Unsupervised Segmentation of 3D and 2D Seismic Reflection Data

Klaus Köster and Michael Spann

School of Electronic and Electrical Engineering

The University of Birmingham, UK

Edgbaston, Birmingham B15 2TT,UK

E-Mail: koesterk@eee.bham.ac.uk, M.Spann@bham.ac.uk

Telephone: ++44 121 414 4328 Fax: ++44 121 414 4291

Abstract

This paper describes an unsupervised method to extract 2D and 3D inner earth structures from seismic reflection measurements. The application is a typical texture segmentation problem, which can be split up into a feature extraction stage and a segmentation stage. As a texture feature, the locally emergent frequency is estimated by a Gabor filter bank. The instantaneous frequency (IF) has already been successfully used for seismic trace analysis [1] and will be compared with the results of the filter bank. The second stage of the algorithm involves a region-growing method to compute the final object structure. The extremely flexible segmentation scheme is appropriate for application to 2D and 3D images of arbitrary vectorial dimension. The merging decision is based on the mutual inlier ratio of two adjacent regions. This ratio is computed by robust regression techniques [2] to avoid noise artifacts. A mutual inlier ratio discrimination function to recognise identical Gaussian distributions, guaranteeing a 97.5% certainty, is derived. This method is compared with the Kolmogorov-Smirnov test and results of the application in a segmentation algorithm are shown. The segmentation stage is also tested with different benchmark data sets from other computer vision problems to demonstrate its general flexibility.

1 Introduction

This work describes an unsupervised method to extract 2D and 3D inner earth structures from seismic reflection measurements in order to support expert interpretation of the data. The data sets are acquired by a seismic reflection method. This method maps the structure of subsurface formations by measuring the times required for a seismic wave to return to the surface. The waves are reflected by interfaces between formations having different physical properties. The reflections are measured by an array of ground motion responsive instruments. A successive migration process converts the measured data into 2D or 3D data sets. By interpreting this data, a geophysicist is able to recognise structural features from characteristic variations in the reflection times. The aim of this work is to support and improve the performance of the time consuming interpretation process (esp. for 3D data).

Direct approaches to this problem in the published literature are rare in particular for 3D data. Ng *et al* [3] proposed a supervised method for 2D data. This approach used 3×3 sized bandpass filters to extract characteristic features for a trained segmentation algorithm. The necessary training step limits the number of distinguishable regions and becomes very complex for 3D data. Love and Simaan [4] propose a segmentation system using a supervised template matching approach and statistics generated from greylevel runlength for texture classification. A rule-based second stage then performs the segmentation. The algorithm is tested on a single 2D seismic section only and the applied rules are specific to this set. Zhang and Simaan [5, 6] describe a similar approach, but apply a texture energy measure [7] as their feature to extract characteristic seismic patterns. Huang and Fu [8], Huang *et al* [8] and Kubichek and Quincy [9] propose methods for detection of special seismic events (bright spots, line structures and stratigraphic traps) in 2D reflection data. Pitas and Kotropoulos [10] propose a supervised multi-feature technique (features derived from Hilbert transform, rule-based seismic horizon picking and greylevel-runlength description) for 2D data. A minimum entropy approach selects thresholds for texture classification based on the extracted features. A region-growing stage then computes the final segmentation. This approach has been tested on a single 2D seismic set. Shen *et al* [11] describe a hierarchical region clustering technique for 2D and 3D segmentation. The approach is applied to a 3D variance-filtered seismic reflection measurement and works in an unsupervised way.

From a computer vision point of view the task is a general texture segmentation problem. Firstly characteristic patterns for seismic events are extracted (feature extraction stage). Secondly, in a segmentation stage, data elements with similar patterns are grouped to geologic structures (segmentation stage). Taner *et al*[1] discussed using a local frequency estimate as a useful feature to describe seismic events. There are several approaches to extract this quantity in the literature [12, 13, 14]. In this work, Gabor filters are applied since they offer the best possible response localisation in the time and frequency domains[15]. Particularly useful tools for geologic interpretation of seismic traces are complex seismic trace analysis [1] and the extraction of the instantaneous frequency (IF). The theory of IF, and a discussion of its interpretation, can be found in [16]. An important advantage of this quantity is, especially for large 3D data sets, its simple in-place computation. The feature extraction will be elucidated in Section 2. Results of both estimators are compared in Section 4.

The segmentation stage of this work uses a region-growing technique based on statistical decision rules. This general technique segments the entire image into small cells (say 2x2, 4x4 or 8x8 pixels) and merges them based on a statistical measure. A problem of the approach is the lack of data in the beginning, due to the small initial cell size, which imposes a high uncertainty on statistical results. Also a regular grid of initial cells is likely to contain cells lying on the boundary of two or more regions, which makes them difficult to handle in the later merging stage. Rousseeuw [2] proposed the least median of squares (LMedS) estimation technique, which results in robust estimates even when the data set contains up to 50% of outliers. Introducing this technique into computer vision, Meer [17] observed more robust results for smaller sets compared to traditional least squares methods. By using the LMedS properties, together with a shifting technique for overlapping cells, the problem of cells placed on multiple regions is solved in this work. After this tessellation step, the algorithm starts with initial cells containing minimum outliers and a maximum certainty for the estimated distribution.

The merging stage compares in general the distributions of the initial regions using a hypothesis test and merges them according to the result. An established method for testing is the Kolmogorov-Smirnov test which does not restrict the underlying distribution, but is known to be unreliable for small data sets and computationally expensive for larger sets. Meer [17] uses a mutual inlier ratio, which in his work must be larger than a specified fixed value to allow two regions to merge. Smaller sets imply larger associated statistical uncertainties. Therefore the Meer specified value must be a function of the set size, to ensure that the probability of merging two regions is invariant of the set size. Section 3 describes this idea and the consequences for the segmentation technique. Finally Section 4 displays results for seismic data and for some other 2D and 3D applications of the algorithm. Conclusions are given in Section 5.

2 Feature Extraction

This section describes the feature extraction stage. Two methods for estimating the local frequency in seismic data will be described in detail. Section 2.1 describes an approach using a Gabor filter bank and Section 2.2 elucidates the estimation of instantaneous frequency, a tool applied for seismic data interpretation proposed by Taner[1].

2.1 Gabor Filter Banks

Gabor filters are widely used for extraction of local frequency[12, 15]. A real 2D-Gabor filter averaging in the z direction, can be mathematically described as

$$\begin{aligned} m(x,y,z) = \quad & e^{-\frac{1}{2}\left[\frac{x^2}{\sigma^2}+\frac{y^2}{\sigma^2}+\frac{z^2}{\sigma_z^2}\right]} \cdot \\ & \cos\left(2\pi\mu_0(x\cos\theta + y\sin\theta)\right), \end{aligned} \tag{1}$$

where μ_0 denotes the radial centre frequency and θ the orientation of the Gabor wavelet. As Gabor filters are not separable in the spatial domain, they are computationally very expensive if implemented as convolution masks in the spatial domain. Equation 1 uses the third dimension for smoothing, but assumes that frequency components in this direction are negligible against the contributions in the x- and y-directions. This assumption is valid in the presented data sets, where changes in the z-direction happen to be very smooth. The advantage is the separability of the convolution mask in z-direction.

