UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

DÉBORA CRISTINA CORRÊA

Inteligência artificial aplicada à análise de gêneros
musicais

São Carlos
2012

DÉBORA CRISTINA CORRÊA

Inteligência artificial aplicada à análise de gêneros
musicais

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.

Área de concentração: Física Computacional

Orientador: Prof. Dr. Luciano da Fontoura Costa

Versão corrigida
Versão original disponível na Unidade que aloja o Programa
São Carlos
2012

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pelo Serviço de Biblioteca e Informação do IFSC,
com os dados fornecidos pelo(a) autor(a)

Corrêa, Débora Cristina
Inteligência artificial aplicada à análise de
gêneros musicais / Débora Cristina Corrêa;
orientador Luciano da Fontoura Costa - versão
corrigida -- São Carlos, 2012.
175 p.

Tese (Doutorado - Programa de Pós-Graduação em
Física Aplicada Computacional) -- Instituto de Física
de São Carlos, Universidade de São Paulo, 2012.

1. Gêneros musicais. 2. Ritmo. 3. Dígrafos. 4.
Detecção de comunidades. I. da Fontoura Costa,
Luciano , orient. II. Título.

Aos meus queridos pais.

Agradecimentos

A toda minha família, em especial aos meus pais, pelo amor, incentivo e apoio incondicional para que este objetivo fosse alcançado.

Ao Prof. Dr. Luciano da Fontoura Costa, pelos ensinamentos, opiniões, paciência, correções, pelos conhecimentos compartilhados, pela dedicação, e por tornar possível a concretização desta etapa importante.

A todo pessoal do grupo de pesquisa, pela amizade, consideração, discussões sobre os mais variados temas e assuntos, pelas risadas e convivência ao longo destes anos. Em especial, aos meus queridos amigos Matheus Viana, Mauro Miazaki, Renato Pimentel, Lucas Antiqueira, Krissia Zawadzki, Renato Fabbri, Vilson Vieira, André Barbieri e Carlos Doro Neto. E em especial ao meu amigo Prof. André Angelis pelas correções durante a etapa de qualificação. E também em especial a minha amiga Monica Campitelli, pois além destas considerações, realizou com carinho e dedicação a revisão deste documento.

Ao meu amigo e companheiro Moussa Mansour pelo apoio, incentivo, paciência, compreensão e companherismo, principalmente na reta final do doutorado.

A todos meus amigos, que direta ou indiretamente fizeram parte desta etapa de quatro anos. Muitos eu já conhecia, muitos conheci neste período. Todos sabem a forma especial que os admiro e o tamanho do meu carinho. Sei que ao nomeá-los corro um risco grande de omitir alguém, mas ainda sim quero deixar um agradecimento especial aos meus queridos: Gisele Philippsen, Daniele Jacinto, Alexandre Levada, Michele Horta, Tatiana Bonifácio, Juliana Kurokawa, Lilian Gambi, Giovana Tangerino, Marcelo Suetake, Marcel Cezare, Amilcar Querubini, Jéssica Santos, Renata Batista, Carol Silvério, Fernanda Silva, Ziza Gaspar, Aline Gaspar, Beatriz Marcolino, Silvia Figueiredo, Talita Perciano, Elaine Ap. de Oliveira, Raphael Togniollo, Cris Ramos e pessoal da escola Ritmo.

A todos os funcionários do Instituto de Física de São Carlos, em especial aos professores pelos ensinamentos durante as disciplinas, aos funcionários que contribuem para a limpeza do ambiente de trabalho, aos funcionários da secretaria e serviços de pós-graduação, e as funcionárias Maria Cristina Dziabas e Maria Neusa de Aguiar Azevedo, pela correção cuidadosa da formatação e bibliografia deste documento.

Ao CNPq pelo apoio financeiro nos meses iniciais da realização do projeto.

A Fapesp pelo apoio financeiro durante a realização deste projeto (processo: 2009/50142-0).

Music is a moral law.
It gives soul to the universe,
wings to the mind,
flight to the imagination,
and charm and gaiety to life,
and to everything.

