

Métodos no domínio do espectro para análise e síntese de sinais musicais

Licio H. Bezerra, Saulo Castilho

UFSC - Departamento de Matemática

88040-900, Florianópolis, SC

E-mail: licio@mtm.ufsc.br, sauloc@inf.ufsc.br.

Resumo: Neste trabalho, apresentamos uma análise crítica de um método de análise espectral de sinais que podem ser descritos por um modelo senoidal variável no tempo. Utilizamos esse método para analisar e ressintetizar sons polifônicos.

Palavras-chave: Short-Time Fourier Transform, Modelo senoidal variável no tempo

1 Introdução

A limitação de métodos no domínio do tempo na determinação, em tempo real, de parâmetros em problemas com modelagem senoidal com várias parciais levou-nos a estudar métodos de análise espectral ([1], [2]). Por exemplo, o tempo necessário para análise e ressíntese de 50 parciais reais de um sinal de 1 segundo, correspondente à nota A4 do piano, pelos métodos no domínio de tempo, ultrapassa 10 minutos (ver [3]). Baseado no PARSHL ([1]), implementamos em MATLAB um programa que analisou e ressintetizou o problema acima em torno de 40 milissegundos (usando síntese por sobreposição). Nosso programa analisa sinais quase harmônicos do ponto de vista espectral, com parâmetros definidos pelo usuário. Estão disponibilizados dois tipos de síntese: sobreposição e aditiva. A comparação de resultados é feita via espectrogramas, uma técnica que tende a ser padrão na literatura ([4]). Esse programa também pode ser utilizado por pesquisadores de outras áreas além da música, pois modelagem exponencial com decaimento surge naturalmente no estudo de vibrações mecânicas, ressonância magnética nuclear etc. Cada passo do algoritmo é discutido detalhadamente a seguir, utilizando ferramentas de Análise de Fourier; ao final do trabalho, mostramos resultados obtidos na ressíntese de sinais musicais com um som polifônico; concluímos com uma observação sobre resultados obtidos, tanto por um método no domínio do tempo descrito em [5] como pelo nosso programa, em um problema de identificação de parâmetros em vibrações mecânicas.

2 STFT

Seja $x(t)$, $t = 0, \dots, P - 1$, um sinal discreto, com frequência de amostragem f_s em Hz (pontos por segundo). $T = \frac{1}{f_s}$ é o intervalo de tempo, em segundos, entre as amostras. Seja M ímpar, $M < P$. Definimos $M_h = \frac{M-1}{2}$. Vamos dividir as amostras em Q quadros, $Q = \frac{P-M}{R} + 1$, em que R é o número de pontos entre um quadro e outro. Vamos definir

$$x_m(n) = x((m-1)R + n + M_h),$$

em que $m = 1, \dots, Q$, $n = -M_h, \dots, M_h$. Note que o intervalo de tempo entre dois quadros será de RT segundos. Se o número total de quadros, não for inteiro, acrescentamos zeros ao sinal $x(t)$, gerando um novo sinal $x'(t)$ de tamanho P' . Esse procedimento não tem influência no domínio das frequências do sinal, pois não muda o tamanho do quadro. Sua influência é em seu domínio

de tempo, correspondendo ao acréscimo de silêncio ao sinal. O modo como esses zeros serão incluídos será discutido mais adiante. Podemos aplicar um pré-processamento nesses dados para que as características que devemos extrair (frequência, amplitude e fase) fiquem mais evidentes. Por exemplo, operando sobre o quadro uma janela $w(n)$, obtemos:

$$\tilde{x}_m(n) = x_m(n)w(n)$$

Posteriormente, acrescentaríamos zeros ao resultado, de forma que o quadro passe a ter um tamanho N . Para facilitar o cálculo da FFT, escolhe-se N como sendo uma potência de 2. Definimos $N_h = \frac{N}{2}$. Após o acréscimo de zeros, temos:

$$\tilde{x}'_m(n) = \begin{cases} \tilde{x}_m(n), & |n| \leq M_h \\ 0, & M_h < n \leq N_h \text{ ou } -N_h + 1 \leq n < -M_h \end{cases} \quad (1)$$

