

ANÁLISE DE FLUTUAÇÕES DE VARIÂNCIA EM SINAIS DE ÁUDIO AGRUPADOS POR GÊNERO MUSICAL

DIRCEU DE FREITAS PIEDADE MELO*

*IFBA - Departamento de Ciências Exatas
Rua Emídio Santos s/n, Barbalho
Salvador, Bahia, Brasil

Email: dirceu@ifba.edu.br

Resumo— Neste artigo será apresentado o estudo das flutuações de variância de sinais de áudio, classificados em gêneros musicais, utilizando um algoritmo derivado do *Detrended Fluctuation Analysis*, denominado *Detrended Variance Fluctuation Exponent*. Através do cálculo deste expoente, denotado aqui por α_{DFA} , será avaliada a *danceability* do sinal de áudio, que está relacionada com o padrão de atividade de instrumentos percussivos nos mesmos. A conclusão é que uma atividade percussiva muito intensa e repetitiva, enfraquece as correlações de longo alcance da série temporal, obtida a partir das flutuações de variância do sinal musical. Assim, ao serem computadas as suas flutuações destendenciadas, observa-se que, quanto maior atividade percussiva de padrão repetitivo no sinal de áudio, menor será o expoente α_{DFA} . Desta maneira é possível estabelecer uma associação entre o α_{DFA} médio de um gênero musical e o sua respectiva característica rítmica.

Palavras-chave— Detrended Fluctuation Analysis, Detrended Variance Fluctuation Exponent, Gêneros Musicais, Séries Temporais, Classificação de Sinais de Áudio.

1 Introdução

O crescente número de arquivos de música em formato digital, disponíveis na Web, em coleções particulares, ou em acervos públicos, tem motivado o estudo da classificação automática de arquivos musicais. Parte dessa iniciativa pode ser atribuída ao fato da classificação de arquivos musicais, ainda ser feita de forma manual. A classificação automática pode então criar uma nova alternativa para realizar a manipulação e organização destes arquivos. Uma das formas mais utilizadas na organização de arquivos de música é a classificação em gêneros musicais. Quando realizada de forma automática, a classificação em gêneros conta com os descritores musicais, para a realização de uma de suas etapas mais importantes, a extração de características. Assim, o estudo de novos descritores musicais, pode colaborar para o desenvolvimento do processo de classificação automática. Os modelos de classificação automática se baseiam na extração de parâmetros ou vetores de características, que devem fornecer o máximo possível de informação relativa aos sinais de áudio (PEREIRA, 2009). Um novo descritor para esse campo de pesquisa, denominado de *Detrended Variance Fluctuation Exponent* (DVFE), foi introduzido por JENNINGS (2004), e ao quantificar as propriedades de escala de variâncias locais em séries não-estacionárias, foram encontradas diferenças quantitativas, associadas a gêneros musicais. Essas diferenças têm relação com propriedades rítmicas do sinal musical.

2 Metodologia

2.1 DFA

O DFA (*Detrended Fluctuation Analysis*) é um método estatístico que permite eliminar a tendência de uma série temporal em diferentes escalas, analisando as flutuações intrínsecas da série. O método DFA foi proposto por PENG (1994). e tem se consolidado como uma ferramenta de grande importância, na detecção de correlações de longo alcance em séries temporais não-estacionárias. Esse método é baseado na teoria de caminhos aleatórios (*random walk theory*) (MANDELBROT, 1983), e é um aperfeiçoamento do método *Fluctuation Analysis* (FA) (PENG, 1994).

O algoritmo para a determinação do DFA é descrito a seguir. Seja $\{x(t)\}$ uma série temporal de tamanho N e $\langle x \rangle$ sua média. Em primeiro lugar determina-se o perfil através da equação:

$$y(t) = \sum_{k=1}^N (x(k) - \langle x \rangle) \quad (1)$$

O perfil é então dividido em caixas de tamanho n . Para cada caixa, é removida a tendência, subtraindo um polinômio de grau k , y_n^k , através do método dos mínimos quadrados. Calcula-se então a flutuação do sinal

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N (y(t) - y_n^k(t))^2}, \quad (2)$$

Se a série original apresenta correlações de longo alcance, então os valores de $F(n)$ seguem uma lei de potência:

$$F(n) \propto n^\alpha, \quad (3)$$

Através de um ajuste linear em escala $\log x \log$ podemos calcular o expoente α , que nos dá informação a respeito dos padrões de correlação da série.

2.2 DVFE

O procedimento para calcular o *Detrended Variance Fluctuation Exponent* (DVFE), ou α_{DFA} , descrito a seguir, segue a metodologia descrita em JENNINGS (2004).

