

4

Análise Tempo – Frequência de Sinais Sísmicos

Partindo do princípio de que a sísmica é a resposta ao impulso gerado por uma fonte sísmica da sub-superfície, sinais sísmicos podem ser interpretados como uma representação em baixa frequência da refletividade da sub-superfície.

Especificamente, sinais sísmicos são caracterizados por transições devido aos contrastes de impedância acústica e podem ser sintetizados, através da operação de convolução da função refletividade com uma onda sísmica padrão. Na sísmica, as reflexões fortes são usualmente associadas aos limites das seqüências estratigráficas e a partir da identificação das transições de refletividade suavizadas pela fonte sísmica são gerados os horizontes dos eventos com o mesmo comportamento geológico. Já a forma de onda entre as reflexões fortes também é influenciada pela composição intermediária entre as unidades estratigráficas.

A análise com dados e amostras reais da sub-superfície mostra que a refletividade não é necessariamente regular e uniforme. Conseqüentemente, as possíveis variações nas funções de refletividade geram grandes diferenças nas formas de ondas associadas, dificultando muitas vezes o processo de identificação das transições e das fácies sísmicas. Portanto, a análise do sinal sísmico, somente no domínio do tempo, pode induzir a erros de interpretação e classificação de fácies.

A análise no domínio da frequência através do espectro de potência é bastante utilizada na interpretação sísmica, principalmente para detecção de camadas delgadas (Partyka et al., 1999) (Mafurt & Kirilin, 2001). Entretanto, o espectro de potência não revela como o conteúdo de frequência varia ao longo do tempo e muito menos a localização das transições do sinal. Assim, para auxiliar a caracterização das propriedades das diversas camadas da sub-superfície, deve-se utilizar algoritmos para a análise da variação conjunta do comportamento de sinais em tempo e em frequência que enfatizem características específicas da sub-superfície.

4.1.Representações tempo – freqüência

A forma mais intuitiva para realizar a análise de sinais em tempo e em freqüência, simultaneamente, é obtida dividindo-se o sinal em segmentos menores e determinando o espectro de cada um dos segmentos através da transformada discreta de Fourier. Esta representação, descrita pela eq.(25), é chamada de Transformada de Fourier de Curto Tempo (Short Time Fourier Transform – STFT) e a seleção dos segmentos no tempo é realizada por uma janela deslizante no tempo.

$$STFT(t, \omega) = \int [x(\tau) \cdot W(\tau - t)] \cdot e^{-j\omega\tau} d\tau \quad (25)$$

onde W é a janela utilizada para segmentar o sinal.

Portanto, através do cômputo da STFT, são obtidos os espectros do sinal para diferentes deslocamentos no tempo t , segmentados pela janela W . Tomando-se o quadrado do módulo da STFT, obtém-se uma representação tempo – freqüência do sinal chamada de espectrograma. O espectrograma mostra como a energia do sinal está distribuída conjuntamente em tempo e em freqüência. Ou seja, através do espectrograma, o sinal representado originalmente em uma dimensão passa a ser representado em duas dimensões: tempo, τ , e freqüência, w . O espectrograma, ou qualquer outra representação tempo – freqüência quadrática, é tipicamente visualizado como uma imagem, onde a intensidade representa a energia e os eixos x e y são o tempo e a freqüência, respectivamente.

Como o espectro referente a uma localização t , corresponde à transformada de Fourier de um sinal segmentado pela janela W em t , conclui-se que o tamanho da janela tem papel fundamental na resolução em tempo e em freqüência da análise realizada. A figura 34 ilustra quatro espectrogramas obtidos para um mesmo sinal sintetizado com quatro diferentes freqüências localizadas no tempo. O sinal utilizado como exemplo possui 256 amostras, sendo composto por quatro senóides concatenadas, cada uma com 64 amostras e freqüências de 0,5 Hz, 0,250 Hz, 0,125 Hz e 0,0625 Hz respectivamente. Nos espectrogramas ilustrados nas figuras 34.a e 34.b foram utilizadas janelas no tempo com 15 e 31 amostras respectivamente, e pode-se observar que o resultado obtido usando uma janela de análise curta no tempo, possui uma boa localização dos eventos no tempo sem, no entanto, obter uma boa resolução em freqüência. Já nas figuras 34.c e 34.d, as análises foram realizadas com janelas de análise mais longas no tempo, com 65 e 127 amostras respectivamente, e o

resultado obtido possui uma boa resolução em frequência, mas uma localização ruim dos eventos no tempo.

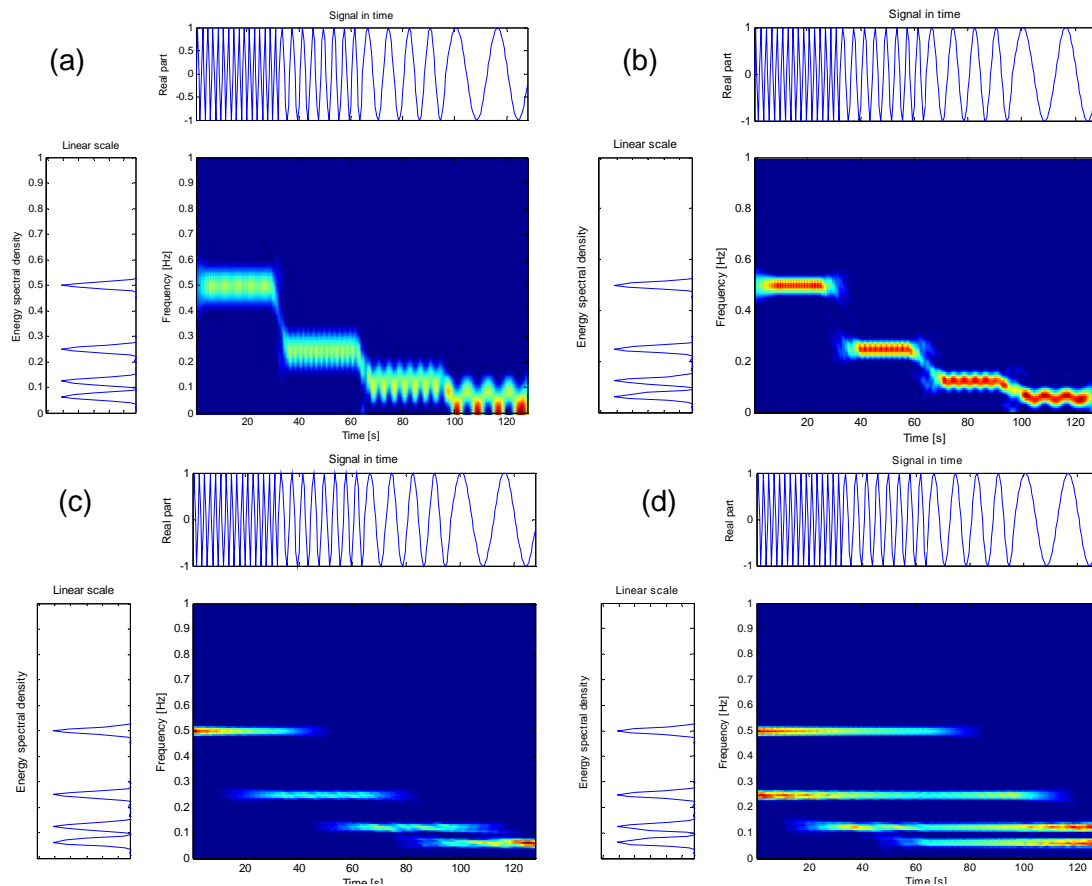


Figura 34 Análise tempo – frequência através do espectrograma de um sinal de teste composto por quatro senóides concatenadas de frequências de 0,5 Hz, 0,250 Hz, 0,125 Hz e 0,0625 Hz respectivamente. (a) espectrograma utilizando janela de análise com 15 amostras; (b) com 31 amostras; (c) com 63 amostras e (d) com 127 amostras. (Auger et al., 1996).

O tipo de janela utilizada também é bastante importante na análise e normalmente são utilizadas janelas do tipo gaussianas, de Hamming, de Hanning, de Kaiser, etc (Auger et al., 1996).

Levando-se em conta o problema da falta de resolução conjunta no tempo e em frequência da STFT devido ao tipo e tamanho da janela utilizada, uma forma efetiva para minimizar este efeito é utilizar uma janela que seja associada ao sinal. Quando a janela utilizada é o próprio sinal invertido, diz-se que a distribuição formada pela eq.(26) é de Wigner-Ville. A figura 35 mostra as distribuições de Wigner-Ville de dois sinais senoidais, modulados em amplitude por um sinal gaussiano, separadamente, enquanto a figura 36 mostra a distribuição dos dois sinais somados. Percebe-se que para uma única oscilação

modulada por uma gaussiana a distribuição tempo – frequência é ótima (Qian, 2002). Entretanto, para os dois sinais gaussianos somados existem termos cruzados que não representam informação alguma a respeito do sinal.

$$WV(t, \omega) = \frac{1}{2\pi} \int x^* \left(t - \frac{1}{2} \tau \right) x \left(t + \frac{1}{2} \tau \right) e^{-j\omega\tau} d\tau \quad (26)$$

A distribuição de Wigner-Ville é de fundamental importância, apesar dos problemas existentes quanto aos termos cruzados, pois serve de base para outras representações tempo – frequência e possui diversas propriedades matemáticas importantes, como a que preserva o deslocamento no tempo e em frequência, sempre tendo como resultado valores reais.