Aplicando-se a Transformada de Fourier em sua forma discreta, obtemos:

$$X_m(k) = \sum_{n=-N_h}^{N_h-1} \tilde{x}'_m(n) e^{-\frac{j2\pi nk}{N}} = \sum_{n=-N_h}^{N_h-1} \tilde{x}'_m(n) e^{-jw_k nT}, \quad (2)$$

em que $w_k = \frac{2\pi k f_s}{N}$ e k é o índice das amostras discretas no domínio de frequências (bins). Torna-se necessário, então, escolher adequadamente as variáveis envolvidas na STFT: tipo de janela ($w(n)$), tamanho do salto (R), tamanho do quadro (M) e tamanho da FFT (N).

1. Janelas

Dois parâmetros devem ser considerados na escolha da janela: a largura do morro principal e altura dos morros laterais. Morros principais estreitos e morros laterais baixos seriam ideais. Morros secundários altos podem ser vistos, equivocadamente, como morros principais, resultando na adição de frequências indesejadas (ruídos) ao sinal. Morros principais muito largos diminuem a precisão na detecção da frequência; morros principais próximos, se forem muito largos, podem ser interpretados como um único morro principal. A diferença de largura também pode ser utilizada como critério para classificar os morros como principais ou secundários, já que os primeiros são muito mais largos que os segundos.

2. Tamanho do salto

O tamanho do salto, R , é o número de pontos entre dois inícios de quadros consecutivos. Um salto pequeno resulta em mais quadros e maior qualidade de análise. Porém, demanda maior esforço computacional. Novamente, deve-se buscar um equilíbrio entre esses fatores. A escolha do R ideal também depende da janela escolhida. Dada uma janela $w(n)$, tal que $w(n) = 0, \forall |n| > |M_h|$, o tamanho do salto R deve satisfazer a seguinte equação:

$$(\forall t) A_w(t) = \sum_{m=-\infty}^{\infty} w(t - mR) = C, \quad (3)$$

em que C é uma constante. Se isso ocorrer, todas as amostras recebem o mesmo peso na soma dos quadros e a análise do sinal não sofre distorções. Para a janela retangular de largura M , o salto pode ser de $\frac{M}{k}$, para qualquer inteiro k tal que $\frac{M}{k}$ seja também inteiro.

3. Tamanho do quadro

O tamanho do quadro, M , é outro parâmetro do algoritmo. Se usarmos um quadro muito grande, perderemos as características instantâneas do som. Por outro lado, um quadro muito curto dificulta a detecção de frequências mais graves, pois o quadro pode ser menor que o período de tais frequências. Recomenda-se que o tamanho do quadro seja, aproximadamente, o tamanho do período da frequência mais grave presente no sinal. Para

aplicações de áudio, como o ouvido humano não reconhece frequências abaixo de 20 Hz, é possível um quadro de tamanho 0,05s (50ms). Por outro lado, o ouvido humano também não identifica frequências acima de 20KHz. Logo, podemos usar $f_s \cong 40KHz$. Assim,

$$T = \frac{1}{f_s} = \frac{1}{4 \cdot 10^4} = 0,25 \times 10^{-4} s.$$

Como a duração do quadro, MT , é igual a 0,05s, $M = \frac{0,05}{0,25 \times 10^{-4}} = 2000$. Se o sinal não possui frequências tão graves, pode-se diminuir o tamanho de M . Idem, para sinais com f_s menor. Por outro lado, M deve ser ímpar para que o quadro possua um ponto central $x_m(0)$, no qual a janela que será aplicada ao quadro deverá assumir seu valor máximo. Esse fato reflete-se no espectro do quadro, acentuando seus picos principais.

4. Tamanho da FFT

O tamanho da FFT, N , também é um parâmetro arbitrário. N deve ser um número maior ou igual a M , para preencher o quadro com $N - M$ zeros. Também deve ser uma potência de 2, para facilitar a aplicação da FFT. Uma escolha padrão para N é a primeira potência de 2 maior que $2M$. A importância desse procedimento é que, se um quadro recebe acréscimo de $\frac{N}{M}$ zeros, seu espectro terá $\frac{N}{M}$ vezes mais pontos do que se não houvesse acréscimo de zeros. Assim, a resolução do espectro aumenta, gerando maior precisão na hora de encontrar a frequência na qual ocorre o pico de maior magnitude.