To estimate local frequencies, multiple Gabor filters tuned to different frequencies and orientations have to be applied on the data. The filter bank has to cover the frequency bands of the expected signal. The arrangement of filters for this work is derived from [15]. For the application in seismic data, a filter bank of 10×10 filters has been designed with frequencies μ_0 between 0 and π and orientations θ between 0 and 180 degrees in the Fourier domain. The values for μ_0 are $\{$0.03, 0.057, 0.086, 0.11, 0.14, 0.169, 0.196, 0.22, 0.25, 0.28$\}$ and for the orientations θ $\{$0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°$\}$. To estimate the local emergent frequency the three largest filter responses are stored. If $m_1(\vec{x})$, $m_2(\vec{x})$ and $m_3(\vec{x})$ are the responses of these three filters, $\mu_1(\vec{x})$, $\mu_2(\vec{x})$ and $\mu_3(\vec{x})$ and $\theta_1(\vec{x})$, $\theta_2(\vec{x})$ and $\theta_3(\vec{x})$ are the corresponding centre frequencies and orientations. The local emergent frequency $\omega(\vec{x})$ and orientation $\Theta(\vec{x})$ are then estimated by:

$$\omega(\vec{x}) \approx \frac{m_2(\vec{x})\mu_1(\vec{x}) + m_2(\vec{x})\mu_2(\vec{x}) + m_2(\vec{x})\mu_3(\vec{x})}{m_1(\vec{x}) + m_2(\vec{x}) + m_3(\vec{x})} \tag{2}$$

$$\Theta(\vec{x}) \approx \frac{m_1(\vec{x})\theta_1(\vec{x}) + m_2(\vec{x})\theta_2(\vec{x}) + m_3(\vec{x})\theta_3(\vec{x})}{m_1(\vec{x}) + m_2(\vec{x}) + m_3(\vec{x})}. \tag{3}$$

2.2 Instantaneous Frequency Estimation

The model of the real seismic signal is assumed to be of the form

$$s(t) = A(t)\cos\phi(t), \tag{4}$$

where the amplitude $A(t)$ varies slowly with respect to $\cos\phi(t)$. Gabor[18] proposed a method to generate a unique complex signal, via a Hilbert transform, generally referred to as the analytic signal:

$$\begin{aligned} z(t) &= s(t) + jH[s(t)] \\ &= A(t)(\cos\phi(t) - j\sin\phi(t)) \\ &= A(t)e^{-j\phi(t)}, \end{aligned} \tag{5}$$

where $H[\,]$ denotes the Hilbert transform. To allow this computation

$$H[s(t)] = H[A(t)\cos\phi(t)] = A(t)H[\cos\phi(t)] \tag{6}$$

has to be valid. The necessary constraint of a slowly varying amplitude has been expressed more specifically by Bedrosian's product theorem for Hilbert transforms[19]. Having computed the analytic signal, the IF $f_i(t)$ can be extracted via:

$$f_i(t) = \tfrac{1}{2\pi}\tfrac{d}{dt}\phi(t). \tag{7}$$

The assumed signal model in equation 4 is a generally accepted model for locally narrowband signals[15]. The computation of the IF does not return meaningful results if this narrow band constraint and Bedrosian's theorem are not satisfied[19, 16]. To get an idea of the signal's local bandwidth, the instantaneous bandwidth (IB) for discrete times can be computed via[14]:

$$b_i(t) = \frac{|z(t)|^2 - |z(t+1)z^*(t-1)|}{|z(t)|^2 + |z(t+1)z^*(t-1)|}, \tag{8}$$

where $z^*(t)$ denotes the conjugate complex value of the analytic signal $z(t)$. $b_i(t)$ has a range of values between 0 and 1. In Taner's paper[1] the results of the IF computation of seismic data contain sharp peaks. Taner uses a weighted average of the IF to emphasise the IF of stronger reflection events and to smooth noise artifacts. Figure 1 illustrates a real seismic signal with computed IF and IB. The peaks in the IF estimate always occur with a peak in the IB showing that the narrow band constraint is violated and the IF estimate is erroneous at this time.

The IF value is the time derivative of the instantaneous phase $\phi(t)$. By taking the x- and y-axis (and z-axis for 3D data) of the input set as the time axis two (three) values are computed per pixel (voxel). The final derived feature is therefore a vector.

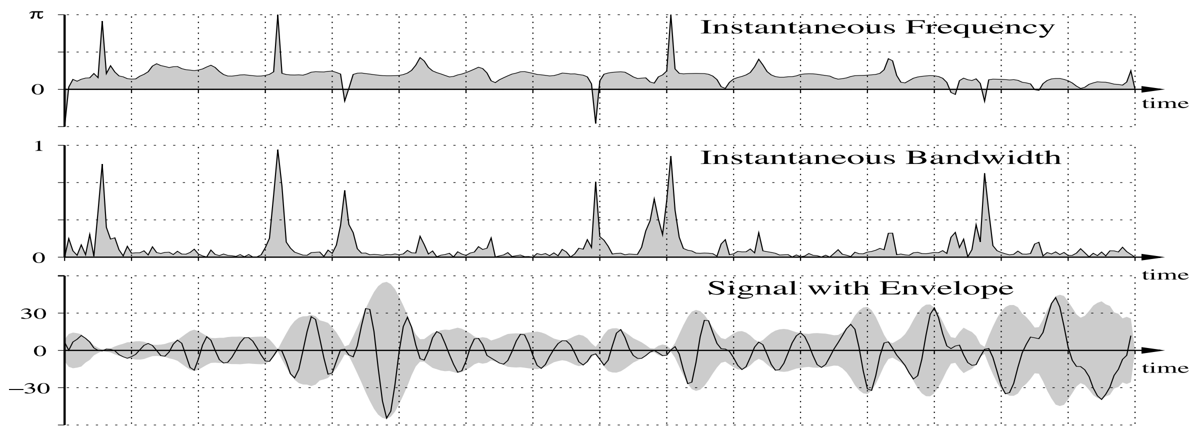


Figure 1: Seismic signal over time with envelope, instantaneous bandwidth (IB) and instantaneous frequency (IF) displays a strong correlation of peaks in the IF with high IB values.

3 The Segmentation Algorithm

The region-growing algorithm works in a bottom-up way generating an irregular graph pyramid to store neighbourhood relations between regions. Each graph node represents a region in the image and contains the parameters of a data model described in Section 3.1. The initial cell tessellation is the lower of two hierarchies and its generation will be described in Section 3.2. Starting from this level, higher hierarchies of primitive regions (PR) are computed by merging the cells in the level below. The merging decision is described in Section 3.3. Based on this decision rule, Section 3.4 describes the generation of the first PR hierarchy and Section 3.5 elucidates an iterative merging strategy leading to the final 2D or 3D region map.

One of the design goals of the algorithm was to keep its mathematics and methods as free as possible from constraints imposed by the spatial dimension of the input image or the data dimension (scalar or vectorial data). This allows the same technique to be applied to a range of completely different segmentation problems, demonstrating to the reader the flexibility of this technique.

3.1 The Data Model

A polynomial description of the region data has been chosen. The implementation supports zero- and first-order models for 2D and 3D images of scalar or arbitrary dimensional, vectorial data. As space is limited, only the model for 3D images, first order models and vectorial input data will be described. Other configurations can easily be derived from these results.

Given a set of voxels $V = \{\vec{v}_1, \vec{v}_2, \ldots\}$, the data model is specified by

$$\vec{v}_i(x, y, z) = \vec{a}_x x + \vec{a}_y y + \vec{a}_z z + \vec{a}_0 + \vec{e}_i, \tag{9}$$

where the $\vec{a}_{\{x,y,z,0\}}$ denote the model parameters and $\vec{e}_i$ is an independent additive normally distributed noise term. For this case the model parameters can be estimated by minimum least squares (LS) techniques. Unfortunately, in the presence of outliers caused by the neighbourhood overlapping a pair of adjacent cells, least squares breaks down. However, the LMedS estimator described in [2, 17] is capable of rejecting up to 50% of outliers, which is the basis of the estimation procedure in this work.

The noise term $\vec{e}_i$ in the model is taken into consideration by a covariance matrix $\boldsymbol{C}$ for each region. For the case of scalar data the matrix becomes simply the variance of the distribution. The smoothness of the estimate is computed by:

$$\text{Smoothness} = \sqrt{Det(\boldsymbol{C})}, \tag{10}$$

where $Det()$ denotes the determinant. Since a covariance matrix can be seen as a general description of an ellipsoid, the smoothness value can also be interpreted as the volume of the ellipsoid[2].

