Tu_R06_06

Automatic Quality Control of Denoise Processes Using Support-Vector Machine Classifier

M. Bekara[1]*

[1] Petroleum Geo-Services

Summary

Noise attenuation is a key processing step in a typical seismic imaging flow. Assessing the quality of the denoise process to ensure that noise has been sufficiently attenuated without signal distortion is important. With the increase in the volume of seismic data, the QC is becoming a bottleneck for many processing project. This abstract proposes a machine learning solution that can automate the QC. Given the availability of some training seismic data lines that have been through optimal, mild and harsh filtering, the task of identifying where in the data the filtering is sub-optimal is similar to a supervised classification problem. Attributes are computed from the training lines and are then used to train a support vector machine (SVM) classifier. The trained classifier is used to find the filtering class for any seismic ensemble in the survey leading to a fast QC report. The solution is tested on a full-scale processing project to QC the denoise step prior to wavefield separation in marine seismic data. Tests show encouraging results and the solution was able to predict locations in the data with clear residual noise. False positives are an issue but their rate can be reduced by using informative attributes.

Introduction

Noise attenuation is an important and a recurrent step in a typical seismic data processing sequence. The purpose of noise attenuation is to remove unwanted seismic energy that can negatively interfere with a given seismic processing step or can reduce the resolution of the final seismic image. For example, the application of a de-bubbling and designature filter on a marine seismic data with moderate swell noise will cause visible frequency artefacts in the output. Therefore, adequately removing swell noise is a prerequisite before such process. As for every important step, performing quality control (QC) after each noise attenuation process is essential to make sure that there is no residual noise left in the data, but also and more importantly, to ensure that there is no signal leakage.
The QC process takes a considerable amount of resources in any seismic processing project. Current QC practices are done by a geophysicist and relay primarily on the visual inspection of some seismic products such as (i) selected seismic gathers before and after the filtering, (ii) 2D/3D stacks before and after the filtering and (iii) maps of generic seismic attributes (e.g., RMS amplitudes, coherency slices) computed before and after the filtering. The QC process can be time consuming for large datasets. Sampled inspection of the seismic is not conclusive and the geophysicist rely more and more on the assessment of global attribute maps that can give an overall picture of the performance of the filtering process. However, this picture is compressed and local problems can be easily missed. Simultaneous visual assessment of multiple attributes is difficult, empirical and subjective. It is evident that improving the efficiency and the turnaround of seismic processing projects can be achieved by automating the QC process.

This abstract proposes a framework that can help the geophysicist to perform a quick QC for a denoise processing production. The proposed framework will generate a QC report that will highlight the areas in the data that need further seismic inspection. This paper expands on previous work by the author (Spanos et al., 2013), where a non-supervised, outlier detection approach was used. Here, we propose to use a supervised learning approach with training data to build an automatic classifier to predict the type of the filtering (mild, optimal or harsh). The framework is tested on a full scale production to QC the results of noise attenuation prior to wavefiled separation for towed streamer dual-sensor marine data.

Methodology

The proposed framework is based on the concept of supervised classification and consists of the following building blocks:

1. ***Labelled training data***

The denoise parameterization are often optimised, through testing, on a subset of data called the test lines. Here, we assume that in addition to the optimal results on the test lines, we have the results of a mild and a harsh parameterization on the same test lines to construct training data for the case of residual noise and signal leakage. Figure 1-a shows a sample shot gather from one test line going through the different type of filtering to attenuate swell noise. The difference between optimal (Figure 1-b) and mild filtering (Figure1-c) is very subtle for this shot (and also throughout the test lines) and this corresponds to a typical mild classification by the user. For the harsh filtering, the output looks cleaner (Figure 1-d) and signal leakage is visible in the difference section (Figure 1-g). It is very important that the mild and harsh filtering should correspond to a typical mild and harsh filtering and not excessively mild and harsh in order to construct a reliable training dataset.

2. ***Attributes computation***

Ensemble based (e.g. shot) statistical attributes that measure the level of similarity between the output of the filtering and the difference are computed from the seismic samples at a targeted time gate for each of the three types of filtering considered above (optimal, mild and harsh). In the case of the optimal filtering, these attributes will not show any similarity (i.e. correlation) between the output and the difference. This assumption comes from the fact that the signal (i.e. output) have nothing in common with the noise (residual) in the case of an ideal filtering. When there is signal leakage or residual noise, the level of similarity will increase as some noise is also present in the output or some signal is also present in the difference and this will be picked up by the attributes. The attributes are multi-dimensional and form a cloud in the attribute space when computed for all the ensembles in the training lines.

