

# **Segmentação e Extração de Descritores de Expressividade em Sinais Musicais Monofônicos**

**Mauricio Loureiro, Tairone Magalhães, Rodrigo Borges, Thiago Campolina, Davi Mota, Hugo de Paula**  
**Universidade Federal de Minas Gerais - UFMG**  
e-mail: mauricioloureiro@ufmg.br

## **1. Introdução**

O estudo da expressividade musical tem focalizado a busca por mecanismos ou princípios que possam estar por trás das diferenças entre uma performance de uma música e sua codificação na partitura, assim como entre performances distintas da mesma música. A singularidade de cada uma está encapsulada em desvios expressivos aplicados a determinados parâmetros musicais, como durações rítmicas, variações de intensidade, afinação e timbre, que podem ser observados em performances de diferentes intérpretes e mesmo entre duas interpretações do mesmo músico. Estes desvios são percebidos com uma clareza surpreendente, mesmo para ouvintes não especializados, o que faz com que uma performance tecnicamente perfeita mas inexpressiva seja menos apreciada que uma interpretação expressiva da mesma partitura que contenha alguns erros ou imprecisões.

As primeiras investigações sobre performance musical feitas por Seashore (1938) sugeriram que a compreensão deste complicado fenômeno envolveria não apenas o comportamento do instrumentista frente ao texto que interpreta, mas também os mecanismos de escuta do ouvinte. Sloboda (1985) investigou aspectos emocionais da performance, evidenciando uma coincidência entre intenções expressivas do intérprete e percepção dos ouvintes. O problema ressurgiu na década de 70, quando passou a despertar interesse de musicólogos, psicólogos, cientistas de computação, engenheiros e físicos. Um grande número de pesquisa quantitativa em diferentes aspectos da expressividade musical vem sendo realizada com base em medições de parâmetros como tempo, dinâmica e timbre, procurando descrever e reconhecer classes de padrões que possam elucidar a influência destes parâmetros na expressividade percebida.

Estudos tais como os de Gabrielsson (1995) e Juslin (1997; 2000) demonstraram consistentemente que músicos utilizam pequenas variações de durações, articulações, intensidade, altura e timbre para comunicar ao ouvinte, aspectos da música que eles interpretam. Repp

(1990) mediu as diferenças entre performances de diferentes pianistas, mas evidenciou em estudo posterior (Repp, 1992) também inúmeros fatores comuns entre elas, que puderam ser relacionados à estrutura da obra interpretada. Uma excelente revisão bibliográfica sobre estudos feitos em performance musical foi elaborado por Palmer (1997) e Gabrielson (1999; 2003), este último incluindo mais de 800 referências.

Canazza, De Poli *et al.* (1997) estudaram a correlação entre a intenção expressiva do intérprete percebida em testes auditivos e a análise de parâmetros acústicos extraídos de 7 diferentes performances do Concerto em Lá Maior para clarineta e orquestra de Mozart, cada uma com uma intenção expressiva distinta, inspiradas em adjetivos sensoriais (duro, mole, pesado, leve, brilhante, escuro) com o objetivo de determinar quais parâmetros estariam sujeitos a variações devido a cada intenção expressiva. As análises revelaram parâmetros acústicos determinantes de diferenciações específicas entre as diferentes execuções, além de evidenciar uma coincidência entre dados das intenções do intérprete e a percepção dos ouvintes.

Partindo de estudos conduzidos por Palmer (1996a; 1996b), De Poli e colegas (De Poli, Rodà *et al.*, 1998) buscaram focalizar um contexto musical específico. Utilizando o conceito de *desvio expressivo* como qualquer desvio dos valores destes parâmetros em relação a uma referência ou *norma*, definida como “plana” ou “sem expressão”, De Poli foi capaz de identificar duas fontes de motivação que levariam o intérprete a realizar este desvios para transmitir suas intenções expressivas: (1) aspectos estruturais da partitura, tais como estrutura hierárquica de frases, estruturas harmônica e melódicas, que é comum a todas as performances e traduziriam o conteúdo expressivo codificado pelo compositor na partitura; (2) intenções de expressividade do intérprete, que são específicas para cada performance. Além de identificar os desvios relacionados a cada uma destas fontes, o estudo mostrou também que a escolha do instrumento musical é determinante na definição dos parâmetros a serem utilizados, já que cada instrumento dispõe de recursos de controle de produção sonora específicos.

Os resultados destes estudos mostraram a complexidade do problema frente às inúmeras possibilidades que o intérprete pode escolher para transmitir sua intenção expressiva. Além disso diferentes motivações expressivas podem ser transmitidas por efeitos acústicos similares, do mesmo modo que diferentes efeitos acústicos podem levar à mesma idéia expressiva, dificultando ainda mais a formalização do problema.

### **1.1 Descritores da intenção expressiva**

A busca por métodos de quantificação destes desvios parece ser a chave para um maior entendimento dos processos envolvidos na estruturação e na percepção de uma performance expressiva. Quantificar as intenções expressivas do intérprete a partir destes desvios envolve a identificação e medição de um conjunto de parâmetros físicos que possam representar os recursos utilizados pelo intérprete para comunicar sua intenção expressiva, os quais são definidos e calculados a partir de

informação extraída do sinal de áudio de gravações da performance. O primeiro passo é determinar que tipo de informação será utilizada e como será extraída do sinal, para em seguida definir os parâmetros descritores da performance ou de aspectos dela a serem investigados. A partir desta quantificação pode-se construir um modelo de análise de expressividade que produza resultados próximos aos obtidos pela percepção humana dentro de um contexto musical.