To test whether a sample p is an inlier within the model, the decision rule[1],

$$\left(\vec{p} - \hat{\vec{v}}_i\right) \boldsymbol{C}^{-1} \left(\vec{p} - \hat{\vec{v}}_i\right)^T \leq \chi^2_{p,0.975} \tag{11}$$

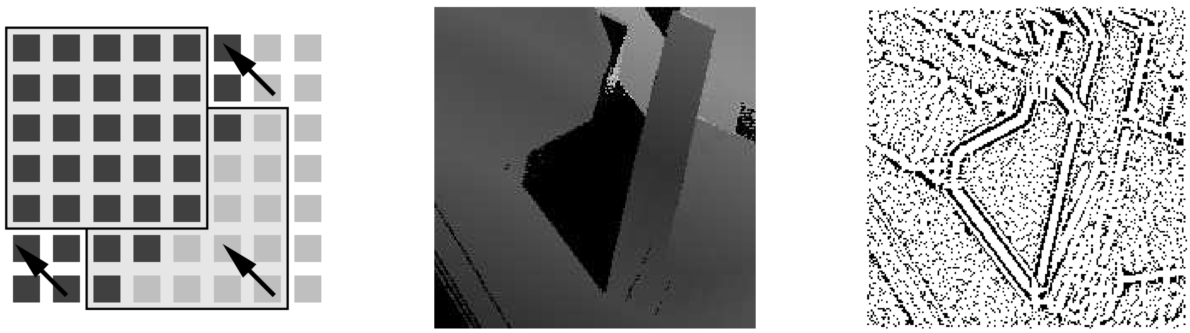
has to be true. $\hat{\vec{v}}$ denotes the LMedS computed estimation. The value $\chi^2_{p,0.975}$ is from chi-square statistics, where p denotes the degrees of freedom, which is equal to the vectorial dimension of the data in this case. The left hand side of the relation expresses also a relative distance of the point $\vec{p}(x, y, z)$ from the data model. Expression 11 is true for 97.5% of the samples of a p-dimensional Gaussian distribution. The advantage of this decision rule, and the above defined measure of smoothness is that both are applicable for arbitrary polynomial

[1] The coordinates have been dropped for convenience.

degrees and vectorial dimensions.

3.2 Initial Segmentation Stage

The original data set is tessellated with cells of $3 \times 3 \times 3$ voxels for the 3D case and 5×5 pixels for the segmentation of 2D images. The initial size is a trade-off between statistical robustness, which requires larger sets when estimating model parameters, and good precision for detecting small regions, which needs small basic elements for later merging. The LMedS estimator is applied to each cell to estimate the model parameters. The cells are overlapping by 2 voxels (4 pixels in 2D) in each dimension. A regular tessellation is likely to contain cells placed on a boundary of two or more regions. To avoid this the initial cells are shifted. The smoothness criterion described in equation 10 will return large values for cells placed on a boundary. Therefore each cell is shifted by at most one voxel (2 pixels in 2D) per direction



a. b. c.

Figure 2: Shifting of the initial cells. The cells, marked by a frame in figure a., are shifted towards a local variance minimum. Figure b. shows a range image example and figure c. the shift result, where the centre pixel of each 5x5 cell has been marked by a black point. Clearly the cells shift away from boundaries.

to a position resulting in a local minimum of the smoothness criterion. Figure 2 illustrates this step for the case of a 2D range image. The initial cells are overlapping to avoid any voxels (pixels in 2D) not being covered by any cell after the shifting procedure. This cannot be avoided for a few pixels at the image boundary. By shifting the cells, the initial tessellation can be adapted to almost any region structure and the final segmentation will not contain the "block-structure" of the initial cells.

3.3 Region Merging based on the Mutual Inlier Ratio

In general, the merging stage of a region-growing algorithm compares statistical measures of two adjacent regions and merges them or not according to the test result. This algorithm uses the mutual inlier ratio to come to a decision, assuming the set elements are Gaussian-distributed. Let R_A be the set of elements of region A. The set of inliers I_A of region A is

$$I_A = \left\{ \vec{v} \in R_A : (\vec{v} - \vec{v}_A)\boldsymbol{C}_A^{-1}(\vec{v} - \vec{v}_A)^T \leq \chi^2_{p,0.975} \right\}, \tag{12}$$

where $\vec{v}_A$ denotes the result of the model of Region A and $\boldsymbol{C}_A^{-1}$ is the inverse of its covariance matrix.[2] To determine the mutual inlier ratio M_{AB} of two regions A and B, the inliers of A which are also inlying in B have to be determined:

$$I_{AB} = \left\{ \vec{v} \in I_A : (\vec{v} - \vec{v}_B)\boldsymbol{C}_B^{-1}(\vec{v} - \vec{v}_B)^T \leq \chi^2_{p,0.975} \right\}. \tag{13}$$

Now the mutual inlier ratio M_{AB} becomes:

$$M_{AB} = \frac{\mathfrak{card}(I_{AB})}{\mathfrak{card}(I_A)}, \tag{14}$$

where the operator $\mathfrak{card}()$ returns the number of elements of a set.

Having computed the ratio, one has to compare the result with a discriminating value. If the ratio is larger than this value, then the two distributions are deemed to be similar. A safe or reliable test result can only be obtained if the similarly computed M_{BA} is also larger than the discriminating value. This technique avoids the situation where regions with large variances accept too many adjacent regions with smaller variances as inliers leading to an over-merged result. Certainly the discriminating value is a crucial value and affects the size of merged regions significantly. In the sequel an expression for the discriminating value will be determined.

It is a well-known fact that the quality of a statistical estimate of a data set's distribution parameters depends on the size of the set. High levels of both precision and reliability can

[2]The coordinates of the vectors have been dropped for convenience.

be achieved by increasing the sample size. For a region-merging algorithm, where the initial region size is in general limited, the exact estimation of the model parameters suffers from the standard deviation of the model σ and from an additional finite sample size error, which decreases with growing set size n. The relationship between the exact, unknown, mean of a distribution μ and the estimated mean $\bar{X}$ can be expressed by [20]:

$$\mu = \bar{X} \pm d, \qquad d = t_{0.975}\frac{s}{\sqrt{n}}, \tag{15}$$

where s denotes the estimated standard deviation, n is the set size and $t_{0.975}$ is based on the Student's t distribution with a 97.5% confidence interval, allowing us to use the estimated standard deviation instead of the true (unknown) standard deviation. A confidence interval for the estimated standard deviations s can be constructed using the chi-square distribution and the standard deviation σ [20, p. 325]:

$$\frac{\sigma^2\chi^2_{(n-1),\alpha/2}}{n-1} \leq s^2 \leq \frac{\sigma^2\chi^2_{(n-1),1-\alpha/2}}{n-1}. \tag{16}$$

Having two estimates of the same distribution, the worst case occurs if estimate $\bar{X}_B$ equals $\mu + d_B$ and estimate $\bar{X}_A$ is computed as $\mu - d_A$. A discrimination value M_{DISC} finding the two estimates of the same distribution to be similar in this worst case can theoretically be determined by solving the integral over the overlapping inlier area of the two estimated distributions:

$$M_{\text{DISC}} = \frac{1}{\sqrt{2\pi}s_B}\int\limits_{\bar{X}_B - s_B\sqrt{\chi^2_{p,0.975}}}^{\bar{X}_A + s_A\sqrt{\chi^2_{p,0.975}}} e^{-\frac{(x-\bar{X}_B)^2}{2s_B^2}}\,dx. \tag{17}$$

In figure 3 the true and unknown underlying distribution of the sets A and B is displayed in the centre (light grey). The estimated distributions being imposed by a finite set size error with their corresponding inlier intervals ($\bar{X} \pm s\sqrt{\chi^2_{p,0.975}}$) are illustrated on the left and right hand side. The integral over the overlapping inlier area in equation 17 is then displayed as the dark grey area.

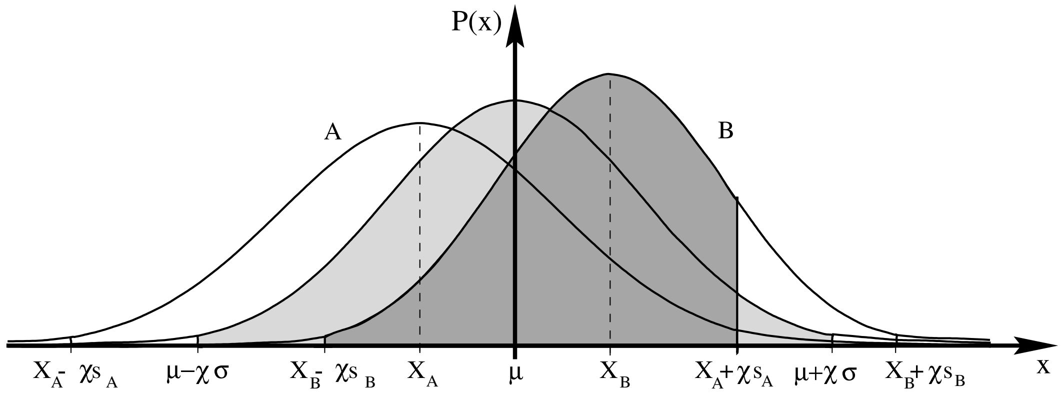


Figure 3: Integral over the overlapping inlier area. The symbol χ in the drawing is an abbreviation for $\sqrt{\chi^2_{p,0.975}}$.