Plato

Resumo

CORREA, D. C. **Inteligência artificial aplicada à análise de gêneros musicais**. 2012. 175p. Tese (Doutorado em Física Aplicada) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.

O crescimento constante dos dados musicais na Internet tem encorajado diversos pesquisadores a desenvolver ferramentas adequadas para a análise e a classificação destes dados. O objetivo principal de tais ferramentas é extrair a informação de forma compacta e representativa ao conteúdo dos bancos de dados. Dentro deste contexto, os gêneros musicais apresentam descrições importantes para o desenvolvimento destas ferramentas. Além dos mesmos serem usados frequentemente para organizar coleções musicais e refletirem a interação entre culturas, resumem características (padrões) comuns entre as peças musicais. Em face ao exposto, a principal motivação deste projeto de pesquisa é propor uma maneira original, e de baixo esforço computacional, para representar os gêneros musicais e investigar a contribuição desta representação em aplicações e estudos que estão inseridos no contexto de pesquisas que envolvem a recuperação da informação musical. A representação proposta refere-se aos padrões rítmicos das músicas, uma vez que o ritmo configura um aspecto musical signifiicante na discriminação dos gêneros. Os padrões rítmicos são estabelecidos pela dependência temporal das notas musicais presentes na percussão, de forma que cada música é representada por um vetor de probabilidades condicionais entre pares e trios de notas computadas pelo uso de cadeias de Markov de primeira e segunda ordem. Os padrões rítmicos de diversos gêneros são explorados em aplicações como: classificação, síntese musical, recomendação musical, humor/emoção em música, e análise de aspectos evolutivos. Constatou-se que estes, como estabelecidos neste estudo, são sensíveis à discriminação dos gêneros, evidenciando sequências de notas que são comuns aos mesmos, e sequências que são distintas e características de cada um. Uma segunda motivação deste projeto é o uso de medidas topológicas de redes e dígrafos de músicas para a análise dos dados. Comunidades obtidas nestas redes proporcionaram a definição de uma abordagem não supervisionada, a qual apresentou taxas de desempenho superiores ao agrupamento hierárquico. A determinação das características rítmicas de cada música motivou o desenvolvimento de estratégias para a composição automática e para a geração de listas de reprodução, assim como para a averiguação da relação destes padrões com aspectos emotivos. Por fim, uma análise estatística da evolução do ritmo de diferentes gêneros é desempenhada, na qual verificou-se a presença de mecanismos de inovação e recuperação. Estes mecanismos parecem ser consequência da competição entre fatores que favorecem a inovação de material musical, e fatores que a previnem, como, por exemplo, a obediência às regras de composição que mantém as características fundamentais de cada gênero.

Palavras-chave: Gêneros musicais. Ritmo. Dígrafos. Detecção de comunidades.

Abstract

CORREA, D C. **Artificial intelligence applied to musical genres analysis**. 2012. 175p. Tese (Doutorado em Física Aplicada) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.

Musical databases have increased in number and size continuously, paving the way to large amounts of online music data, including discographies, biographies and lyrics. The constant growth of data on the Internet has attracted musical research for developing tools to analyze and classify music data. The main objective of such tools is to extract reliable information to adequately represent and compact music content in databases. In this context, musical genres are particularly interesting descriptors, since they have been used for years to organize music collections, reflect interaction between cultures and summarize common features (or patterns) between musical pieces. The main motivation of this study is to propose a original and low cost framework to represent musical genres, as well as investigate the contribution of this representation in applications and studies that are placed in the context of music information retrieval researches. The representation of music content is referred to the rhythmic patterns, since rhythm configures a significant aspect in the discrimination of musical genres. The rhythmic patterns are determined by the temporal dependency of the musical notes present in the percussion, so that each song is represented by a vector of conditional probabilities between pairs and triples of notes, computed by the use of first and second order Markov chains. The rhythm patterns from distinct genres are investigated in applications such as: classification, music synthesis, music recommendation, mood/emotion in music, and analysis of evolutionary aspects. The main finding is that the rhythmic patterns as established in this study are sensitive to the genre discrimination, suggesting that there are sequences of notes common to all genres, and sequences that are distinct and characteristics of each one. A second motivation for this study is the use of topological measures of music networks and music digraphs for the data analysis. Communities obtained from these networks contributed to the definition of an unsupervised approach that provided performance rates superior to the hierarchical clustering. The rhythmic patterns also motivated the development of strategies for automatic composition, for the generation of playlists, and the analysis of the relationship between these patterns and emotional aspects. Finally, a statistical analysis of the rhythm evolution is performed, in which the principal finding is the presence of innovation and retrieval mechanisms for all genres. These mechanisms seem to be the result of the competition between factors that promote the innovation, and factors that prevent it, as, for example, the obedience to composition rules that retains the fundamental characteristics of each genre.