Uma grande diversidade de métodos de extração e processamento de informação de conteúdo musical pode ser encontrada na literatura e a grande divergência de abordagem entre eles mostra que não existe ainda um padrão metodológico para este tipo de procedimento. Alguns destes parâmetros descritores podem ser estimados a partir de informação extraída no domínio do tempo, tais como: *Duração Local* (tempo local da nota); *Duração Local Nominal* (extraído da partitura); *Duração da Nota*, *Duração de Ataque*, *Tempo Médio* (valor médio do metrônomo em batidas por minuto); *Tempo Local*, *Índice de Articulação*, *Sustentação*, *Vibrato*. Por outro lado, variações intencionais de afinação e timbre, também utilizadas por intérpretes para transmitir suas intenções expressivas, estão relacionadas a informação acústica espectral. Estas variações são mais salientes em instrumentos nos quais a ação do instrumentista participa durante toda a produção do som, como ocorre nos instrumentos de cordas e de sopros. Descritores espectrais podem mostrar correlação com variações expressivas intencionais de timbre, entre eles: *Altura*, *Centróide Espectral*, *Irregularidade Espectral*, *Fluxo Espectral*, *Decaimento Espectral*, *Achatamento Espectral*.

Este texto busca ilustrar uma variedade de métodos adotados na extração de informação do sinal de áudio de interpretações na clarineta, focalizando alguns procedimentos envolvidos na segmentação do sinal e na definição e extração de descritores espectrais e de envelope de energia. Os procedimentos aqui descritos integram um projeto de pesquisa que tem por objetivo construir um modelo de análise da expressividade de interpretações em instrumentos musicais, a princípio monofônicos, mas que deverá se expandir para outros tipos de sinais musicais polifônicos ou eletroacústicos.

## 2. Segmentação

### 2.1 Níveis de Segmentação

Uma questão fundamental na análise da performance musical é a definição dos critérios de segmentação do material musical, já que a definição e extração destes descritores pode se dar em diferentes níveis de segmentação do sinal analisado. Apesar de ser a nota musical a unidade de segmentação que define a maioria das características expressivas de uma performance, ela não é a única unidade de medição e análise de parâmetros descritores. Parâmetros tais como descritores espectrais e medições de movimento físico do interprete, podem se referir a níveis de segmentação inferiores à duração da nota, por exemplo, igual

à duração das janelas temporais usadas para calcular o envelope de energia, ou o conteúdo espectral. Do mesmo modo, a análise de descritores temporais tais como tempo local, pode ser mais significativa a partir de níveis métricos de segmentação, que incluam grupos de notas, como por exemplo o nível da batida (unidade de tempo do compasso), como utilizado por Shaffer e colegas (1985), ou do compasso, como utilizado por Todd (1985) e Repp (1992).

Quatro níveis de segmentação podem servir adequadamente como referência para as medições e análise de vários tipos de descritores, cada um com durações específicas e relacionados a diferentes tipos de descritores:

1. *Quadros de Tempo*, de duração constante. Com uma taxa de amostragem de 44,1 kHz com 16 bits, o tamanho do quadro escolhido para o cálculo de descritores de energia RMS e de conteúdo espectral foi de 1024 amostras (23,2 ms), com um passo de 256 amostras (5,8 ms).
2. *Nota*, de duração igual à duração de nota. Uma adequada segmentação neste nível depende da precisão na detecção do início e fim da nota, assim como de seu envelope. Inúmeros descritores são facilmente estimados a partir desta segmentação, tais como duração da nota e de ataque, índice de articulação, índice de sustentação, etc.
3. *Batida*, de duração igual à unidade de tempo do compasso, de acordo com a partitura. A normalização de descritores do envelope da nota neste nível de segmentação pode facilitar a análise da evolução temporal destes parâmetros, de acordo com Timmers e colegas (2000), que observaram que o intérprete normalmente “planeja” suas intenções expressivas em relação à pulsação local da música.
4. *Global*, de duração mais longa, relacionada a unidades de maior nível hierárquico, tais como frases e sub-frases. Alguns descritores de envelope ou espectrais podem apresentar significado quando calculados ao longo de toda a performance analisada, tais como tempo global médio, intensidade global média, centróide global médio, etc.

## **2.2 Metodologia de segmentação de notas**

A precisão e robustez do processo de segmentação do sinal é fundamental para a definição e estimação adequada dos descritores de expressividade. Além da detecção dos instantes de início e fim das notas que segmentam o sinal em notas individuais, é necessário detectar também instantes intra-notas que segmentam o envelope da nota em ataque, sustentação e decaimento. Tais métodos são descritos abaixo.

### **Envelope de energia RMS**

O envelope de amplitude corresponde ao contorno de amplitude de um som é geralmente dividido em três segmentos: ataque, sustentação e decaimento. Sua estimação é mais comumente feita através da medida RMS - *root mean square* - RMS (raiz da média dos quadrados), que está relacionada com a potência média de curta duração do sinal. Não existe uma metodologia padronizada de estimação do envelope de amplitude da nota, que pode variar muito de acordo com o tipo de dado do problema investigado. O cálculo do envelope a partir do valor de pico de amplitude, baseia-se em variações muito bruscas do valor da amplitude ao longo de intervalos de tempo curtos demais para que possam ser percebidos. Por outro lado o valor RMS é perceptualmente mais relevante, se aproximando mais da maneira como percebemos a intensidade de um som. O valor RMS de um sinal,  $x$  é dado por:

$$RMS(n) = \sqrt{\frac{1}{N} \sum_{m=-\frac{N}{2}}^{\frac{N}{2}-1} x[n-m]^2 w[m]}$$

onde  $n$  é o número da amostra (instante de tempo) e  $N$  o tamanho da janela  $w$  de cálculo da média, o qual determina a resolução temporal do envelope. Janelas mais longas suavizam o envelope e podem dificultar a detecção correta da região do ataque. É nesta região que ocorrem grandes flutuações de energia conhecidas como transientes, sendo por isso muito relevante para a estimação de parâmetros relacionados à articulação das notas. Janelas mais curtas salientam a região do ataque, mas podem diminuir a precisão da detecção do instante final do ataque e de início de decaimento. Uma suavização do envelope através de um filtro passa baixa de frequência de corte adequada pode também facilitar a detecção de cada um dos instantes relevantes do envelope.