According to the chosen χ^2-value, $M_{\text{DISC},97.5\%}$ should be 0.975 for a set of infinite size as the finite set error d disappears. Using the worst case estimates of the standard deviation in equation 16 the discrimination value M_{DISC} can be determined as

$$M_{\text{DISC}} = \frac{1}{\sqrt{2\pi} s_B} \int_{\bar{X}_A - s_A \chi_{p,97.5\%}}^{\bar{X}_B + s_B \chi_{p,97.5\%}} e^{-\frac{(x-\bar{X}_B)^2}{2s_B^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{b_{low}}^{\chi_{p,97.5\%}} e^{-\frac{x^2}{2}} dx,$$

$$\text{where} \quad b_{low} = s_{ratio}\Big(\frac{t_{(n_A-1),\alpha/2}}{\sqrt{n_A}} - \chi_{p,97.5\%}\Big) + \frac{t_{(n_B-1),\alpha/2}}{\sqrt{n_B}}$$

$$\text{and} \quad s_{ratio} = \frac{\sqrt{n_B - 1} \cdot \chi_{(n_A-1),1-\alpha/2}}{\sqrt{n_A - 1} \cdot \chi_{(n_B-1),\alpha/2}}. \tag{18}$$

The detailed steps between these two expressions for the discrimination value can be found in [21].

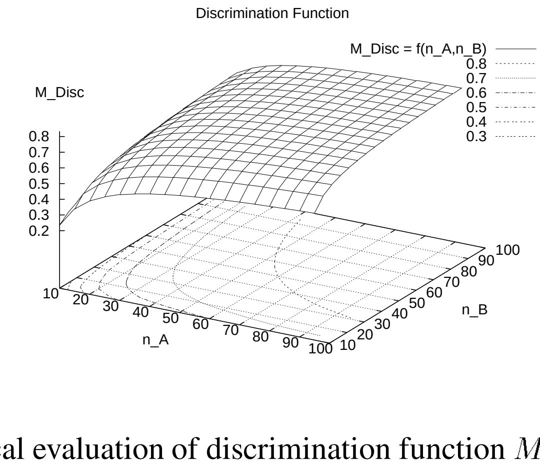


Figure 4: Numerical evaluation of discrimination function $M_{\text{DISC},97.5\%}(n_A, n_B)$.

Equation 18 expresses the discrimination value as a function of the set sizes n_A and n_B, allowing a simple hypothesis test which is ideal for a region-merging algorithm. Figure 4 shows the numerical evaluation of the discrimination function. The test supports higher-order models as well as vectorial data by adapting the outlier/inlier rule in expression 11. The sole method's parameter is the width of the confidence intervals in equations 15 and 16. In [21] we have compared the MIR-method with the established Kolmogorov-Smirnov (KS) test. The MIR-method resulted in similar classification errors, but the error when comparing two distributions with different variances was significantly lower for smaller set sizes. Figure 5

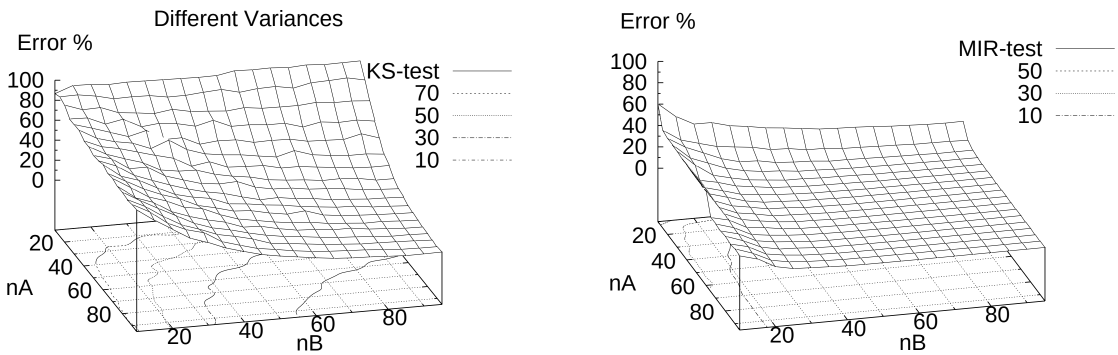


Figure 5: Comparison of KS-test and MIR-test. Average classification error when examining sets with identical mean and different variance of the underlying distribution.

shows the average classification error measured when comparing two sets of sizes n_A and n_B. The figures shows a smaller classification error for larger test sets for both tests, but the MIR-result converges faster to lower error rates for small sets. However, it should be kept in mind that the KS-test does not assume any distribution, while the inlier ratio assumes Gaussian statistics to detect outliers.

3.4 Primitive Region Generation Stage

This stage of the algorithm generates a first PR hierarchy by merging the initial cells. The stage starts by sorting the cells according to their smoothness ("volume"). A first PR is created and the smallest-volume-cell is added to this region. Next all spatially adjacent neighbours of this cell are visited recursively, becoming a member of the PR if they are inliers within the model of all of the other cells of the PR. The recursion stops if an adjacent cell does not satisfy the merging rule. The inlier decision is found by the described mutual inlier ratio. If no further adjacent cells can be added to the first PR, a new PR is started with the next available smallest-volume cell. Again the neighbours of this cell are visited recursively. This step is repeated until every cell is associated with a PR. The sorting step in the beginning of this stage avoids over-merging since cells with a large variance tend to accept other cells more easily as inliers because of the “large volume”- covariance matrix.

3.5 Iterative Region-Merging Stage

This final stage of the algorithm iteratively merges the PRs until no adjacent region pair satisfies the merging rule. Each iterative merging step starts by constructing a link to the neighbour, which results in a mutual satisfaction of the decision rule (equation 11) with the *highest* sum of mutual inlier rations MIR:

$$MIR = M_{AB} + M_{BA}. \tag{19}$$

A merge test, based on all elements of the two regions, would only merge two regions if all the elements strictly belong to the estimated model. By using only data close to the boundary of the two examined regions (local data), smooth model changes within a region are also tolerated. The local data consists of elements of cells placed at the boundary of the two PRs. The boundary cells are recognised in the graph hierarchy since they possess a neighbouring cell whose parent is the other PR considered in the merge test. In figure 6(a) these boundary

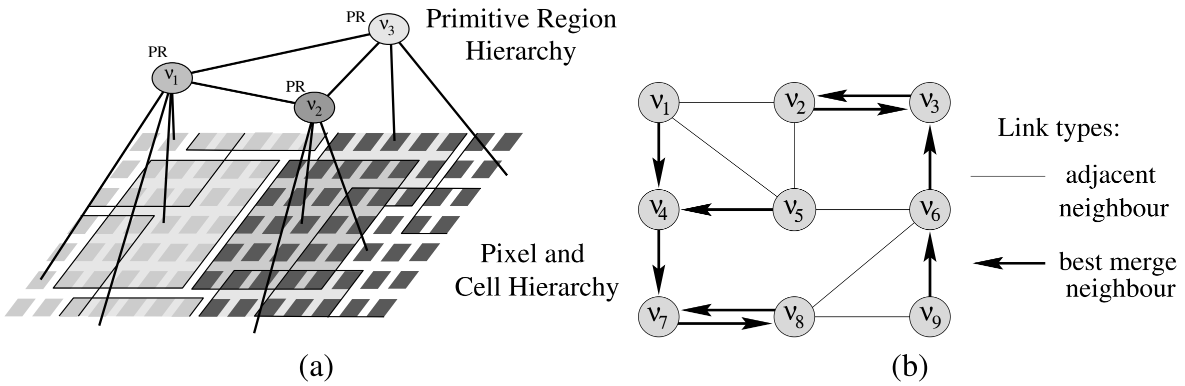


Figure 6: Data hierarchy of the algorithm (a). Displayed are the initial cells with parent links to the PRs. Each PR contains a set of spatially adjacent neighbours links. The pixels of the two regions are symbolised by squares with different greylevel. Figure (b) displays the regions ν_i having a link to their neighbour being most qualified for merging.

cells are marked by a darker background.

