

UNIVERSIDADE FEEVALE

GREGÓRIO CHALINSKI KUSOWSKI

REPRESENTAÇÃO VISUAL DE SEGMENTOS DE ÁUDIO

(Título Provisório)

Anteprojeto de Trabalho de Conclusão

Novo Hamburgo
2016

GREGÓRIO CHALINSKI KUSOWSKI

REPRESENTAÇÃO VISUAL DE SEGMENTOS DE ÁUDIO

(Título Provisório)

Anteprojeto de Trabalho de Conclusão de
Curso, apresentado como requisito parcial
à obtenção do grau de Bacharel em
Ciência da Computação pela
Universidade Feevale

Orientador: João Batista Mossmann

Novo Hamburgo
2016

RESUMO

Com o aumento constante no montante de informações geradas diariamente nas mais diversas áreas do conhecimento humano, juntamente com a ineficiência do ser humano na interpretação de grandes volumes de dados, verifica-se uma dificuldade para análise e gestão do conteúdo capturado de forma rápida. Neste sentido, a utilização de gráficos para exibição das informações favorece o reconhecimento e a percepção de características. No contexto de coleções musicais, algumas ferramentas exibem representações de ondas do conteúdo de músicas em listas, possibilitando ao usuário uma prévia visual dos conteúdos de áudio. Entretanto, as formas de ondas de um segmento de som não são capazes de representar muita informação sobre o mesmo. Neste contexto, o trabalho aborda o estudo da interpretação de segmentos digitais de áudio e técnicas de visualização de dados. Sendo assim, pretende-se investigar um modelo de representação visual de áudio a partir da interpretação de sinais e técnicas de visualização.

Palavras-chave: Visualização Computacional. Recuperação de Informação de Música.

SUMÁRIO

MOTIVAÇÃO	5
OBJETIVOS	9
METODOLOGIA	10
CRONOGRAMA	11
BIBLIOGRAFIA	12

MOTIVAÇÃO

Com o aumento constante da automação dos processos nas mais diversas áreas do conhecimento humano, verifica-se a geração de um volume muito grande de informação. Sabe-se da ineficiência do ser humano na interpretação de grandes volumes de dados, sejam na forma numérica ou textual. Nesse sentido, a exibição de dados em forma gráfica facilita o reconhecimento e a percepção de um conjunto grande de características, por estarem apresentados de maneira mais organizada (BOTELHO, 2002).

Então, a **visualização de informação** pode estimular naturalmente a percepção e inteligência humana. Ela permite que os seres humanos, ao interpretarem um grande volume de dados, apliquem uma combinação de flexibilidade, criatividade e conhecimento geral para apoio em tomadas de decisão. Logo, o resultado a partir de grandes bases de dados, quando exibido em forma gráfica, faz com que a capacidade de análise em gráficos do ser humano seja aplicada na descoberta de informações (BOTELHO, 2002).

Dentre as técnicas utilizadas para a exibição de dados, uma das mais aplicadas é a projeção geométrica, que consiste na plotagem de atributos e valores em um plano bidimensional. Para atuar com vários atributos, também existe a técnica de apresentação icônica, que visa mapear cada item em um ícone que será exibido em conjunto com as outras entidades, porém este formato possui uma limitação relacionada à quantidade de dados (BOTELHO, 2002). Semelhante à forma icônica, existe a exibição orientada a pixel, que se baseia na separação de uma tela em janelas para cada dimensão/atributo, onde cada registro é representado como um pixel em cada janela, exibindo seu atributo correspondente (BRANCO, 2003).

Conforme Freitas et al. (2001), a consideração inicial para a escolha de uma técnica de visualização é a caracterização dos dados. Essa etapa é composta por três critérios, o primeiro trata sobre a classe de informação a qual um atributo pertence, de forma a possibilitar o enquadramento da entidade em uma categoria ou grupo. O segundo, se refere ao tipo de dado que um atributo possui, seja numérico, alfanumérico ou simbólico. O terceiro critério diz respeito à dimensão e natureza de onde os atributos são definidos, seja pela dimensionalidade, no caso de domínio ou pela classificação de dados quantitativos em discretos ou contínuos, no caso de natureza.

Além da importância da escolha da técnica de visualização, baseada na classificação ou categorização do conjunto de dados, a escolha das cores utilizadas nas representações de dados é importante, conforme destaca Branco (2003). Tutida (1998 apud BRANCO, 2003) afirma que a seleção de cores adequadas é um aspecto crítico em visualização. Branco (2003) ainda cita um exemplo no qual as cores que compõem a representação visual são capazes de determinar o sucesso dos resultados de uma técnica.

Uma forma muito comum de visualizar dados é através de gráficos baseados em tempo. Alencar (2007) define **Séries Temporais** como um conjunto de observações sequenciadas no tempo. Ehlers (2009) ainda ressalta que uma característica relevante nesse tipo de representação é a dependência entre os dados que estão adjacentes, uma vez que a ordem dos mesmos é crucial para uma série temporal.

