AUTOMATIC DETECTION AND IMAGING OF DIFFRACTION POINTS USING PATTERN RECOGNITION

J. J. S. de Figueiredo, J. Schleicher, F. Oliveira, E. Esmi, L. Freitas, A. Novais, P. Sussner, and S. Green

email: *jadsomjose@gmail.com and js@ime.unicamp.br*
keywords: *Diffraction location, diffraction operator, patterns recognition*

ABSTRACT

In this work, we further develop a method for the detection of diffractor points in a common-offset-gather domain. In our approach, the diffraction imaging is based on the diffraction operator, which can be used in both the time and depth domains, in dependence of the complexity of the area. The method applies pattern recognition to the amplitudes along the diffraction operator. While the method, in principle, requires knowledge of the migration velocity field, i.e., root-mean-square (rms) or interval velocities, it is very robust with respect to an erroneous model. Numerical example using synthetic seismic and real Ground-Penetrating-Radar (GPR) data demonstrate the feasibility of the technique and show its usefulness for automatically mapping the diffraction points in a seismic section.

INTRODUCTION

Hydrocarbon reservoirs are commonly located in geological structures that are difficult to image with seismic methods, particularly if a high resolution is desired. These structures include common hydrocarbon traps, such as faults, pinch-outs, unconformities, salt domes, and other structures the size of which is in the order of the wavelength (Trorey, 1970). One common characteristic of wavefields recorded in these kinds of media is a typical signature associated with diffracted waveforms. Thus, this feature is very important for developing methods that are capable of identifying these complex structures.

Because of the importance of these types of structures, several methods for imaging diffractions have been developed in the recent past. The first authors to look into the topic were Landa et al. (1987) and Landa and Keydar (1998), who proposed and refined a detection method related to specific kinematic and dynamic properties of diffracted waves. The latter authors also showed the importance of diffracted energy in order to monitor activity in the subsurface. In the same context, Liu et al. (1997) investigated the influence of large crack regions in a fractured reservoir on the scattered waveforms. Khaidukov et al. (2004) suggested a method based on the focusing of diffracted waves and defocussing of reflected events in order to obtain superresolution images. Using several processing work-flows in the pre-stack data domain, Bansal and Imhof (2005) suggested ways to suppress reflections and to enhance diffractions events. In a similar approach, Moser and Howard (2008) proposed to suppress specular reflections to image diffractions in the depth domain. Most recently, Zhu and Wu (2010) developed a method based on the local image matrix (LIM), which uses an image condition in the local incident and reflection angles for source-receiver pairs to detect diffractions.

In the context of seismic migration processing, diffractions are of particular importance because they carry immediate velocity information. While reflectors are only relocated to the correct position in depth or time, time diffracted energy are collapsed into points, if the velocity model is correct, allowing for direct recognition of incorrect velocities. Using the collapsing energy property of diffractions, Sava et al. (2005) deployed migration velocity analysis (MVA) to estimate velocity models updates based on diffraction fo-

cusing analysis. Based on the same idea, Novais et al. (2006, 2008) and Fomel et al. (2006, 2007) developed a method to extract a lateral varying velocity model for GPR processing by velocity continuation. Reshef and Landa (2009) and Fomel et al. (2007) have shown the effect of a velocity error on post-stack migration.

In this work, we further develop one of the diffraction-detection methods proposed in de Figueiredo et al. (2010). It is based on an amplitude analysis along the elementary diffractions (Tabti et al., 2004). This method does not require any knowledge apart of from the migration velocity field, i.e., root-mean-square (rms) velocities or interval velocities depending on the complexity of the area. It applies pattern recognition to the amplitudes along the diffraction operator. Numerical examples on synthetic and ground penetrating radar (GPR) field data demonstrate the feasibility of the method.

THEORY

The diffraction-operator

Tabti et al. (2004) introduced amplitude analysis along elementary diffractions for Fresnel aperture specification. As illustrated in Figure 1a, the traveltime of an elementary diffraction associated with a reflection point is tangent to the reflection traveltime at the stationary point (location of the specular reflection event). More specifically, in limited bandwidth situations, this tangential point becomes a tangential contact region, which defines the minimum aperture for true-amplitude Kirchhoff migration (Schleicher et al., 1997). Tabti et al. (2004) named it the Fresnel aperture due to its close relationship to the Fresnel zone. For an image point other than a reflector or diffractor, below referred to as "void image point", there is no such region. The traveltime of the elementary diffraction associated with a void image point may cross some reflection events, but will not be tangential to any events (see Figure1b), except for extremely rare coincidental situations.

The diffraction-operator at an image point M is defined as the amplitude vector

$$D(M,\xi) = W(M,\xi)\partial_t U(\xi,t)|_{t=\tau_D(M,\xi)} \tag{1}$$

as a function of position ξ. Here, $U(\xi,t)$ is the seismic data measured at ξ, $\tau_D(M,\xi)$ is the traveltime of the elementary diffraction of M and $W(M,\xi)$ is a weight function. This operator (1) derives from the Kirchhoff depth migration integral (Schleicher et al., 1993), which represents the image $I(M)$ as

$$I(M) = \int\limits_{A_f} d^2\xi\; W(M,\xi)\partial_t U(\xi,t)|_{t=\tau_D(M,\xi)} = \int\limits_{A_f} d^2\xi\; D(M,\xi)\;. \tag{2}$$

Here, A_f is the migration aperture, ideally given by the Fresnel aperture. Integration variable ξ is the horizontal coordinate of the seismic section to be migrated, for instance the midpoint coordinate for a common-offset section or the receiver coordinate for a common-shot section.

Instead of performing the summation, the diffraction operator $D(M,\xi)$ at an image point M collects a single-valued curve (or surface, in the 3D case), defined as a function of the integration variable ξ. Its value at ξ is the amplitude the stack in equation 2 will consider at ξ. This value defines the amplitude of the elementary diffraction measured at ξ. For simplicity, we omit the weight function in the present study, i.e., from now on $W(M,\xi) = 1$. Then, the diffraction operator becomes

$$D(M,\xi) = \partial_t U(\xi,t)|_{t=\tau_D(M,\xi)}. \tag{3}$$

In this paper, we restrict ourselves to a simplified version of the diffractor operator proposed by Tabti et al. (2004), based on the Kirchoff time migration integral for the common-offset configuration. Therefore, the elementary diffraction traveltime $\tau_D(M,\xi)$ is given by (Landa et al., 1987)

$$\tau_D(M,\xi) = \sqrt{\left(\frac{t_0}{2}\right)^2 + \left(\frac{\xi-h}{2v}\right)^2} + \sqrt{\left(\frac{t_0}{2}\right)^2 + \left(\frac{\xi+h}{2v}\right)^2} \tag{4}$$

where t_0 is the zero-offset time for any point in the subsurface, h is the half-offset and v is the medium velocity.