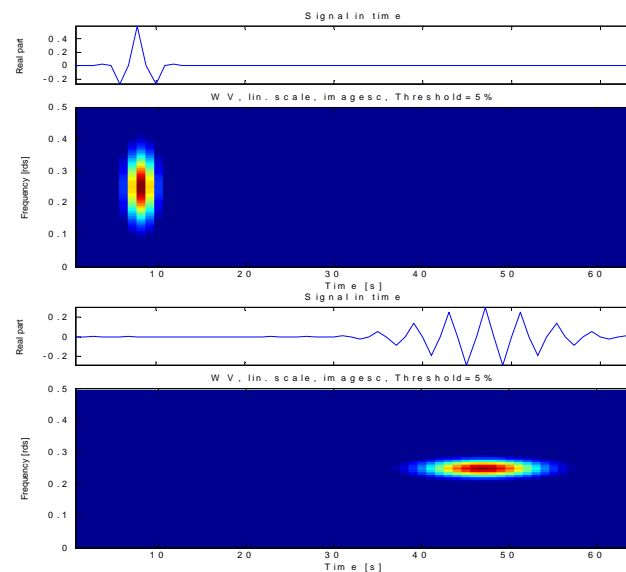


Figura 35 Distribuição de Wigner-Ville de dois sinais modulados por uma gaussiana.

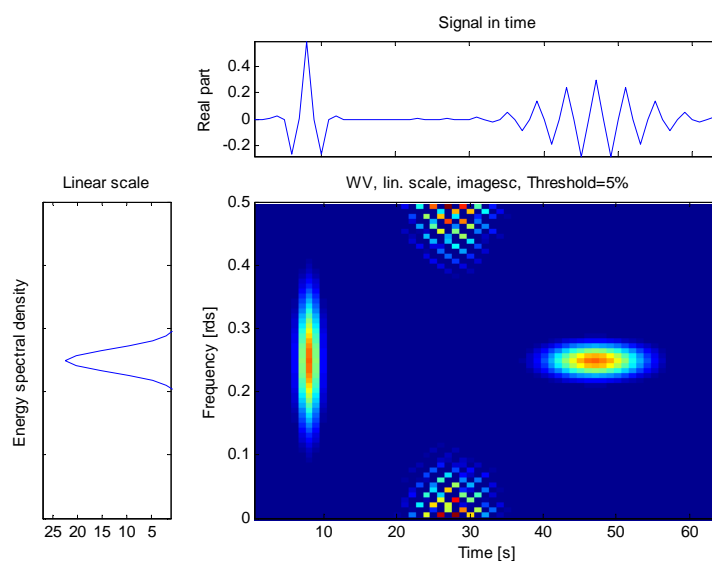


Figura 36 Distribuição de Wigner-Ville para um sinal modulado por duas gaussianas deslocadas no tempo.

Uma das formas de reduzir os termos cruzados introduzidos é através da suavização da distribuição ao longo do tempo e da frequência como representado pela eq.(27).

$$C(t, \omega) = \frac{1}{4\pi} \int \int \int x^* \left(u - \frac{1}{2} \tau \right) x \left(u + \frac{1}{2} \tau \right) \phi(\theta, \tau) e^{-j\theta t - j\tau \omega + j\theta u} du d\tau d\theta \quad (27)$$

Existem diferentes funções de suavização, $\phi(\theta, \tau)$, chamadas de núcleo ou kernel da representação tempo – frequência (Auger et al., 1996). Outras distribuições tempo – frequência, como por exemplo o espectrograma, podem ser consideradas como caso particular da representação conhecida como classe de Cohen (Cohen, 1995). Nesta tese, estas formas de representação não serão utilizadas diretamente.

Outra forma de se realizar a análise tempo – frequência pode ser obtida decompondo-se adaptativamente o sinal em partes com distribuição tempo – frequência bem conhecidas.

4.2. “Matching pursuit” com átomos de Gabor

Uma forma bastante flexível para se analisar sinais em tempo e em frequência conjuntamente é através do algoritmo de “matching pursuit”. Segundo Mallat (Mallat e Zhang, 1993), qualquer sinal pode ser decomposto através do algoritmo de “matching pursuit” em uma expansão linear de formas de onda devidamente selecionadas de um conjunto de funções redundantes chamado de dicionário. Basicamente, o dicionário é um conjunto de funções pertencentes a um espaço de Hilbert H , $D = \{g_1, g_2, \dots, g_n\}$ e normalizadas tal que $\|g_i\| = 1$. Supõe-se que o dicionário utilizado é completo, o que significa que as expansões lineares finitas de elementos em D são densas em H (Mallat, 2001). Cada função g_i do dicionário também é chamada de átomo.

Uma família geral de átomos com boa representação tempo – frequência pode ser gerada através do escalonamento, translação e modulação de uma função janela simples $g(t)$, desde que a função $g(t) \in L^2(\mathbb{R})$ seja real, contínua, diferenciável, $O(1/(t^2+1))$, possua $\|g\| = 1$, a integral de $g(t)$ seja maior que zero e $g(0) \neq 0$. Partindo dessas premissas, define-se então, para qualquer escala $s > 0$, frequência de modulação ξ e translação u , o átomo $g_\gamma(t)$:

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{j\xi t} \quad (28)$$

onde $\gamma = (s, u, \xi)$ e o fator $1/\sqrt{s}$ normaliza em 1 a norma de $g_\gamma(t)$.

Como, usualmente, $g(t)$ é uma função par, conseqüentemente, a função $g_\gamma(t)$ é centralizada em u com a maior parte de sua energia distribuída em função da escala s . Tomando a transformada de Fourier de $g_\gamma(t)$, obtém-se:

$$\hat{g}_\gamma(w) = \sqrt{s} \hat{g}(s(w - \xi)) e^{-j(w - \xi)u} \quad (29)$$

Como $g(t)$ é real, então $|\hat{g}(w)|$ é par e $|\hat{g}_\gamma(w)|$ é centralizada em $w = \xi$ com a sua energia distribuída em torno de ξ com um espalhamento proporcional à $1/s$. Logo, cada átomo gerado $g_\gamma(t)$ é localizado no tempo em u com espalhamento relacionado à escala s , enquanto em freqüência é localizado em ξ com espalhamento relacionado ao inverso da escala s , ou seja, independente da função $g(t)$ arbitrada, não é possível uma localização perfeita em tempo e em freqüência simultaneamente (Cohen, 1995).

O dicionário formado $D = (g_\gamma(t))_{\gamma \in \mathcal{R}^+ \times \mathcal{R}^2}$ é extremamente redundante e qualquer função $f(t)$ pode ser representada eficientemente através da seleção apropriada de um subconjunto finito de átomos $(g_{\gamma_n}(t))_{n \in N}$, onde $\gamma_n = (s_n, u_n, \xi_n)$, como a combinação linear dos M átomos com maior proximidade em relação ao sinal:

$$f(t) = \sum_{n=0}^M a_n g_{\gamma_n}(t) \quad (30)$$

Dependendo da escolha do dicionário, certas propriedades do sinal poderão ser realçadas ou não. Por exemplo, pode-se interpretar a STFT como um caso particular de decomposição através de um dicionário de funções. Basta considerar que todos os átomos do dicionário possuem a mesma escala s , que para este caso em especial, é o comprimento da janela utilizada para o cômputo da transformada de Fourier. Logo, como apresentado anteriormente através da figura 34, a STFT não representa bem sinais com suporte muito menor ou muito maior que a escala s . Cabe ressaltar, que para o caso da STFT o dicionário formado para análise é mínimo e constitui uma base ortonormal em $L^2(\mathcal{R})$.

Outro caso especial de base ortonormal ou bi-ortogonal que pode ser interpretado como um dicionário mínimo é a transformada wavelet, que ao contrário da STFT, decompõe o sinal em átomos que variam a escala. A família de funções wavelet $(g_{\gamma_n}(t))_{n \in N}$ é construída relacionando o parâmetro freqüência ξ_n à escala s_n como $\xi_n = \xi_0/s_n$. O dicionário é composto por dilatações e deslocamentos de uma função base, chamada de wavelet mãe, multiplicada por um parâmetro de fase complexo. Os coeficientes a_n obtidos após a expansão

com a família de funções wavelet caracterizam bem o comportamento em diferentes escalas do sinal, como as transições, tendências e comportamento fractal. Entretanto, devido à restrição do parâmetro de frequência ξ_n , a expansão em coeficientes wavelet não provê estimativa precisa do conteúdo em frequência das formas de onda analisadas (Mallat & Zhang, 1993).

Dentre os dicionários de funções existentes, o dicionário de Gabor provê localização ótima em tempo e frequência (Jaffard et al., 2001) e é construído a partir da definição da função $g(t)$ como uma função gaussiana. Portanto, o dicionário de Gabor é formado a partir da dilatação, translação e modulação, representadas, respectivamente, pelos parâmetros s , u e ξ na eq.(31), de uma função base gaussiana:

$$g_{\gamma}(t) = \frac{1}{\sqrt{s}} e^{-\pi \left(\frac{t-u}{s}\right)^2} e^{j\xi t} \quad (31)$$

Nesta tese, o dicionário de Gabor foi escolhido para análise de sinais sísmicos devido a sua representação ótima em tempo e em frequência e também devido a sua semelhança com sinais gerados por fontes sísmicas.