Keywords: Musical genres. Rhythm. Digraphs. Community detection.

Lista de Tabelas

Tabela 2.1 -	Associação das notas musicais com as letras do alfabeto	42
Tabela 2.2 -	A relação entre as notas de valores	42
Tabela 3.1 -	Desempenho da classificação de acordo com o coeficiente Kappa .	58
Tabela 3.2 -	Métricas de distância e suas respectivas definições	60
Tabela 3.3 -	Critérios de <i>linkage</i> e suas respectivas definições	61
Tabela 4.1 -	Disposição das notas de valores e batidas correspondentes ao trecho em destaque na faixa de percussão da Figura 4.2	73
Tabela 5.1 -	Taxonomia de gêneros para o sistema de composição de sequências rítmicas	79
Tabela 6.1 -	Taxonomia de gêneros para o sistema de classificação	91
Tabela 6.2 -	Resultados de classificação na matriz de dados $\mathbf{P}^p$	97
Tabela 6.3 -	Resultados de classificação na matriz de dados $\mathbf{P}^p$	98
Tabela 6.4 -	Matrizes de confusão dos resultados em destaque nas Tabelas 6.2 e 6.3	98
Tabela 6.5 -	Resultados de classificação na matriz de dados $\mathbf{P}^s$	99
Tabela 6.6 -	Matriz de Confusão para a classificação com kappa = 0.82	101
Tabela 6.7 -	Matriz de confusão para o agrupamento hierárquico aglomerativo	101
Tabela 6.8 -	Resultados da classificação multi-categorias para as músicas de blues e mpb	107
Tabela 6.9 -	Resultados da classificação multi-categorias para as músicas de reggae e rock	108
Tabela 6.10 -	Taxonomia de estilos para o sistema de classificação	109
Tabela 6.11 -	PCA - Resultados de classificação na matriz de dados $\mathbf{P}^p$ na taxonomia de 10 estilos.	111
Tabela 6.12 -	LDA - Resultados de classificação na matriz de dados $\mathbf{P}^p$ na taxonomia de 10 estilos.	111
Tabela 6.13 -	Matriz de confusão para o agrupamento hierárquico aglomerativo dos 10 estilos	112
Tabela 7.1 -	Taxonomia de gêneros para detecção de comunidades em redes de música.	116
Tabela 7.2 -	Grupos de acordo com a distância do cosseno.	118
Tabela 7.3 -	Grupos de acordo com o algoritmo por Girvan e Newman.	119
Tabela 7.4 -	Grupos de acordo com a distância Euclidiana.	119
Tabela 7.5 -	Grupos de acordo com o particionamento espectral.	121
Tabela 8.1 -	Taxonomia de gêneros para o sistema de recomendação musical .	125

Tabela 9.1 -	Expressões de emoção/humor de acordo com diferentes fatores musicais: ritmo, melodia e forma musical. Adaptado de (106)	. . . 135
Tabela 9.2 -	Taxonomia de gêneros para a análise de motivos musicais	 137
Tabela 9.3 -	Ilustração da detecção de motivos melódicos.	 138
Tabela 9.4 -	Ilustração da detecção de motivos rítmicos.	 139
Tabela 10.1 -	Taxonomia adotada na investigação da evolução dos gêneros	. . . 143