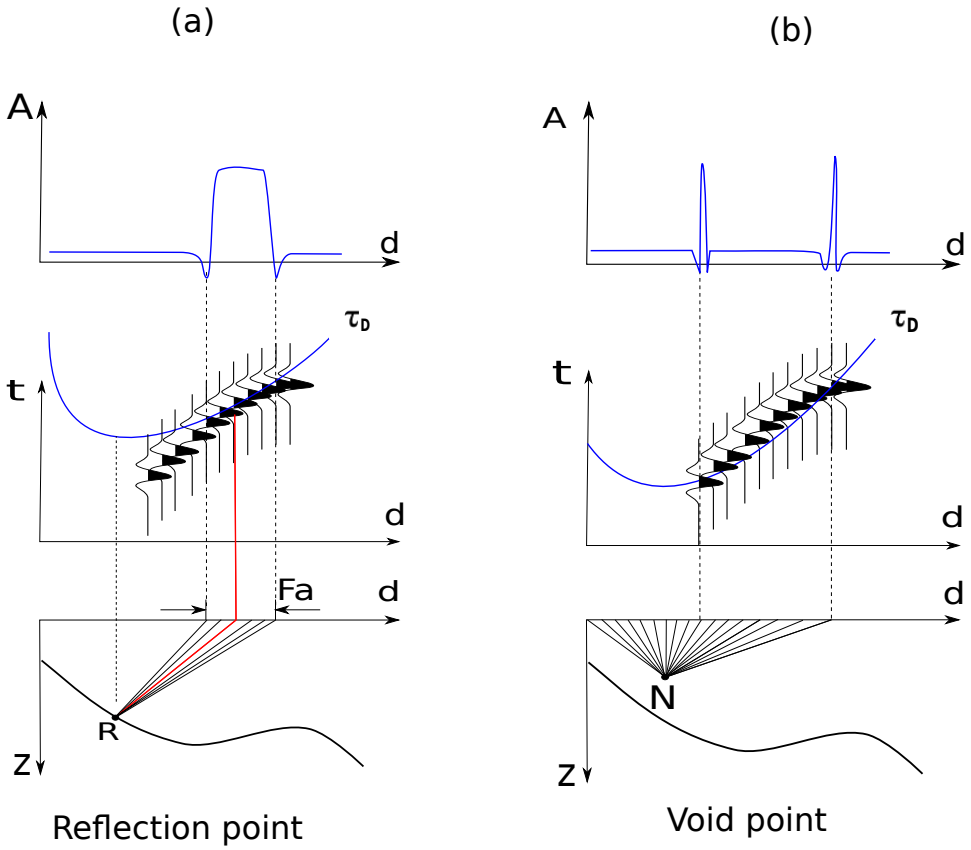


Figure 1: Illustration of the diffraction operator for (a) a reflection point and (b) a void image point. Top: amplitude along the diffraction operator; center: diffraction traveltime and seismic event; bottom: image point and ray family.

Tabti et al. (2004) also noted that in the case of a diffractor point (either a point scatterer or an edge) at M, the associated elementary diffraction time corresponds to the scattered seismic event. The Fresnel aperture then theoretically extends to infinity, regardless of the frequency content of the source pulse. Figure 2 illustrates the diffraction operator at a diffraction point (Figure 2a) and diffraction edge (Figure 2b).

The parts of Figure 2 form the basis of our diffraction imaging algorithm explained in the next section. The main idea of our detection method is to classify every point M in the image domain as a diffractor, reflector or void point by means of the characteristics of its diffraction operator $D(M, \xi)$. The approach consists in a straightforward classification using a well-established pattern-recognition technique called k nearest neighbors (kNN).

Diffraction imaging by pattern recognition

Pattern recognition aims at classifying data (patterns) based either on a priori knowledge or on statistical information extracted from the patterns (Duda and Hart, 1973). Pattern recognition techniques have found applications in various areas, for instance, decision making, inspection of objects, and automatic character recognition (Theodoridis and Koutroumbas, 1999).

The mathematical tool to achieve this goal is called a classifier. Suppose we are faced with the problem of dividing a certain set of patterns into N classes, $w_1, \cdots, w_n$. Let $x_1, \cdots, x_p \in I\!R^n$ be samples of patterns the classes of which are already known, and let $C_i \subseteq \{x_1, \cdots, x_p\}$ be a subset of patterns associated with class w_i such that $C_j \cap C_i = \emptyset$ for $i \neq j$, i.e., there are no patterns that fall into two different classes at the same time. Given an arbitrary pattern x, a classifier aims at associating x with one of the N classes w_i. In this work, we are only concerned with imaging diffractions. Therefore we use two classes ($N = 2$):

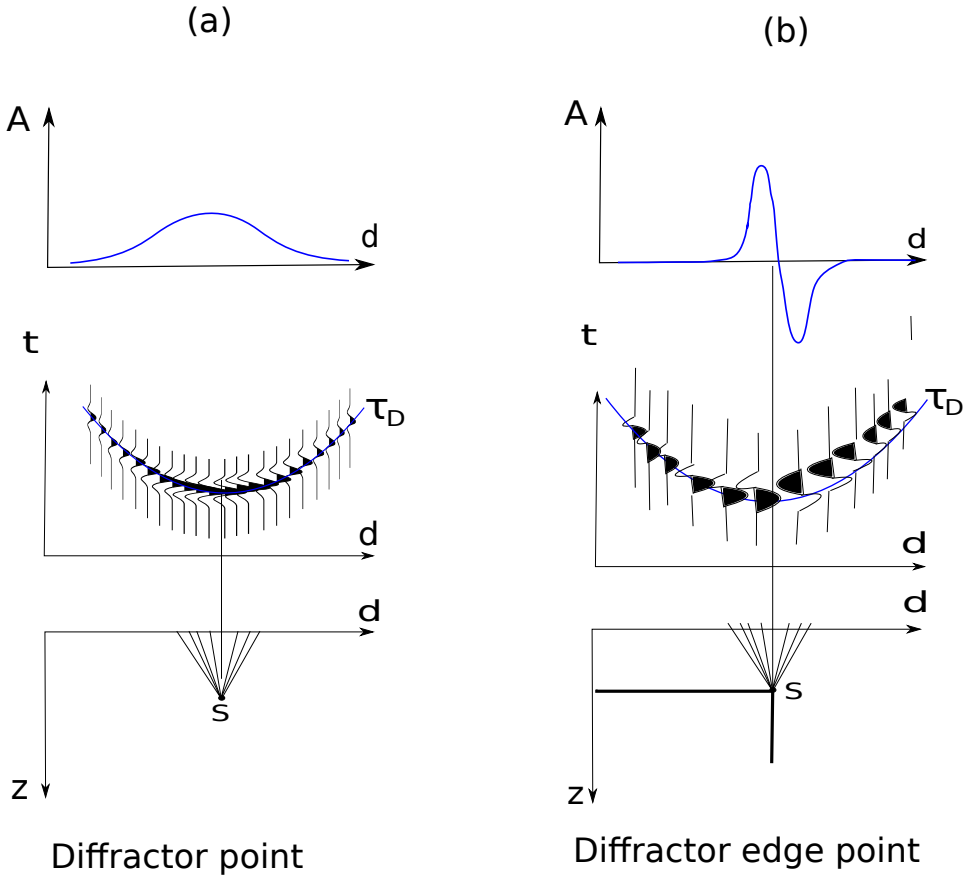


Figure 2: Illustration of the diffraction operator for a (a) diffraction point and (b) diffraction edge region. Top: amplitude along the diffraction operator; center: diffraction traveltime and seismic event; bottom: image point and ray family.

the diffraction class C_0 (that includes point and edge diffractions) and the non-diffraction class C_1 (that includes both reflections and void points). We also restrict ourselves to the so-called k-nearest-neighbor (kNN) classifier, because of its simple implementation.