Definido o dicionário, o algoritmo de “matching pursuit” busca de uma forma adaptativa encontrar uma aproximação ótima ao sinal f . Considerando, que a aproximação é obtida a partir da expansão das M funções escolhidas apropriadamente do dicionário, o algoritmo busca minimizar o erro de aproximação dado pela eq.(32):

$$\varepsilon = \left\| f(t) - \sum_{i=1}^M w_i g_{\gamma_i}(t) \right\| \quad (32)$$

onde $\{g_{\gamma_i}\}_{i=1,2,\dots,M}$ representa os índices das funções escolhidas no dicionário g_{γ_i} .

O algoritmo “matching pursuit” busca passo a passo pela função, ou átomo, g_{γ_i} que melhor se aproxima do sinal $f(t)$. A primeira etapa consiste em buscar no dicionário, qual o átomo g_{γ_0} , ou função, que mais se aproxima do sinal. O átomo com maior similaridade, medida através do produto interno no espaço $L^2(R)$ do sinal pelos átomos do dicionário, é escolhido e o sinal passa a ser representado pela projeção de f no átomo vencedor acrescido de resíduo Rf :

$$f = \langle f, g_{\gamma_0} \rangle g_{\gamma_0} + Rf \quad (33)$$

Como g_{γ_0} é ortogonal à Rf , então:

$$\|f\|^2 = \left| \langle f, g_{\gamma_0} \rangle \right|^2 + \|Rf\|^2 \quad (34)$$

Após a primeira iteração, o processo passa a buscar o átomo do dicionário que mais se aproxima da função residual obtida após a subtração do sinal original pelo resultado obtido nas iterações anteriores. Logo, o algoritmo de “matching pursuit” é um processo iterativo que decompõe o resíduo obtido após cada iteração em função da melhor similaridade com as funções pré-definidas no dicionário. Definindo a função f como o resíduo de nível zero, o processo pode ser descrito através da eq.(35):

$$\begin{cases} R^0 f = f \\ R^n f = \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} f \\ g_{\gamma_n} = \arg \max_{g_{\gamma_i} \in D} \left| \langle R^n f, g_{\gamma_n} \rangle \right| \end{cases} \quad (35)$$

Pode ser provado, por indução matemática (Mallat & Zhang, 1993), que se o dicionário for completo, o procedimento descrito acima converge e face à ortogonalidade existente entre $R^{n+1}f$ e g_{γ_n} em cada iteração do processo, o algoritmo conserva a energia do sinal, como descrito na eq.(36):

$$\|f\|^2 = \sum_{n=0}^{m-1} \left| \langle R^n f, g_{\gamma_n} \rangle \right|^2 + \|R^m f\|^2 \quad (36)$$

Como o algoritmo preserva a energia, a distribuição tempo – frequência do sinal pode ser descrita como a soma das distribuições de Wigner-Ville das funções selecionadas do dicionário, que dependendo do dicionário arbitrado, podem ser individualmente livres de termos cruzados, como descrito na eq.(37):

$$Ef(t, w) = \sum_{n=0}^{m-1} \left| \langle R^n f, g_{\gamma_n} \rangle \right|^2 W_{g_{\gamma_n}}(t, w) \quad (37)$$

Portanto, para o caso especial do dicionário de Gabor, a representação tempo – frequência é livre de termos cruzados, sendo a soma das representações individuais de cada átomo.

Mallat e Zhang (Mallat & Zhang, 1993) também demonstraram que o algoritmo de “matching pursuit” se aplica a espaços finitos, sendo, basicamente, implementado escolhendo os átomos que mais se aproximam dos resíduos da função f . Um algoritmo similar para decomposição tempo – frequência foi desenvolvido de forma independente por Qian e Chen (Qian & Chen, 1988).

Em suma, após m iterações o sinal f pode ser representado por uma combinação linear das m projeções do sinal nos átomos mais representativos do

dicionário em relação ao sinal a menos de um erro devido ao resíduo que tende a zero, quando M tende a infinito:

$$f = \sum_{n=0}^{m-1} \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} + R^m f \quad (38)$$

Quanto menor o número de átomos escolhidos do dicionário para a aproximação com um determinado erro dado pela eq.(38), melhor a coerência do sinal com o dicionário escolhido. O conceito de coerência para o algoritmo de “matching pursuit” foi introduzido por Mallat e Zhang (Mallat & Zhang, 1993) para verificar o nível de aderência do sinal ao dicionário. O mesmo conceito foi utilizado como base para a aplicação do algoritmo como ferramenta de remoção de ruído. Assim, quando o dicionário possui uma grande coerência com o sinal analisado, o erro obtido ao representar sinais com poucos átomos é aceitável e o resíduo pode ser desprezado. Este conceito também pode ser aplicado a compressão de sinais (Yukhananov et al., 2000).

Para exemplificar o processo de decomposição descrito acima, foram tomados como sinais de exemplo, três sinais sísmicos sintéticos, retirados cada sinal de uma diferente fície do modelo sintético utilizado como exemplo no Capítulo 3 e apresentado na figura 28. As figuras 37, 39 e 41 ilustram para cada um dos três sinais as projeções a_i dos sinais em cada átomo do dicionário, o traço reconstruído fr com uma aproximação de 90% do sinal original e o traço original f . O número de átomos utilizados para a reconstrução de cada sinal com a exatidão estipulada foi de 9, 8 e 10, respectivamente, o que mostra uma grande coerência entre os sinais sintetizados e o dicionário.

Outra forma para verificação da coerência do sinal com o dicionário é através da propriedade de conservação de energia do algoritmo de “matching pursuit”, expressada pela eq.(36). As figuras 38, 40 e 42 ilustram, a contribuição de cada átomo em relação à energia total do sinal, $\left| \langle R^n f, g_{\gamma_n} \rangle \right|^2 / \|f\|^2$, via gráfico de barras, e a participação acumulada destes átomos também em relação à energia do sinal, via gráfico de linha. Da eq.(36), sabe-se que a energia acumulada em relação à energia total tende ao valor unitário, e observa-se, que para os três exemplos, apenas quatro átomos representam pelo menos 90% da energia total do sinal. Na figura 43, o traço sintético que representa a fície 1 do modelo utilizado é analisado através do algoritmo de “matching pursuit” com 9 e 4 átomos respectivamente. Observa-se que apesar das reconstruções dos sinais não serem perfeitas com poucos átomos, as representações tempo – frequência são bastante similares, indicando um evento

em baixa frequência localizado próximo a amostra 34. Tal fato indica que apenas alguns átomos definem a forma do sinal, enquanto os restantes definem os detalhes do sinal.

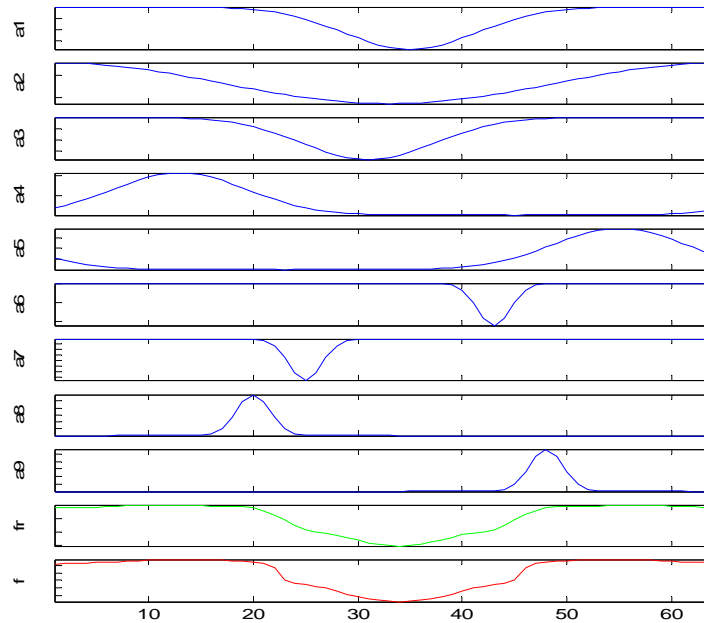


Figura 37: Decomposição utilizando algoritmo “matching pursuit” de um sinal sísmico sintético, f , ilustrando as 09 principais projeções a_i do sinal nos respectivos átomos do dicionário, o traço reconstruído fr com uma aproximação de 90% do sinal original e o traço original.

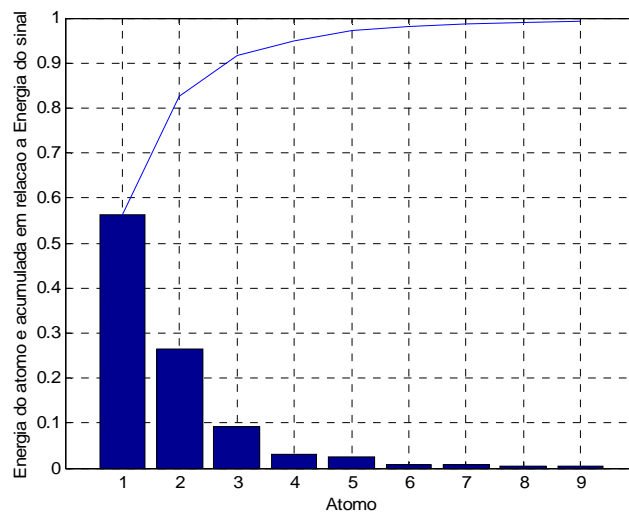


Figura 38 O gráfico de barras indica a contribuição de energia de cada átomo em relação à energia total do sinal, enquanto o gráfico de linha indica a contribuição de energia acumulada pelos átomos mais importantes em relação à energia total do sinal.