Neste contexto, um exemplo de Séries Temporais com um grande volume de dados, são os segmentos de mídia, como músicas. Ono (2015) afirma que a representação visual de músicas visa facilitar a interpretação, exibindo tons, notas, acordes, harmonia e outras propriedades. Entendendo que a extração de características é a base para qualquer tipo de análise de áudio, Tzanetakis et al. (2002 apud ONO, 2015) explicam que o sinal sonoro pode ser expresso através de um cálculo de representação numérica. Percebe-se, então, que com a representação numérica de um conjunto temporal de sinais, é possível gerar uma visualização do resultado.

O sinal de áudio, de forma analógica pode ser definido como uma função real sobre uma variável que representa o tempo, porém em forma digital ele é definido como uma função que mapeia apenas um intervalo de tempo relacionado à amostragem utilizada. O processamento de um conjunto de sinais é feito sequencialmente, e cada valor dessa sequência recebe o nome de amostra. Para tal, existe o ciclo de processamento do sinal digital, ciclo DSP do inglês *Digital Signal Processing*, que consiste na obtenção e processamento de amostras com o intuito de que para cada valor de entrada, exista um de saída (BIANCHI, 2014).

A representação visual de um segmento de áudio pode ser realizada em forma de gráfico, tendo tempo e frequência especificados em seus eixos, e utilizando cores para a magnitude das frequências. Essa representação é conhecida como espectrograma (ONO, 2015). Chen et al. (2010) ainda explicam que a simples exibição de onda no domínio do tempo não demonstra muito sobre o seu conteúdo, em vista que a distribuição de frequência é o principal portador de informação dos sinais para análise musical.

Conforme Paulus , Müller e Klapuri (2010), a diferença entre sons aleatórios e música não é clara, pois o que pode ser uma sequência de ruídos para alguns, pode ser uma engenhosa composição musical para outros. Por outro lado, os autores entendem como um comum acordo que eventos sonoros isolados não são capazes de representar um significado musical. Para isso, é necessária uma estrutura, ou então um relacionamento entre vários eventos de som.

Smith e Williams (1997) reconhecem que a forma mais popular de visualizar música, é através da notação musical. Através dela, músicos experientes são capazes de conhecer o resultado sonoro sem de fato ouvir uma melodia. Um formato digital que se aproxima dessa notação, no sentido de conter cada símbolo e seu respectivo tempo de ocorrência, é o formato **MIDI**. Embora seja um dos formatos mais populares, o MIDI (*Musical Instrument Digital Interface*) não é utilizado para a representação de performances, tal como uma gravação sonora, em específico, mas sim para o mapeamento da informação existente naquele segmento definido.

Para a representação visual de arquivos MIDI, *softwares* como o *MIDITrail* e o *Music Annotation Machine* podem ser utilizados (ONO, 2015). Os símbolos, ou notas são exibidos em uma plotagem baseada em tempo. Dado que o formato MIDI também possibilita a separação de instrumentos por canais, as ferramentas de visualização podem, então, representar visualmente informações sobre as notas contidas no arquivo em uma disposição semelhante a uma partitura, dado que utiliza um domínio discreto com informação booleana.

Em outro sentido, aplicativos como o *Sonic Visualiser* (CANNAM; LANDONE; SANDLER, 2010) buscam exibir as informações de sinais de áudio. Nele, formatos digitais, geralmente formados pela digitalização de conteúdo analógico, são exibidos através de técnicas de visualização em representações de espectrograma, formas de ondas e picos de frequências, e a possibilidade de utilização de *plugins* para acrescentar mais métodos de análise aos já existentes.

Com o objetivo de fazer uma exibição durante a execução de um segmento de áudio, o **MAM**, *The Music Animation Machine* (MALINOWSKI; TURETSKY, 2016), traz uma evolução histórica nas abordagens deste assunto. Iniciando em 1974, quando é relatada uma alucinação em que notas dançavam em uma página enquanto ouvia-se música, passando pela disponibilização de vídeos de músicas com suas respectivas animações gráficas no *YouTube* e um aplicativo para *iPad*. Além deste, algumas ferramentas, inclusive *players* comerciais possuem opções para visualização momentânea do que está sendo reproduzido sonoramente.

Atuando na representação de áudio em forma de objetos reais, existem projetos como os do "*Sculpteur de Voix*" (escultor de voz), Gilles Azzaro (AZZARO, 2016), que reproduzem discursos através de impressões em três dimensões, e o projeto *Frozen Waves* (VAN MELLE; MARKS, 2014), que faz uma ligação artística entre som e objeto.

Com base neste contexto, o trabalho propõe investigar um modelo para a representação visual de áudio baseada na interpretação de sinais e técnicas de visualização, possibilitando assim, a apresentação dos dados de áudio de forma gráfica mais agradável e rica em informação ao usuário final.

OBJETIVOS

Objetivo geral

Investigar um modelo de representação visual de áudio a partir da interpretação de sinais e técnicas de visualização.