The kNN classifier is a supervised approach towards pattern classification that employs a certain distance measure and a fixed set of samples in the feature space for which the associated class label is already known (Cover and Hart, 1967; Duda and Hart, 1973). The development of the kNN classifier was inspired by the technique for the estimation of a non-parametric probability density function (PDF) called k-nearest-neighbor density estimation, which is basically a variation of the histogram approximation of an unknown PDF. Moreover, although no assumptions about PDFs need to be made, the strategy used by the kNN model to classify patterns reminds of the well-known Bayes classification rule Duda and Hart (1973); Theodoridis and Koutroumbas (1999).

Let $k \leq p$ be a positive fixed integer and $dist$ a distance measure in $\mathbb{R}^n$. Then, the kNN classification process is given by the following rules:

- Find the k nearest neighbors of x in the set $\{x_1, \cdots, x_p\}$ in terms of their distances $dist(x, x_i)$, for $i = 1, \cdots, p$. Let the symbols $\tilde{x}_1, \cdots, \tilde{x}_k \in \{x_1, \cdots, x_p\}$, with $\tilde{x}_i \neq \tilde{x}_j$ for $i \neq j$, denote those k nearest neighbors.
- Identify the number k_i of patterns $\tilde{x}_i$ among these k nearest neighbors that belong to class w_i for $i = 1, \cdots, N$.
- Assign x to the class w_j with the maximum number $k_j = \max_{i=1,\ldots,N} k_i$ of samples.

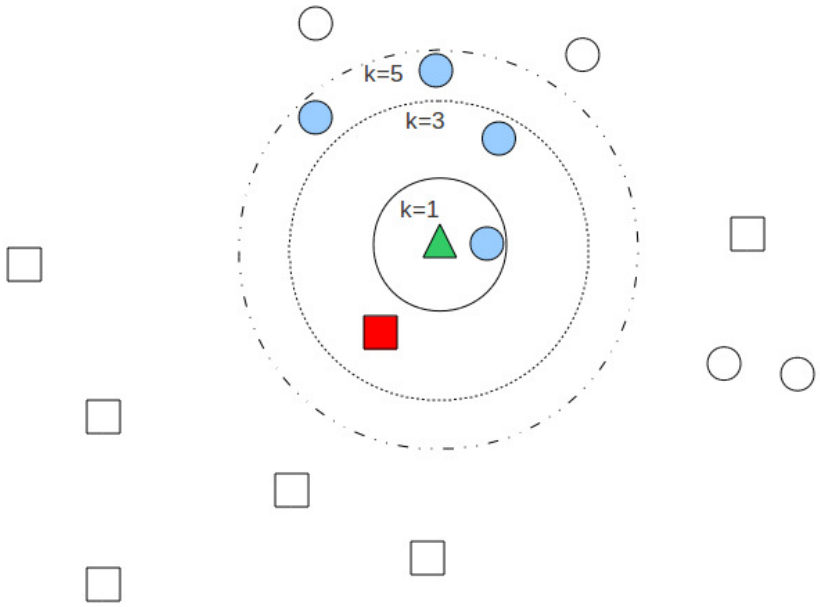


Figure 3: Graphical illustration of kNN classification. For $k = 1$ (one nearest neighbor) the element $\triangle$ is classified as $\circ$. For $k = 3$ as well as $k = 5$ the $\triangle$ is also classified as $\circ$ because most of its neighbors are $\circ$.

As an example, Figure 3 visualizes a set of patterns that are known to belong to either one of the classes 'circle' ($\circ$) or 'square' ($\square$). Suppose that our goal is to assign the pattern marked by $\triangle$ to one of two classes using the kNN classifier. Generally it is more useful to use odd numbers for k because they don't allow ties. In our example, for any choice $k = 1, 3, 5$, the kNN classifier maps the $\triangle$ pattern into class $\circ$ because the majority of nearest neighbors belongs to this class (see Figure 3).

Note that a classification problem consists of finding a map ϕ from the pattern space to the set of class labels. As a visual example of the map ϕ produced by the kNN model, we use a well-known synthetic dataset that consists of data samples from two classes (Ripley, 1996). Each sample has two features that provide the axes of the data space in Figure 4. The data are divided into a training set and a test set consisting of 250 and 1000 samples, respectively, with the same number of samples belonging to each of the two classes. Thus, we obtain a binary classification problem in $I\!R^2$. Figure 4 shows the decision surfaces produced by the kNN model for the choices $k = 1, 3, 5$. The patterns marked by $\square$ or $\circ$ symbols represent the training data. The kNN classifier divides the model space into a gray region, where all patterns are classified as "squares' and a white region, where all patterns are classified as "circles"

Figure 4 reveals that some training patterns are misclassified if k equals 3 or 5 in this example because there are outliers in the trainig data. No misclassifications of training data can occur for $k = 1$.

The results of a kNN model depend on the choice of the distance measure and the number k of nearest neighbors. Techniques to select an appropriate parameter k such as cross-validation can be employed (Devijver and Kittler, 1987). Usually, Euclidean distance is used as the distance metric. As with all supervised models, the accuracy of the kNN classifier also depends on the training data. If non-representative samples of classes are used as training data, the performance of kNN classification can be severely degraded. In the simulations described in the Numerical Tests section, we have employed a value of $k = 1$ and the Euclidean distance measure.

