

Danilo Makoto Ikuta

TCC Título

São Paulo – Brasil

2016

Danilo Makoto Ikuta

TCC Título

Modelo canônico de trabalho monográfico acadêmico em conformidade com as normas ABNT apresentado à comunidade de usuários \LaTeX .

Centro Universitário Senac

Orientador: Gabriel de Faria Andery

São Paulo – Brasil

2016

Resumo

Segundo a [ABNT \(2003, 3.1-3.2\)](#), o resumo deve ressaltar o objetivo, o método, os resultados e as conclusões do documento. A ordem e a extensão destes itens dependem do tipo de resumo (informativo ou indicativo) e do tratamento que cada item recebe no documento original. O resumo deve ser precedido da referência do documento, com exceção do resumo inserido no próprio documento. (...) As palavras-chave devem figurar logo abaixo do resumo, antecidas da expressão Palavras-chave:, separadas entre si por ponto e finalizadas também por ponto.

Palavras-chaves: latex. abntex. editoração de texto.

Abstract

This is the english abstract.

Key-words: latex. abntex. text editoration.

Lista de ilustrações

Lista de tabelas

Lista de abreviaturas e siglas

Fig. Area of the i^{th} component

456 Isto é um número

123 Isto é outro número

lauro cesar este é o meu nome

Sumário

1	Introdução	8
2	Revisão Bibliográfica	9
2.1	Obtenção do Espectrograma Através de Transformadas de Fourier	9
2.2	Separação de Vocal/Melodia e Acompanhamento	9
2.2.1	Estimativa de Tremolo/Vibrato	9
2.2.2	Harmonic Coded Structure(HCS)	10
2.2.3	Estimativa de Frequências Fundamentais(F0)	10
2.2.4	Somatória Sub-harmônica Normalizada	10
2.2.5	Rastreamento Harmônico	10
2.3	Separação de Elementos Harmônicos e Percussivos	11
3	Proposta do Trabalho	12
3.1	Lectus lobortis condimentum	12
3.2	Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae	12
4	Expectativas	13
	Referências	14

1 Introdução

Cifras, partituras, covers, vídeo aulas e outras fontes podem auxiliar um músico a aprender a tocar uma determinada música em um instrumento musical. Para músicas mais conhecidas, tal material de apoio pode ser encontrado em livros, revistas ou na internet. Entretanto, esse material pode não existir ou estar em outra língua, dependendo da música e do instrumento que se deseja aprender, podendo exigir certa habilidade para discernir instrumentos e padrões dentro da música.

Para músicos menos experientes, aprender uma música utilizando-se apenas da habilidade auditiva(“de ouvido”) pode ser uma tarefa árdua. As dificuldades podem incluir a identificação de acordes, identificação de um instrumento ou melodia em músicas onde há múltiplos instrumentos, simultaneamente, e identificar qual ou quais notas estão sendo tocadas em dado momento da música. Possíveis formas de auxiliar esse processo de aprendizado incluem a extração de uma certa característica e a exclusão de características indesejadas.

O campo de recuperação de informação musical(MIR, do inglês Music Information Retrieval) trata, dentre outros assuntos, da identificação e extração de instrumentos(fontes sonoras), melodia, percussão, acompanhamento entre outras características de uma música.

Para a análise de áudio, costuma-se usar o espectrograma, que representa o som através de um gráfico frequência X tempo, onde cada coordenada possui uma certa intensidade (amplitude).

Com a finalidade de auxiliar músicos menos experientes, a geração de arquivos de áudio derivados da análise e extração de componentes de uma música, como um instrumento específico, vozes, percussão, melodia e acompanhamento é descrito.

Serão comparadas técnicas tanto voltadas à instrumentos e características específicos quanto métodos com foco na abrangência de instrumentos e características.

Dentre os desafios a serem ponderados, estão a relação entre componentes harmônicos e de ruído dos instrumentos e a interação entre múltiplos instrumentos em um curto intervalo de tempo, no qual a frequência e/ou timbre pode afetar a qualidade da análise.