### **Detecção de início e final de nota pela energia RMS**

Estes instantes são comumente detectados a partir da curva de energia RMS do sinal. De Poli (2005) sugeriu um método em que inícios de notas são calculados a partir de uma curva média de energia RMS calculada em um intervalo de duração bem maior do que a do quadro de tempo de cálculo da energia RMS. Esta curva estabelece um limiar adaptativo, que intercepta a curva de energia RMS nas regiões de transição entre notas consecutivas, nos pontos  $P_1$  e  $P_2$  da Figura 1, que mostra a curva da energia RMS (curva contínua) e a curva média de RMS de longa duração (curva pontilhada). Os pontos de início de nota (*onsets*) podem ser encontrados a partir do valor mínimo de energia na região abaixo do limiar adaptativo.

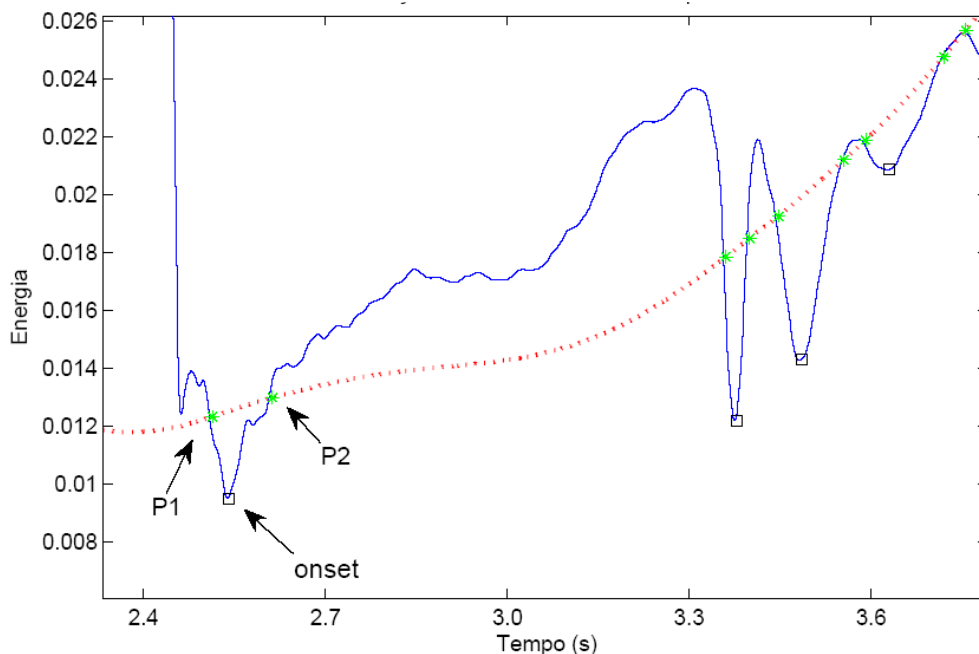


Figura 1: Detecção de inícios de notas a partir de um limiar adaptativo (média da energia RMS de longa duração).

Este método estima apenas um ponto entre duas notas consecutivas, mesmo no caso de duas notas separadas por pausa. A estimação dos instantes de início e final de nota foi possível a partir de uma alteração simples do método do De Poli, que consiste em considerar apenas um percentual do RMS médio de longa duração, aplicando um fator multiplicador  $\lambda$  de valor menor do que 1 ao limiar adaptativo. Assim, quando a energia do sinal ultrapassar esse percentual, considera-se o ponto de início da nota. O fim da nota acontecerá quando a energia cair para abaixo desse percentual (Figura 2). O valor da constante  $\lambda$  escolhido empiricamente, foi de 0,2.

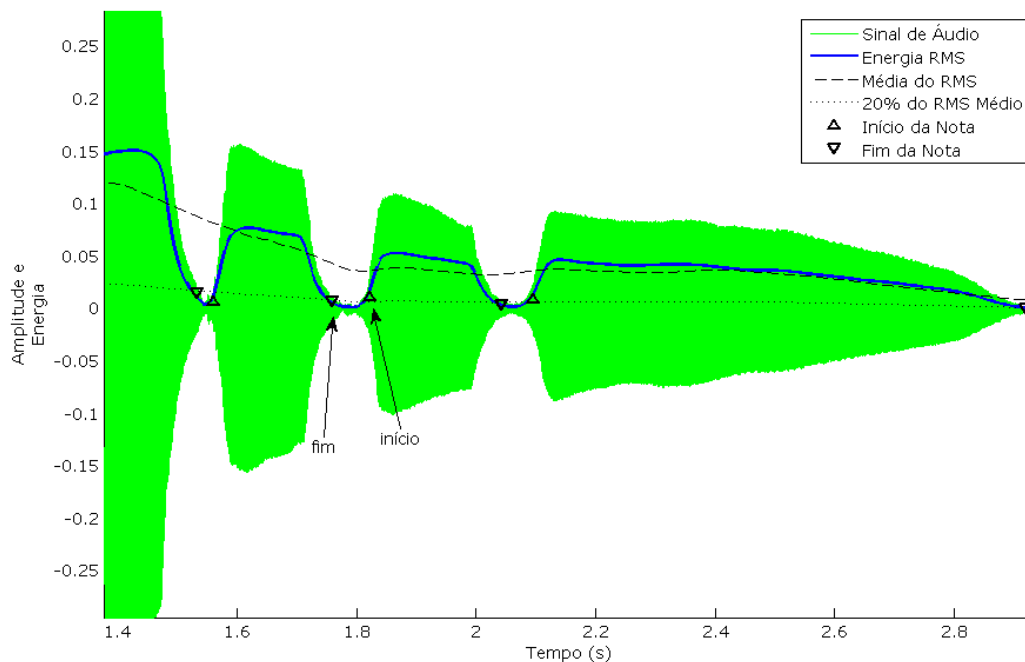


Figura 2: Detecção de inícios de notas a partir de um limiar adaptativo (20% da média da energia RMS de longa duração).

### Detecção de início e final de nota a partir da frequência fundamental

A segmentação de notas consecutivas pode não ser viável apenas a partir de níveis de amplitude, como por exemplo de notas ligadas, gravadas em ambiente com reverberação, em que o instante de final de nota pode coincidir com o instante de início da nota seguinte. Nesses casos a estimação da frequência fundamental ou mesmo outros parâmetros como o *fluxo espectral* ou *achatamento espectral* (descritos nas seções 3.3 e 3.6), pode auxiliar na detecção destes instantes.