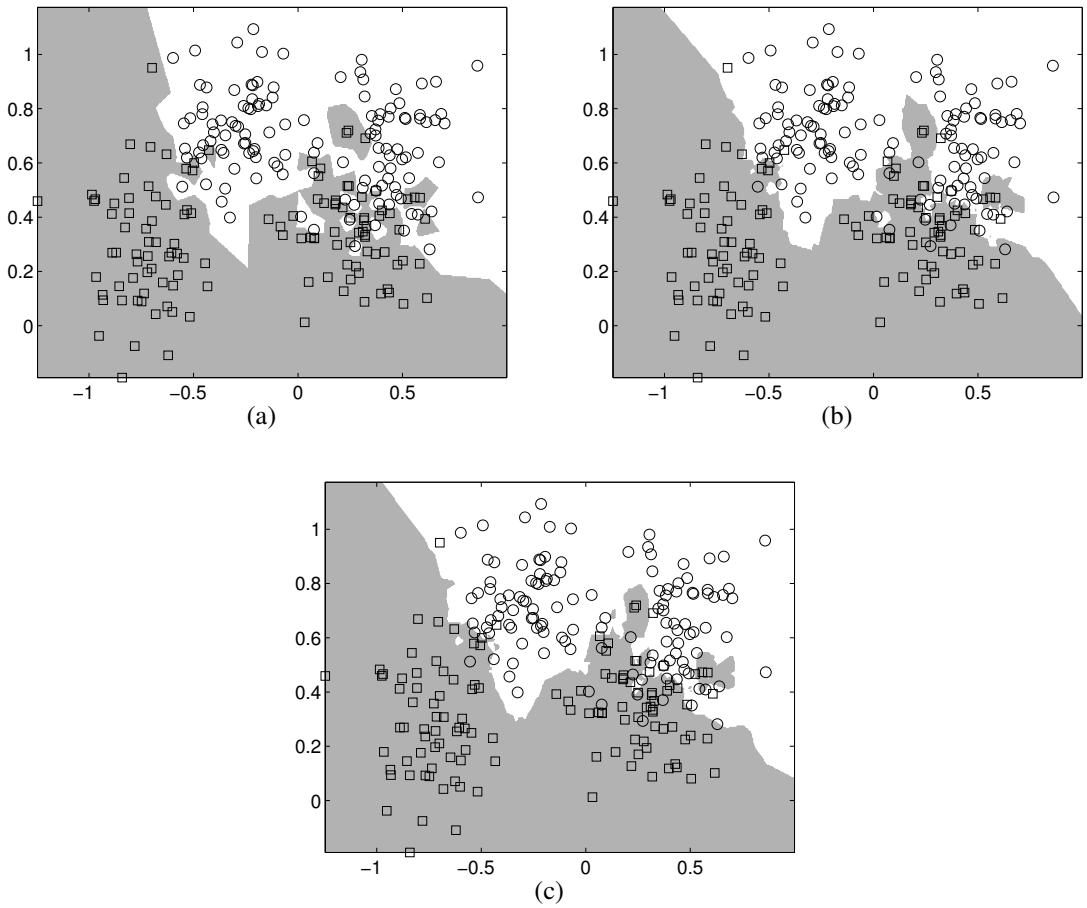


Figure 4: The training patterns of the Ripley (1996) dataset and the corresponding level surfaces produced by kNN for (a) $k = 1$, (b) $k = 3$, and (c) $k = 5$.

RESULTS

In this section, we present numerical result of our diffraction-imaging technique on one complex synthetic data set and one real GPR data set.

Synthetic example: Model with thirteen diffractors

In continuation of the tests presented in de Figueiredo et al. (2010), we evaluate the application this method to the model consisting of 13 diffraction points (see Figure 5). For completeness, we include the basic results for this model already reported in de Figueiredo et al. (2010). There are 4 isolated point scatterers at depths 0.3 km and 0.5 km, 4 tips of reverse faults at 0.7 km and 1.0 km, and 5 tips of normal faults at 2.0 km and 2.3 km depth. They are buried in a constant-velocity background with $v = 2000$ m/s covering an extension of 7 km in the x direction. The maximum depth of the model is 4 km.

We simulated the zero-offset data by Kirchhoff modeling, using 700 source-receiver pairs spaced by 10 m with a Ricker wavelet of dominant frequency 15 Hz. Additionally, we added random noise with a signal-to-noise ratio of 50. Then we normalized the data by their envelope in order to increase the diffraction signature. The normalized synthetic zero-offset data are depicted in Figure 6.

Figure 7 shows the Kirchhoff time-migrated image. While the edge diffractors are clearly identifiable, the point diffractors at the top of the model are rather weak and will certainly be invisible in a noisy data set.

Figure 8 shows the result of the diffraction imaging approach using pattern recognition. For the training

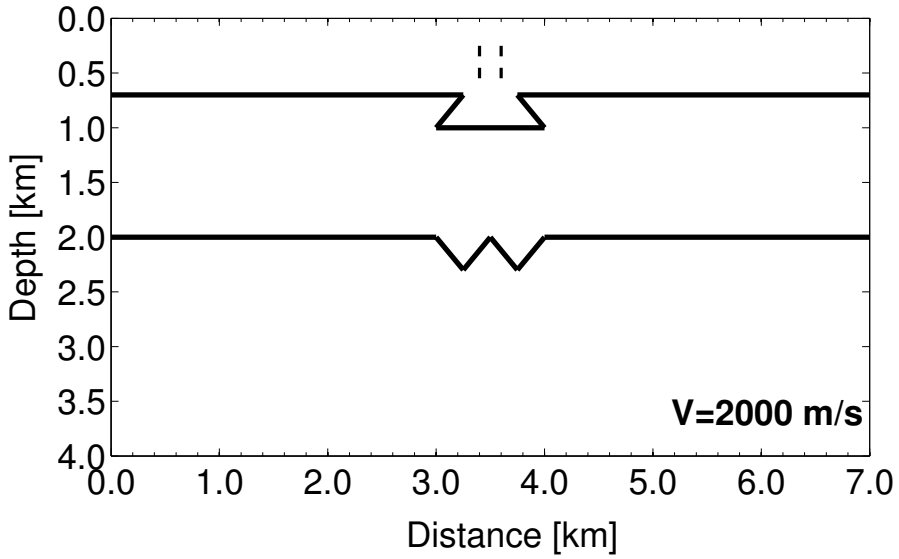


Figure 5: Model with constant velocity of 2000 m/s and with thirteen diffractors. Four of them are scattering points and other nine are the edge diffractors.

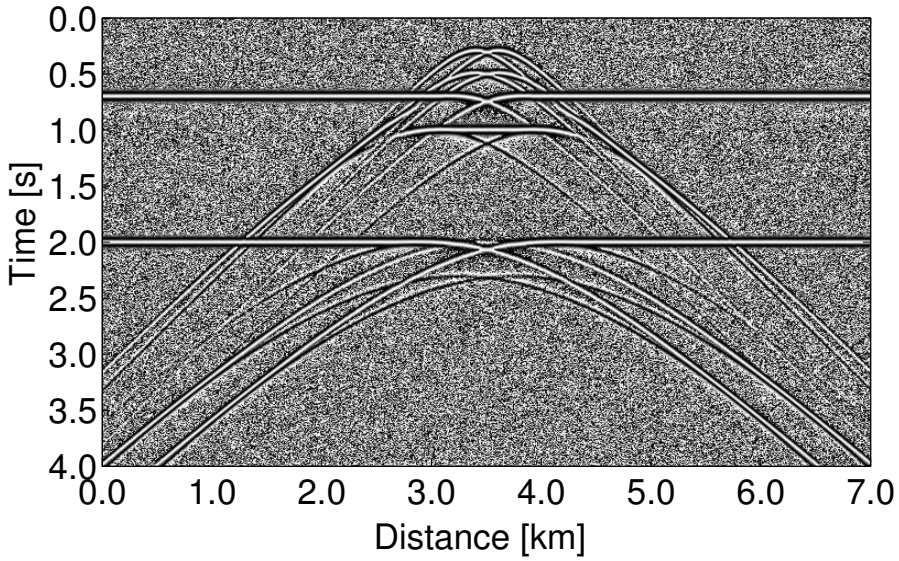


Figure 6: The normalized data show increased diffraction amplitudes.

of the kNN classifier we used the simple model discussed in de Figueiredo et al. (2010). It is important to note that no additional training was performed using the data of this model. Note that the kNN classifier correctly identified all 13 diffraction points and did not produce any false positives.

It is important to note that the location of the diffractions generated by the normal faults have lower resolution when compared to the others (some smearing can be seen). Other numerical tests (not shown here) indicate that this resolution loss is systematic and can be related to the dip of the faults.

