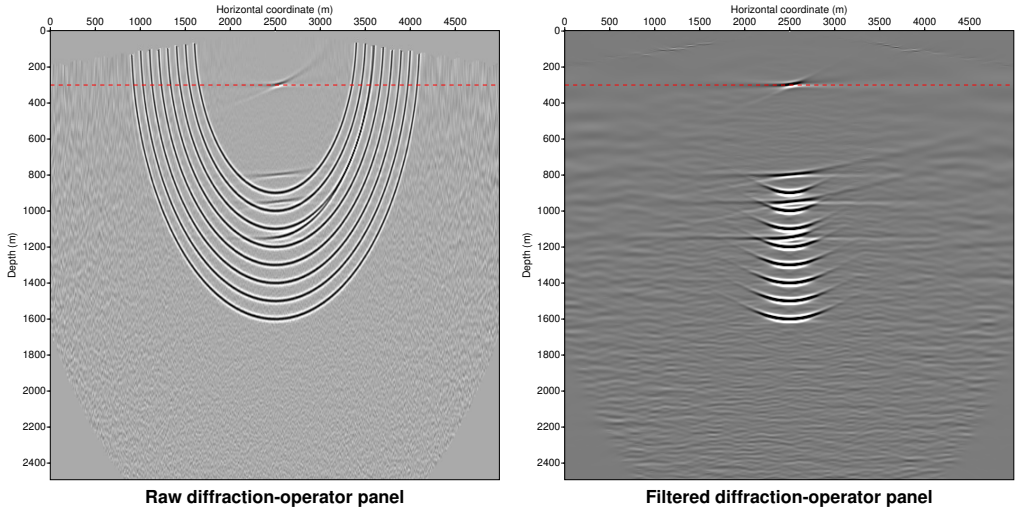


Figure 7: Diffraction Corrected CO gather before (left) and after (right) dip-filtering.

Validation

From a common-offset section, the method consists on building a binary map of subsurface indicating where small scatterers are located. Validation of the method should be performed by "demigration" style, using our results to model a seismic acquisition using the binary map as a reference of scatterers position, using the same velocity model used to calculate the diffraction operators as background. The velocity model can be updated in an iterative fashion until the difference between the original section and the modeled one is reasonable enough.

We used the algorithm provided by Cohen and Stockwell (2013) package to estimate the migration residual statics using the technique presented by Tjan et al. (1994), to check if detected scatterers appear at correct times on the input session. The program takes a migrated seismic section and a set of travel time tables for a specific background velocity model and generates synthetic seismic data in the form of common shot gathers. A demigration algorithm with reliable amplitudes is still needed to implement the iterative version of the method.

RESULTS

For numerical experiments, we use a simple synthetic data generated by Kirchhoff modelling, simulating a zero-offset section with 500 source-receiver pairs spaced by 10 m covering an extension of 5000 m, presented in Figure 1(a). We normalize the data set trace-by-trace using its envelope in order to increase magnitude of diffraction events.

For each point on the common-offset section, we extract the diffraction operator and calculate its first six central moments to compose the input data for the classification phase.

The algorithm was also applied to a GPR data set, that consists on a radargram acquired with the purpose of checking the existence and distribution of tie bars between concrete slabs of a rigid pavement structure along a section of BR-101/NE highway, in the Brazilian state of Pernambuco (Silva et al., 2013) (see Fig. 8). Classification results of SVM classification on GPR dataset is shown in Figure 9, where it is possible to see that evident diffractions from the tie bars were correctly positioned, and less evident diffractions formed by the contact between gravel and soil are detected by the algorithm.

CONCLUSIONS

We presented a new set of descriptors for imaging points on a seismic section and a method for classification of scatterers using Support Vector Classifiers. It consists in an application of well-stablished Machine Learning techniques to distinguish diffraction events from reflection events and noise areas by