Na análise de sinais musicais monofônicos, que não utilizam técnicas contemporâneas de execução, a altura (frequência fundamental) permanece muito estável durante a execução de uma nota. Desenvolvemos um algoritmo para detecção do início e fim da nota, correspondentes aos instantes de início e fim de uma estabilização da variação da frequência fundamental, permitida dentro de um limiar considerado de 5 %, ou seja, pouco menor que um semitom, como mostra a Figura 3. A detecção de frequência fundamental foi feita a partir da estimação dos 3 valores máximos de magnitude do espectro, para cada quadro. A frequência fundamental é escolhida como aquele candidato que apresentar o maior valor de energia somada às energias de seus 2 primeiros componentes harmônicos.

Este método pode falhar no caso de duas notas de mesma altura consecutivas, entretanto, nestas situações o método do RMS é mais eficaz na detecção de início da nota seguinte, já que há uma articulação entre elas. A combinação entre estes dois métodos, estabilização da frequência fundamental e limiar de energia RMS, resulta num método mais robusto para a detecção de início e fim de nota.

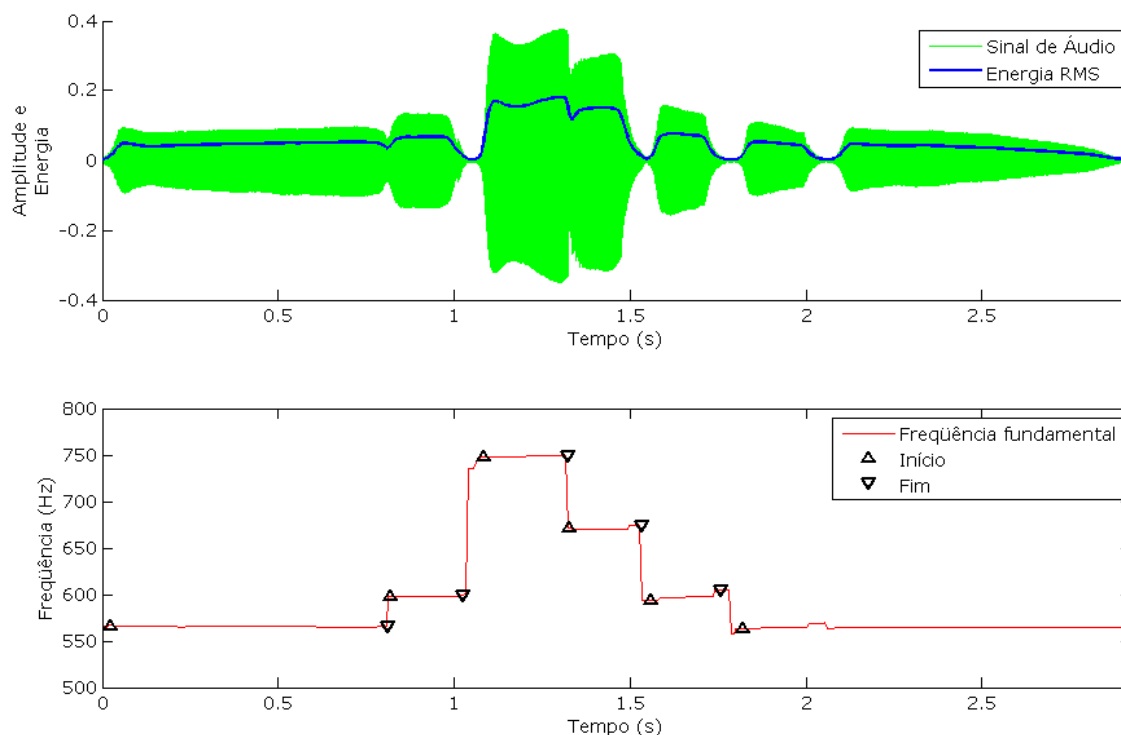


Figura 3: Detecção de inícios de notas a partir da frequência fundamental.

### Detecção de início e final de nota por fluxo espectral

Durante a transição entre duas notas acontecem quedas abruptas no valor do *fluxo espectral* (descrito na seção 3.3), devido à grande presença de transientes no sinal. Portanto, o fluxo pode ser utilizado para estimar pontos de início e final de notas, ou até mesmo de final de ataque e início de repouso.

O fluxo espectral tende a estabilizar quando existe uma harmonicidade entre as componentes espectrais de maior magnitude. Portanto, a estabilização do fluxo espectral está geralmente associada ao início de uma nova nota. Porém, o que se constatou é que a estabilidade total do fluxo espectral (correlação máxima entre quadros adjacentes) só é atingida após o ponto do início perceptivo na nota. Assim, um valor limiar em torno de 0.97 foi estabelecido após alguns testes empíricos, a partir do qual o fluxo é considerado como estável. Para a detecção do início de uma nota, essa estabilidade deve se manter por pelo menos um intervalo de tempo correspondente à duração da menor nota possível de ser detectada pelo método (aproximadamente 30 ms).