Em primeiro lugar o conjunto $\{U(i)\} = \{U(1), \dots, U(N)\}$ é segmentado em m blocos ou caixas não-sobrepostas de tamanho λ . Aqui foi adotado um tamanho de caixa λ igual 110 *samples*, que está associado a um tempo de 10ms.

Para cada caixa $j = 1 \dots m$ é calculado o desvio-padrão.

Na a j -ésima caixa temos:

$$V(j) = \sqrt{\frac{\sum_{(j-1)\cdot\lambda+1}^{j\lambda} (U(i) - \bar{U}_j)^2}{\lambda - 1}}, \quad (4)$$

onde a média é dada por

$$\bar{U}_j = \frac{\sum_{(j-1)\cdot\lambda+1}^{j\lambda} (U(i))}{\lambda} \quad (5)$$

Com isso cria-se uma nova série $V(j) = \{V(1), V(2), \dots, V(m)\}$, com N/λ amostras. Efetua-se então a integração de $V(j)$,

$$Y(m) = \sum_{j=1}^m V(j), \quad (6)$$

A sequência gerada por $Y(m)$ é também dividida em subsequências sobrepostas de tamanho τ . Cada subsequência é deslocada de uma única amostra em relação à subsequência anterior (JENNINGS, 2004).

Para cada bloco de comprimento τ é removida a tendência linear \hat{y}_k . Em seguida é computada a média do quadrado residual para cada bloco k (STREICH and HERRERA, 2005),

$$D(k, \tau) = \frac{1}{\tau} \cdot \sum_{m=0}^{\tau-1} (y(k+m) - \hat{y}_k(m))^2 \quad (7)$$

A flutuação DFA é dada por:

$$F(\tau) = \sqrt{\frac{1}{K} \sum_{i=1}^K (D(k, \tau))} \quad (8)$$

Aqui foi utilizada uma variação de τ de 31 a 909. Essa variação corresponde a uma janela temporal de 0,31s a 9,09s, denominada de janela de interesse.

$$\alpha(i) = \frac{\log_{10} F(\tau_{i+1}) - \log_{10} F(\tau_i)}{\log_{10}(\tau_{i+1} + 3) - \log_{10}(\tau_i + 3)} \quad (9)$$

Finalmente, o α_{DFA} é calculado através da Equação 10.

$$\alpha_{DFA} = \frac{\sum_{i=1}^{32} \alpha(i)}{32} \quad (10)$$

A batida periódica da música, considerada como uma superposição de tendências periódicas do sinal acústico, ocasiona desvios importantes na escala da lei de potência, dentro da janela de interesse (JENNINGS, 2004). Assim, quanto menor correlação de longo alcance das flutuações de *loudness*, menor será a estabilidade de escala, devido a um maior desvio da escala da lei de potência. Se temos um maior desvio, teremos, em consequência, uma maior predominância de tendências periódicas. Com isso associa-se um menor valor do α_{DFA} a sinais de áudio com atividade percussiva mais regular. Esses sinais musicais estão associados a gêneros mais dançantes, ou com maior *danceability* segundo (STREICH and HERRERA, 2005). Por outro lado, quanto maior a correlação nas flutuações de *loudness*, menor será a predominância de tendências periódicas, resultando em um maior valor para o α_{DFA} . Desta maneira, sinais musicais com menor *danceability* estão associados a um maior valor do α_{DFA} .

3 Contribuição deste artigo

4 Resultados

Foram calculados os valores do α_{DFA} aplicando-se o algoritmo da *Detrended Variance Fluctuation Exponent* nos arquivos de áudio do banco BRMUSIC Genre Collection <http://www.4shared.com/folder/KsxAP6fL/brmusicdatabase.html>. Suas características são apresentadas em (Tabela 1).

Tabela 1: BRMUSIC GENRE COLLECTION

formato	wave
taxa de amostragem	44.100 Hz
taxa de quantização	16 bit
modo	stereo
# gêneros	10
# tracks/gênero	20
# total de tracks	200
# duração/track	completo

O banco BRMUSIC é constituído por 10 gêneros da música popular brasileira, com um total de 200 tracks: Axé Music, Funk Carioca, Samba, Forró, Mangue Beat, Rock Nacional, Clube da

Esquina, Tropicalismo, Vanguarda Paulistana e Bossa-Nova. Foi utilizado também um *dowsampling* com 11.025 Hz e 8 bit. Neste banco foram preservadas as durações originais de cada música e foram capturadas diretamente do formato digital. O procedimento computacional adotado para implementar a função $\alpha(t)$, que calcula o *Detrended Variance Fluctuation Exponent*, ou simplesmente α_{DFA} segundo STREICH and HERRERA (2005), segue a metodologia descrita em JENNINGS (2004).