Lista de Figuras

Figura 1.1 -	Esquema resumido dos fatores principais que influenciam a evolução dos gêneros musicais.	37
Figura 1.2 -	Visão geral das direções exploradas neste trabalho.	39
Figura 2.1 -	As sete notas musicais	41
Figura 2.2 -	Representação das notas musicais em duas claves	43
Figura 2.3 -	Ilustração de (b) motivo tonal e (c) rítmico das primeiras quatro notas da 5a. Sinfonia de Beethoven em (a)	44
Figura 2.4 -	As notas musicais, suas frequências e respectivos números MIDI .	46
Figura 2.5 -	Representação em pentagrama de um arquivo MIDI	47
Figura 2.6 -	Ilustração de um pentagrama de percussão	47
Figura 3.1 -	Ilustração das técnicas PCA e LDA.	53
Figura 3.2 -	Dendrograma para uma situação com 8 amostras, adaptado de (7)	58
Figura 3.3 -	Um exemplo de grafo conectado em (a), e uma árvore geradora mínima em (b), adaptado de (61)	64
Figura 3.4 -	Exemplo de (a) rede randômica (b) rede de pequeno mundo (c) rede livre de escala (59)	65
Figura 3.5 -	Ilustração do agrupamento de vértices para detecção de comunidades.	66
Figura 3.6 -	Ilustração do conceito de arestas com centralidade de intermediação	67
Figura 4.1 -	Visão geral dos passos principais para o desenvolvimento do trabalho	72
Figura 4.2 -	Parte de uma faixa de percussão	73
Figura 4.3 -	Exemplos de matrizes de transição de primeira e segunda ordem.	74
Figura 4.4 -	Dígrafos de primeira ordem de quatro músicas.	77
Figura 4.5 -	Dígrafos de segunda ordem de quatro músicas.	78
Figura 5.1 -	Visão geral das etapas realizadas para a síntese de sequências rítmicas.	80
Figura 5.2 -	Síntese de música: reggae	82
Figura 5.3 -	Síntese de música: blues	83
Figura 5.4 -	Dígrafos de gêneros para o processo de síntese.	85
Figura 5.5 -	Dígrafos de gêneros resultantes da síntese.	86
Figura 5.6 -	Sequências rítmicas geradas pela síntese.	87
Figura 6.1 -	Diagrama dos passos principais para classificação automática de gêneros musicais	90
Figura 6.2 -	Diagrama de Venn para o banco de dados multi-categorias	93
Figura 6.3 -	Espaço de atributos PCA	94
Figura 6.4 -	Espaço de atributos LDA	95