Velocity dependence. To verify the robustness of our method with respect to the accuracy of the velocity model, we applied it using a velocity error of $\pm 3\%$, i.e., velocities ± 60 m/s different from the correct veloc-

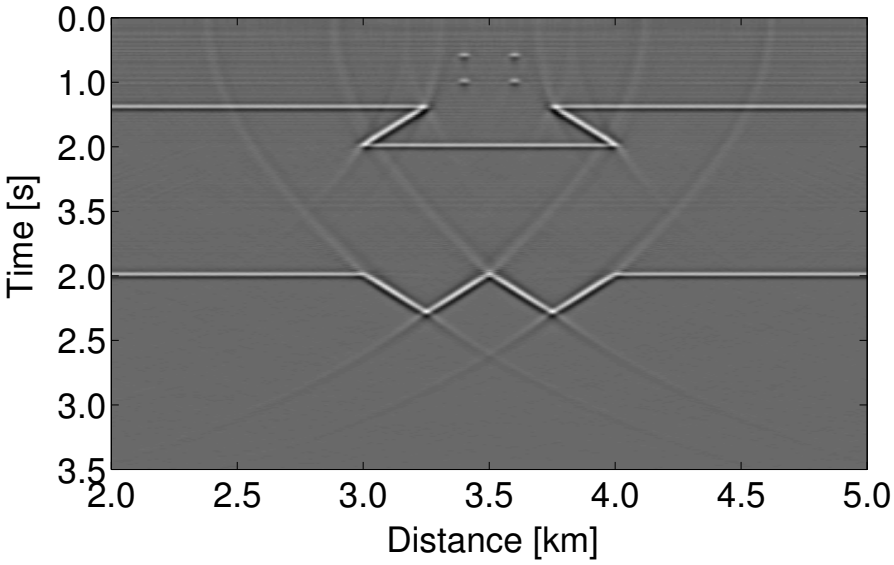


Figure 7: Time-migrated image of the data shown in Figure 6a.

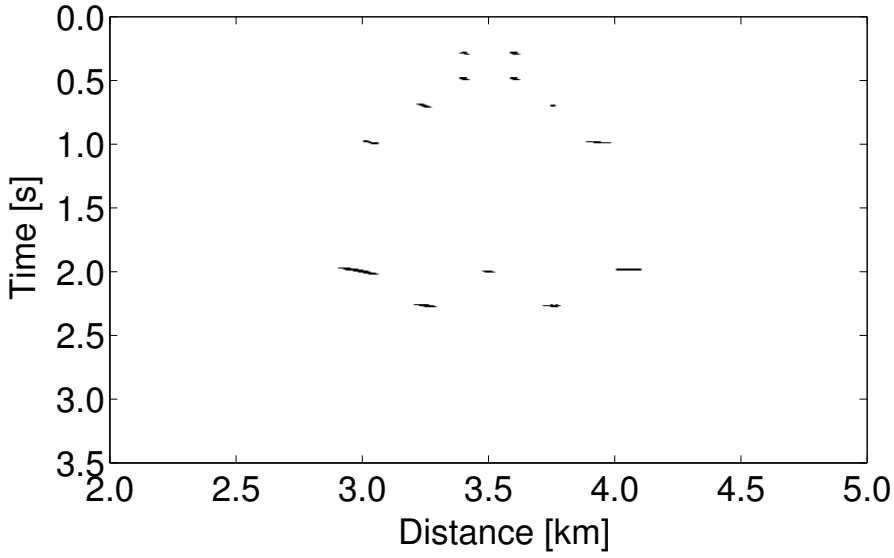


Figure 8: Diffraction locations by kNN classifier. We can see that all 13 diffraction points were well indentified automatically.

ity of 2000 m/s. When migrating the data shown in Figure 6a with these incorrect velocities (see Figure 9) the migrated sections show upward (higher velocity) or downward (lower velocity) smiles. This makes it even harder to correctly interpret the diffractions. On the other hand, the kNN procedure (Figure 10) was able to localize all diffraction points and did not generate false positives.

This capability to detect diffractions even using approximate velocities demonstrates that our method is useful in practice, where usually only velocity estimates are available. Although the lateral position of the diffractors found by the method is not exact, the fact that diffraction points where detected is already very useful. It not only points toward an error in the velocity model, but also indicates that some complex structure is present in the imaged region that cannot be correctly visualized in the migrated section.

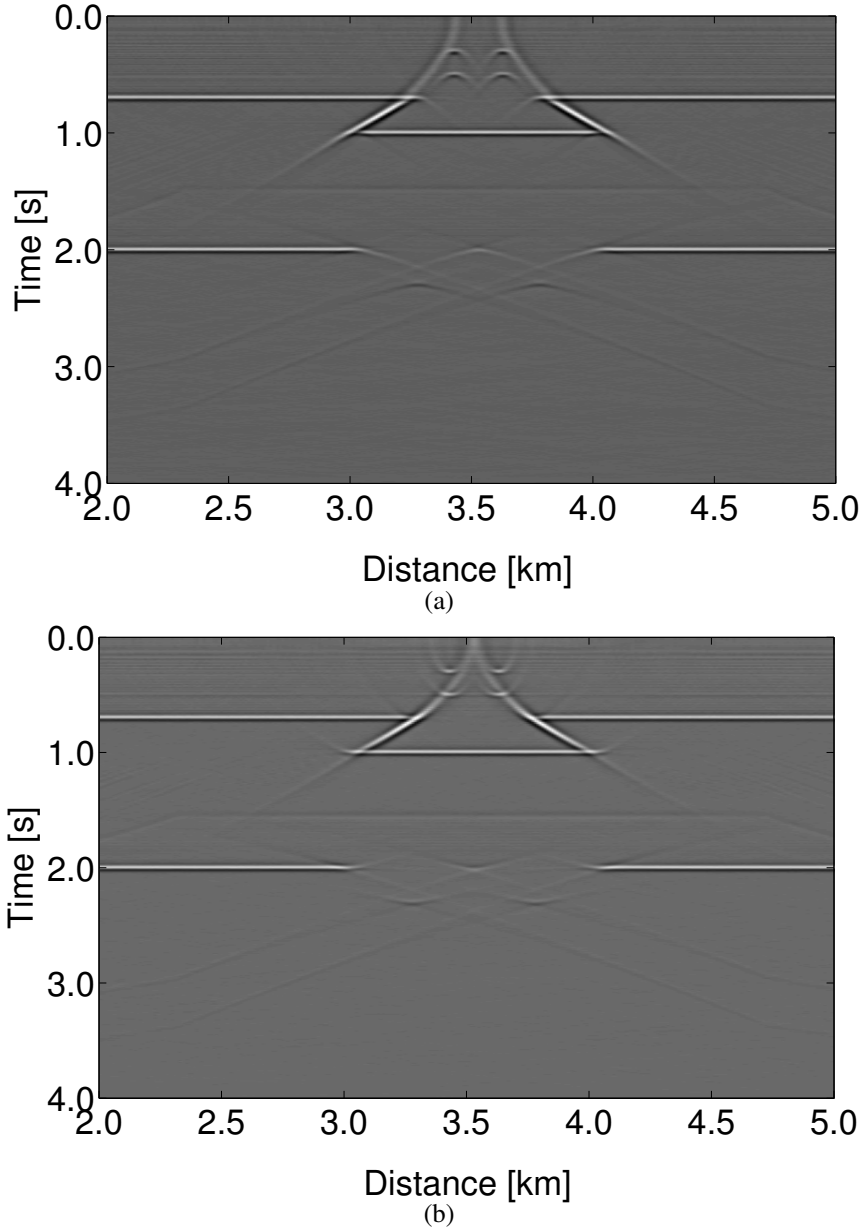


Figure 9: Time-migrated image of the data shown in Figure 6a with velocities: (a) $v = 1940$ m/s and (b) $v = 2060$ m/s. The diffractions are hardly identifiable.