The stage starts merging the PRs after the best neighbour for merging has been determined for each PR. Each PR satisfying the merge test with one of its neighbours, is linked to that neighbour which results in the maximum MIR. However neighbouring PR's are *not* necessarily mutually optimal, e.g. in figure 6(b) PR ν_6 optimises the MIR for PR ν_9 even though ν_9 does not optimise the MIR for ν_6. The links build up "chains" ending in a mutually best-neighbour pair. Only these pairs are merged in each iteration to avoid over-merging. Others might be connected in later iterative steps so long as the merging criteria are still satisfied. In figure 6(b) only regions ν_2 and ν_3 and regions ν_7 and ν_8 would qualify for merging according to this rule.

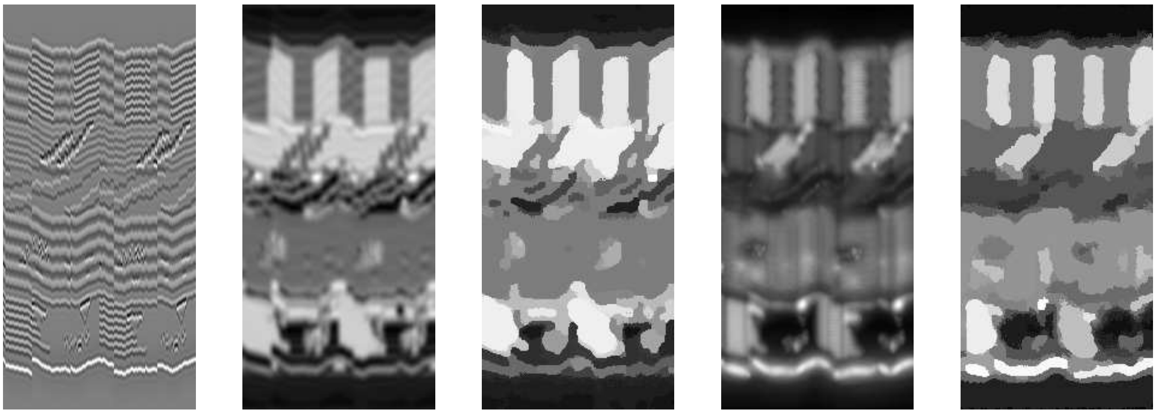
This iterative merging strategy has the advantage of being independent of the order of visiting each PR. Therefore the iterative step is implemented in parallel code using multiprocessor architectures.

4 Experimental Results

This section demonstrates experimental results of the stages for feature extraction and segmentation. Section 4.1 shows results of the feature extraction stage on synthetic and real seismic data. The properties of the segmentation stage are examined in Section 4.2. Section 4.2.1 evaluates the segmentation stage and compares it with an algorithm for segmentation of 3D scalar data published in [11]. A quantitative evaluation of the segmentation stage, using real seismic data, suffers from the lack of objective ground truth data. For this reason we have applied the segmentation stage to other data sets where the generation of ground truth data is more objective. Results of this work are described in Section 4.2.2. Finally we show geologic structures extracted from real seismic reflection data.

4.1 Feature Extraction in Seismic Data

Figure 7(a) displays a 2D synthetic data set and the results obtained from the IF estimation (b) and the bank of Gabor filters (d). A brighter greylevel codes a higher estimated local frequency. The frequency measurements are, qualitatively, strongly in agreement with the



(a) (b) (c) (d) (e)

Figure 7: 2D synthetic seismic image (a) and results obtained from the IF estimator in vertical direction (b) and segmentation (c). Result of bank of 10×10 Gabor filters (d) and segmentation (e).

pattern in the top part of the image. Although the Gabor filter responds with small deviations at the pattern boundaries, details in the bottom part of the image are distinguished slightly better than by the IF approach. To compute the IF the data set has to be Fourier transformed, multiplied and inversely Fourier transformed, while each Gabor filter requires a convolution of the whole data set. Therefore the IF offers a significant computational advantage, especially for larger 3D data sets.

A major problem of the IF estimation in seismic data are the described peaks in the estimates due to the violation of the narrow band constraint of the assumed model (equation 4). Although the Gabor filter approach assumes a similar model [15], the results do not contain these artifacts and provide, especially for real data, a more robust base for the unsupervised segmentation stage.

Figure 8 shows a slice of a 3D seismic measurement and the corresponding magnitude of the frequency estimate obtained from Gabor filters. The layer structure of the reflection measurement is highlighted in the displayed feature map.

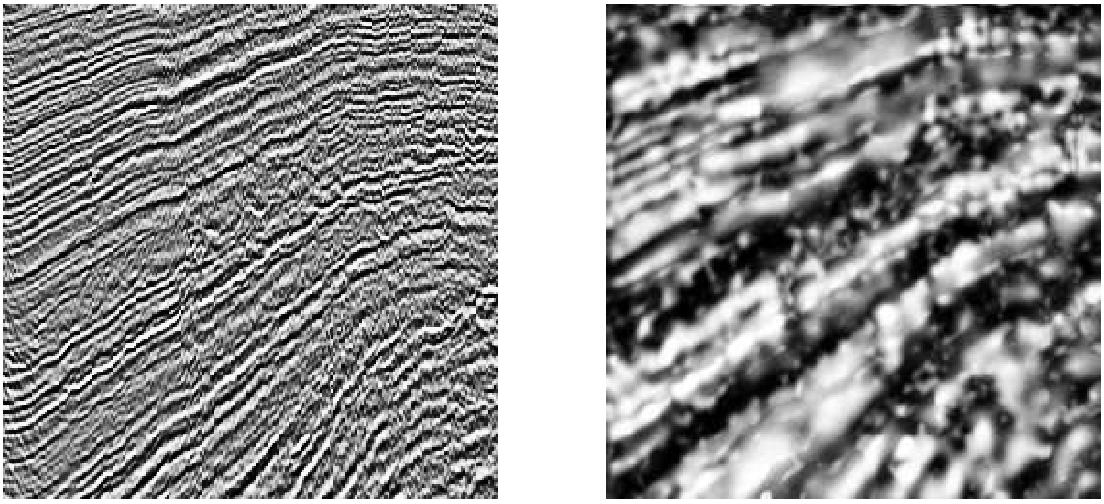


Figure 8: Slice of a 3D synthetic seismic image (left) and estimated local frequency obtained from a bank of 10×10 Gabor filters.

4.2 Segmentation Stage

4.2.1 Evaluation of the Segmentation Stage

This section describes a quantitative comparison of the segmentation stage described here and the work of Shen *et al* [11]. This technique is chosen as a reference because it is one of the few segmentation techniques for 3D seismic data in the published literature. Although both algorithms are supposed to segment 2D as well as 3D data sets, this evaluation will be performed using several 2D slices of a 3D feature map obtained from a 3D seismic data set.

The algorithm of Shen *et al* is a hierarchical bottom-up method successively clustering regions at each level until no further changes are possible according to the clustering conditions. The base level of the hierarchy is the original image data, while the top level describes the result of the segmentation. The iterative calculation of the next irregular graph hierarchy is performed by clustering regions of the lower hierarchy. Clustering at each level of the hierarchy is performed only if two adjacent regions are mutually nearest neighbours (NN), while the segmentation method of this paper is referred to as the *MIR-method* due to the applied merge criterion. For this reason we refer to this algorithm from now on as the *NN-method*. The nearest neighbour of a region is the adjacent region which minimises the model and the variance difference measured by exploiting Student's *t-test*. The data of a homogeneous region is represented by a linear model.

