

UNIVERSIDADE FEEVALE

GUSTAVO HENRIQUE STRASSBURGER

TRANSCRIÇÃO AUTOMÁTICA DE MÚSICA DE ÁUDIOS
MONOFÔNICOS

(Título Provisório)

Anteprojeto de Trabalho de Conclusão

Novo Hamburgo
2016

GUSTAVO HENRIQUE STRASSBURGER

TRANSCRIÇÃO AUTOMÁTICA DE MÚSICA DE ÁUDIOS
MONOFÔNICOS

(Título Provisório)

Anteprojeto de Trabalho de Conclusão de
Curso, apresentado como requisito parcial
à obtenção do grau de Bacharel em
Sistemas de Informação pela
Universidade Feevale

Orientador: Gabriel Simões

Novo Hamburgo
2016

RESUMO

O campo da recuperação de informação de música tem crescido de forma contundente nos últimos anos, tendo o crescimento da disponibilidade de recursos como um dos principais responsáveis para tal evento. Uma de suas ramificações, a transcrição automática de música, também tem sido, cada vez mais, objeto de pesquisa de artigos científicos e teses. Em razão disto, novos métodos têm surgido a fim de solucionar as lacunas ainda existentes. Sendo objetivo da área transcrever sinais de áudio de música em notações de música conhecidas, as pesquisas, em grande parte, objetivam melhorar o processo de detecção de melodias, de tal forma que seja possível produzir softwares mais confiáveis e com menos restrições de uso. Sendo assim, este trabalho tem como objetivo analisar o estado da arte da transcrição automática de música e, a partir desta pesquisa, desenvolver um protótipo de software capaz de detectar, extrair e converter sinais de áudio empregando as técnicas que apresentarem melhores resultados nas etapas da detecção de harmonias em trabalhos relacionados.

Palavras-chave: Transcrição Automática de Música. Recuperação de Informação de Música. Sinais de Áudio Monofônicos.

SUMÁRIO

MOTIVAÇÃO.....	5
OBJETIVOS.....	8
METODOLOGIA.....	9
CRONOGRAMA.....	10
BIBLIOGRAFIA.....	11

MOTIVAÇÃO

O crescimento da pesquisa na área da recuperação da informação da música (MIR, do inglês *Music Information Retrieval*) deve-se, sobretudo, em razão do aumento de disponibilidade de obras musicais em formato digital (SANTINI; SOUZA, 2007). Apesar de sua existência desde a década de 1960, apenas recentemente, com maior poder computacional e desenvolvimento de novos algoritmos, a área consegue produzir resultados satisfatórios (ELLIS, 2006).

O processo de conversão de sinais de áudio em notações musicais estabelecidas, estudada pela área da transcrição automática de música (AMT do inglês *Automatic Music Transcription*) – subárea da MIR – ainda não está inteiramente solucionado. Atualmente, a transcrição de sinais monofônicos já pode ser considerada compreendida, no entanto, a transcrição de sinais polifônicos e sem restrição dos instrumentos ainda requer maior aprofundamento (BENETOS et al, 2012).

Extração de *pitch* (altura da nota), *onset* (início da execução de uma nota) e duração de uma nota são aspectos essenciais para recriar ou transcrever uma determinada melodia. Aspectos físicos do sinal, tais como taxa de repetição, amplitude da onda e formato da onda, correlacionam-se com aspectos menos abstratos, sendo eles, respectivamente, *pitch*, volume e timbre e, portanto, são utilizados pelos softwares de AMT no processo (BELLO; MONTI; SANDLER, 2000).

Na transcrição de áudios monofônicos, somente uma nota pode ser tocada por vez, tendo em vista que podem haver múltiplos instrumentos na faixa, contanto que os sons não se sobreponham. Trabalhar apenas com sinais monofônicos limita o número de aplicações das ferramentas, no entanto, aumentam as chances de ser obtido um retorno confiável e rápido (BELLO; MONTI; SANDLER, 2000).

Ao contrário da transcrição de sinais monofônicos, transcrever áudios polifônicos exige detectar e identificar múltiplos *pitches* concorrentes que podem se sobrepor. Esta etapa permanece sendo o principal desafio a ser resolvido nesta área. Detecção de *onsets* e *offsets*, reconhecimento de instrumentos, extração de ritmo e tempo também são problemas a serem solucionados na área (BENETOS et al, 2013).

Sistemas de AMT, a fim de detectar o *pitch* nos sinais de áudio, precisam calcular a frequência fundamental (menor frequência de um som) das notas no sinal a fim de estimar

suas alturas. Métodos como a autocorrelação (TREVILATTO JUNIOR; BARBEDO; LOPES, 2005) ou o método de Klapuri podem ser utilizados para identificar as frequências fundamentais (DINIZ, 2009).