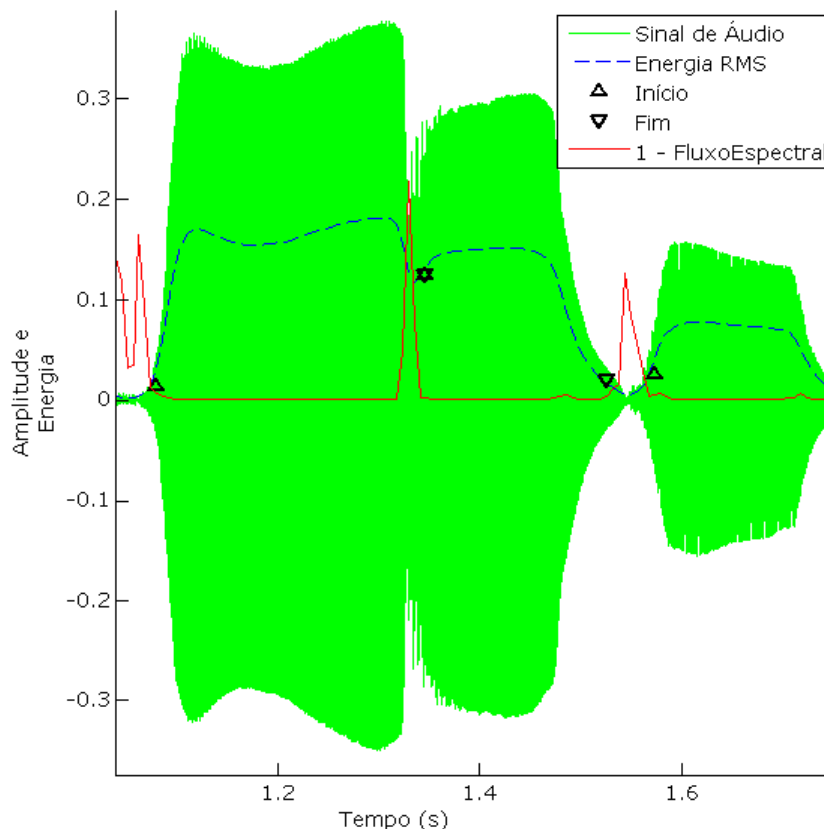


Figura 4: Detecção de inícios de notas a partir do fluxo espectral.

A detecção do final da nota através do fluxo espectral não é tão trivial, pois o fluxo tende a sofrer uma queda significativa próximo ao início do decaimento da nota, e a partir dessa queda seu comportamento é muito imprevisível (devido à grande presença de transientes nessa região). Portanto a combinação do método do fluxo espectral com o método do RMS pode gerar resultados mais satisfatórios.

### **Detecção de final de ataque e início de decaimento a partir da energia RMS**

Não existe na literatura um método de medição que possa descrever inequivocamente o ataque (Park, 2004). Em muitas situações, o instante de final de ataque refere-se ao instante de amplitude máxima da nota, mas sua detecção pode demandar procedimentos mais complexos, como, por exemplo em instrumentos não percussivos ou não pinçados, tais como sopros, cordas e voz, nos quais a amplitude máxima poder ser atingida ao longo da parte sustentada da nota, bem depois do ataque. Pode-se contornar o problema estabelecendo-se um valor adequado de limiar máximo de amplitude, mas que pode não apresentar resultados consistentes para notas mais longas tocadas nestes instrumentos.

A detecção dos instantes de final do ataque e o início do decaimento pode ser feita a partir da estimação dos máximos das taxas de variação de energia dentro da nota. O método é baseado na análise do contorno do envelope de energia da nota. Uma vez que os inícios e os finais de notas

foram detectados, observa-se a curvatura do envelope da nota, com a finalidade de identificar os três segmentos distintos: ataque, sustentação e decaimento. Considerando o envelope de energia como uma função diferenciável contínua no tempo, os pontos de máxima curvatura são considerados variações máximas locais da derivada primeira do sinal, e conseqüentemente, máximos e mínimos da derivada segunda. Os pontos correspondentes aos valores mínimos da derivada segunda (curva superior da Figura 5) determinam os pontos candidatos para os instantes de final de ataque e início de decaimento na curva de energia RMS (curva inferior da Figura 5). O candidato escolhido para o final de ataque será aquele que apresentar maior inclinação positiva medida entre ele e o início da nota, enquanto que a maior inclinação negativa com o final da nota indicará o início de decaimento.

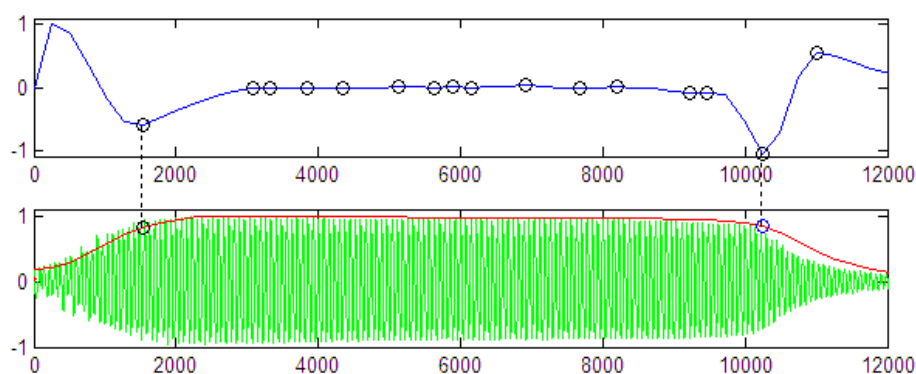


Figura 5: Detecção de final de ataque e início de decaimento a partir da energia RMS.

O ataque é definido como o intervalo de tempo entre os instantes de início da nota e o final de ataque, enquanto o decaimento é definido como o intervalo de entre os instantes de início de decaimento e final da nota. Quando o final de ataque coincide com o início do decaimento, a nota é considerada como sem sustentação.

### 3. Descritores Espectrais

Uma grande variedade de parâmetros derivados da distribuição espectral para descrever o timbre vem sendo recentemente proposta e testada por vários grupos de pesquisa inseridos em áreas tais como, percepção e cognição musical, psicologia da música e extração e processamento de informação musical (*MIR* - Music Information Retrieval). Através de testes subjetivos de similaridade, Stephen McAdams e colaboradores evidenciaram a correlação entre uma série de parâmetros espectrais e dimensões de espaços gerados pela percepção de timbre, corroborando a adequação destes parâmetros para a descrição deste atributo (McAdams, Winsberg *et al.*, 1995; Hajda, Kendall *et al.*, 1997; Misdariis, Smith *et al.*, 1998; Loureiro, de Paula *et al.*, 2001).