We apply an adapted version of the framework of Hoover *et al* [22] for the evaluation of range image segmentation algorithms to compare the segmentation techniques. The framework compares the *machine segmentation* (MS) with a corresponding set of *ground truth segmentations* (GT). An objective automated compare tool[3] performs the evaluation using specified metrics for *correct detection*, *over-segmentation* and *under-segmentation*. The definitions for deciding classifications are based upon a tolerance T having a range of $0.5 < T \leq 1.0$, reflecting the strictness of the definitions. Informally, Hoover *et al* [22] defined the metrics as follows:

correct detection: A pair of regions R_n in the GT image and R_m in the MS image are classified as an instance of correct detection if

1. at least T percent of the pixels in R_m are marked as pixels of R_n *and*
2. at least T percent of the pixels in R_n are marked as pixels of R_m.

over-segmentation: A region R_n in the GT image and a set of regions $R_{m1}, \ldots, R_{mx}$ in the MS image are classified as an instance of over-segmentation if

1. at least T percent of the pixels in each region R_{mi} are marked as pixels of R_n *and*
2. at least T percent of the pixels in R_n are marked as pixels of the union of $R_{m1}, \ldots, R_{mx}$.

under-segmentation: A set of regions $R_{n1}, \ldots, R_{nx}$ in the GT image and a region in the MS image R_m are classified as an instance of under-segmentation if

1. at least T percent of the pixels in region R_m are marked as pixels of the union of $R_{n1}, \ldots, R_{nx}$ *and*
2. at least T percent of the pixels in R_{ni} are marked as pixels of R_m.

Here the ground truth segmentations for the slices are obtained by computing an average segmentation of multiple independently obtained manual segmentations for each slice. The averaging process was introduced to avoid the influence of errors in the individual segmentations and to come to a more objective result. To compute the average segmentation of multiple

[3] The compare tool and a large number range image test sets is publicly available via http://marathon.csee.usf.edu/range/seg-comp/SegComp.html

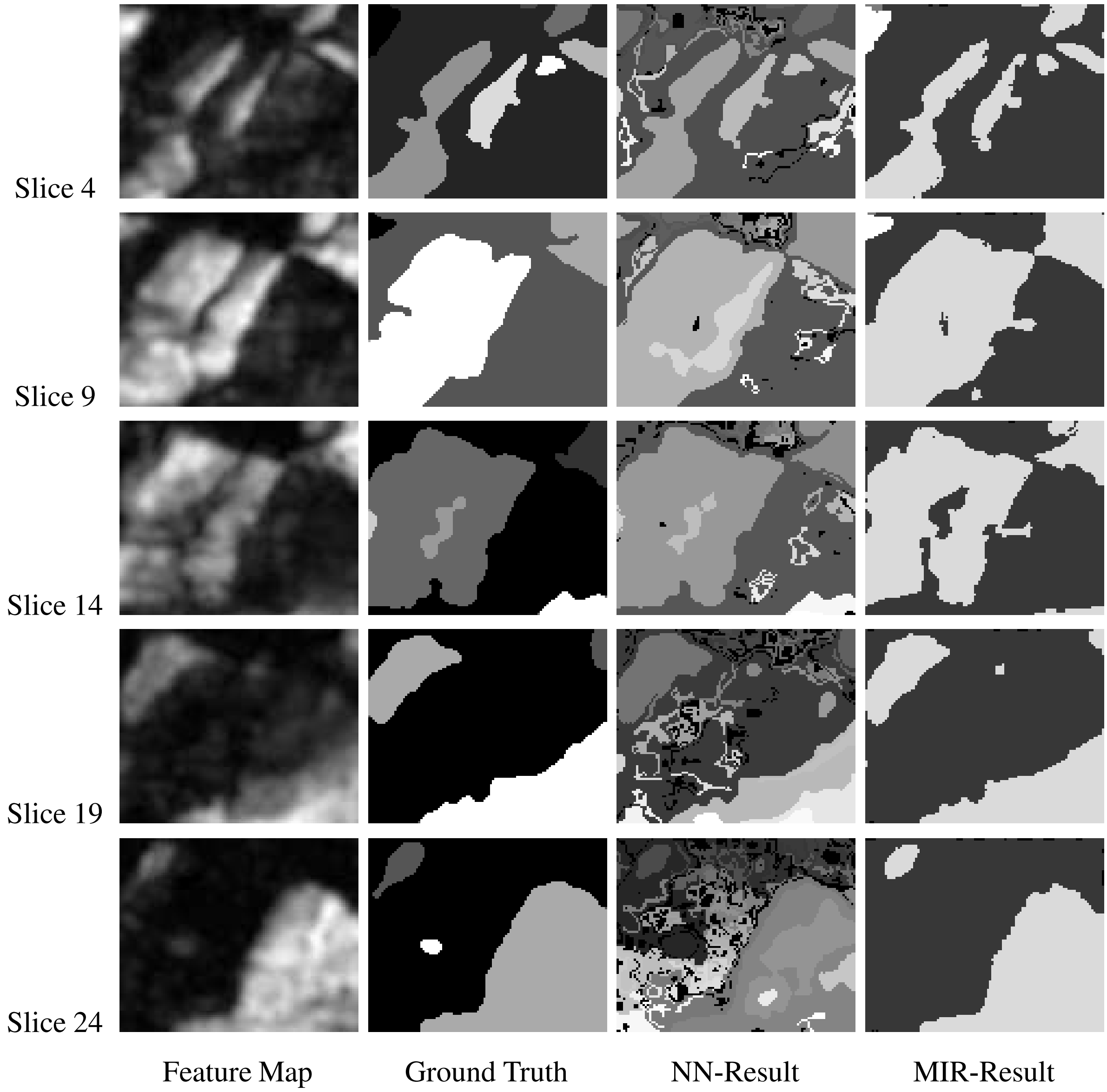


Figure 9: 5 equi-distant test 2D slices of a 3D set and their corresponding average segmentations, segmentations result obtained by the NN-method and the result of the technique described in this paper (MIR-method).

manual segmentations each slice was initially partitioned into pixel groups which belong to one region according to all manual segmentations. Adjacent partitions were then merged if they belong to the same region in at least 50% of the manual segmentations. Figure 9 shows slices of the 3D feature map and the corresponding segmentation results of the averaging process and the two compared algorithms. The total size of the chosen data set was $110 \times 90 \times 30$ voxels.

Figure 10 shows the performance graphs obtained from the application of the compare tool. A good segmentation algorithm would obtain a high score for correctly classified regions and low values in the error measures of over- and under-segmentation. The results would not deteriorate for larger compare tool tolerances.

The five test slices contain on average 4.2 regions, which sets the ideal result in the graph showing the number of correctly classified regions. Here the performance of the MIR-method is slightly superior to the NN-method. However, both segmenters' performance drops for compare tool tolerance larger than 80%. This may be interpreted as a lack of precision, but one has to keep in mind that the reference segmentations are obtained from *subjective* manual segmentations, which might themselves contain inaccuracies. The NN-method returns a larger value for the over-segmentation measure. In the displayed results in Figure 9 the algorithm returns a very detailed segmentation of the background causing this result. The graph documenting the average number of under-segmented regions shows a higher error measure for the MIR-method. The NN-method returns here a zero error graph. However, the high over-segmentation result and the low under-segmentation measure show the the trade-off between under- and over-segmentation is not completely balanced in the NN-method. Further parameter tuning might provide segmentations closer to the reference segmentations.

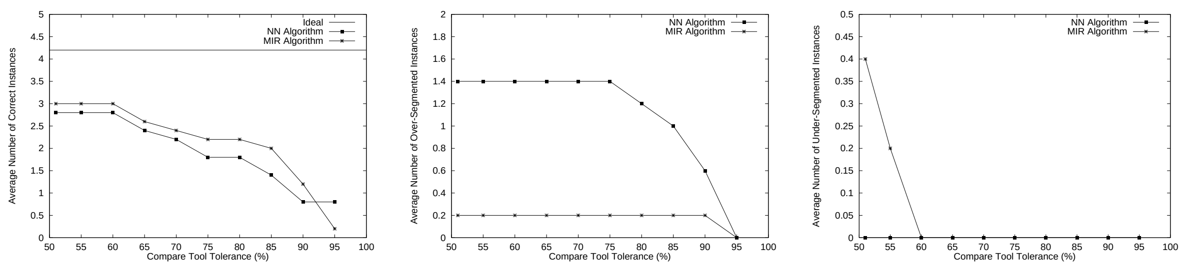


Figure 10: Average number of correctly detected regions (left), of over-segmented instances (center) and of under-segmented instances (right) for NN-method and MIR method.