Para os 10 gêneros do banco BRMUSIC foram calculados o α_{DFA} bem como o $\langle \alpha_{DFA} \rangle$ e os seus respectivos desvios-padrão. (Tabela 2)

Realizou-se então a análise de variância (ANOVA way1) para avaliar se o DFA expoente médio $\langle \alpha_{DFA} \rangle$, seria representativo ou não para estabelecer diferenças entre os gêneros. Foi aplicado então o teste *Shapiro – Wilk* e foi verificado que, com um nível de 95% de confiança, os agrupamentos do banco atenderam as condições de aderência à distribuição normal. Segundo o teste *Tukey*, a um nível de 95% de confiança, foi que, para alguns pares de gêneros, a diferença entre as médias não são significantes. O gráfico 3 mostra as diferenças entre as médias, segundo o teste *Tukey* e os resultados apresentados na Figura 1.

Tabela 2: Valores médios do α_{DFA} do banco BRMUSIC.

	$\langle \alpha \rangle$	σ
Axé Music	0,677	0,094
Funk Carioca	0,688	0,096
Samba	0,695	0,050
Forró	0,702	0,119
Mangue Beat	0,754	0,178
Rock Nacional	0,811	0,181
Clube da Esquina	0,831	0,088
Tropicalismo	0,883	0,163
Vanguarda Paulistana	1,041	0,175
Bossa-Nova	1,114	0,080

5 Análise de Resultados e Conclusão

O DVFE apresentou respostas satisfatórias na descrição da atividade rítmica nos gêneros musicais do banco BRMUSIC, pois os resultados numéricos apresentaram, de um modo geral, coerência com a experiência do senso comum, onde os gêneros que possuem uma presença percussiva mais intensa e repetitiva retornam um α_{DFA} menor que gêneros que apresentam esse tipo de atividade em menor escala. Isso é mostrado na Figura 1, onde os gêneros são apresentados crescente de *danceability*, e ordem decrescente de α_{DFA} . Na Figura 2 são mostradas as flutuações de α_{DFA} nos gêneros, onde fica evidente uma grande e intensa atividade

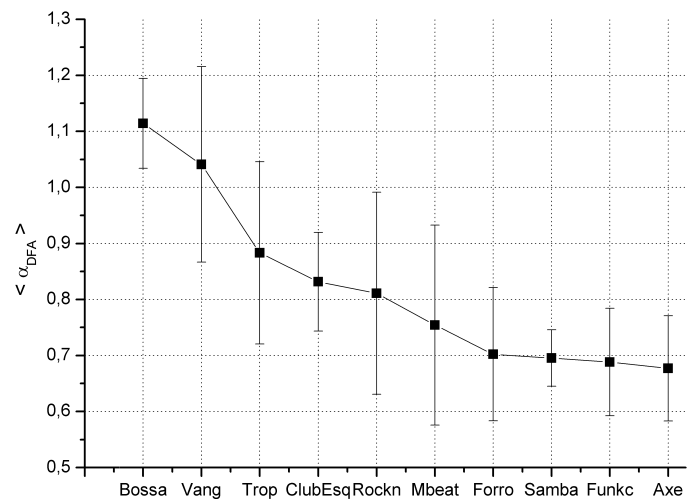


Figura 1: α_{DFA} médio do banco BRMUSIC.

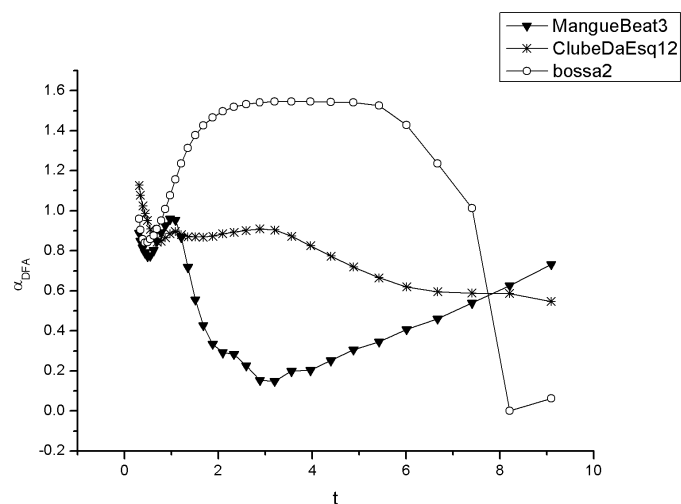


Figura 2: Flutuações do α_{DFA} dos tracks Bossa-Nova 2, Clube da Esquina 12 e Mangue Beat 3.

de instrumentos percussivos no track do gênero *Mangue Beat*, uma baixa atividade em *Bossa-Nova*, e uma atividade intermediária em *Clube da Esquina*.