Os descritores espectrais utilizam a informação sobre as magnitudes das componentes espectrais de um sinal, obtidos através da Transformada de Fourier de Tempo Curto (*STFT - Short Time Fourier Transform*):

$$X[n, k] = \sum_{m=-\frac{N}{2}}^{\frac{N}{2}-1} x[hn+m]w[m]e^{-\frac{2j\pi mk}{N}}$$

onde  $w[m]$  representa uma janela de Hamming de tamanho  $N$ , e  $h$  o tamanho do passo.  $X[n, k]$  é a STFT do  $n$ -ésimo quadro em uma frequência  $k$ .

### 3.1 Centróide espectral

O centróide espectral (SC) é um parâmetro conhecido por sua proeminente correlação com o “brilho” do som, desde as primeiras pesquisas sobre percepção de timbre e é calculado como o “centro de gravidade” do espectro de amplitude das componentes de frequência do sinal:

$$SC[n] = \frac{\sum_{k=1}^N kX[n, k]}{\sum_{k=1}^N X[n, k]}$$

Durante a execução de uma nota, o centróide pode se modificar, especialmente se há uma alteração de intensidade ao longo da nota, que pode gerar variações assíncronas dos valores das magnitudes das parciais, alterando a posição do centróide e conseqüentemente a sensação de timbre associada a ele. Sons com qualidades “escuras” tendem a ter um maior conteúdo de baixa frequência, e aqueles com som mais “brilhante” tendem a ter maior predominância de harmônicos superiores, o que pode ser inferido a partir do valor do centróide espectral.

### 3.2 Irregularidade espectral

Descreve o grau de irregularidade espectral, comparando a amplitude de um componente espectral com seus adjacentes. Pode ser medido para cada quadro ou a partir das amplitudes médias ao longo de um período determinado:

$$IE[n] = 20 \sum_{k=2}^{N-1} \left| \log(X[n, k]) - \frac{\log(X[n, k-1]) + \log(X[n, k]) + \log(X[n, k+1])}{3} \right|$$

A irregularidade espectral está inversamente associada à suavidade de um espectro, a qual promove a integração das parciais a uma mesma fonte, enquanto que uma única parcial de grande intensidade tende a ser percebida como um som independente, o que contribuiria para um maior valor deste índice.

### 3.3 Fluxo espectral

O Fluxo espectral é obtido através da correlação entre componentes de frequência correspondentes de quadros espectrais adjacentes, e mede o grau de variação espectral ao longo de um sinal.

$$SF = \frac{1}{T} \sum_{p=1}^T |r_{p,p-1}|$$

onde  $T$  é a duração total do sinal e  $r_{p,p-1}$  o coeficiente de correlação de Pearson entre as amplitudes dos espectros calculados nos instantes  $t_p$  e  $t_{p-1}$ .

### 3.4 Decaimento espectral (*roll off*)

Mede a frequência abaixo da qual reside 85% da energia do sinal.

$$\sum_{k=1}^{DE} A_k = 0,85 \sum_{k=1}^N A_k$$

### 3.5 Inarmonicidade

É uma medida do erro entre a frequência dos harmônicos medidos e os seus valores de frequência esperados. A inarmonicidade é comumente observada em instrumentos como o piano, nos quais a rigidez das cordas causa inarmonicidade especialmente nas notas com as parciais mais altas.

$$INH[n] = \sum_{k=1}^{P-1} \frac{|f_k - kf_0|}{kf_0}, n=2, \dots, P$$

onde  $f_k$  é a frequência do k-ésimo harmônico,  $P$  é o número de parciais harmônicas e  $n$  é o número quadro analisado.

### 3.6 Achatamento espectral

É definido como a razão entre a média geométrica ( $Gm$ ) e a média aritmética ( $Am$ ) das amplitudes das frequências. Este parâmetro estima o conteúdo de ruído do sinal: o índice tende a zero em sinais mais senoidais e tende a 1 em sinais que se aproximam de ruído, ou seja, mais achatados e com pouca correlação. Este índice não infere sobre a harmonicidade do sinal, mas sugere a falta ou a presença de componentes senoidais.

$$AE[n] = 10 \log \left( \frac{Gm}{Am} \right) = 10 \log \left[ \frac{\left( \prod_{k=1}^N X[n, k] \right)^{\frac{1}{N}}}{\frac{1}{N} \sum_{k=1}^N X[n, k]} \right]$$

Inúmeros descritores de timbre já foram utilizados em estudos em extração e processamento de informação musical, principalmente aqueles voltados para o reconhecimento automático de instrumento. Tae Hong Park (2004) detalha um vasto elenco de descritores espectrais, entre eles, *Expansão/Compressão Harmônica*, *Inclinação Harmônica*, *Shimmer*, *Jitter*, *Envelope Espectral*, *Sincronia Harmônica*, *Tristimulus*, *Espalhamento Espectral*, *Fase*.

## 4. Descritores do Envelope

### 4.1 Duração

Descritores de duração são definidos a partir de instantes medidos no envelope da nota: início e final de nota, final de ataque e início de decaimento. Estes descritores podem ser facilmente estimados a partir destes instantes.