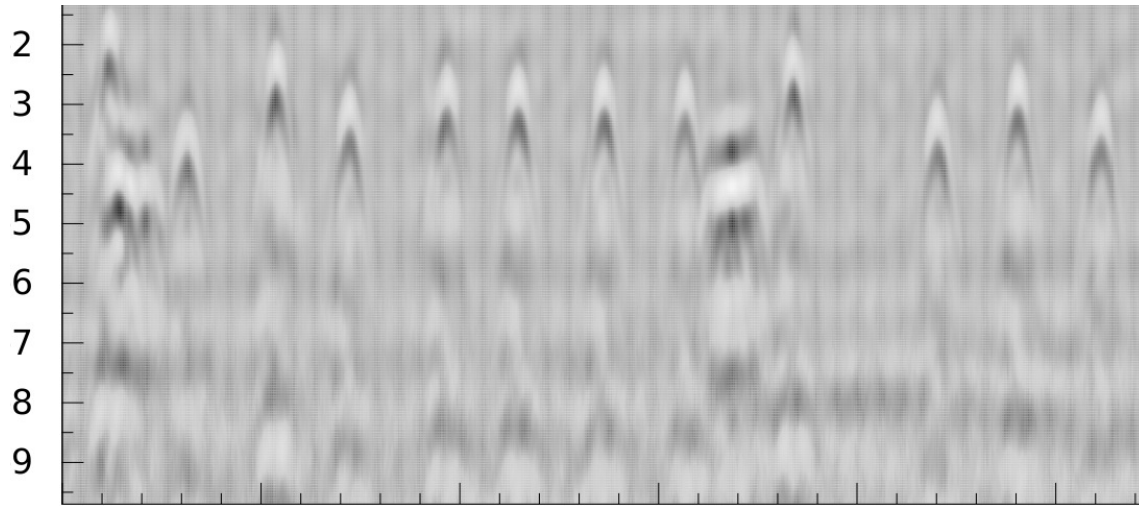


Figure 8: Radargram showing diffractions from metal bars buried on the subsurface.

their kinematical and dynamical patterns. We discussed a new approach for arrangement of input data on pattern recognition methods for diffraction detection using central moments that optimizes computation and generates new images for interpretation. Application on a GPR dataset presented successful results.

As a mather of fact, the economy on computational time in the classification step in our procedure comparing to Figueiredo et al. (2013) approach is irrelevant, since the computation of central moments is significantly time consuming. The biggest differential of the two approaches is the fact that for supervised learning, the proposed descriptor allows the user to build one single training set for several datasets to be classified. Another advantage is that unsupervised learning already gives good results with the proposed descriptor.

Our preliminary results indicate that pattern recognition methods are a wide field of research that opens new possibilities to create suitable tools for detection of diffractions. Further studies include development of combining machine learning techniques, diffraction imaging, local velocity analysis and completness of the validation workflow, consisting on modeling the classified image in order to measure effectiveness of the method.

Another idea to be further expanded is the use of fuzzy classification systems instead of binary classification. The transition between a real diffraction and a reflection event is smooth, so binary classification might provide random outcomes on regions where the size of the scatterer is about the size of the wavelength. There are already implemented versions of fuzzy SVM and k-means that can be used to this end. Weights from fuzzy classification might be used for weightning stacking procedures in order to obtain modified versions of reflection or diffraction images.

ACKNOWLEDGMENTS

This work was kindly supported by the sponsors of the *Wave Inversion Technology (WIT) Consortium*.

REFERENCES

Bona, A. and Pevzner, R. (2015). Using Fresnel zone to characterise and image different types of diffractors in low S/N situations. In *77th EAGE Conference & Exhibition*.

Burnett, W., Klokov, A., Fomel, S., Bansal, R., Liu, E., and Jenkinson, T. (2015). Seismic diffraction interpretation at Piceance Creek. *Interpretation*, 3:SF1–SF14.

Cohen, J. K. and Stockwell, J. J. W. (2013). *CWP/SU: Seismic Un*x Release 43R5: an open source software package for seismic research and processing*. Colorado School of Mines.