4.2.2 Application to other Data

The previous section showed a quantitative evaluation of the segmentation stage using real seismic data. Although the algorithm is designed for segmentation of seismic data, an evaluation using different data is more appropriate when it comes to an examination of the segmentation precision. The applied reference segmentations were obtained from manual segmentations which can contain inaccuracies and inconsistencies due to the effect of the human interpreter. The evaluation can therefore only document general properties.

For this reason, and to demonstrate the flexibility of the approach, we have applied the algorithm to a set of 80 range images[4] and to 3D MRI images. The displayed range images have been measured with a structured light range finder.

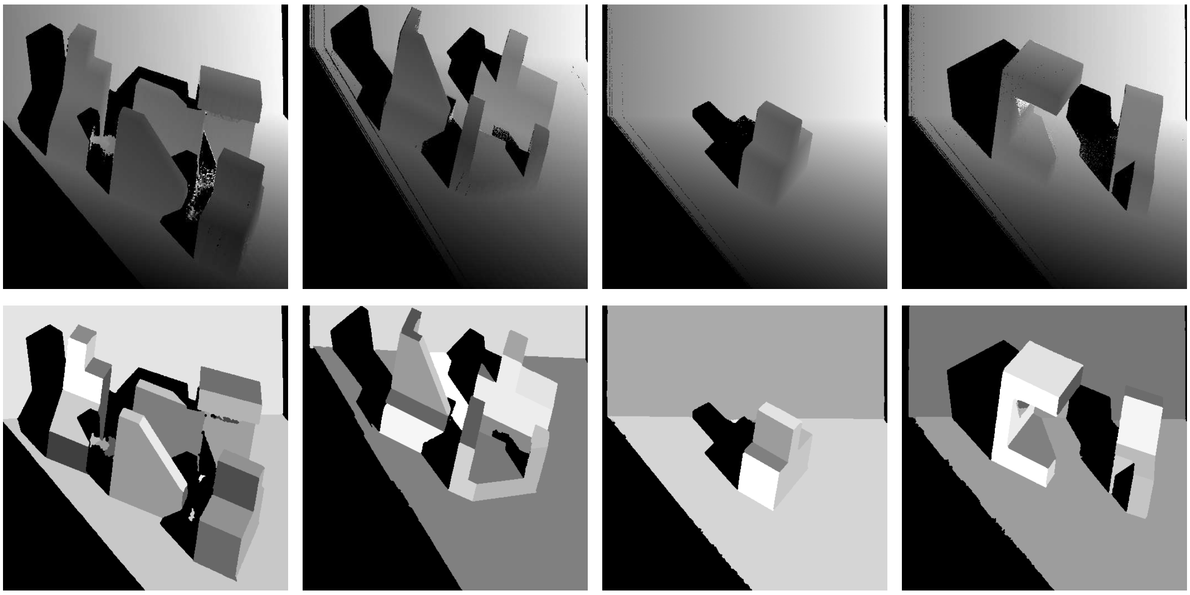
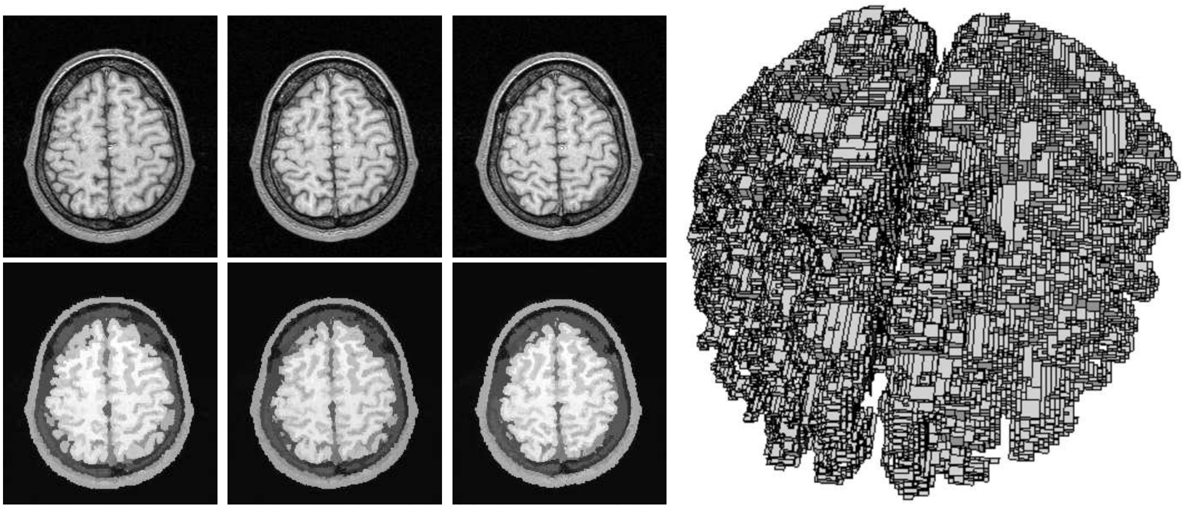


Figure 11: Segmentation results of range data. The top shows the range images and displays the segmentation result obtained from the MIR algorithm.

For the range image segmentation we have used a first order polynomial model. Using a final boundary refinement based on the model and the outlier discrimination rule (expression 11) we obtain the result displayed in figure 11. The boundaries have been detected with a high precision. It was found to be difficult to detect long thin regions (< 3 pixels) with the here described approach due to the PC mask size. However, these regions can be detected as outliers within the resulting regions. Such a step was not necessary for seismic data. A complete evaluation of our segmentation technique and a comparison with four other techniques based on the results obtained from 80 real range images is described in [21]. The images were measured by two different range finders (structured light and laser range finder) and in particular the segmentations of the noisier laser range finder results benefitted from the application of the LMedS estimator and showed a superior performance compared to the other algorithms in the comparison.

[4] Available from *http://marathon.csee.usf.edu/range/seg-comp/SegComp.html*

Figure 12(a-c) shows three slices of a 3D MRI data set and the corresponding segmented slices. The result was obtained using a constant data model. The example contains some long thin regions, which have been merged to adjacent regions in the result. To be detected a region has to cover a least half of an initial cell due to the LMedS estimator. The initial cells will shift away from regions, which are too thin to fulfil this constraint. Nevertheless these regions can be recovered as their voxels (pixels in 2D) are outliers within the adjacent region being merged to. Finally figure 12(d) illustrates a 3D voxel model of the segmented image.



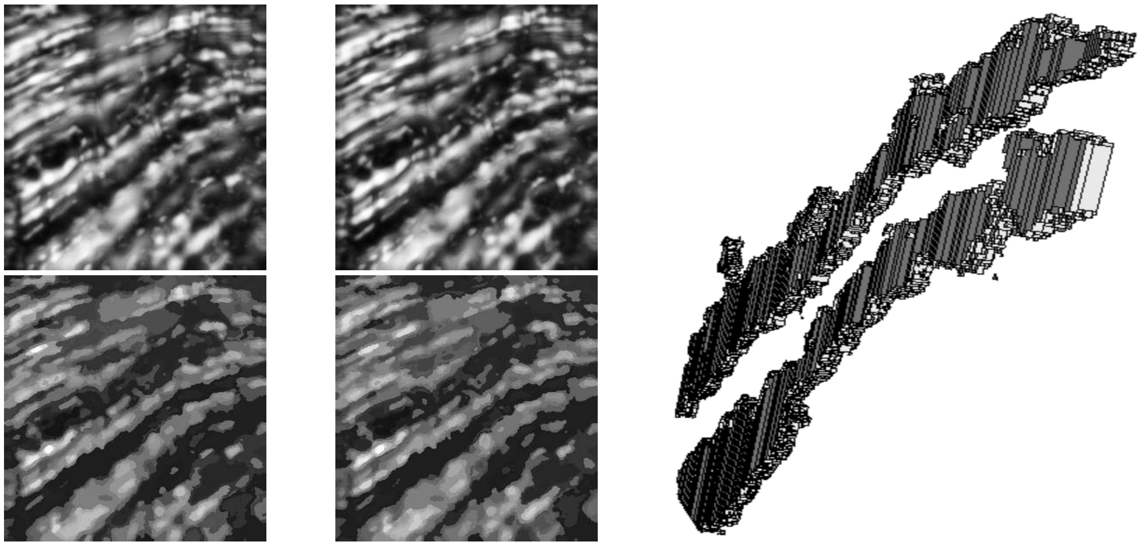
(a) (b) (c) (d)

Figure 12: Slices of 3D MRI data and resulting segmentations (a-c). 3D voxel visualisation of the brain obtained from a 256x256x45 voxel data set.