3 Análise Individual do Quadro

Nesta etapa, nosso foco é um quadro isolado. Interpretamos seu espectro de forma a encontrar as frequências presentes naquele instante, junto com suas respectivas amplitudes e fases. Denotemos por $X_m(k)$, $k = 0, \dots, N - 1$, o espectro complexo do quadro x_m . Tomando apenas os dados relevantes, consideramos $k = 1, \dots, N_h + 1$. Assim, nos pontos discretos do domínio de frequências:

$$\begin{aligned} \text{Frequência: } f_m(k) &= \frac{k}{NT}; \\ \text{Amplitude: } A_m(k) &= |X_m(k)|; \\ \text{Fase: } \theta_m(k) &= \angle X_m(k). \end{aligned}$$

Sabemos que as frequências que compõem o sinal raramente encontram-se presentes com exatidão entre as frequências representadas pelos bins e, com isso, sua energia é dividida entre as frequências vizinhas. Portanto, para aproximarmos as reais frequências, devemos analisar os morros encontrados no módulo do espectro, ou seja, onde se encontram as maiores amplitudes $A_m(k)$ em seu espectro de potência. Como o espectro de potência é discreto, a maior magnitude (eixo y) nele encontrada corresponderá a uma frequência (eixo x) com precisão de meio bin, para mais ou para menos. Para uma janela retangular, uma precisão de 0,1% seria obtida apenas com fator de acréscimo de zeros superior a 1000, o que é impraticável. Porém, utilizando interpolação quadrática nos 3 pontos de maior magnitude em um mesmo morro, pode-se atingir a mesma precisão com um fator de acréscimo de zeros igual ou superior a 5. Resultados empíricos indicam que, se a interpolação for feita com os dados em decibéis (escala logarítmica), a precisão dobra em relação à interpolação com magnitude linear [1]. Porém, resta-nos a questão de quantos (ou quais) máximos locais podem ser considerados como componentes senoidais do sinal. Utilizaremos um número fixo de picos a serem detectados em cada quadro. Assim, procura-se o ponto de maior magnitude no espectro, utilizam-se seus pontos vizinhos para a determinação de uma parábola e estima-se seu pico. Traduzem-se, então, suas coordenadas em amplitude e frequência associadas a uma senóide. Depois disso, retiram-se os pontos associados a esse morro do espectro de potência e procura-se o próximo ponto de maior magnitude. Faz-se isso até que o número estipulado de picos por quadro tenha sido detectado.

4 Análise Entre Quadros

Analisando cada quadro, separadamente, obtemos as características instantâneas do som. Relacionando as informações encontradas em um quadro com as do quadro seguinte, podemos obter uma interpretação dinâmica do sinal. Primeiro, vamos relacionar o espectro de um quadro ao do quadro seguinte, analisando como amplitude e fase variam entre os bins; depois, relacionamos apenas os dados referentes aos picos encontrados na etapa anterior.

4.1 Variações de parâmetros entre quadros

Como M e N são constantes, assim como o intervalo T entre as amostras, o domínio das frequências é sempre o mesmo, ou seja, os bins são fixos. Logo, $X_{m-1}(k)$ e $X_m(k)$ correspondem a duas amostras subseqüentes do mesmo bin. Para ver como cada bin se comporta com o tempo, consideraremos as funções $X_m(k)$ (uma função para cada quadro cuja variável é a frequência) como funções de frequência fixa, variando no tempo. Temos, assim, para cada bin, as funções de amplitude - $A_k(m) = |X_m(k)|$ - e fase, $\theta_k(m) = \angle X_m(k)$, em que a variável de tempo é o índice do quadro. Logo, cada bin é associado a uma **faixa** ou **oscilador** com frequência constante. Osciladores são dispositivos que geram um movimento senoidal. Diremos que duas frequências pertencem à mesma faixa (ou que são geradas pelo mesmo oscilador) se forem os parâmetros instantâneos de uma mesma senóide em tempos diferentes. Se não se considerarem todos os bins, e sim, apenas as frequências obtidas no processo de detecção de picos, haverá faixas sem frequências constantes. A pergunta é como associar dois picos entre dois quadros diferentes.