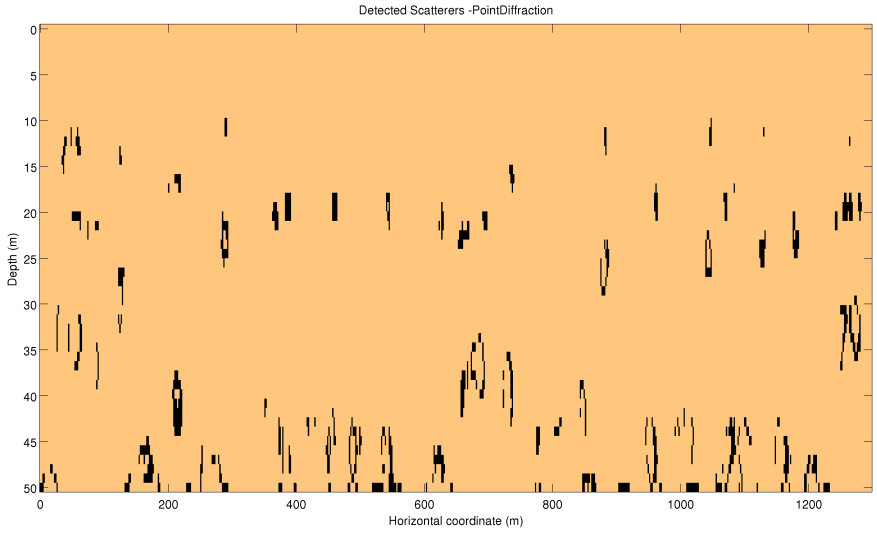


Figure 9: Classified point diffractions imaging points for GPR data.

Cramer, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton, USA, 2nd edition.

Figueiredo, J., Oliveira, E., Esmi, E., Freitas, L., Schleicher, J., Novais, A., Sussner, P., and Green, S. (2013). Automatic detection and imaging of diffraction points using pattern recognition. *Geophysical Prospecting*, 61:368–379.

Fomel, S., Landa, E., and Taner, T. (2007). Postack velocity analisys by separation and imaging of seismic diffractions. *Geophysics*, 72:U89–U94.

Grasmueck, M., Moser, T., Pelissier, M., Pajchel, J., and Pomar, K. (2015). Diffraction signatures of fracture intersections. *Interpretation*, 3:SF55 – SF68.

Hilterman, J. (1970). Three-dimensional seismic modeling. *Geophysics*, 35:1020 – 1037.

Jones, I. (2010). *An introduction to: velocity model building*. EAGE Publications bv.

Keller, J. B. (1962). Geometrical theory of diffractions. *Journal of the Optical Society of America*, 52:116–130.

Khaidukov, V., Landa, E., and Moser, T. J. (2004). Diffraction imaging by focusing-defocusing: An outlook on seismic superresolution. *Geophysics*, 69:1478–1490.

Klem-Musatov, K. (1994). *Theory of seismic diffractions*. Society of Exploration Geophysicists.

Kotsiantis, S. B., Zaharakis, I. D., and Pintelas, P. E. (2006). Machine learning: A review of classification and combining techniques. *Artificial Intelligence Review*, 26:159–190.

Kowalski, H., Godlewski, P., Kobusinski, W., Makarewicz, J., Podolak, M., Nowicka, A., Mikolajewski, Z., Chase, D., Dafni, R., Canning, A., and Koren, Z. (2014). Imaging and characterization of a shale reservoir onshore Poland, using full-azimuth seismic depth imaging. *First Break*, 32:101 – 109.

Landa, E., Shtivelman, V., and Gelchinsky, B. (1987). A method for detection of diffracted waves on common-offset sections. *Geophysical Prospecting*, 35:359–374.

Reshef, M. and Landa, E. (2009). Post-stack velocity analisys in the dip angle domain using diffractions. *Geophysical Prospecting*, 57:811–821.

Sava, P., Biondi, B., and Etgen, J. (2005). Wave-equation migration velocity analisys by focusing diffractions and reflections. *Geophysics*, 70:U19–U27.

Schleicher, J., Hubral, P., Tygel, M., and Jaya, M. S. (1997). Minimum aperutres and Fresnel zones in migration and demigration. *Geophysics*, 62:183–194.

Silva, L. A., Borges, W. R., Cunha, L. S., Branco, M. G. C., and Farias, M. M. (2013). Use of GPR to identify metal bars and layer thickness in a rigid pavement. *Geotechnical and Geophysical Site Characterization 4*, pages 1341–1346.

Sturzu, I., Popovici, A., Pelissier, M., Wolak, J., and Moser, T. (2014). Diffraction imaging of the Eagle Ford shale. *First Break*, 32:49–59.

Tabti, H., Gelius, L., and Hellmann, T. (2004). Fresnel aperture prestack depth migration. *First Break*, 22:39–46.

Theodoridis, S. and Koutroumbas, K. (2009). *Pattern Recognition*. Academic Press Elsevier, San Diego, USA.

Tjan, T., Lamer, K., and Audebert, F. (1994). Prestack migration for residual statics estimation in complex media. *SEG Annual Meeting Abstracts*.

Trorey, A. W. (1970). A simple theory for seismic diffractions. *Geophysics*, 35:762–784.