Figura 6.5 -	As dez sequências de pares de notas mais importantes de acordo com os primeiros autovetores do PCA e LDA.	96
Figura 6.6 -	Valores do kappa de acordo com o número de componentes utilizadas na classificação.	100
Figura 6.7 -	Dendrograma da formação dos quatro grupos (coloridos em azul - reggae, verde - MPB, roxo- rock e vermelho - blues)	102
Figura 6.8 -	Dendrograma detalhado do primeiro grupo, em azul.	103
Figura 6.9 -	Dendrograma detalhado do segundo grupo, em verde.	103
Figura 6.10 -	Dendrograma detalhado do terceiro grupo, em roxo.	104
Figura 6.11 -	Dendrograma detalhado do quarto grupo, em vermelho.	104
Figura 6.12 -	Linhas de contorno das densidades condicionais e o espaço de atributos.	106
Figura 6.13 -	Espaço de atributos (a) PCA e (b) LDA	110
Figura 7.1 -	Diagrama dos passos principais para a detecção de comunidades em redes de música	115
Figura 7.2 -	Comunidades da rede 10-regular com a distância do cosseno . . .	118
Figura 7.3 -	A rede 10-regular de gêneros musicais. Distância do cosseno entre vetores de atributos. Grupos formados pelo algoritmo de Girvan e Newman	119
Figura 7.4 -	Comunidades da rede 10-regular com a distância Euclidiana . . .	120
Figura 7.5 -	A rede de gêneros de acordo com o vetor Fiedler	121
Figura 7.6 -	Coeficientes do vetor de Fiedler	121
Figura 7.7 -	Distribuição dos coeficientes do vetor de Fiedler	122
Figura 8.1 -	Diagrama dos passos principais para a geração automática de listas de reprodução	126
Figura 8.2 -	Uma ilustração do algoritmo <i>jumping walk</i>	128
Figura 8.3 -	Árvore geradora mínima para a rede de gêneros considerando o algoritmo de Prim (85) e distância do cosseno entre os vértices . .	129
Figura 8.4 -	Distâncias entre músicas subsequentes para as primeiras 100 músicas da lista de reprodução obtida pelo algoritmo JW e das listas de estabelecidas por diferentes variações da busca em largura . .	131
Figura 8.5 -	Alternância de gêneros entre músicas subsequentes para as primeiras 100 músicas da lista de reprodução obtida pelo algoritmo JW e das listas de estabelecidas por diferentes variações da busca em largura	132
Figura 9.1 -	Visão geral das etapas realizadas para a análise da presença de motivos musicais em gêneros e relação com expressões de emoção/humor.	136
Figura 9.2 -	Representação da melodia através de vetores NA e NV.	137
Figura 9.3 -	O vetor NA de duas melodias.	139
Figura 9.4 -	O vetor NV de duas melodias.	140
Figura 9.5 -	Configuração de motivos rítmicos e melódicos de sete gêneros. . .	141
Figura 10.1 -	Diagrama da metodologia proposta para a investigação da evolução dos gêneros musicais.	144
Figura 10.2 -	Correlogramas de duas séries temporais.	147

Figura 10.3 - Evolução do ritmo de blues.	149
Figura 10.4 - Evolução do ritmo de rock.	149
Figura 10.5 - Evolução do ritmo de country.	150
Figura 10.6 - Evolução do ritmo de reggae.	150
Figura 10.7 - Evolução temporal dos atributos principais do ritmo para quatro gêneros.	151
Figura 10.8 - Ampliação das trajetórias dos quatro gêneros.	152

Sumário

1	Introdução	27
1.1	Contextualização e importância dos tópicos abordados	28
1.1.1	Classificação automática de gêneros musicais	29
1.1.2	Sistemas de recomendação	35
1.1.3	Síntese automática	36
1.1.4	Evolução de gêneros musicais	36
1.1.5	Relação entre gênero e aspectos emotivos	38
1.2	Principais contribuições	39
1.3	Organização	40
2	Fundamentação teórica I - teoria musical e o formato MIDI	41
2.1	Notação musical	41
2.2	Motivos musicais	43
2.3	Gêneros musicais	44
2.3.1	Razões para uso de gêneros em sistemas de classificação	45
2.4	MIDI	45
2.4.1	Razões para o uso do formato MIDI	48
2.5	Considerações finais	48
3	Fundamentação teórica II - métodos de análise	51
3.1	Extração de atributos	52
3.1.1	Análise de componentes principais	54
3.1.2	Análise discriminante linear	54
3.2	Funções discriminantes lineares e quadráticas	55
3.2.1	Coefficiente Kappa	56
3.3	Agrupamento hierárquico aglomerativo	58
3.3.1	Métricas de distâncias	59
3.3.2	Critérios de <i>linkage</i>	59
3.4	Representação de sistemas por grafos	61
3.4.1	Árvores geradoras mínimas	64
3.4.2	Redes complexas	64
3.4.3	Deteção de comunidades em redes	65