Objetivos específicos

- Estudar os principais formatos digitais de áudio, considerando parâmetros que possam ser diferenciais na representação visual de sua análise.
- Apresentar uma abordagem de representação visual que utilize os conceitos estudados
- Desenvolver uma representação visual através da análise de áudio.
- Validar os resultados obtidos durante os experimentos

METODOLOGIA

Baseado nos conceitos metodológicos apresentados por Prodanov e Freitas (2013), o trabalho caracteriza-se, quanto a sua natureza, como **pesquisa aplicada**, visto que seu produto visa solucionar um problema específico.

Devido ao método avaliativo que será utilizado sobre a conformidade da solução com o problema proposto, baseado em outros estudos relacionados, este estudo é qualificado como **pesquisa qualitativa**.

Com base na investigação, levantamento bibliográfico e análise de exemplos e experiências práticas para a formulação da solução, quanto aos seus objetivos pode-se caracterizar o trabalho como **pesquisa exploratória**.

O trabalho também caracteriza-se como **pesquisa bibliográfica** e **experimental** no que se refere a procedimentos técnicos, pois aborda o estudo das técnicas e conceitos que serão utilizados como referencial teórico e utiliza de experimentação para desenvolvimento e observação dos resultados.

CRONOGRAMA

Trabalho de Conclusão I

Etapa	Meses				
	Mar	Abr	Mai	Jun	Jul
Entrega do Aceite de Orientação	■				
Elaboração do Anteprojeto	■	■			
Estudo sobre Visualização de Dados		■	■		
Estudo sobre Formatos de Áudio Digital		■	■		
Estudo sobre Interpretação e Decodificação de Áudio Digital		■	■		
Estudo de projetos relacionados com visualização de áudio			■	■	
Estudo e definição de um modelo			■	■	
Desenvolvimento e entrega do Trabalho de Conclusão I		■	■	■	■

Trabalho de Conclusão II

Etapa	Meses			
	Ago	Set	Out	Nov
Desenvolvimento do modelo	■	■		
Aplicação do modelo		■	■	
Realizar um experimento de validação			■	
Desenvolvimento e entrega do Trabalho de Conclusão II	■	■	■	■
Apresentação do trabalho na Banca				■

BIBLIOGRAFIA

ALENCAR, Aretha Barbosa. **Mineração e visualização de coleções de séries temporais**. 2007. Tese de Doutorado. Instituto de Ciências Matemáticas e de Computação.

AZZARO, Gilles. **Gilles AZZARO - Sculpteur de Voix**. Internet: <http://www.gillesazzaro.com/>, [31 de Abril de 2016], 2016.

BIANCHI, André Jucovsky. **Processamento de áudio em tempo real em dispositivos computacionais de alta disponibilidade e baixo custo**. 2014. Tese de Doutorado. Universidade de São Paulo.

BOTELHO, Elisângela. **Visualização de dados como ferramenta de Classificação em sistemas de Bases de Dados para Data Mining**. 2002. Tese de Doutorado. Instituto de Ciências Matemáticas e de Computação.

BRANCO, Vinícius Marques Alves. **Visualização como suporte à exploração de uma base de dados pluviométricos**. 2003. Tese de Doutorado. Instituto de Ciências Matemáticas e de Computação.

CANNAM, Chris; LANDONE, Christian; SANDLER, Mark. **Sonic visualiser**: An open source application for viewing, analysing, and annotating music audio files. In: Proceedings of the 18th ACM international conference on Multimedia. ACM, 2010. p. 1467-1468.

CHEN, Qing et al. **Analysis of Music Representations of Vocal Performance Based on Spectrogram**. In: Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on. IEEE, 2010. p. 1-4.

EHLERS, Ricardo S. **Análise de séries temporais**. Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná, 2007.

FREITAS, Carla Maria Dal Sasso et al. **Introdução à visualização de informações**. Revista de informática teórica e aplicada. Porto Alegre. Vol. 8, n. 2 (out. 2001), p. 143-158, 2001.

MALINOWSKI, Stephen; TURETSKY, L. **Music animation machine**. Internet: <http://www.musanim.com/>, [31 de Abril de 2016], 2016.

ONO, Jorge Henrique Piazzentin. **Visualização de similaridades em bases de dados de música**. Tese de Doutorado. Universidade de São Paulo.

PAULUS, Jouni; MÜLLER, Meinard; KLAPURI, Anssi. **State of the Art Report: Audio-Based Music Structure Analysis**. In: ISMIR. 2010. p. 625-636.

PRODANOV, Cleber Cristiano; DE FREITAS, Ernani Cesar. **Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico-2ª Edição**. Editora Feevale, 2013.

SMITH, Stephen M.; WILLIAMS, Glen N. **A visualization of music**. In: Visualization'97., Proceedings. IEEE, 1997. p. 499-503.

VAN MELLE, Gerbrand; MARKS, Stefan. **Frozen waves: exploring the transformation between sound and object**. 2014.