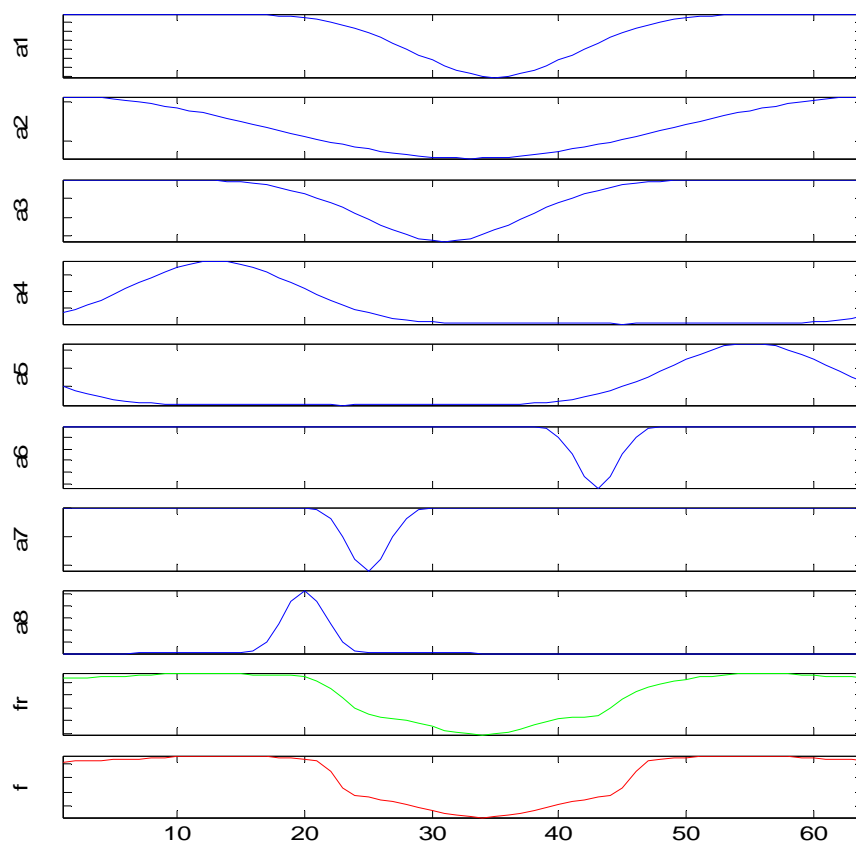


Figura 39 Decomposição utilizando algoritmo “matching pursuit” de um sinal sísmico sintético, f , ilustrando as 08 principais projeções a_i do sinal nos respectivos átomos do dicionário, o traço reconstruído fr com uma aproximação de 90% do sinal original e o traço original.

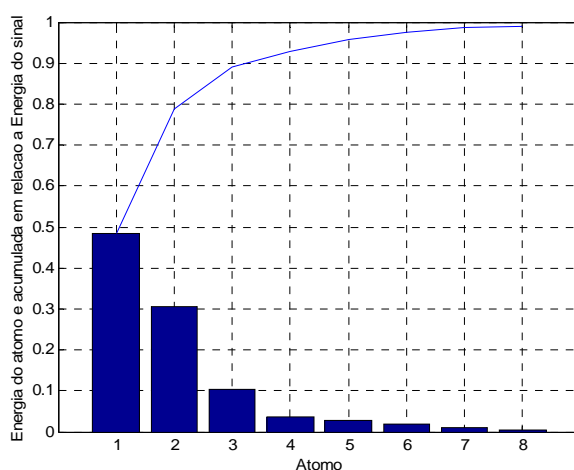


Figura 40 O gráfico de barras indica a contribuição de energia de cada átomo em relação à energia total do sinal, enquanto o gráfico de linha indica a contribuição de energia acumulada pelos átomos mais importantes em relação à energia total do sinal.

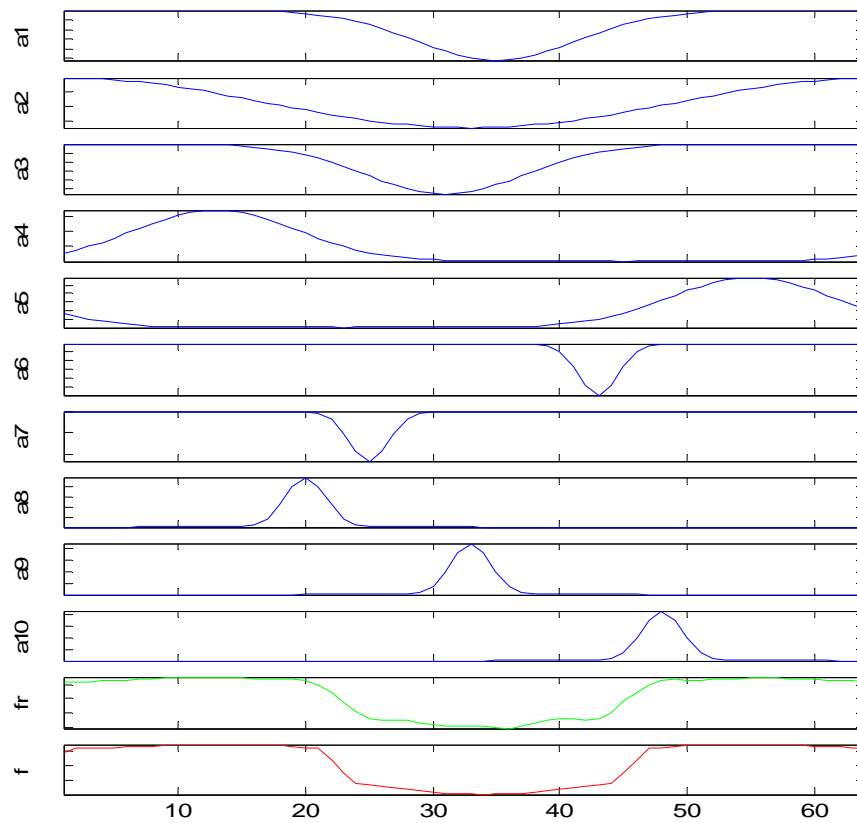


Figura 41 Decomposição utilizando algoritmo “matching pursuit” de um sinal sísmico sintético, f , ilustrando as 09 principais projeções a_i do sinal nos respectivos átomos do dicionário, o traço reconstruído fr com uma aproximação de 90% do sinal original e o traço original.

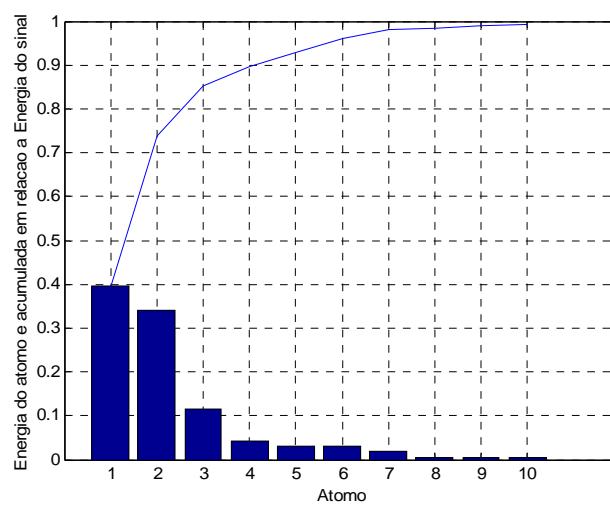


Figura 42 O gráfico de barras indica a contribuição de energia de cada átomo em relação à energia total do sinal, enquanto o gráfico de linha indica a contribuição de energia acumulada pelos átomos mais importantes em relação à energia total do sinal.