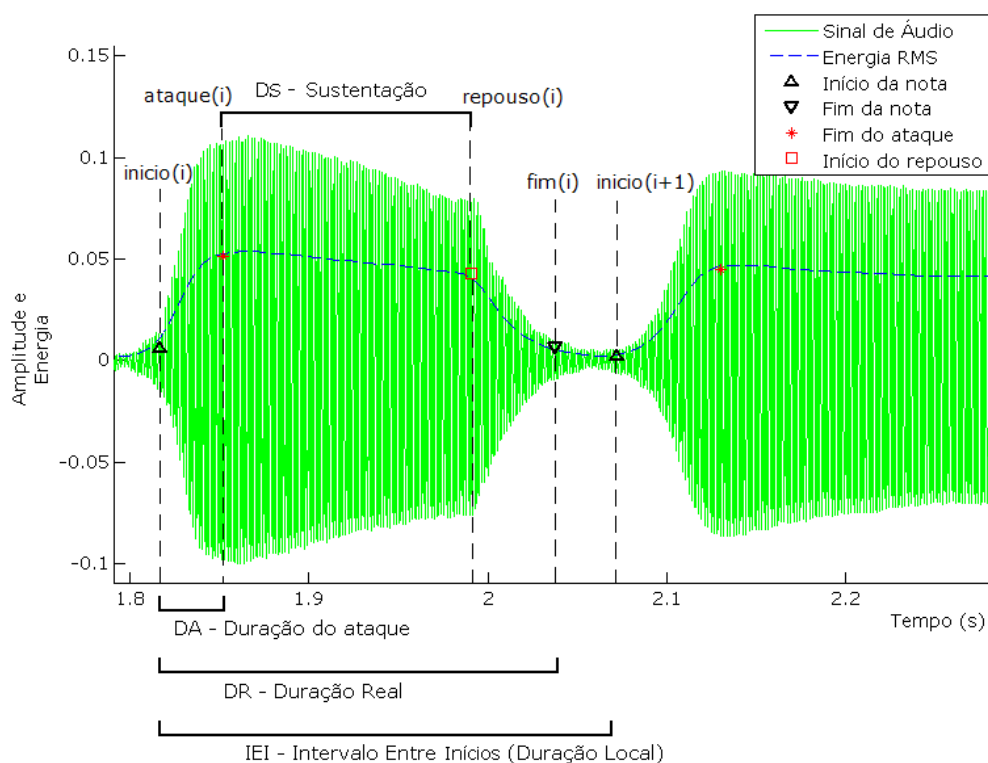


Figura 6: Descritores de duração definidos a partir dos instantes de início e final de nota, de final de ataque e início de decaimento.

- **Duração Local** - Intervalo Entre Inícios:  $IEI(i) = \text{início}(i+1) - \text{início}(i)$   
Intervalo de tempo medido entre os inícios de notas sucessivas, conhecido como *intra-onset-interval* (IOI), correspondente às especificações da partitura.

- *Duração Real da Nota*:  $DR(i) = fim(i) - início(i)$   
Intervalo de tempo entre o início e o fim da nota
- *Duração do Ataque*:  $DA(i) = ataque(i) - início(i)$   
Intervalo de tempo entre o início da nota e o final do ataque. Stephen McAdams e colegas mostraram que o valor do logaritmo da duração do ataque corresponde mais à nossa percepção de timbre, sendo portanto mais adequada quando o foco da descrição é a variação de timbre (McAdams, Winsberg *et al.*, 1995; Misdariis, Smith *et al.*, 1998).
- *Duração da Sustentação*:  $DS(i) = decaimento(i) - ataque(i)$

## 4.2 Inclinação do ataque

A inclinação do ataque pode fornecer informação sobre a qualidade do ataque e de articulações tais como *staccatti* e *sforzatti*. A inclinação do ataque pode ser calculada a partir da variação da energia RMS entre o ponto de início da nota e o ponto de final do ataque dividida pelo tempo de duração do ataque.

$$IA(i) = \frac{RMS(ataque(i)) - RMS(início(i))}{DA(i)}$$

## 4.3 Índice de legato

O índice de legato pode inferir sobre a qualidade da articulação entre notas consecutivas e é medido com a razão entre os valores de duração da nota e do tempo local:

$$IL(i) = \frac{DR(i)}{IEI(i)}$$

O valor de  $IL(i)$  é menor ou igual a 1, sendo igual a 1 para notas ligadas e está intimamente relacionado com o tempo de reverberação do ambiente de execução.

## 5. Resultados

A seguir, um exemplo dos procedimentos descritos acima em uma sequência de 7 notas do tema principal do primeiro movimento do Concerto para clarineta e orquestra em Lá Maior KV 622 de Mozart (notas em destaque na Figura 7). O envelope do trecho é mostrado abaixo juntamente com os pontos de início e fim de nota, fim de ataque e início de decaimento.



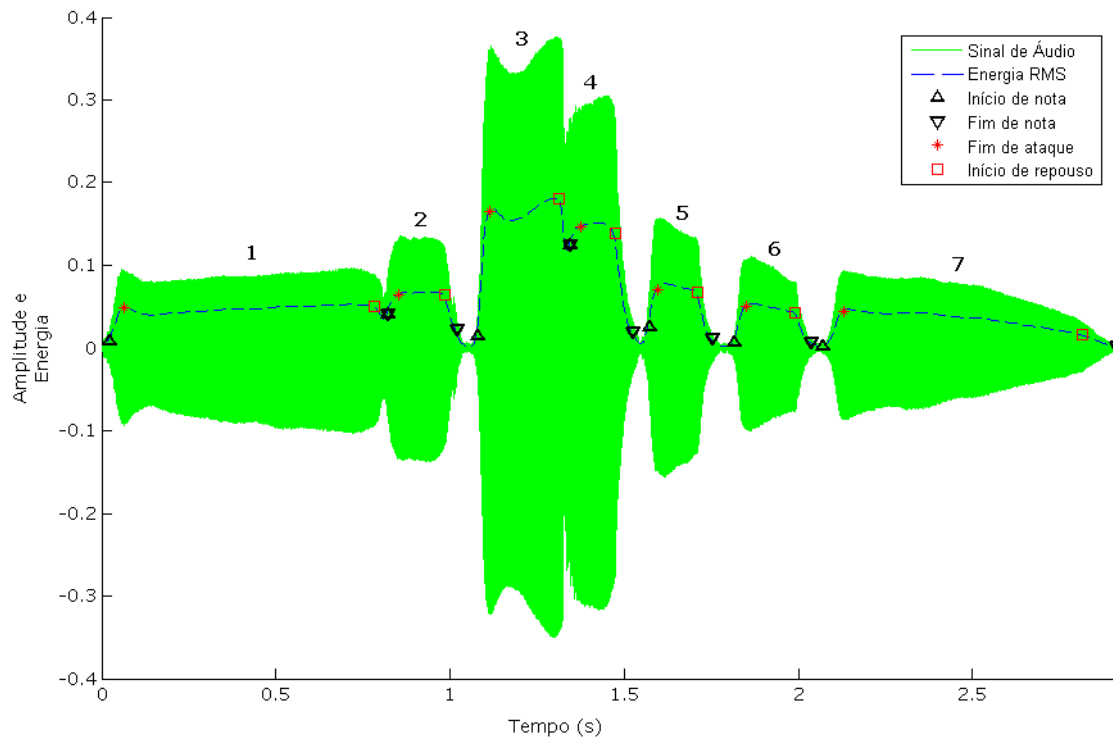


Figura 7: Exemplo de segmentação de uma sequência de 7 notas do tema principal do primeiro movimento do Concerto para clarineta e orquestra em Lá Maior KV 622 de Mozart: início e final de nota, final de ataque e início de decaimento.