4.2 Interligação de picos

De um quadro para outro podem ocorrer as seguintes situações com as frequências encontradas:

- (a) uma frequência pode se manter de um quadro para o outro ou sofrer modulações leves;
- (b) uma frequência presente no quadro anterior pode não estar no seguinte;
- (c) uma frequência não presente no quadro anterior pode surgir no seguinte.

Vamos supor que o sinal é gerado por osciladores, cada um gerando uma senóide com frequência e amplitude variáveis no tempo. Assim, de um quadro para outro, cada oscilador pode continuar a gerar a senóide, parar de gerá-la, ou começar a gerá-la. Denotemos as frequências de um quadro por f_1, \dots, f_p e as do quadro seguinte por g_1, \dots, g_r . Dado Δf , se houver g_j tal que

$$|f_i - g_j| \leq \Delta f, \quad (4)$$

para algum f_i , podemos considerar g_j como continuação de f_i , gerados pelo mesmo oscilador ((a)), e dizemos que a frequência g_j pertence à mesma faixa que f_i . Se algum f_i não satisfaz (4), então consideraremos que a faixa correspondente a f_i é desligada de um quadro para outro ((b)). Assim, a faixa correspondente a f_i manterá a frequência de um quadro para outro, porém sua amplitude diminuirá suavemente até 0 no quadro seguinte. Se algum g_j não satisfaz (4) para nenhum f_i , então consideraremos que uma nova faixa, de frequência inicial g_j foi criada de um quadro para o outro ((c)). Assim, inclui-se no quadro anterior a faixa com frequência inicial g_j e amplitude inicial 0. Caso haja conflito, isto é, existam g_{j_1} e g_{j_2} que satisfaçam (4), a faixa será prolongada para a frequência mais próxima, enquanto a outra procurará uma nova faixa para se ligar. Esse processo continua até que todas as faixas se encaixem em uma das 3 situações acima. Sendo p o número de picos por quadro fixo, se utilizarmos todos os picos detectados em ambos os quadros, o número de faixas entre dois quadros estaria entre p (todas as faixas são prolongadas) e $2p$ (nenhuma faixa obtém uma continuação). Entretanto, o caso de nenhum pico ser combinado de um quadro para outro é indesejado, pois acarreta uma grande descontinuidade no processo de síntese. Para evitar que isso aconteça, é comum a detecção de um número de picos superior ao número dos picos que serão efetivamente utilizados. Assim, prioriza-se a ligação das faixas já existentes de forma que, para cada faixa desligada, cria-se uma nova faixa para o pico de maior intensidade que surgir no quadro seguinte.

5 Síntese

Após o processo de análise, podemos montar uma linha de tempo mostrando como cada faixa varia sua fase, amplitude e frequência com o tempo para reconstruir o som original. Temos duas alternativas para a reconstrução do sinal: síntese por sobreposição e síntese aditiva.

5.1 Síntese por Sobreposição

Aplicamos *FFT* inversa a cada quadro e, a seguir, faz-se a sobreposição, somando pontos comuns

a mais de um quadro: $\hat{x}(t) = \sum_{m=1}^Q \hat{x}'_m(n)$, em que $n = t - (m - 1)R - M_h$ e $n \in [-M_h, M_h]$.

Os processos de análise e síntese por sobreposição são inversos, caso as janelas obedeçam a condição (3), com $C = 1$. Se (3) não é satisfeita, há uma modulação de amplitude, de período R , distorcendo o sinal original, ou seja, uma distorção periódica de frequência $\frac{f_s}{R}$.