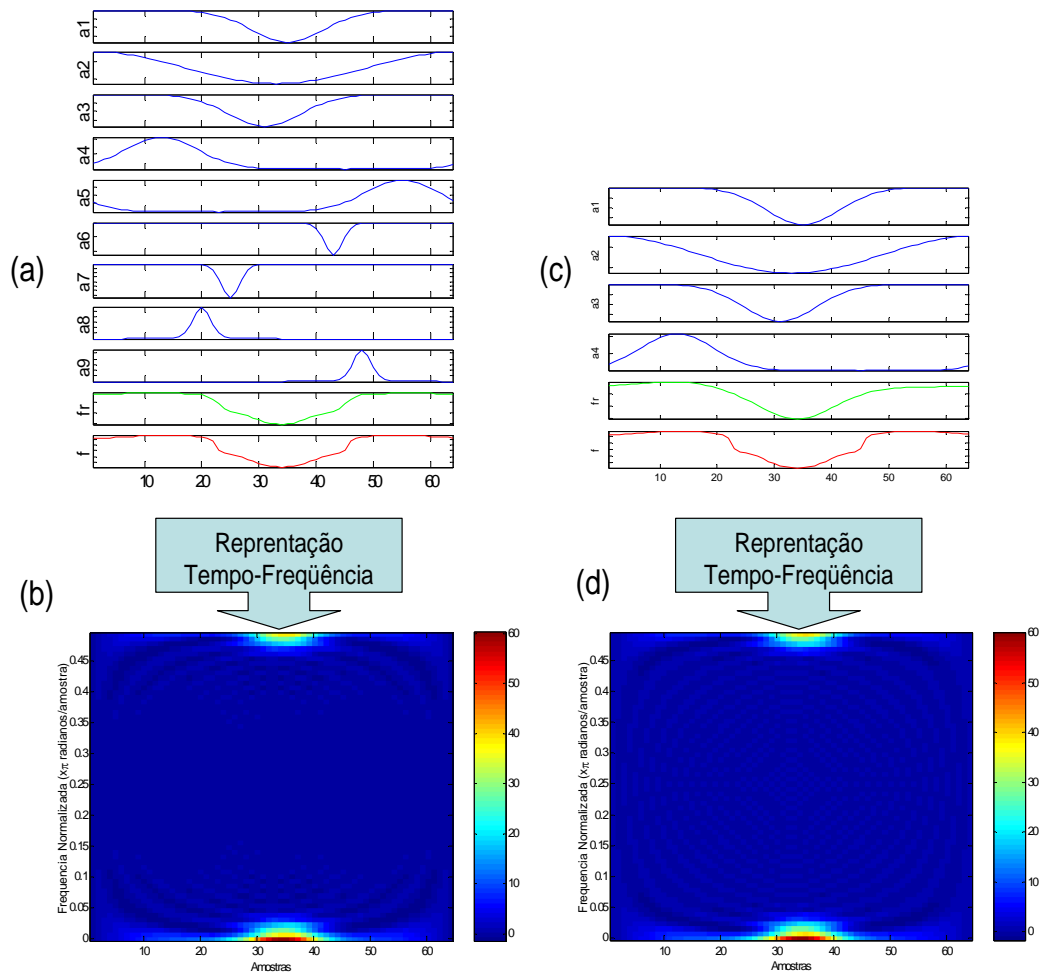


Figura 43 Decomposição utilizando “matching pursuit” com 9 e 4 átomos e suas respectivas representações tempo – freqüência.

Portanto, para a localização no tempo e em freqüência de eventos sísmicos, como os sintetizados, e para a obtenção de uma aproximação da forma de onda, basta utilizar alguns átomos obtidos através da decomposição utilizando o algoritmo de “matching pursuit”. A figura 44 ilustra a reconstrução dos mesmos três traços sintéticos utilizando apenas quatro átomos do dicionário. Pode-se observar que os sinais reconstruídos com apenas quatro funções aproximam, de forma bastante razoável, os eventos de interesse. Verifica-se que os sinais reconstruídos têm uma forma bastante semelhante ao traço original.

A compactação de energia em poucos átomos observada com o algoritmo de “matching pursuit” sugere uma possível semelhança com a transformada de Karhunen-Loève ou análise de componentes principais. Entretanto, enquanto no algoritmo de “matching pursuit” o processo de otimização é realizado amostra a amostra, sugerindo uma ação localizada em função das similaridades do sinal

com o dicionário de funções utilizado, na transformada de Karhunen-Loève a transformação é baseada na estatística estacionária do sinal, o que normalmente não acontece para sinais com poucas amostras (Goyal, 1995), (Goyal et al., 1995). Portanto, segundo Goyal (Goyal, 1995), geralmente, a transformada de “matching pursuit” consegue uma maior compactação de energia que a transformada de Karhunen-Loève.

Outra característica importante do algoritmo de “matching pursuit” é a capacidade de localização de eventos deslocados no tempo. No exemplo ilustrado na figura 45, verifica-se que se um sinal for deslocado no tempo, a sua representação tempo – frequência também será deslocada. Logo, como esperado, os eventos de interesse podem ser localizados tanto no tempo quanto em frequência.

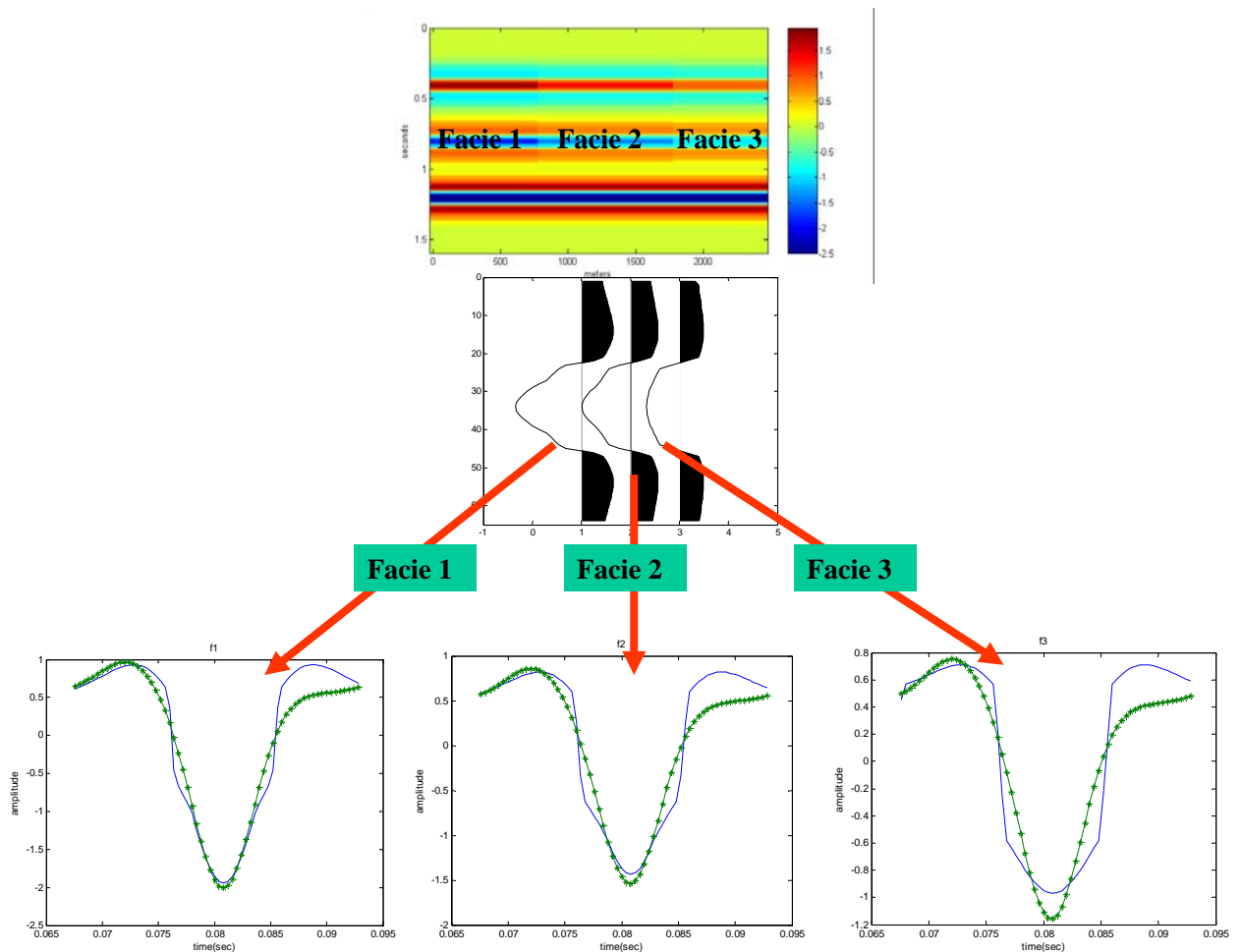


Figura 44 Reconstrução dos traços sintéticos com três fácies diferentes usando o “matching pursuit” com os quatro átomos de Gabor mais significativos em verde.

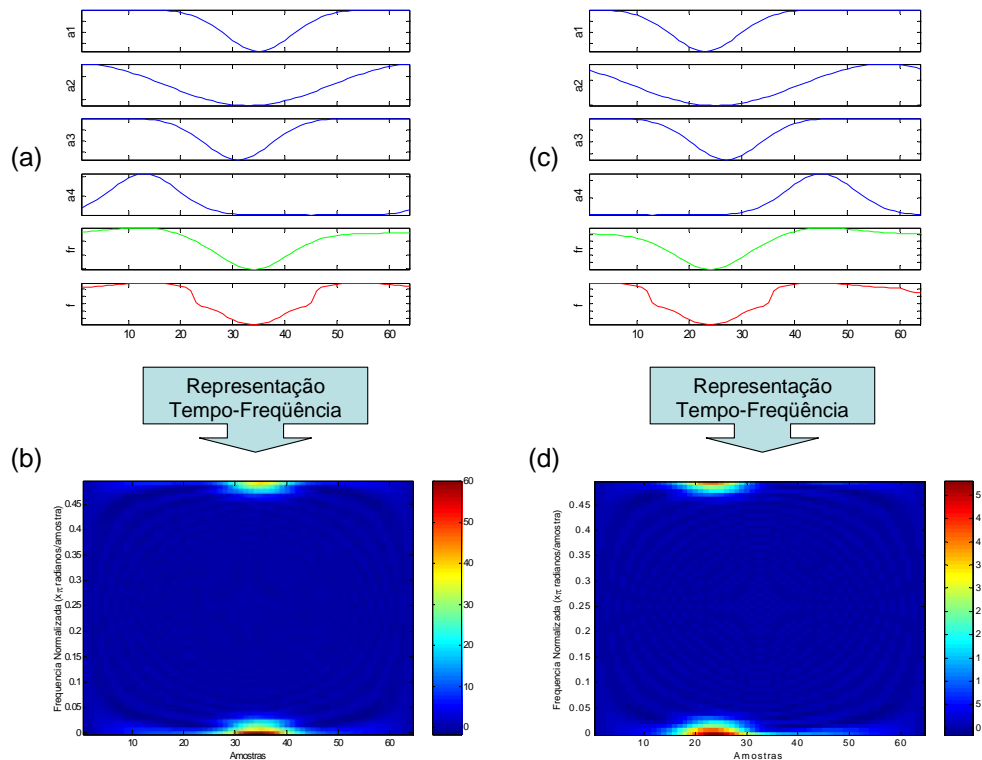


Figura 45 Representações tempo – freqüência obtidas através do algoritmo de “matching pursuit” de um traço sintético e o seu correspondente deslocado no tempo.