Quando se tratou do $\langle \alpha_{DFA} \rangle$ para representar um gênero musical, concluiu-se que, é necessária a realização de testes de análise de variância. Nesta pesquisa foi realizado o Teste *Tukey*, que mostrou quando existe diferença significativa entre α_{DFA} médios dos gêneros estudados. Isto pode sugerir que tais gêneros possuem ou não uma maior afinidade rítmica; se compartilham ou não elementos estéticos; ou se podem ou não ser reagrupados com um rótulo comum. Segundo esse ponto de vista, pelos critérios estatísticos do Teste *Tukey*, não existiriam diferenças

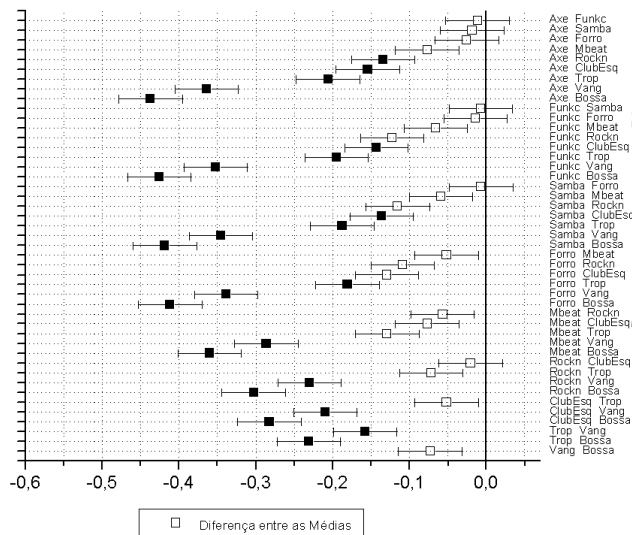


Figura 3: Comparação entre $\langle \alpha_{DFA} \rangle$ do banco BRMUSIC, segundo o teste *Tukey*.

quantitativas significativas entre determinados pares de gêneros, a partir do seu α_{DFA} médio. Essas conclusões se contrapõem as de JENNINGS (2004), que considera o $\langle \alpha_{DFA} \rangle$ como suficiente para estabelecer diferenças quantitativas entre os gêneros musicais. Se utilizadas as diferenças significativas, segundo o teste *Tukey* o banco BRMUSIC pode ser classificado em três grupos segundo a *danceability*. Um grupo formado por Bossa-nova e Vanguarda Paulistana, que tem $\langle \alpha_{DFA} \rangle$ com baixa *danceability*; outro grupo formado por Tropicália e Clube da Esquina, ocupando uma posição intermediária; e Rock Nacional, Mangue Beat, Forró, Samba, Funk Carioca e Axé Music, com alta *danceability*. Esse tipo de informação é muito importante em modelos de classificação automática, e se utilizada em associação a outros descritores, num vetor de características, pode alcançar resultados bastante satisfatórios, como em TERMENS (2009).

Referências

- JENNINGS, H. e. a. (2004). Variance fluctuations in nonstationary time series: a comparative study of music genres, *Physica A* (336): 585–594. DOI: [10.1016/j.physa.2003.12.049](https://doi.org/10.1016/j.physa.2003.12.049)
- MANDELBROT, B. (1983). *The Fractal Geometry of Nature*, Freeman and Company, New York.
- PENG, C. e. a. (1994). Mosaic organization of dna nucleotides, *Physical Review* (E49): 1685–16895.
- PEREIRA, E. M. (2009). *Estudos Sobre uma Ferramenta de Classificação Musical*, Disserta-

ção de mestrado, Faculdade de Engenharia Elétrica e de Computação - UNICAMP, São Paulo.

STREICH, S. and HERRERA, P. (2005). Detrended fluctuation analysis of music signals: Danceability estimation and further semantic characterization, in D. Keller and Z. Budas (eds), *Audio Engineering Society Convention, 118th*, Barcelona, Spain.

TERMENS, E. G. (2009). *Audio Content Processing for Automatic Music Classification: Descriptors, Databases, and Classifiers*, Doctor in computer science and digital communication, Universitat Pompeu Fabra, Barcelona.