3.5	Considerações finais	69
4	Materiais e métodos	71
4.1	Obtenção do vetor de notas de valores	71
4.2	Obtenção dos atributos	73
4.3	Representação em dígrafos	75
4.4	Considerações finais	76
5	Síntese de sequências rítmicas	79
5.1	Taxonomia adotada	79
5.2	Estratégia	80
5.3	Resultados e discussões	81
5.3.1	Síntese por música	81
5.3.2	Síntese por gênero	84
5.4	Considerações finais	86
6	Classificação de gêneros musicais	89
6.1	Taxonomia adotada	89
6.2	Estratégia	90
6.3	Resultados e discussões	93
6.3.1	Extração de atributos	93
6.3.2	Classificador Gaussiano	97
6.3.3	Agrupamento hierárquico aglomerativo	101
6.3.4	Classificação multi-categorias	105
6.4	Análises com uma taxonomia de estilos	109
6.5	Considerações finais	112
7	Detecção de comunidades em redes de música	115
7.1	Taxonomia adotada	116
7.2	Estratégia	116
7.3	Resultados e discussões	117
7.3.1	Inclusão de medidas topológicas	122
7.4	Considerações finais	123
8	Sistema de recomendação musical	125
8.1	Taxonomia adotada	125
8.2	Estratégia	126
8.2.1	Algoritmo <i>Jumping-walk</i>	127
8.3	Resultados e discussões	128
8.4	Considerações finais	133

9	Detecção de motivos e relação com expressões de emoção/humor . . .	135
9.1	Taxonomia adotada	136
9.2	Estratégia	137
9.2.1	Detecção automática dos padrões temporais	137
9.3	Resultados e discussões	139
9.4	Considerações finais	142
10	Evolução de gêneros musicais	143
10.1	Taxonomia adotada	143
10.2	Estratégia	144
10.2.1	Análise da dispersão global do ritmo através de séries temporais . .	145
10.2.2	Evolução dos componentes principais do ritmo	147
10.3	Resultados e discussões	148
10.3.1	Correlação temporal dos atributos do ritmo	148
10.3.2	Evolução temporal dos atributos principais do ritmo	151
10.4	Considerações finais	154
11	Conclusões	157
11.1	Trabalhos futuros	161
	Referências	163

Capítulo 1

Introdução

De forma geral, a música pode ser entendida como a arte de organizar e combinar sons no espaço e no tempo (1). O espaço está relacionado com a harmonia, isto é, com a relação vertical ou simultânea entre os sons. O tempo, por outro lado, estabelece uma relação horizontal ou sequencial dos sons. A cadência de acordes é um exemplo de organização espacial, enquanto que as notas de uma melodia configuram uma organização temporal (2).

A música permite a expressão de sentimentos e é utilizada para um vasto conjunto de objetivos, entre eles: o instrutor aeróbico busca por uma música com um certo ritmo; o produtor de filme busca por uma trilha musical que estimula uma determinada emoção (por exemplo, tensão); o motorista pode precisar de uma música que o mantém em alerta. Uma vez que a música possui papéis importantes dentro dos contextos sociais e psicológicos, formas de representação eficazes devem permitir buscas por conteúdo que contemplem estes aspectos (113). Tais formas são tipicamente baseadas em quatro tipos de informações: gênero, emoção, estilo e similaridades (113). Estas informações estabelecem as principais aplicações em pesquisas que envolvem recuperação da informação musical (do inglês, *Music Information Retrieval - MIR*).

Gêneros e estilos musicais são, portanto, importantes descritores, uma vez que resumam características (ou padrões) comuns entre peças. Estes descritores são usados há anos para organizar coleções e refletem interação entre culturas. No contexto etnográfico, gêneros e estilos são interessantes porque expressam identidades das fundações culturais às quais eles pertencem (3), além de refletirem trajetórias e interações entre artistas.

Apesar do amplo uso, gêneros musicais não formam um conceito claro e definido, e suas fronteiras são controversas (124). Como consequência, o desenvolvimento de uma taxonomia se torna confusa e redundante, representando um desafio. Foi demonstrado que não há um acordo geral sobre taxonomias de gêneros musicais, as quais podem, inclusive, ser influenciadas por fatores culturais, convenções históricas, interações entre artistas e estratégias de *marketing* (124). Até mesmo termos como “rock”, “jazz”, “blues” e “pop” não são totalmente claros e definidos.