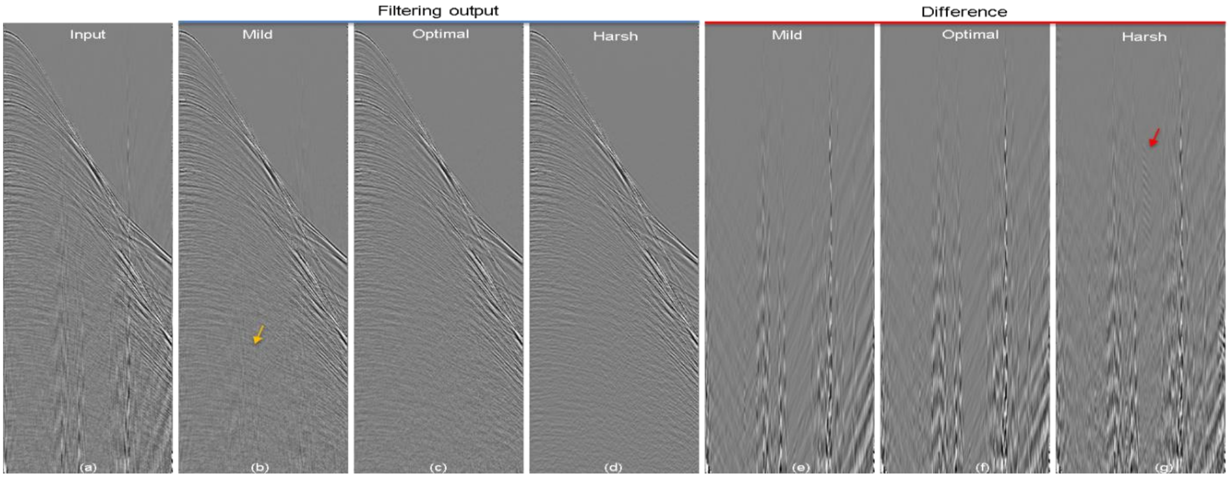


***Figure 1** Example of a single marine shot gather going through mild, optimal and harsh filtering to construct the training data. Residual noise and signal leakage are indicated by arrows. Difference is multiplied by 3.*

Examples of attributes include Pearson cross-correlation, Kendell cross-correlation and mutual information to name few (Chen, 2005). The cross-plots of five different attributes computed from three test lines are shown overlaid for the three types of filtering in Figure 2-a. A tri-colour code is adopted for the display (mild = blue, optimal = green and harsh = red). Some attributes (e.g., TLAMBD, in Figure 2-a) show a good level of visual separation between the different types of filtering, particularly the harsh one. The clusters of attributes for the mild and the optimal filtering are close as they reflect the observation made earlier about the subtle differences between the two types of filtering.

3. *Feature extraction*

There will be always a hidden correlation between the individual attributes due to their common origin. Their dimension can also be extremely large, making the sub-sequent classification problem difficult. The task of de-correlating the attributes to extract useful structures in them is called feature extraction. It is a mapping that transforms each vector of attributes into, an optionally lower dimensional, vector of features. Often the features tend to have a better cluster-discrimination power compared to the attributes. Key linear feature extraction procedures are principle component analysis (PCA) and independent component analysis (ICA) (Hyvärinen et al, 2001). Figure 2-b shows the cluster of features after applying non-linear mapping followed by PCA on the cluster of attributes in Figure 2-a.