5.2 Síntese Aditiva

A síntese aditiva é baseada nas faixas que se estendem ao longo do tempo. Vamos considerar dois casos: cada bin pode ser interpretado como uma faixa; ou resintetizamos um número de faixas menor, com as informações mais importantes do sinal. No primeiro caso, a síntese aditiva se equivale à síntese por sobreposição, com a desvantagem de não usar a *FFT* inversa para diminuir a quantidade de operações. Para cada faixa, temos sua amplitude, frequência e fase ($\hat{A}_m^r, \hat{f}_m^r, \hat{\theta}_m^r$) em cada quadro. Na STFT, é o ponto central que assume o valor máximo e, portanto, esse será o ponto escolhido para ser o representante do quadro: $A_r((m - 1)R + M_h) = \hat{A}_m^r$, em que \hat{A}_m^r é a amplitude da faixa r no quadro m e $A_r(t)$, a amplitude da faixa r no ponto t . Como o ponto central de cada quadro será seu representante, podemos acrescentar M_h zeros no início do sinal, para que o primeiro quadro tenha como representante o primeiro ponto de amostragem. Caso contrário, o ponto inicial receberia peso inferior aos outros e a condição (3) não seria satisfeita, a não ser que uma janela retangular fosse usada. Idem para o último ponto. Acrescentamos, portanto, ao final do sinal, M_h zeros mais a quantidade necessária para que $\frac{P'-M}{R}$ seja inteiro. Logo, todos os pontos do sinal original recebem mesma ponderação, satisfazendo a equação (3).

5.3 Interpolação entre quadros

Para cada faixa r , temos ($\hat{A}_{m-1}^r, \hat{f}_{m-1}^r, \hat{\theta}_{m-1}^r$), referentes às amplitude, frequência e fase do quadro anterior, e ($\hat{A}_m^r, \hat{f}_m^r, \hat{\theta}_m^r$), referentes às do quadro posterior. Vamos definir esses parâmetros para os pontos intermediários: $n = 1, \dots, S - 1$, em que S é o número de pontos entre os inícios de dois quadros da etapa de síntese (se não há a intenção de alterar a velocidade do sinal, $S = R$). A amplitude pode ser interpolada linearmente, de forma que o volume do som seja alterado gradativamente de um quadro para outro. Portanto,

$$\hat{A}_m(n) = \hat{A}_{m-1} + \frac{\hat{A}_m - \hat{A}_{m-1}}{S} n. \quad (5)$$

Já a frequência e a fase não podem ser interpoladas separadamente desta forma, pois a fase instantânea é dependente das fases e frequências iniciais e finais, enquanto a frequência instantânea é a derivada da fase. Assim, procuramos uma função para gerar uma interpolação suave de fase e frequência. Como temos 4 parâmetros disponíveis, utilizaremos uma polinomial cúbica:

$$\Theta(n) = \zeta + \gamma n + \alpha n^2 + \beta n^3. \quad (6)$$

Apesar de $\Theta(n)$ ser discreta, vamos tratá-la como uma função contínua e derivável. Como a frequência instantânea é a derivada da fase, $\Theta'(n) = \gamma + 2\alpha n + 3\beta n^2$. Notemos que:

$$\begin{aligned} \hat{\theta}_{m-1} &= \Theta(0) = \zeta; & \hat{f}_{m-1} &= \Theta'(0) = \gamma; & \hat{f}_m &= \Theta'(S) = \hat{f}_{m-1} + 2\alpha S + 3\beta S^2; \\ \hat{\theta}_m + 2\pi\eta &= \Theta(S) = \hat{\theta}_{m-1} + \hat{f}_{m-1}S + \alpha S^2 + \beta S^3. \end{aligned}$$

Como $\hat{\theta}_M$ está em um intervalo de comprimento 2π , não sabemos quantas voltas a fase percorreu até chegar no ponto atual. Portanto, o fator $2\pi\eta$ é necessário como forma de desencapsulamento da fase. Assim, para cada $\eta \in \mathbb{Z}$, temos uma função de interpolação:

$$\begin{cases} S^2\alpha(\eta) + S^3\beta(\eta) &= \hat{\theta}_m - \hat{\theta}_{m-1} - \hat{f}_{m-1} + 2\pi\eta \\ 2S\alpha(\eta) + 3S^2\beta(\eta) &= \hat{f}_m - \hat{f}_{m-1} \end{cases}.$$