Diniz (2009) afirma que o modelo de execução de uma nota, normalmente, divide-se em quatro etapas: ataque (intervalo de tempo no qual a amplitude da onda aumenta), transitório (período em que o sinal se comporta de forma imprevisível), sustentação (trecho no qual a amplitude da nota se mantém próxima de constante) e relaxamento (momento em que a amplitude da nota diminui de forma acentuada). Tendo definido *onset* como sendo o início da execução de determinada nota, este fica menos abstrato se considerado um cenário onde a nota é precedida por silêncio e não por outra nota.

Três etapas geralmente constituem a detecção de *onsets*, sendo elas: pré-processamento, redução e detecção de picos. A primeira etapa é opcional no processo, pois somente é responsável por enaltecer e atenuar aspectos relevantes e irrelevantes do sinal. Sendo etapa chave de uma grande variedade de métodos de detecção de *onsets*, a etapa de redução, também chamada de função de detecção, é responsável por transformar o sinal em uma amostra menor, porém com maior expressividade (BELLO et al., 2005). Por fim, a última etapa é responsável por detectar os picos da função de detecção, a fim de filtrar somente *onsets* reais, ignorando falsos picos gerados a partir de fenômenos como reverberação e eco (DINIZ, 2009).

Além de detectar os *onsets*, os sistemas de AMT devem ser capazes de identificar as pausas na música. A pausa é caracterizada por um decrescimento exponencial do sinal, ocorrido na etapa de dissipação do sinal, em que o mesmo tende a se tornar desprezível. Este fenômeno recebe o nome de *offset* (SIMOES, 2006).

A partir da extração das principais informações dos sinais torna-se possível efetivamente convertê-las em alguma notação conhecida, neste momento pode ser considerado o MIDI. O protocolo MIDI (*Musical Instrument Digital Interface*) permite a transmissão de mensagens contendo informações como o *pitch*, intensidade das notas e sinais de controle. O sinal do áudio não é transmitido pelas mensagens, ficando a cargo do receptor converter as informações recebidas a partir de seus próprios recursos. Assim, é possível interpretar este protocolo como uma espécie de partitura (representação mundialmente conhecida para músicas) virtual (DINIZ, 2009).

Apesar de demonstrar um avanço significativo nos últimos anos, as áreas da transcrição de sinais monofônicos e, sobretudo, de sinais polifônicos, ainda dependem de resultados mais significativos. Assim, este estudo terá como foco de pesquisa a transcrição de sinais monofônicos, abordando todas as etapas do processo. Serão abordados e discutidos os principais entraves para transcrição de áudios polifônicos, trazendo à tona seu estado da arte.

Portanto, o objetivo deste trabalho é desenvolver um protótipo de software capaz de interpretar arquivos de áudio e convertê-los em arquivos MIDI. Os resultados desse protótipo serão avaliados a partir da análise da taxa de acertos e percentual de falsos positivos, com base em áudios sintetizados e de instrumentos reais. Por fim, o desempenho será comparado com o obtido por trabalhos relacionados, a fim de validar a abordagem utilizada.

OBJETIVOS

Objetivo geral

Desenvolver um protótipo de software capaz de transcrever para uma notação estabelecida as notas e suas durações a partir de sinais de áudios monofônicos.

Objetivos específicos

- Analisar métodos de detecção de *pitch*, *onsets* e *offsets* para sinais monofônicos;
- Desenvolver um protótipo de software capaz de converter sinais de áudios para arquivos MIDI;
- Comparar resultados gerados pelo software com resultados de trabalhos relacionados;
- Avaliar estado da arte da transcrição de áudios polifônicos.

METODOLOGIA

O presente trabalho, em relação a sua natureza, caracteriza-se como pesquisa aplicada, pois tem como objetivo desenvolver um protótipo a fim de solucionar um determinado problema. A abordagem utilizada será quantitativa, à medida que serão analisados os resultados e quantificados os percentuais de acertos e erros na detecção, interpretação e conversão das notas. Posteriormente, serão comparados os resultados do software desenvolvido com os de outros trabalhos, de tal forma que seja possível detectar a validade da abordagem adotada nesta pesquisa.

Por tratar do estudo de métodos a fim de melhorar o processo de transcrição automática de áudios monofônicos, reduzindo a taxa de erros nas etapas de detecção de notas e duração destas, esta pesquisa possui objetivo exploratório. Sendo material base para este estudo, em sua maioria, artigos científicos, teses e livros, ou seja, constituído a partir de material previamente publicado (PRODANOV; FREITAS, 2013), este trabalho irá utilizar como procedimento técnico a pesquisa bibliográfica.