4.3. Atributos sísmicos e classificação de padrões sísmicos utilizando “matching pursuit”

A caracterização de sinais sísmicos através da identificação de atributos sísmicos gerados a partir de representações tempo – freqüência é um tópico ainda não muito explorado. Uma das abordagens, propõe a geração de atributos sísmicos, tais como freqüência instantânea, freqüência média e outros similares, através dos momentos em freqüência da representação tempo – freqüência, não sendo, necessariamente, gerados a partir do algoritmo de “matching pursuit” (Steeghs & Drijkoningen, 2001), (Baraniuk et al., 2001).

Uma outra forma para caracterização de sinais sísmicos utilizando o algoritmo de “matching pursuit” foi proposta por Verhelst (Verhelst, 1998), (Verhelst, 2000) e consiste em utilizar os parâmetros obtidos, após apenas uma iteração do algoritmo de “matching pursuit”, como atributos sísmicos. Para tal, o dicionário usado também foi limitado a funções com parâmetro de modulação em freqüência, ξ , igual a um. Em suma, o referido autor propõe aproximar o traço sísmico pelo átomo com maior semelhança do dicionário com freqüência $\xi=1$, sugerindo também a utilização do parâmetro fase como um possível

discriminador de fácies sísmicas. Entretanto, sabe-se que, tipicamente, sinais sísmicos possuem grande parte do seu conteúdo em baixas frequências, o que é exatamente o oposto do sugerido.

Conforme descrito no item anterior, qualquer sinal analisado através do algoritmo de “matching pursuit” com dicionário de Gabor pode ser representado apenas pelos parâmetros $\gamma_n = (s_n, u_n, \xi_n)$ e $\langle R^n f, g_{\gamma_n} \rangle$ dos átomos selecionados. A representação do sinal através destes parâmetros carrega toda a informação a respeito do sinal e pode ser usada para representá-lo.

Considerando o parâmetro u_n como um atributo sísmico e analisando as figuras 37, 39, 41, 43 e 45, verifica-se que a localização do primeiro átomo, u_1 , marca o evento de interesse e pode ser utilizada como ferramenta para interpretação de horizontes. Tomando o dado utilizado na figura 31 do Capítulo 3, que foi obtido após a segmentação temporal de um dado sintético utilizando um horizonte acrescido de ruído gaussiano, a figura 46 mostra o atributo deslocamento u_1 do principal átomo, obtido após decomposição com o algoritmo de “matching pursuit”, representado pela linha branca. Verifica-se que o atributo localiza com bastante precisão o horizonte procurado. No referido exemplo, o tamanho da janela utilizada para análise é de 64 amostras, tamanho maior que o usual para caracterização de reservatórios de petróleo. Portanto, conforme descrito e ilustrado na figura 45, a localização de eventos no tempo pode ser utilizada como poderosa ferramenta para interpretação sísmica.

Caso se deseje levar em conta também os outros átomos para a obtenção de um atributo relacionado a localização no tempo, pode-se utilizar o tempo médio da representação tempo – frequência (Flandrin, 1999).

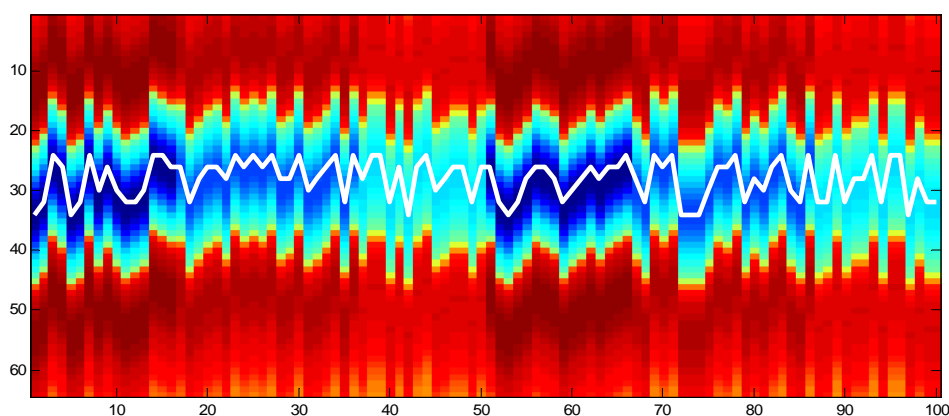


Figura 46 Interpretação sísmica utilizando a localização do átomo mais importante da decomposição

O procedimento para localização de eventos no tempo pode ser estendido para os outros parâmetros obtidos através do algoritmo de “matching pursuit”. Mas, como sugerido para o caso particular da decomposição de sinais sísmicos sintéticos, com a utilização de apenas quatro principais átomos é possível representar 90% da energia do sinal. Tal fato sugere que a utilização conjunta dos quatro parâmetros como atributos sísmicos, associada ao algoritmo de agrupamento do Mapas de Kohonen, proposto no Capítulo 3, possa ser utilizada como uma ferramenta de análise de fácies sísmicas.

A utilização direta dos parâmetros $\gamma_n=(s_n, u_n, \xi_n)$ e $\langle R^n f, g_{\gamma_n} \rangle$, como atributos sísmicos de entrada para o sistema de análise de fácies sísmicas utilizando o agrupamento do SOM, não é recomendada, pois o parâmetro deslocamento no tempo u_n é sensível a deslocamentos no tempo, como ilustrado na figura 45. Uma forma de se evitar este problema é tomar o deslocamento relativo dos átomos em relação ao primeiro átomo escolhido pelo algoritmo de “matching pursuit”. Assim, fixando a decomposição em 4 níveis, os seguintes atributos em relação ao deslocamento são gerados: $\delta_1=u_2-u_1, \delta_2=u_3-u_1, \delta_3=u_4-u_1$.

A utilização destes atributos não inviabiliza a caracterização do sinal, mas apenas isola o efeito de deslocamento no tempo. Este efeito pode também ser entendido através da representação tempo – frequência de dois sinais idênticos defasados no tempo, como ilustrado na figura 45, cujas diferenças são apenas devidas aos termos referentes aos deslocamentos no tempo u_n e podem ser retirados como sugerido.

Retirado o efeito de deslocamento no tempo, um vetor de atributos que caracteriza totalmente o sinal analisado é montado da seguinte forma:

$$\text{MPA } 1 = [\delta_p; \gamma_n=(s_n, u_n, \xi_n) \text{ e } \langle R^n f, g_{\gamma_n} \rangle]; \quad (39)$$

onde $p=n-1$, e, normalmente, $n=4$.

Considerando que a representação tempo – frequência, obtida após a decomposição utilizando o algoritmo de “matching pursuit”, pode ser considerada como a soma de “bolhas” de energias de gaussianas (Mallat & Zhang, 1993) cujas localizações e variâncias no tempo e em frequência são determinadas pelos parâmetros $\gamma_n=(s_n, u_n, \xi_n)$, outra opção para a construção de um vetor de atributos que caracterizem o sinal sísmico pode ser utilizada substituindo o termo $\langle R^n f, g_{\gamma_n} \rangle$, pela energia de cada átomo $\left| \langle R^n f, g_{\gamma_n} \rangle \right|^2$, como descrito na eq.(40). Esta abordagem assume que a distribuição tempo – frequência funciona como

uma descrição do comportamento dos sinais e pode ser utilizada como um padrão.

$$\text{MPA 2} = [\delta_p; \gamma_n = (s_n, u_n, \xi_n) \text{ e } \left\langle R^n f, g_{\gamma_n} \right\rangle^2]; \quad (40)$$

onde $p=n-1$, e, normalmente, $n=4$.

4.4. Algoritmo proposto utilizando agrupamento do SOM associado ao “matching pursuit”

Como os vetores de atributos representados pelas eq.(39) e eq.(40) podem ser utilizados para caracterizar traços sísmicos, então o agrupamento destes atributos utilizando o SOM pode ser utilizado para análise de fácies sísmicas de forma conjunta, como descrito a seguir:

- a) Segmentação espacial e temporal com orientação geológica;
- b) Decomposição dos traços sísmicos utilizando o algoritmo de “matching pursuit” e montagem do vetor de atributos através da eq.(39) ou eq.(40);
- c) Formação e treinamento do SOM com número de vetores protótipos muito maior que o número esperado de fácies sísmicas;
- d) Utilizando a visualização da U-matrix do SOM em comparação com o índice de Davies-Bouldin para diferentes números de agrupamentos é estimado o número de fácies sísmicas;
- e) Clusterização e rotulação dos vetores protótipos do SOM utilizando o algoritmo partitivo K-means;
- f) Após os elementos do SOM serem rotulados com o número de fácies sísmicas estimada, os atributos sísmicos, para cada ponto do reticulado de entrada, são comparados com os vetores protótipos do SOM e então classificados de acordo com a fácies do agrupamento mais próximo.
- g) Construção e Interpretação dos mapas de fácies sísmicas.