2 Revisão Bibliográfica

2.1 Obtenção do Espectrograma Através de Transformadas de Fourier

- FFT [Roads et al. \(1995\)](#)

Segundo [Roads et al. \(1995\)](#), um espectrograma é uma forma de se visualizar informações sonoras no decorrer do tempo, em um gráfico de tempo e frequência, onde uma escala de cor representa a amplitude(intensidade) em cada intervalo de frequência e tempo. Atualmente, a forma mais comum de se calcular o espectrograma é utilizando a transformada de Fourier de tempo curto(STFT, Short-time Fourier Transform).

Para Fourier como citar [Ivor \(1995\)](#)?, qualquer sinal periódico, independente da complexidade, pode ser representado como uma somatória de diversas ondas simples(em especial ondas senoidais)

- MR-FFT [Hsu e Jang \(2010\)](#)

2.2 Separação de Vocal/Melodia e Acompanhamento

2.2.1 Estimativa de Tremolo/Vibrato

[Hsu e Jang \(2010\)](#) e [Regnier e Peeters \(2009\)](#) partem do princípio de que poucos instrumentos, além da voz humana, apresentam simultaneamente ambos vibrato e tremolo ([VERFAILLE; GUASTAVINO; DEPALLE, 2005](#)). Vibrato se refere à modulação de frequência(FM) ou a micro-variação de tom, ao passo que o tremolo refere-se à modulação de amplitude(AM) ou micro-variações na intensidade. Considerando a taxa de vibrato/tremolo em torno de 6 Hz para a voz humana, calcula-se as amplitudes(extensão) de tremolo/vibrato para cada parcial p_k , com duração de t_i até t_j . Para o cálculo do vibrato, a transformada de Fourier das frequências f_{p_k} é:

$$F_{p_k}(f) = \sum_{t=t_i}^{t_j} (f_{p_k}(t) - \mu_{f_{p_k}}) e^{-2i\pi f \frac{t}{L}}, \quad (2.1)$$

onde $\mu_{f_{p_k}}$ é a frequência média da parcial $p_k(t)$ e $L = t_j - t_i$. Em seguida, é calculada a extensão relativa:

$$\Delta f_{rel p_k}(f) = \frac{F_{p_k}(f)}{L \mu_{f_{p_k}}}, \quad (2.2)$$

enfim, calcula-se a extensão relativa entre 6Hz:

$$\Delta f_{p_k} = \max_{f \in [4,8]} \Delta f_{rel p_k}(f). \quad (2.3)$$

Para o cálculo de tremolo, utiliza-se as mesmas equações para o cálculo de vibrato, mas substituindo f_{p_k} por a_{p_k}

2.2.2 Harmonic Coded Structure(HCS)

O HCS é utilizado por [Joo et al. \(2011\)](#) no processo de encontrar os tons dominantes da melodia em um espectrograma. Uma vez que as amplitudes harmônicas variam com o instrumento e tom [citar fonte?](#), fez-se necessário construir um codebook a partir de amostras de notas em diversos instrumentos e voz humana, agrupando-as de acordo com o algoritmo proposto em ([PELLEG; MOORE, 2000](#)).

Calculam-se os HCSs para cada possível frequência fundamental(F0) seguindo as equação:

$$h_\eta[n] = w[n] \sum_{m=1}^H b_m \cos(m \cdot 2\pi\eta \cdot n + \phi_m), H = \lfloor \frac{f_s}{2\eta} \rfloor, \quad (2.4)$$

onde f_s é a frequência de amostragem, η a frequência fundamental(F0) do HCS, $w[n]$ a janela de análise, b_m a amplitude da m-ésima harmônica, e ϕ_m a fase de b_m .