A primeira etapa deste trabalho será responsável por analisar o cenário atual da transcrição de áudios monofônicos. A partir deste estudo, será iniciado o desenvolvimento de um protótipo. A transcrição de sinais de áudios polifônicos também será abordada neste trabalho, com o levantamento do estado da arte e próximos passos de pesquisa no tema.

A interpretação de áudios digitais e detecção de *onsets* e *offsets* das notas será a etapa inicial do desenvolvimento do protótipo. Em um segundo momento, serão implementados os algoritmos responsáveis por detectar o *pitch* das notas extraídas. A etapa final será responsável por converter os dados para o protocolo MIDI, permitindo integrar as saídas do sistema com softwares que implementem este protocolo.

Durante todo o processo de desenvolvimento serão medidas as taxas de acerto apresentadas pelo sistema. No entanto, somente após concluídas todas as etapas de desenvolvimento, comparativos serão feitos entre os resultados do protótipo e de outros trabalhos. Desta forma, será possível concluir os pontos positivos e negativos em relação à abordagem utilizada.

CRONOGRAMA

Trabalho de Conclusão I

Etapa	Meses			
	Mar	Abr	Mai	Jun
Redação do anteprojeto				
Revisão do anteprojeto				
Entrega do anteprojeto				
Estudo de áudio digital				
Estudo da detecção de <i>pitch</i> , <i>onsets</i> e <i>offsets</i>				
Estudo do protocolo MIDI				
Desenvolvimento do protótipo				
Redação do TCC I				
Revisão do TCC I				
Entrega do TCC I				

Trabalho de Conclusão II

Etapa	Meses			
	Ago	Set	Out	Nov
Desenvolvimento do protótipo				
Estudo de técnicas avançadas				
Testes do protótipo				
Análise dos resultados				
Redação do TCC II				
Revisão do TCC II				
Entrega do TCC II				
Apresentação dos resultados à banca avaliadora				

BIBLIOGRAFIA

BELLO, Juan Pablo; DAUDET, Laurent; ABDALLAH, Samer; DUXBURY, Chris; DAVIES, Mike; SANDLER, Mark B. A Tutorial on Onset Detection in Music Signals. **IEEE Transactions on speech and audio processing**, v. 14, n. 5, p. 1035-1047, set. 2005.

BELLO, Juan Pablo; MONTI, Giuliano; SANDLER, Mark. Techniques for Automatic Music Transcription. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MUSIC INFORMATION RETRIEVAL (ISMIR), 1., 2000, Plymouth. **Anais...** Plymouth: UMass, 2000.

BENETOS, Emmanouil; DIXON, Simon; GIANNOULIS, Dimitrios; KIRCHHOFF, Holger; KLAPURI, Anssi. Automatic music transcription: breaking the glass ceiling. In: PROCEEDINGS OF THE INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL CONFERENCE (ISMIR), 13., 2012, Porto. **Anais...** Porto: INESC TEC, 2012. p. 379-384.

_____. Automatic music transcription: challenges and future directions. **Journal of Intelligent Information Systems**, v. 41, n. 3, p. 407-434, dez. 2013.

DINIZ, Felipe Castello da Costa Betrão. **Transcrição musical automática usando representação frequencial eficiente por banco de filtros de alta**. 2009. 170 f. Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-graduação em Engenharia Elétrica, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 2009.

ELLIS, Daniel P. W. Extracting information from music audio. **Communications of the ACM**, v. 49, n. 8, ago. 2006.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico**. 2. ed. Novo Hamburgo: Universidade Feevale, 2013. 276 p.

SIMÕES, Gabriel; FREITAS, Allan; SOUZA, Hercules. Desenvolvimento de um sistema computacional de transcrição de melodias monofônicas para partitura. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 26., 2006, Campo Grande. **Anais...** Campo Grande: UCDB, 2006. p. 22-27.

SANTINI, Rose Marie; SOUZA, Rosali Fernandez de. Recuperação da informação de música e a ciência da informação: tendências e desafios de pesquisa. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 8., 2007, Florianópolis. **Anais...** Salvador: UFBA, 2007.

TREVILATTO JUNIOR, Narciso; BARBEDO, Jayme Garcia Arnal; LOPES, A. Transcrição Automática de Sinais de Áudio Monofônicos. In: SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO MUSICAL, 10., 2005, Belo Horizonte. **Anais...** Belo Horizonte: PUC Minas, 2005. p. 291-294.