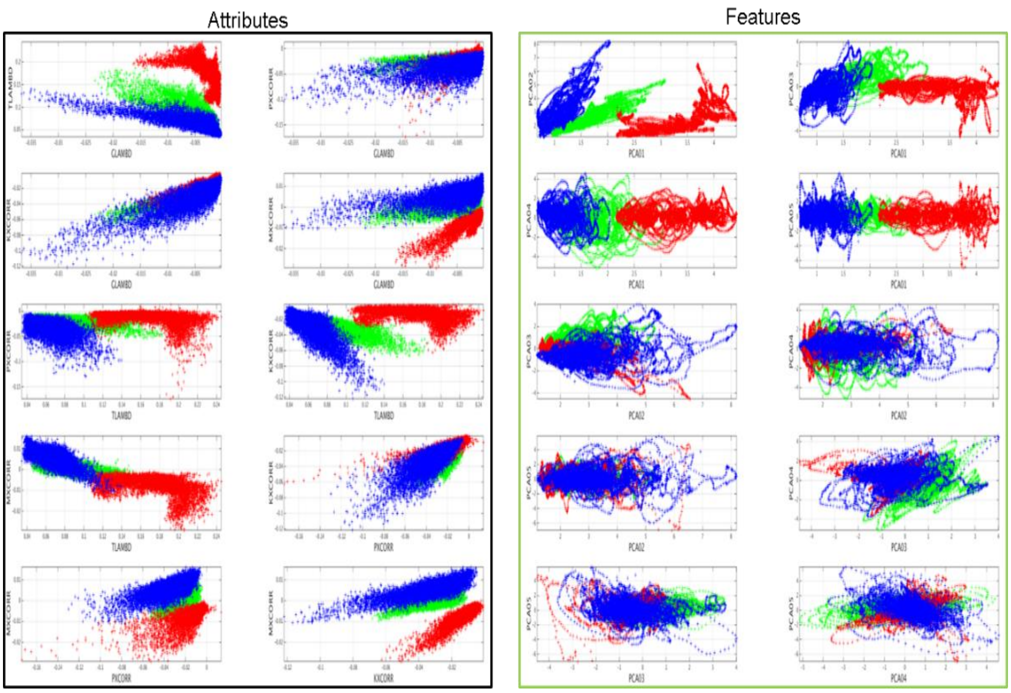


***Figure 2** Cross-plot of 5 attributes and their mapped features computed from three tests lines.*

One can see that the visual separation between clusters of the different filtering have improved for the dominant PCAs and the higher order PCAs are mixed up as they don't convoy any information. They can be ignored reducing the dimension of the input that will be fed to the classifier for the training phase.

4. *Training the classifier*

A supervised classification algorithm based on support vector machines (SVM) (Cristianini et al., 2000) is constructed using the training data (features + filtering type). Building a machine classifier results in partitioning the feature space into regions called decision sub-spaces (DS) that correspond respectively to optimal, mild and harsh filtering. To use the machine classifier to predict the class-type of any new single vector of attributes, it is simply a matter of first mapping the attribute into the feature space and then identifying in which DS it falls. For ease of visualisation, Figure 3-a shows a cross-plot of a 2D PCA based features computed from the training dataset for the three types of filtering. Only the first two dominant PCAs are selected. Figure 3-b shows the partition of the feature space into DSs using SVM with polynomial kernels. The boundaries of the DSs carves the clusters of the features for the different filters. They are optimally constructed to minimize the miss-classification rate. It is worthwhile highlighting that during the design of this framework we found that the attributes are the most important building block. The more informative the attributes, the more robust the outcome of the prediction.

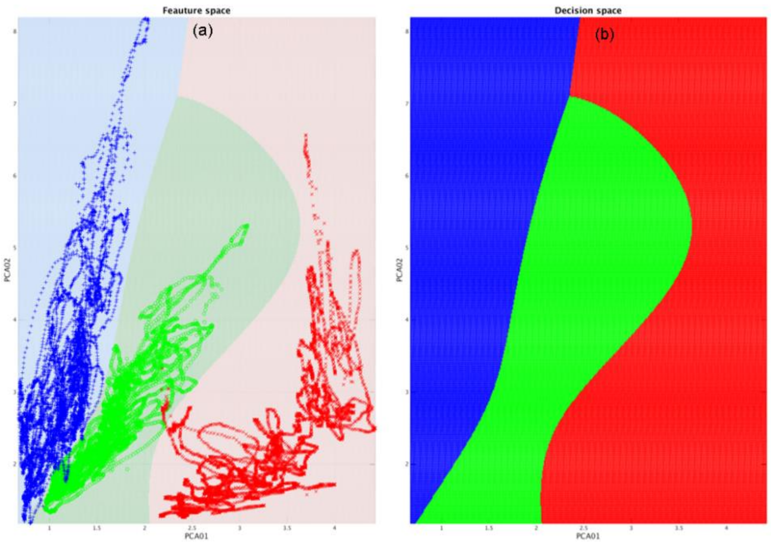


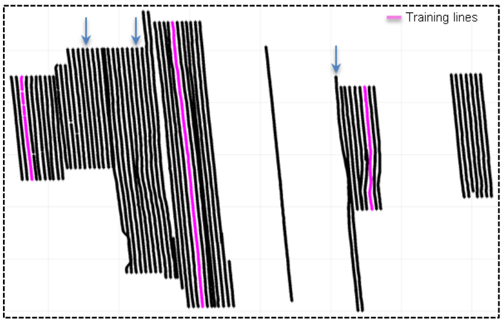
Figure 3 *Cross-plot of 2 features computed from the training data (blue=mild, green=optimal, red=harsh) (a), the corresponding Decision Space partition obtained using SVM with polynomial kernels (b).*