4.4.1. Análise de fácies sísmicas de um dado sintético

Para verificação do método proposto, o mesmo dado sintético obtido através de uma interpretação ruidosa, ilustrado nas figuras 47a e 48a, foi analisado utilizando para geração de atributos as eq.(40) e eq.(39) respectivamente. As figuras 47c e 48c ilustram os mapas de fácies sísmicas obtidos e verifica-se que os resultados foram excelentes se comparados com o resultado obtido anteriormente e também ilustrado nas figuras 47b e 48b.

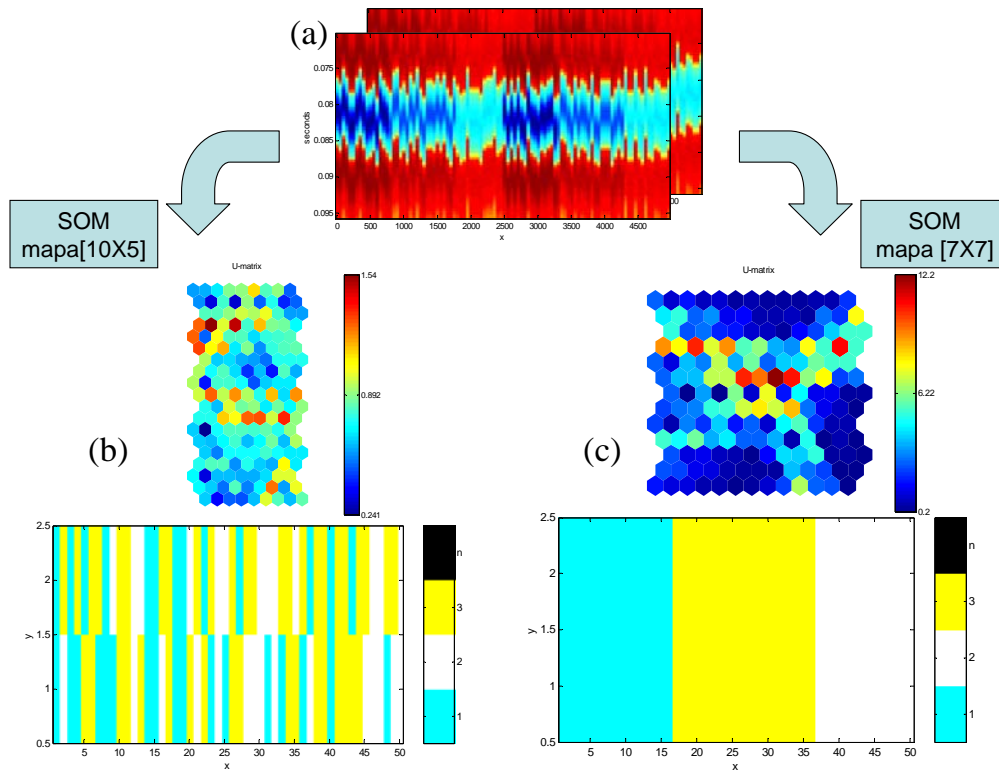


Figura 47: Análise de fácies sísmicas de dados sintéticos obtidos através de um interpretação ruidosa; (a) dado sísmico; (b) análise utilizando forma de onda como atributo; (c) análise utilizando atributos obtidos através da eq.(40) via “matching pursuit”.

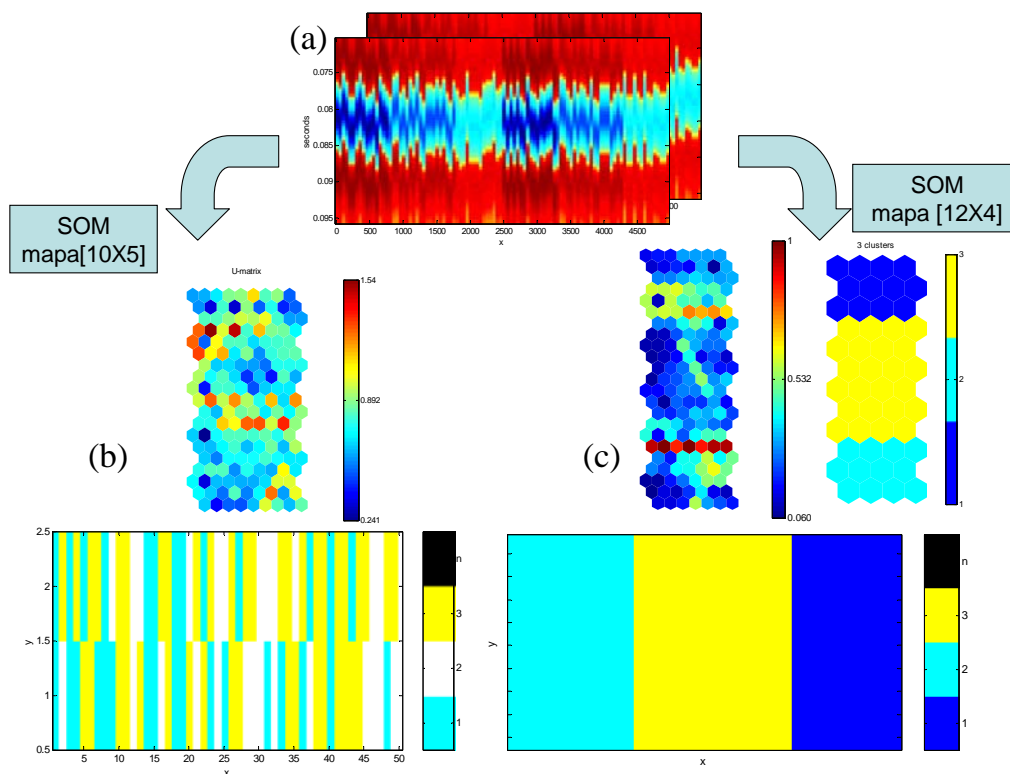


Figura 48: Análise de fácies sísmicas de dados sintéticos obtidos através de um interpretação ruidosa; (a) dado sísmico; (b) análise utilizando forma de onda como atributo; (c) análise utilizando atributos obtidos através da eq.(39) via “matching pursuit”.

Uma vantagem esperada do método e confirmada no resultado ilustrado é a diminuição dos efeitos de localização dos eventos no tempo.

4.4.2. Análise de fácies sísmicas de um dado real

O mesmo dado sísmico 3D real utilizado no item 3.4.2 foi utilizado para verificação do método proposto utilizando “matching pursuit” e SOM.

No exemplo ilustrado na figura 33, foram utilizados, como atributos sísmicos, as amplitudes instantâneas interpoladas dentro de uma unidade estratigráfica delimitada via uma interpretação cuidadosa e baseada em outros dados obtidos em vários poços perfurados na região. Já para a análise utilizando o método proposto com “matching pursuit”, a região foi segmentada sem muita rigidez. Uma região utilizando dezesseis amostras em torno do horizonte que delimita a topo do reservatório foi selecionada, tomando os atributos através da eq.(39). O resultado da análise de fácies sísmicas é apresentado na figura 49.

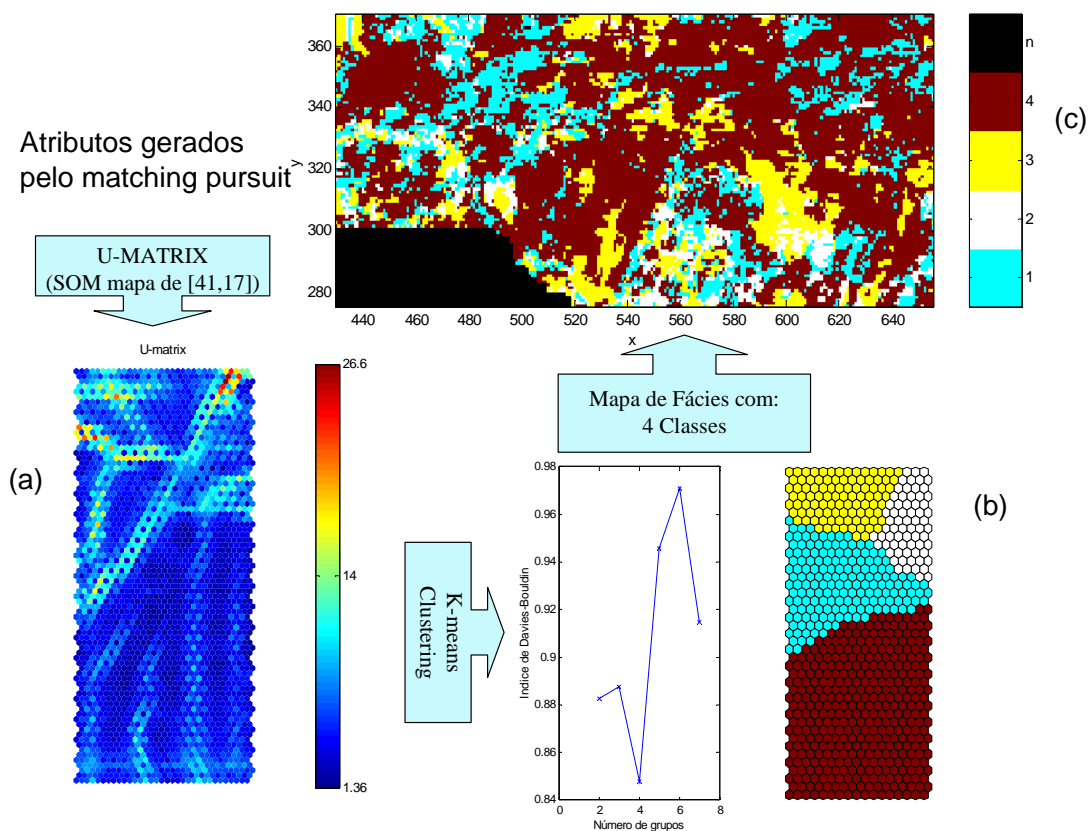


Figura 49: Análise de fácies sísmicas de um dado sísmico real utilizando o método proposto; (a) Matriz-U obtida com o método de agrupamento do SOM; (b) Análise dos grupos formados utilizando o algoritmo K-means; (c) Mapa de fácies sísmicas.