Resolvendo o sistema, obtemos $\alpha(\eta)$ e $\beta(\eta)$. Como estamos tomando intervalos pequenos entre quadros e frequências próximas de um quadro para outro, a tendência é que a frequência tenha variação mínima. No caso ideal, de frequência constante, a derivada da frequência seria nula durante todo o intervalo. Vamos, então, calcular η^* que minimize a função

$$\int_0^S [\Theta''(n, \eta)]^2 dn, \quad (7)$$

em que $\Theta'(n)$ é a frequência no intervalo. Considerando as funções obtidas para ζ , γ , α e β , vamos tomar η^* como sendo o inteiro mais próximo de x^* , o que minimiza a função:

$$\Phi(x) = \int_0^S \left[\frac{d^2\Theta}{dt^2}(t, x) \right]^2 dt = 4\alpha^2(x)S + 12\alpha(x)\beta(x)S^2 + 12\beta^2(x)S^3.$$

Impondo $\Phi'(x) = 0$, obtemos o valor de x^* . Assim, se no intervalo entre os quadros m e $m - 1$ tivermos F_m faixas, esse intervalo será reconstruído por:

$$\hat{x}_m(n) = \sum_{r=1}^{F_m} \hat{A}_m^r(n) \cos(\Theta_m^r(n)), \quad n = 0, \dots, S - 1.$$

No som final, $\hat{x}(t) = \hat{x}_m(n)$, em que m e n são inteiros tais que $t = \frac{m-1}{S} + n \in \{0, \dots, P - 1\}$.

Na Figura 1, vemos espectrogramas de um som polifônico gerado por guitarra, baixo, teclado, voz e bateria (este gera sons não harmônicos). Analisamos um trecho de 4 s do total de 61.24 s ($P = 1352583$), com $M = 1023$, $N = 2048$ e $R = 512$.

6 Conclusões

Notemos, pela Figura 1, que resolvemos de modo eficiente o seguinte problema inverso: calcular os parâmetros de uma função do tipo $x(t) = \sum_{k=1}^{M(t)} [A_k(t) \cos(\Theta_k(t))]$, em que $\Theta_k(t) = \theta_k + 2\pi \int_0^t f_k(t) dt$, a partir de amostras em intervalos de tempo igualmente espaçados. Outro problema científico relevante associado a essa formulação é o de determinar os modos de vibração mecânica de uma estrutura: percuta-se a estrutura (uma ponte, por exemplo) e aparelhos medem a resposta de tempo em tempo. Pelas leis físicas, essa resposta é um sinal do tipo acima com $M(t) = M$, $A_k(t) = A_k$ e $f_k(t) = f_k$, constantes. Há algumas abordagens no domínio do tempo para esse problema, do ponto de vista determinístico (ver [5]). No entanto, nenhum se compara em eficiência (precisão e tempo de computação) ao método no domínio do espectro analisado aqui. O problema de vibrações mecânicas exige poucos parâmetros a serem rastreados se comparado com o de análise e ressíntese de sinais musicais, o qual continua desafiando o matemático numérico, pois os computadores atuais são ainda limitados para lidar com a dimensão do problema (ver [3]). Os autores do PARSHL são um músico e um engenheiro que programaram seu algoritmo em ALGOL, no final da década de 80. O nosso programa foi feito em MATLAB, no qual todo problema numérico é resolvido como um problema de Álgebra Linear Numérica, e é suficientemente flexível para resolver vários outros problemas de identificação de parâmetros: vibrações mecânicas, ajuste de dados exponenciais etc. Por exemplo, um problema inverso tratado em [5] era o de calcular o número de frequências contidas no sinal $h(t) = e^{-0.06.t} .sen 4t + 0.8e^{-0.056.t} .sen t + 1.2e^{-0.09.t} .sen 9t$, a partir de 80 amostras com

$\Delta t = 0.05$. O nosso programa calculou quais eram essas frequências: 3.99, 0.92 e 9.00. A frequência mais grave não foi detectada com a mesma precisão das outras devido ao número pequeno de amostras em relação ao tamanho do Δt utilizado. O programa que ressynetizou o som polifônico e o que calculou as frequências (a escolha de parâmetros é o que faz a diferença entre eles) estão disponíveis em <http://mtm.ufsc.br/~licio/cnmac/> para testar a ressynetese de sons polifônicos, ou resolver problemas inversos de modelagem exponencial.