2.2.3 Estimativa de Frequências Fundamentais(F0)

- Estimativa de tendência de tom/F0 [Hsu e Jang \(2010\)](#), [Cao et al. \(2007\)](#), [Joo et al. \(2011\)](#)

2.2.4 Somatória Sub-harmônica Normalizada

Uma forma alternativa e complementar ao rastreamento harmônico é o método de somatória sub-harmônica, proposto por [Hermes \(1986\)](#) e implementado em [Cao et al. \(2007\)](#) e [Hsu et al. \(2009\)](#) por representar cada quadro com as estruturas harmônicas realçadas, em especial as provenientes de vocais. O algoritmo é dado por:

$$H_t(f) = \sum_{n=1}^N h_n P_t(nf), \quad (2.5)$$

onde $H_t(f)$ é o valor de somatória sub-harmônica da frequência f no quadro t , $P()$ é o espectro calculado via STFT, N a quantidade dos n componentes harmônicos considerados e h_n é o peso do n-ésimo componente harmônico.

2.2.5 Rastreamento Harmônico

- Rastreamento harmônico [Cao et al. \(2007\)](#)
- Remoção parcial de instrumentos [Hsu e Jang \(2010\)](#)
- Cálculo dos ESIs [Hsu e Jang \(2010\)](#), [Hsu et al. \(2009\)](#)

- Extração do contorno/estrutura mais provável [Hsu e Jang \(2010\)](#), [Cao et al. \(2007\)](#), [Joo et al. \(2011\)](#)
- Estrutura harmônica estável [Cao et al. \(2007\)](#)
- Matrizes de similaridade [Rafii e Pardo \(2012\)](#), [Rafii et al. \(2013\)](#)
- Modelo de repetição(acompanhamento) [Rafii e Pardo \(2012\)](#), [Rafii et al. \(2013\)](#)
- Modelos Ocultos de Markov(HMM) [Hsu et al. \(2009\)](#)
- Modelo de voz [Rafii et al. \(2013\)](#) - Divergência de Kullback-Leibler [Rafii et al. \(2013\)](#)
- Máscara tempo-frequência [Rafii e Pardo \(2012\)](#), [Rafii et al. \(2013\)](#), [Driedger, Müller e Disch \(2014\)](#)
- Fatorização de matriz não-negativa(NMF) [Rafii et al. \(2013\)](#)
- Filtros de Wiener [Rafii et al. \(2013\)](#)
- Mel-Frequency Cepstral Coefficients(MFCC) [Hsu et al. \(2009\)](#)
- Suavização de contorno [Joo et al. \(2011\)](#)

2.3 Separação de Elementos Harmônicos e Percussivos

- Filtros de mediana [Driedger, Müller e Disch \(2014\)](#)

3 Proposta do Trabalho (O que vai ser desenvolvido!)

3.1 Lectus lobortis condimentum

3.2 Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae

4 Expectativas

Referências

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 6028*: Resumo - apresentação. Rio de Janeiro, 2003. 2 p. Citado na página 2.

CAO, C. et al. Singing melody extraction in polyphonic music by harmonic tracking. In: *Proceedings of the 8th International Conference on Music Information Retrieval, ISMIR 2007, Vienna, Austria, September 23-27, 2007*. [s.n.], 2007. p. 373–374. Disponível em: <http://ismir2007.ismir.net/proceedings/ISMIR2007_p373_cao.pdf>. Acesso em: 21.05.2016. Citado 2 vezes nas páginas 10 e 11.

DRIEDGER, J.; MÜLLER, M.; DISCH, S. Extending harmonic-percussive separation of audio signals. In: *Proceedings of the 15th International Society for Music Information Retrieval Conference, ISMIR 2014, Taipei, Taiwan, October 27-31, 2014*. [s.n.], 2014. p. 611–616. Disponível em: <http://www.terasoft.com.tw/conf/ismir2014/proceedings/T110_127_Paper.pdf>. Acesso em: 13.05.2016. Citado na página 11.

HERMES, D. J. Measurement of pitch by subharmonic summation. *Journal of Acoustic of Society of America*, v. 83, p. 257–264, 1986. Citado na página 10.