Observa-se na figura 49a que é possível a identificação de grupos na matriz-U, ao contrário do método utilizando somente a forma de onda ilustrada na figura 33. O número de grupos proposto através do índice mínimo de Davies-Bouldin na figura 49b coincide com a análise estratigráfica (Johann, 1999), embora a matriz-U não indique, necessariamente, a existência de apenas quatro grupos.

A capacidade de localização dos eventos no tempo e em frequência do método proposto permite que a segmentação temporal não seja muito rígida, possibilitando a utilização de horizontes obtidos através de métodos rápidos de automapeamento, tipicamente, utilizados na área de exploração.

A análise foi também realizada com dezesseis amostras em torno do horizonte que delimita a base do reservatório utilizando dezesseis amostras. O resultado utilizando a eq.(39) na geração dos atributos está ilustrado na figura 50.

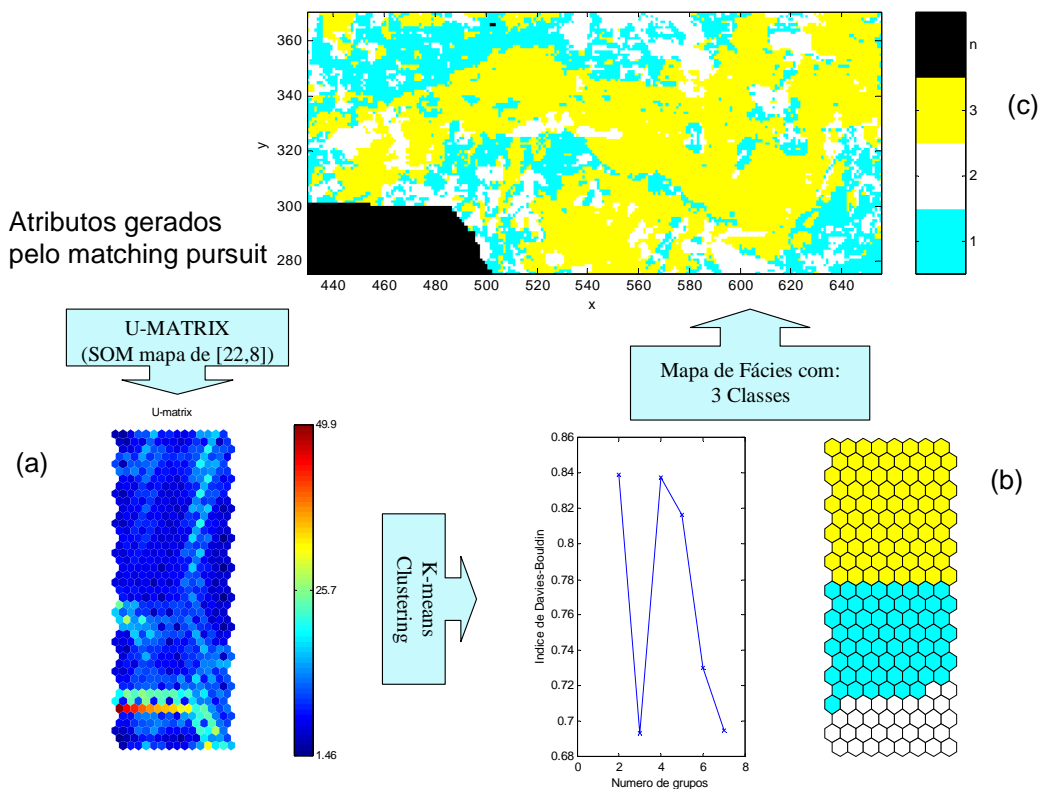


Figura 50: Análise de fácies sísmicas da base do reservatório de um dado sísmico real utilizando o método proposto; (a) Matriz-U obtida com o método de agrupamento do SOM; (b) Análise dos grupos formados utilizando o algoritmo K-means; (c) Mapa de fácies sísmicas.

Tomando o resultado da análise estratigráfica (Johannn, 1999), na figura 51 é ilustrado o mapa de fácies sísmicas com 4 fácies, embora, a matriz-U ilustrada na figura 50a e a análise dos grupamentos formados no SOM ilustrada na figura 50b não sugiram este número de fácies sísmicas.

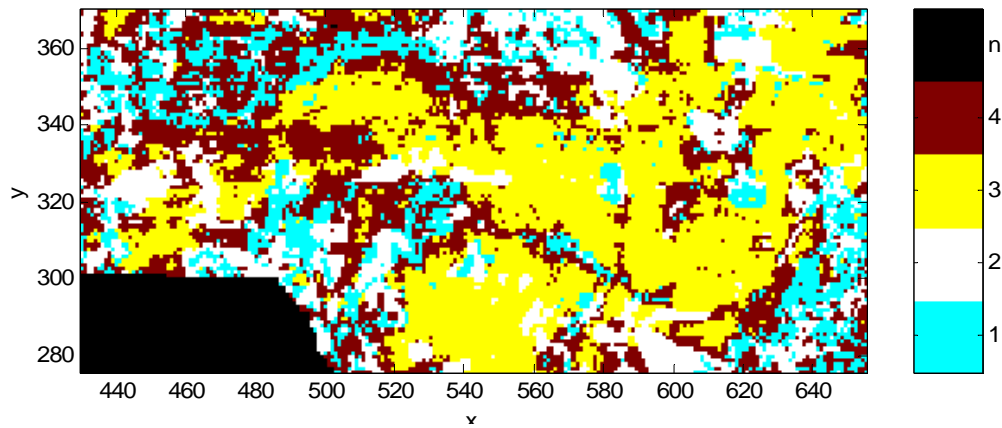


Figura 51: Mapa de fácies sísmicas da base do reservatório do dado real utilizando o método proposto.

A análise de fácies sísmicas foi repetida utilizando a eq.(40) para a geração dos atributos e o resultado está ilustrado na figura 52.

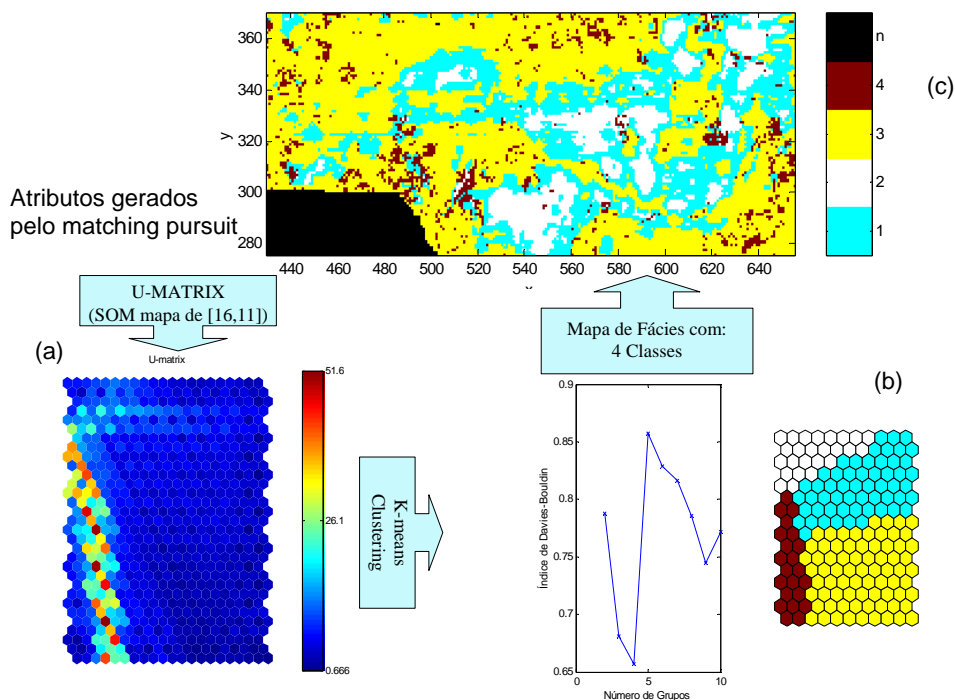


Figura 52: Análise de fácies sísmicas da base do reservatório de um dado sísmico real utilizando o método proposto; (a) Matriz-U obtida com o método de agrupamento do SOM; (b) Análise dos grupos formados utilizando o algoritmo K-means; (c) Mapa de fácies sísmicas.

Observa-se que os mapas de fácies sísmicas formados utilizando os atributos gerados através das eq.(39) e eq.(40), ilustrados nas figuras 51 e 52c, são bastante similares. Entretanto, como as análises de fácies sísmicas realizadas foram não-supervisionadas, no futuro, deve ser realizado um estudo da correlação entre os atributos gerados e os parâmetros geológicos do dado analisado.