A tabela abaixo mostra valores calculados para os descritores de envelope acima mencionados. Para melhor comparação, os valores de duração IEI, DR e DS foram normalizados para o valor da colcheia:

	Nota	IEI (s)	DR (s)	DS (s)	IL	IA
1	Mi5	0.2670	0.2670	0.2399	1	2.2164e-005
2	Fá5	0.2554	0.1974	0.1335	0.7727	1.8074e-005
3	Lá5	0.2670	0.2670	0.1974	1	9.8224e-005
4	Sol5	0.2264	0.1800	0.0987	0.7949	1.6952e-005
5	Fá5	0.2438	0.1800	0.1161	0.7381	4.3569e-005
6	Mi5	0.2554	0.2206	0.1393	0.8636	2.9062e-005
7	Mi5	-	0.2805	0.2283	-	1.6614e-005

## 6. Conclusão

Em um levantamento exaustivo da a pesquisa em performance musical conduzida até o final do século XX, Gabrielsson (2003) constatou que os estudos que focalizam a extração e a medição de parâmetros da

performance correspondem ao grupo mais numeroso deste levantamento. Mesmo encapsulando estruturas complexas e muitas vezes não acessíveis, a extração destes parâmetros envolve procedimentos menos complexos e cresce cada vez mais o número de trabalhos que buscam interpretar e identificar princípios gerais que possam estar por trás destes parâmetros. A análise deste grande volume de dados obtidos nestes estudos tem conduzido à formulação de uma grande variedade de modelos, que buscam descrever como e por que o músico modifica, às vezes inconscientemente, o que é indicado na partitura. Os resultados destas análises não apenas contribuirão para a compreensão deste problema complexo da pesquisa musicológica, mas poderão também possibilitar o surgimento de novos conceitos de abordagens pedagógicas objetivas para o ensino da música, em especial da prática instrumental.

## 7. Bibliografia

- Canazza, S., G. De Poli, S. Rinaldin e A. Vidolin. Sonological Analysis of Clarinet Expressivity. In: M. Leman (Ed.). *Music, Gestalt and Computing: Studies in Cognitive and Systematic Musicology*. Berlin-Heidelberg: Springer Verlag, p.431-440, 1997.
- De Poli, G. Modeling expressiveness in music performance. In: (Ed.), 2005.
- De Poli, G., A. Rodà e A. Vidolin. Note-by-Note Analysis of the Influence of Expressive Intentions and Musical Structure in Violin Performance. *Journal of New Music Research*, v.27, n.3, p.293-321, 1998.
- Gabrielsson, A. Expressive Intention and Performance. Music, Mind and Machine. In: R. Steiner (Ed.). New York: Springer, p.35-47, 1995.
- \_\_\_\_\_. Music Performance. In: D. Deutsch (Ed.). *Psychology of music*. New York: Academic Press, p.506-602, 1999.
- \_\_\_\_\_. Music Performance Research at the Millenium. *Psychology of Music*, v.31, p.221-272, 2003.
- Juslin, P. N. Emotional Communication in music performance: a functionalist perspective and some data. *Music Perception*, v.14, p.383-418, 1997.
- \_\_\_\_\_. Cue utilization in communication of emotion in music performance: relating performance to perception. *Journal of Experimental Psychology: Human perception and performance*, v.26, n.6, p.1797-1813, 2000.
- McAdams, S., S. Winsberg, S. Donnadieu, G. De Soete e J. Krimphoff. Perceptual Scaling of Synthesized Musical Timbres: Common Dimensions, Specificities and Latent Subject Classes. *Psychological Research*, v.58, p.177-192, 1995.
- Misdariis, N. R., B. K. Smith, D. Pressnitzer, P. Susini e S. McAdams. Validation of a Multidimensional Distance Model for Perceptual Dissimilarities Among Musical Timbres. *16th International Congress on Acoustics*, Woodbury, New York: ASA - The Acoustical Society of America. p., 1998.
- Palmer, C. Anatomy of a Performance: Sources of Musical Expression. *Music Perception*, v.13, p.433-454, 1996a.
- \_\_\_\_\_. On the Assignment of Structure in Music Performance. *Music Perception*, v.14 n.1, p.23-56, 1996b.
- \_\_\_\_\_. Music Performance. *Annual Review of Psychology*, v.48, p.115-138, 1997.
- Park, T. H. *Towards Automatic Musical Instrument Timbre Recognition*. (Ph. D. Dissertation). Department of Music, Princeton University, 2004.



- Repp, B. H. Patterns of Expressive Timing in Performances of a Beethoven Minuet by 19 Famous Pianists. *Journal of the Acoustical Society of America*, v.88, p.622-641, 1990.
- . Diversity and Commonality in Music Performance - An Analysis of Timing Microstructure in Schumann's Traumerei. *Journal of the Acoustical Society of America*, v.92 n.5, p.2546-2568, 1992.
- Seashore, C. E. *Psychology of Music*. New York: McGraw-Hill, 1938.
- Shaffer, L. H., E. F. Clarke e N. P. M. Todd. Meter and Rhythm in Piano Playing. *Cognition*, v.20, p.61-77, 1985.
- Sloboda, J. A. Expressive Skill in Two Pianists. *Canadian Journal of Psychology*, v.39, p.273-293, 1985.
- Timmers, R., R. Ashley, P. Desain e H. Heijink. The Influence of Musical Context on Tempo Rubato. *Journal of New Music Research*, v.29, n.2, p.131-158, 2000.
- Todd, N. P. M. A Model of Expressive Timing in Tonal Music. *Music Perception*, v.3, p.33-58, 1985.