Real data example: Ground Penetrating Radar dataset

For a more meaningful test, we applied this method to GPR data. The data set is from a survey conducted over four metal drain pipes crossing under a road at the University of Houston Coastal Center, located in La Marque, Texas, United States (Figure 11a). Since the survey line was perpendicular to the direction of the pipes, prominent diffractions from the pipes occur. Because of the high reflectivity and attenuation of metal, GPR reflections from the pipes occur only from the exterior of the pipes and not from inside de pipes (Zeng and McMechan, 1997). Since the pipes are circular, the reflections behave like true diffraction events.

Figure 11b shows the 250 MHz GPR profile acquired. The distance between the transmitter and reciever antennas was 0.28 m (half-offset 0.14 m), and the interval between traces was 0.05 m. A total of 445 traces

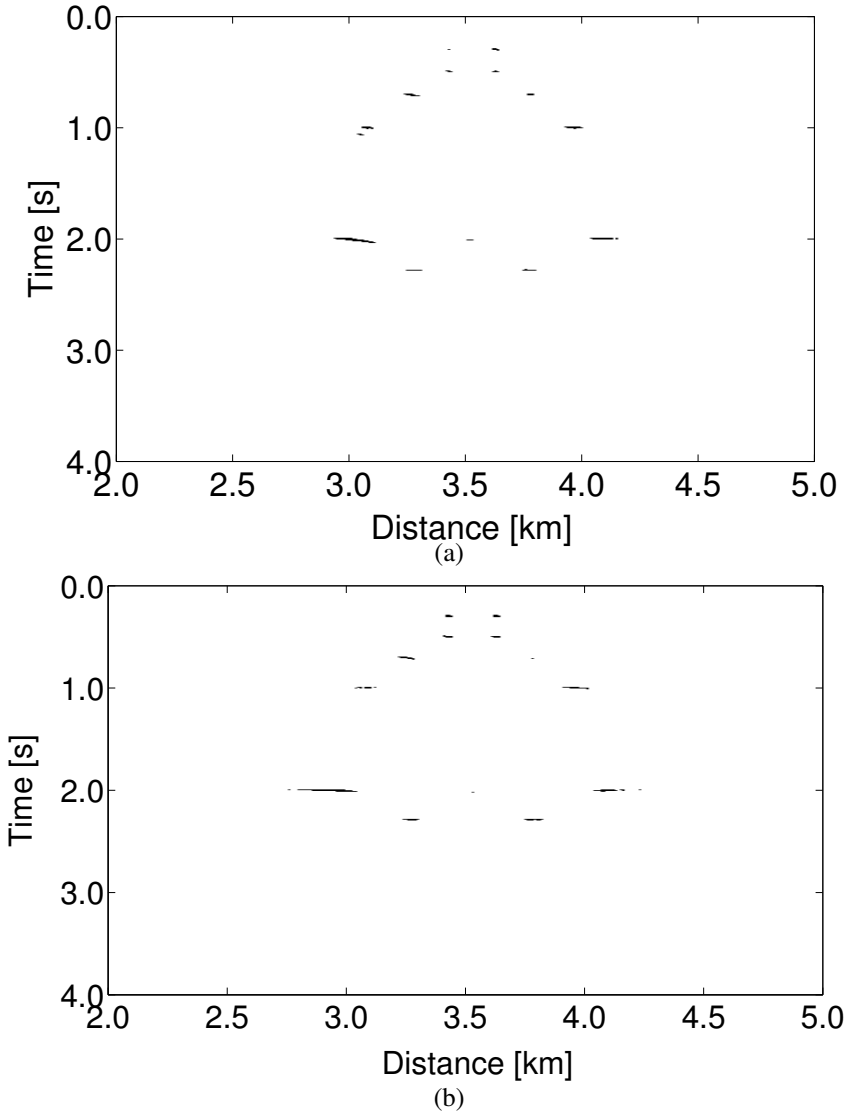


Figure 10: Diffraction locations by kNN classifier from data shown in Figure 6a with velocities: (a) $v = 1940$ m/s and (b) $v = 2060$ m/s. All diffractions are correctly identified.

were collected along the 22.25 m survey line. The length of the time window was 99 ns and the number of samples per trace was 247, resulting in a time sampling rate of 0.4 ns.

Figure 12a depicts the diffraction panel for the migrated trace at $x = 10$ m obtained from the data (see Figure 11b) for a migration velocity of 0.143 m/ns (0.143×10^8 m/s). As we can see, this panel exhibits a straight horizontal line indicating that the correct velocity was used. This velocity is within the expected range for soil mixtures. The corresponding Kirchhoff time-migrated section is depicted in Figure 12b. The migration collapsed all diffractions at the centers of the respective pipes at about 14 ns vertical time.

The normalized data is shown in Figure 13a, and Figure 13b shows the result of the diffraction imaging approach using pattern recognition with the kNN classifier trained on the first synthetic data set.

Figure 13b reveals that the application of the kNN classifier to real GPR field data was successful. Comparing the migrated section (see Figure 12b) with the classified section (see Figure 13b) we can see that the kNN classifier did not produce any false positives and the four diffractions were clearly identified.