The range of different data set shows the flexibility of the algorithm. It is applicable to 2D and 3D images as we have demonstrated with range data (2D), MRI data (3D) and seismic data (3D or 2D). The statistical test applied in the region merging stage supports scalar data as well as vectorial data of any dimension (eg. for texture segmentation problems with multiple features). The polynomial data model can be of any degree, but for larger degrees the technique becomes computational expensive. We have tried linear models (range data) and constant models (MRI data, seismic data).

4.3 Extraction of Geologic Structures

Figure 13 displays two slices of a Gabor filtered real 3D seismic data set and the result obtained from the 3D segmentation stage. A constant polynomial model with scalar input data has been used to obtain the result. The displayed segmentation result contains many detailed



(a) (b) (c)

Figure 13: Segmentation of a real 3D seismic set. Top of figures (a) and (b) display slices of the frequency estimation result of the Gabor bank and the bottom figures the corresponding segmentation results. Figure (c) visualises a 3D model of the two diagonal regions in the bottom part of the image.

regions leaving the final decision to merge regions to the interpreter. The feature map obtained from the Gabor filter bank returns many details, which are difficult to ignore if only the major structures are of interest. In the intended application the financial consequences of over-merging are possibly much worse than for under-merging. Therefore a greater amount of autonomy is allowed for the human interpreter in this case. The displayed 3D model of the two major structures in the bottom half of the image in figure 13(c) is obtained by manually merging the algorithm's result.

Figure 14 displays a second example for a geometric model of a geologic structure. The structure can be found in the top left corner of the 2D slices displayed in figures 13 and 8.

5 Conclusions

This paper has described a system for 2D and 3D processing of seismic reflection data. The system starts by extracting information on geologic structures by applying a local frequency estimator and uses the result in a segmentation step to obtain the final region map allowing a 3D object visualisation.

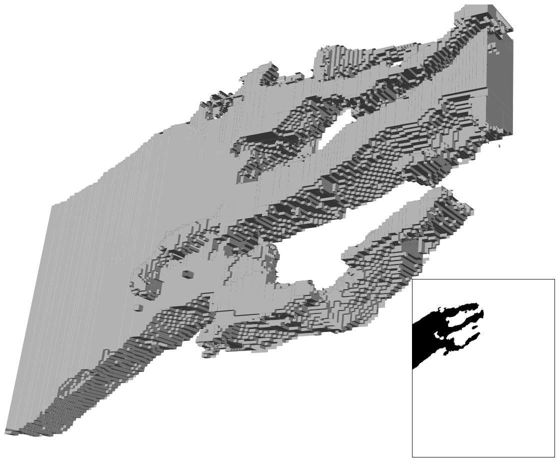


Figure 14: 3D model of geologic structure and map to find object position in figure 13(a).

Two methods to extract geological structure have been examined. A bank of Gabor filters extracts the geological features with good precision but is computationally demanding since Gabor filter masks are not separable spatially. The second method, instantaneous frequency, is easier to compute and has already proven to be a useful tool for interpreting seismic traces [1, 10]. Nevertheless in real seismic data the assumed narrow band constraint of the data model is violated frequently, resulting in sharp "jumps" in the estimate. This effect makes an unsupervised segmentation difficult using this feature.

The second stage of the system generates a 2D or 3D structure map based on the extracted feature by applying a robust-region growing technique. A new shift technique for the initial cells using the LMedS estimator enables the initial tessellation to adapt to any 2D or 3D region shape. The algorithm does not need any thresholds and is completely free from any assumptions on region number, shape or size. A merging technique has been proposed offering a high robustness especially at the crucial initial merging stage of the algorithm, where statistical estimates possess a high uncertainty due to small set sizes.

The merging technique applies a robust statistical similarity test evaluating the mutual inlier ratios of the two tested sets. The ratios are compared to a theoretically derived discrimination value to come to a merging decision. The technique is not limited to scalar data as it does not involve any sorting etc. as for example the KS-test. The test takes a parameter estimation error for small data sets into account.

The segmentation stage of the system, as well as the feature extraction, are highly parallel and appropriate for implementations on multi-processor architectures. A parallel implementation of the segmentation stage resulted into a 35% performance gain on a double processor machine compared to the implementation using only a single CPU.

References

[1] M. F. Taner, F. Koehler, and R. E. Sheriff, "Complex seismic trace analysis," *Geophysics*, vol. 44, pp. 1041–1063, June 1979.

[2] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. New York: Wiley, 1987.

[3] I. Ng, J. Kittler, and J. Illingworth, "Supervised segmentation using a multi-resolution data representation," *Signal Processing*, vol. 31, pp. 133–163, 1993.

[4] P. L. Love and M. Simaan, "Segmentation of a seismic section using image-processing and artificial-intelligence techniques," *Pattern Recognition*, vol. 18, pp. 409–419, 1985.

[5] Z. Zhang and M. Simaan, "A rule-based interpretation system for segmentation of seismic images," *Pattern Recognition*, vol. 20, pp. 45–53, 1987.

[6] Z. Zhang and M. Simaan, "A knowledge-based system controlled by an iterative quadtree splitting scheme for segmentation of seismic sections," *IEEE Trans. On Geoscience Remote Sensing*, vol. 26, pp. 518–524, 1988.

[7] K. I. Laws, *Texture image segmentation*. PhD thesis, Department of Engineering, USC, 1980.

[8] K.-Y. Huang and K.-S. Fu, "Syntactic pattern recognition for the recognition of bright spots," *Pattern Recognition*, vol. 18, no. 6, pp. 421–428, 1985.

[9] R. F. Kubichek and E. A. Quincy, "Statistical modeling and feature-selection for seismic pattern- recognition," *Pattern Recognition*, vol. 18, pp. 441–448, 1985.

[10] I. Pitas and C. Kotropoulos, "A texture-based approach to the segmentation of seismic images," *Pattern Recognition*, vol. 25, pp. 929–945, 1992.

[11] X. Shen, M. Spann, and P. Nacken, "Segmentation of 2D and 3D images through a hierarchical clustering based on region modelling," *Pattern Recognition*, vol. 31, September 1998.

[12] A. C. Bovik, "Analysis of multichannel narrow-band filters for image texture segmentation," *IEEE Transactions on Signal Processing*, vol. 39, September 1991.

[13] S. G. Mallat, "A theory for multi-resolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, July 1990.

[14] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The wigner distribution - a tool for time-frequency signal analysis, part 1-3," *Philips Journal of Research*, vol. 35, pp. 217–250, 276–300 and 372–389, March - May 1980.

[15] A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, "Localised measurement of emergent image frequencies by gabor wavelets," *IEEE Transactions on Information Theory*, vol. 38, March 1992.

[16] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal - part 1: Fundamentals," *Proceedings of the IEEE*, vol. 80, pp. 521 – 538, April 1992.

[17] P. Meer, D. Mintz, and A. Rosenfeld, "Least median of squares based robust analysis of image structure," (Pittsburgh, P. A.), pp. 231–254, DARPA, 1990.

[18] D. Gabor, "Theory of communication," *Proceedings IEE*, vol. 93, no. 3, pp. 429–457, 1946.

[19] E. Bedrosian, "A product theorem for hilbert transforms," *Proceedings of the IEEE*, vol. 51, pp. 868–869, May 1963.

[20] L. L. Lapin, *Probability and statistics for modern engineering*. Boston, MA: PWS-Kent, 2nd ed., 1990.

[21] K. Köster and M. Spann, "MIR: An Approach to Robust Clustering - Application to Range Segmentation." submitted for publication in IEEE Trans. on PAMI, April 1998.

[22] A. Hoover, G. Jean-Baptiste, X. Jiang, P. J. Flynn, H. Bunke, D. B. Goldgof, K. Bowyer, D. W. Eggert, A. Fitzgibbon, and R. B. Fisher, "An experimental comparison of range image segmentation algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 673–689, July 1996.