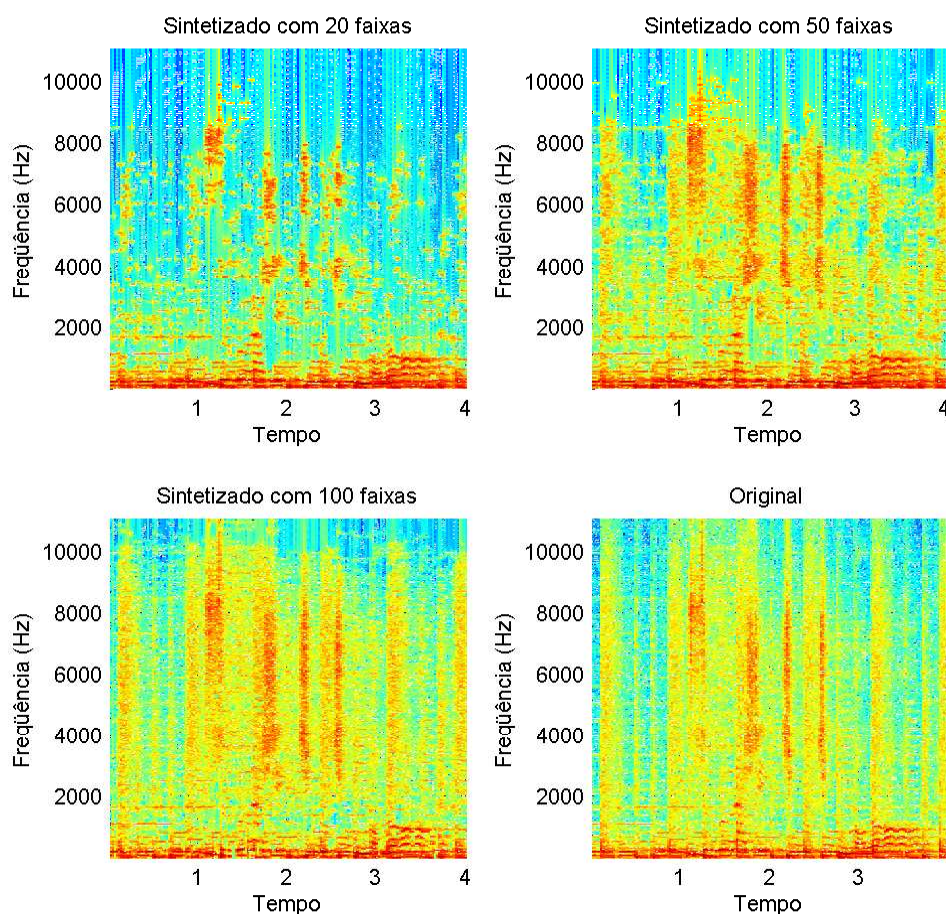


Figura 1: Espectro de uma música polifônica, durante 4 segundos.

Referências

- [1] X. Serra and J. O. Smith, PARSHL: An Analysis/Synthesis Program for Non-Harmonic Sounds Based on a Sinusoidal Representation, *Proceedings of the International Computer Music Conference (ICMC)*, 290-297, Champaign-Urbana, USA, 1987.
- [2] A. Klapuri, Multiple fundamental frequency estimation by harmonicity and spectral smoothness, *IEEE Transactions on Speech and Audio Processing* 11 (6), 804-816, 2003.
- [3] L. H. Bezerra e S. Castilho, Método de rastreamento de subespaços dominantes: aplicação em sinais de áudio, *anais do XXXI CNMAC*, Belém, 2008.
- [4] S. A. Fulop and K. Fitz, A spectrogram for the 21st century, *Acoustics Today*, 26-33, 2006.
- [5] Licio H. Bezerra and Fermin S. V. Bazán, Eigenvalue locations of generalized companion predictor matrices, *SIAM J. Matrix Anal. Appl.*, 19 (4), 886-897, 1998.