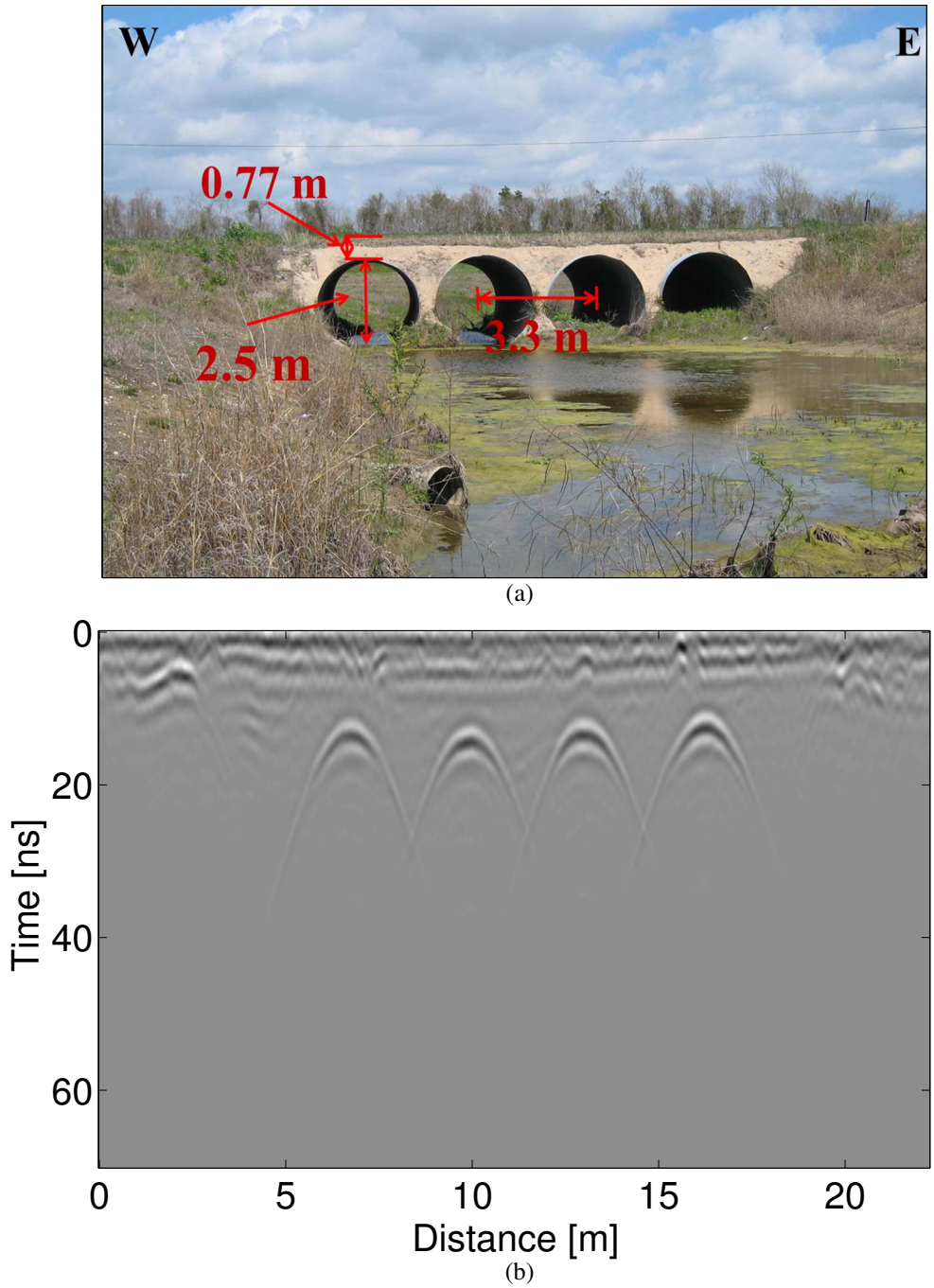
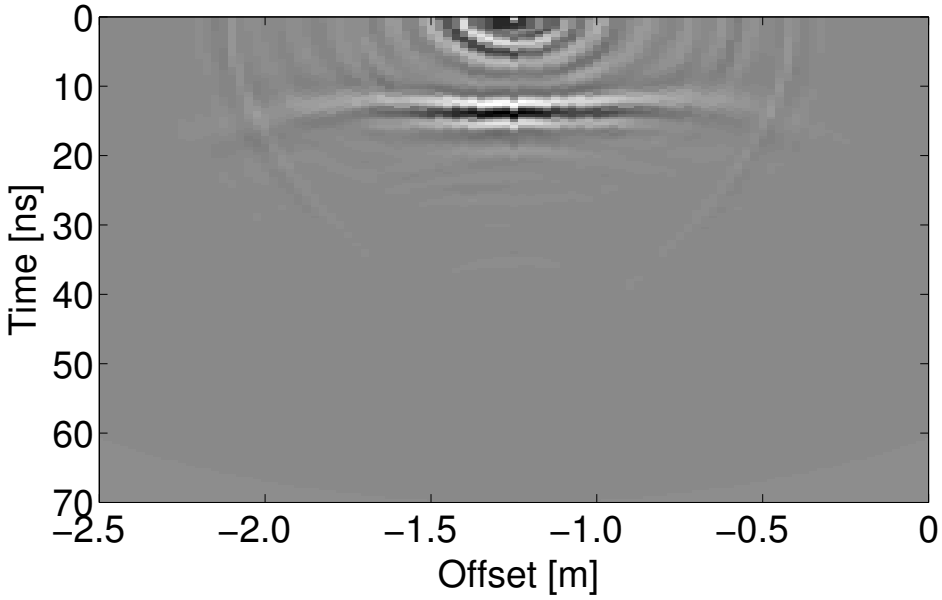


Figure 11: (a) Metal pipes located in La Marque, Houston, Texas, United States. (b) 250 MHz GPR profile showing diffractions from the pipes.

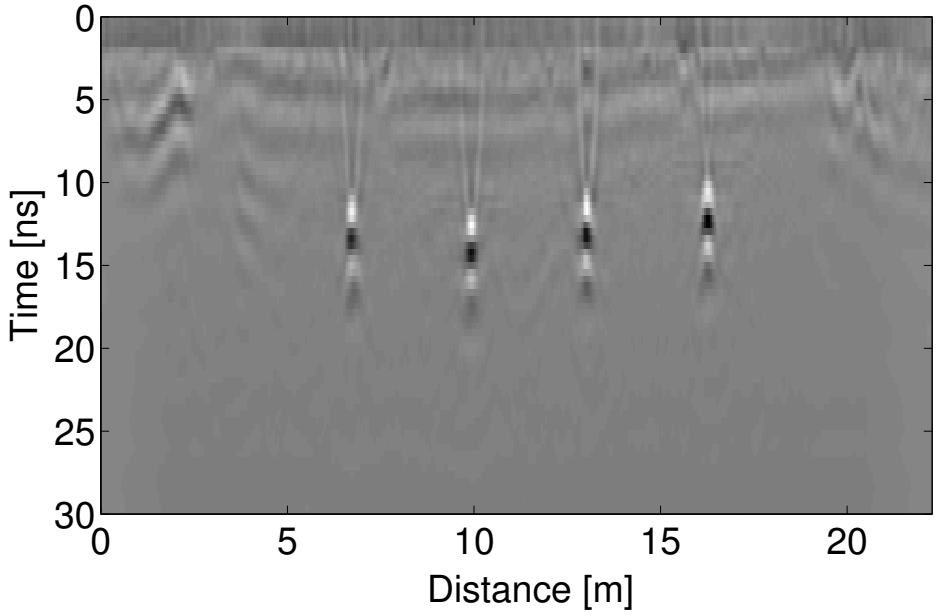
CONCLUSIONS

In this paper, we have discussed a new method for the automatic identification and imaging of diffractor points. For this purpose, we have introduced a new application of pattern recognition in seismic processing, extending its uses beyond the routinely and well established practice in seismic reservoir characterization. The new method open news possibilities to obtain high-resolution images.

The method uses the diffraction operator as a tool for diffraction imaging. It consists in a straight-



(a)



(b)

Figure 12: (a) Time-migrated profile. (b) Diffraction panel at $x = 10$ m

forward application of a pattern recognition technique to identify and distinguish diffraction events from reflection events and noise areas by their amplitude pattern. After training with selected diffraction operators pertaining to a synthetic data set from a very simple synthetic model, the kNN classifier was able to correctly detect all diffraction points in a considerably more complicated model and in a real GPR dataset, not missing a single point and not creating a single false positive.

The satisfactory results obtained in our data examples have shown that the kNN algorithm is a robust method to achieve the goal of identification and localization of diffractor points. As demonstrated in our numerical examples, the method works well even if the velocity model is not exact. A 3.3% velocity error did not affect the diffractor detection at all. Future investigations will be needed to test the algorithm in more noisy environments and for data with interfering events.