Application

The proposed framework is tested on a dual-sensor marine data processing project to QC the denoise production prior to the wavefield separation. For this test, we have access to six lines each with three types of filtering (Figure 4-a). In order to assess the quality of the prediction, the six lines are split into three lines for training the machine learning classifier and the other three lines are used to validate the prediction (validation lines indicated by arrows in Figure 4-a). The training lines are spread through the survey to improve robustness of the learning. The validation error for a given filtering type is the percentage of shots, from the validation lines, that were not correctly classified. Similarly the training error is the same miss-classification rate but computed for the shots that belong to the three training lines. The training error is a measure of how separable the training features are, and the validation error is a measure of how robust the QC system is when predicting the filtering type for a new shot that was not used in the training. Figure 4-b shows the training and validation error rates using the raw attributes (no feature mapping). One can see that the training error for all the types of filtering is small, indicating the attributes are informative. The validation errors are also small, except for the optimal filtering, where it is moderately large. About 20% of the optimal filtering is miss-classified as harsh or mild filtering. This miss-classification error is decreased largely when the features are used instead of the attributes to train the machine learning classifier, as can be seen in Figure 4-b.

The classifier is then used to predict the outcome of the filtering for 68 lines that constitute the entire

(a)

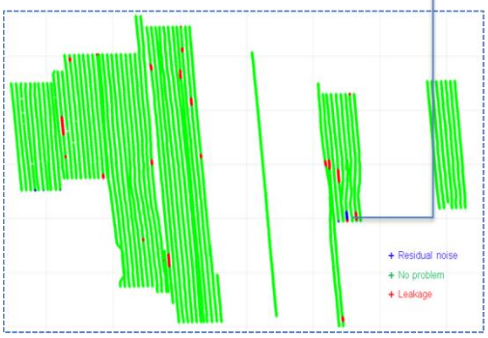


(b)

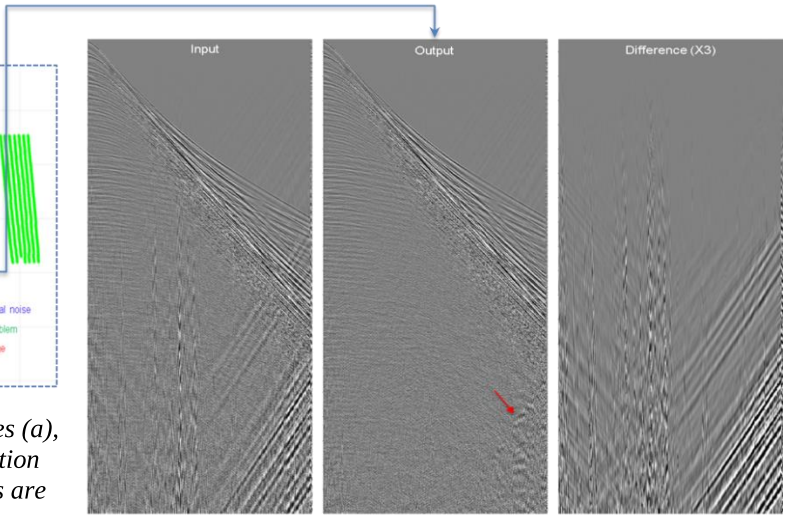
	Using attributes			Using features		
	Prodction	Harsh	Mild	Prodction	Harsh	Mild
Training error (%)	1.42	0.00	2.58	0.00	0.00	0.00
Validation error (%)	20.64	0.01	2.75	1.10	0.00	0.00

Figure 4 *Locations of the training lines and validation lines (arrows) (a), the corresponding training and validation error when using attributes or features (b).*

survey. The result of the automatic QC is shown as a tri-colour decision map for every shot in the survey (Figure 5-a). Inspection of the seismic at the highlighted locations shows some residual turn noise left in the data (Figure 5-b). An extra denoise pass was applied on this line to address this problem. The majority of localised small red anomalies were false positives. They were related to the attenuation of coherent linear noise that was not present in the training lines, but removed by the denoise process.



(a)



(b)

Figure 5 *Locations of the training lines (a), the corresponding training and validation error for both cases of using attributes are features.*

Conclusions

Machine learning can offer a solution to the problem of automating the QC process. It can help the user to focus the attention on the areas that may have potential problems. The challenge is to determine suitable attributes that can capture the success and the failure of the filtering step, to reduce the rate of false positives. These attributes are better if they are also tailored to the type of noise targeted by the filtering process and therefore needs to convey some geophysical meaning. Extracting hidden structures in the set of attributes is useful to improve robustness of the automatic QC process.

References

Chen, C., Chu, T.H, [2005] Similarity Measurement Between Images, IEEE Conference on Computer Software and Applications, 41-42, Edinburgh, UK.

Cristianini, N., Shawe-Taylor, J. [2000] An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press.

Hyvärinen, A., Karhunen, J., Oja, E. [2001] Independent component analysis, John Wiley & Sons, Inc. Spanos A., Bekara, M. [2013] Using statistical techniques to improve the QC process of swell noise Filtering, 75th EAGE Conference & Exhibition.