HSU, C. et al. Singing pitch extraction from monaural polyphonic songs by contextual audio modeling and singing harmonic enhancement. In: *Proceedings of the 10th International Society for Music Information Retrieval Conference, ISMIR 2009, Kobe International Conference Center, Kobe, Japan, October 26-30, 2009*. [s.n.], 2009. p. 201–206. Disponível em: <<http://ismir2009.ismir.net/proceedings/PS2-2.pdf>>. Acesso em: 20.05.2016. Citado 2 vezes nas páginas 10 e 11.

HSU, C.; JANG, J. R. Singing pitch extraction by voice vibrato / tremolo estimation and instrument partial deletion. In: *Proceedings of the 11th International Society for Music Information Retrieval Conference, ISMIR 2010, Utrecht, Netherlands, August 9-13, 2010*. [s.n.], 2010. p. 525–530. Disponível em: <<http://ismir2010.ismir.net/proceedings/ismir2010-89.pdf>>. Acesso em: 18.05.2016. Citado 3 vezes nas páginas 9, 10 e 11.

IVOR, G. *Joseph Fourier: 1768-1830*. [S.l.]: The MIT Press, 1995. Citado na página 9.

JOO, S. et al. Melody extraction based on harmonic coded structure. In: *Proceedings of the 12th International Society for Music Information Retrieval Conference, ISMIR 2011, Miami, Florida, USA, October 24-28, 2011*. [s.n.], 2011. p. 227–232. Disponível em: <<http://ismir2011.ismir.net/papers/PS2-10.pdf>>. Acesso em: 17.05.2016. Citado 2 vezes nas páginas 10 e 11.

PELLEG, D.; MOORE, A. W. X-means: Extending k-means with efficient estimation of the number of clusters. In: *Proceedings of the Seventeenth International Conference on Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000. (ICML '00), p. 727–734. ISBN 1-55860-707-2. Disponível em: <<http://dl.acm.org/citation.cfm?id=645529.657808>>. Acesso em: 09.08.2016. Citado na página 10.

RAFII, Z. et al. Combining modeling of singing voice and background music for automatic separation of musical mixtures. In: *Proceedings of the 14th International Society for Music Information Retrieval Conference, ISMIR 2013, Curitiba, Brazil, November 4-8, 2013*. [s.n.], 2013. p. 41–46. Disponível em: <[http://www.ppgia.pucpr-br/ismir2013/wp-content/uploads/2013/09/63_Paper.pdf](http://www.ppgia.pucpr.br/ismir2013/wp-content/uploads/2013/09/63_Paper.pdf)>. Acesso em: 15.05.2016.

Citado na página 11.

RAFII, Z.; PARDO, B. Music/voice separation using the similarity matrix. In: *Proceedings of the 13th International Society for Music Information Retrieval Conference, ISMIR 2012, Mosteiro S.Bento Da Vitória, Porto, Portugal, October 8-12, 2012*. [s.n.], 2012. p. 583–588. Disponível em: <<http://ismir2012.ismir.net/event/papers/583-ismir-2012.pdf>>.

Acesso em: 15.05.2016. Citado na página 11.

REGNIER, L.; PEETERS, G. Singing voice detection in music tracks using direct voice vibrato detection. In: *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. [s.n.], 2009. p. 1685–1688. Disponível em: <<http://articles.ircam.fr/textes/Regnier09b/index.pdf>>. Acesso em: 08.06.2016. Citado na página 9.

ROADS, C. et al. *The Computer Music Tutorial*. [S.l.]: The MIT Press, 1995. Citado na página 9.

VERFAILLE, V.; GUASTAVINO, C.; DEPALLE, P. Perceptual evaluation of vibrato models. In: *Proceedings of Conference on Interdisciplinary Musicology (CIM05)*. [s.n.], 2005. Disponível em: <http://oicrm.org/wp-content/uploads/2012-03/VERFAILLE_V_CIM05.pdf>. Acesso em: 08.06.2016. Citado na página 9.