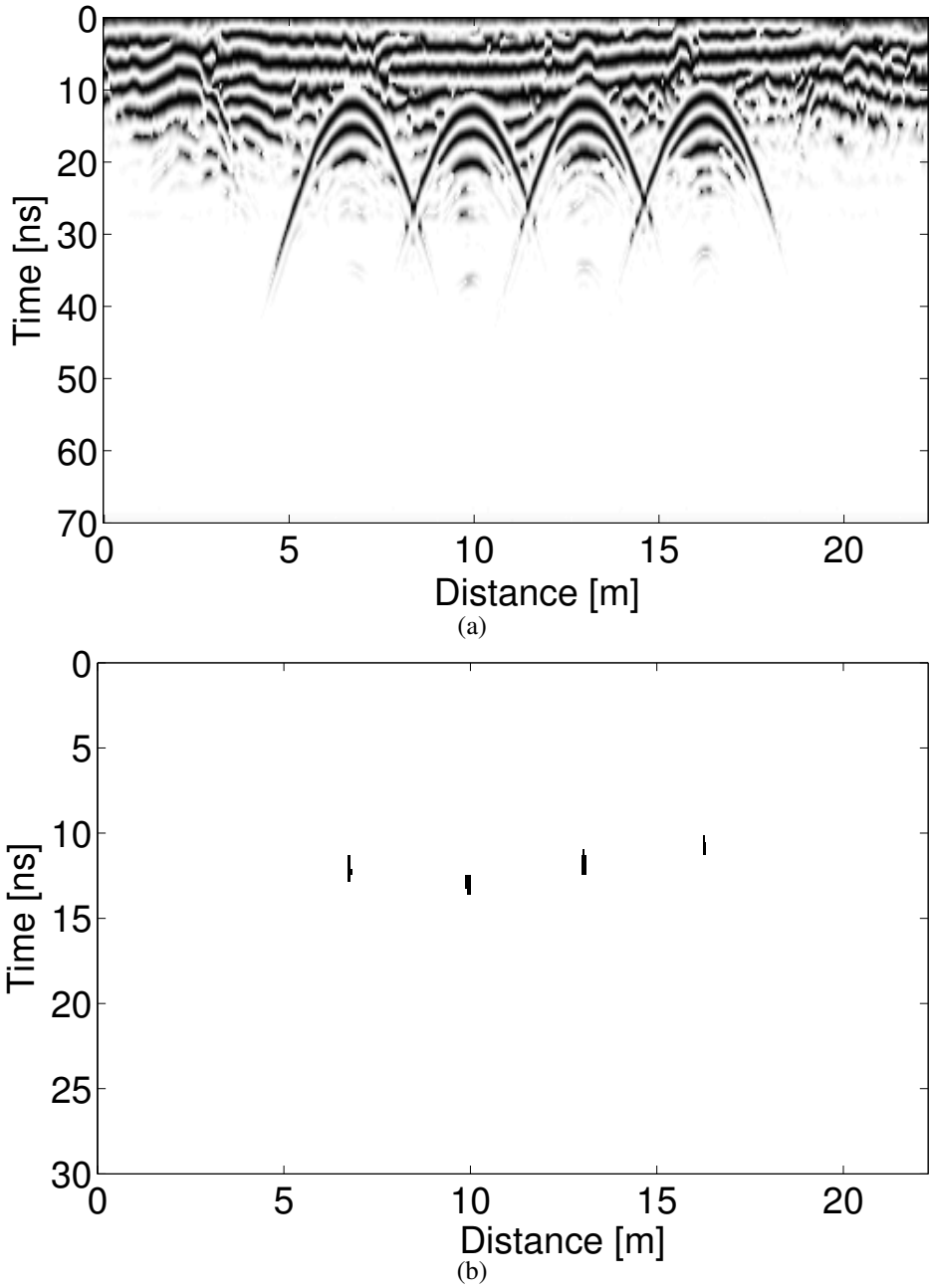


Figure 13: (a) Normalized GPR data. (b) Four diffraction locations identified by the kNN classifier.

ACKNOWLEDGEMENTS

This work was kindly supported by the Brazilian agencies CAPES, FINEP, and CNPq, as well as Petrobras and the sponsors of the *Wave Inversion Technology (WIT) Consortium*.

REFERENCES

Bansal, R. and Imhof, M. G. (2005). Diffraction enhancement in prestack seismic data. *Geophysics*, 70(3):V73–V79.

Cover, T. M. and Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Infor-*

mation Theory, 13(1):21–27.

de Figueiredo, J. J. S., Oliveira, F., Esmi, E., Freitas, L., Novais, A., and Schleicher, J. (2010). Diffraction imaging based on the diffraction operator. *14th Annual WIT Report*, 14:42–55.

Devijver, P. and Kittler, J. (1987). *Pattern recognition: A statistical approach*. Prentice Hall International, Englewood Cliffs, NJ.

Duda, R. and Hart, P. (1973). *Pattern Classification and Scene Analysis*. John Wiley & Sons, New York, NY.

Fomel, S., Landa, E., and Taner, M. T. (2006). Post-stack velocity analysis by separation and imaging of seismic diffractions. In *76th Ann. Internat. Meeting, SEG, Expanded Abstracts*, pages 2559–2563.

Fomel, S., Landa, E., and Taner, M. T. (2007). Poststack velocity analysis by separation and imaging of seismic diffractions. *Geophysics*, 72(6):U89–U94.

Khaidukov, V., Landa, E., and Moser, T. J. (2004). Diffraction imaging by focusing-defocusing: An outlook on seismic superresolution. *Geophysics*, 69(6):1478–1490.

Landa, E. and Keydar, S. (1998). Seismic monitoring of diffraction images for detection of local heterogeneities. *Geophysics*, 63:1093–1100.

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

Liu, E., Crampin, S., and Hudson, J. A. (1997). Diffraction of seismic waves by cracks with application to hydraulic fracturing. *Geophysics*, 62(1):253–265.

Moser, T. and Howard, C. (2008). Diffraction imagining in the depth. *Geophysical Prospecting*, 56.

Novais, A., Costa, J., and Schleicher, J. (2006). Velocity determination by image-wave remigration. In *68th Ann. Internat. Mtg., EAGE*, pages P157:1–4.

Novais, A., Costa, J., and Schleicher, J. (2008). GPR velocity determination by image-wave remigration. *Journal of Applied Geophysics*, 65(2):65–72.

Reshef, M. and Landa, E. (2009). Post-stack velocity analysis in the dip-angle domain using diffractions. *Geophysical Prospecting*, 57(5):811–821.

Ripley, B. D. (1996). *Datasets for Pattern Recognition and Neural Networks*. Cambridge University Press.

Sava, P. C., Biondi, B., and Etgen, J. (2005). Wave-equation migration velocity analysis by focusing diffractions and reflections. *Geophysics*, 70(3):U19–U27.

Schleicher, J., Hubral, P., Tygel, M., and Jaya, M. S. (1997). Minimum apertures and Fresnel zones in migration and demigration. *Geophysics*, 62(2):183–194.

Schleicher, J., Tygel, M., and Hubral, P. (1993). 3-D true-amplitude finite-offset migration. *Geophysics*, 58(8):1112–1126.

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

Theodoridis, S. and Koutroumbas, K. (1999). *Pattern Recognition*. Academic Press.

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

Zeng, X. and McMechan, G. A. (1997). GPR characterization of buried tanks and pipes. *Geophysics*, 62(3):797–806.

Zhu, X. and Wu, R. (2010). Imaging diffraction points using the local image matrices generated in prestack migration. *Geophysics*, 75:S1–S9.