A distinção entre gênero e estilo musical é similarmente controversa e redundante. Enquanto que gênero é geralmente considerado um conceito mais amplo e subjetivo em termos de conteúdo musical (105), não há um acordo geral sobre o que um estilo deve abranger ou se é aceitável o uso indiscriminado desses dois conceitos (4). Neste trabalho, o termo “gênero” está utilizado como aspecto mais abrangente (por exemplo, *blues*), enquanto que o termo “estilo” será empregado como variações dentro do gênero (por exemplo, *country-blues*, *rock-blues*).

1.1 Contextualização e importância dos tópicos abordados

As últimas décadas foram marcadas pelo constante crescimento de dados musicais *online*, o que tem atraído pesquisas para o desenvolvimento de ferramentas adequadas para analisar, indexar, sumarizar e classificar dados musicais em larga escala. O objetivo principal destas ferramentas é extrair informação necessária e compacta para representar o conteúdo musical de forma adequada. Dentro deste contexto, torna-se interessante a aplicação de métodos automáticos de agrupamento de dados, pois estes permitem a redução da dimensão de um determinado conjunto de dados através do uso de características intrínsecas aos mesmos. O desafio na aplicação destes métodos está relacionado com a escolha de métricas adequadas de similaridade, e com a obtenção de formas de representação que sejam ao mesmo tempo compactas e discriminativas.

Dentro deste prisma, o objetivo principal deste doutorado é identificar e explorar características dos gêneros musicais em aplicações inerentes à recuperação da informação musical, através de uso de atributos musicais originais, e do emprego de formas de representação e métodos de agrupamento de dados até então não utilizados no contexto de classificação e modelagem de gêneros musicais. Tais atributos se referem aos padrões rítmicos de cada música. Desta forma, busca-se caracterizar músicas de diferentes gêneros pelos seus respectivos padrões rítmicos.

O ritmo fornece uma percepção de regularidade temporal e pode ser entendido como um padrão específico produzido por notas que se diferem em duração, pausa (silêncio) e intensidade. Através dele é possível distinguir, por exemplo, uma música de rock e uma música de salsa. A motivação para usá-lo como referência está relacionada com o fato de que o ritmo possui regras definidas de composição e é intuitivamente característico dos gêneros musicais, sendo um dos aspectos mais adotados em tarefas de classificação deste contexto (5).

Os principais tópicos nos quais as estratégias desenvolvidas neste trabalho podem ser aplicadas no tocante à recuperação da informação musical estão discutidos a seguir.

A forma de representação e metodologia adotada são originais e estão apresentadas em cada caso. Igualmente, busca-se contextualizar o estado da arte de pesquisas anteriores relacionadas, identificando questões em aberto e destacando a contribuição do presente trabalho para a investigação destas questões.

1.1.1 Classificação automática de gêneros musicais

A caracterização e classificação dos gêneros musicais têm impacto direto em sistemas que organizam e buscam conteúdo musical, sendo de suma importância o desenvolvimento de métodos computacionais para a organização de grandes volumes de dados a partir de representações compactas. Avanços em pesquisas de recuperação da informação musical buscam sumarizar estas informações, de forma que as mesmas ainda sejam adequadas para representar os gêneros.

De forma similar a diversos problemas envolvendo reconhecimento de padrões, o processo de classificação de gêneros musicais pode ser usualmente dividido em três etapas principais: representação, extração de atributos e modelagem do processo de classificação (6, 7). A informação musical pode estar representada de forma simbólica ou através de sinais acústicos (8). A representação simbólica é geralmente utilizada através de formatos que configuram partituras musicais (como, por exemplo, o formato MIDI - *Musical Instrument Digital Interface*), em que cada nota é descrita em termos de altura, duração e intensidade. Por outro lado, as representações provenientes de sinais acústicos são obtidas pela amostragem de onda sonora. Destaca-se que o trabalho se concentra na informação simbólica da informação musical. A justificativa para tal, mais detalhada na seção 2.4.1, refere-se a possibilidade de identificar de forma direta e clara quais elementos musicais estão sendo importantes na análise. Os conceitos desta representação e de notação musical estão detalhados no Capítulo 2.

Atributos musicais utilizados em tarefas de classificação de gêneros estão usualmente associados com o timbre, variação temporal, ritmo, altura de nota (do inglês, *pitch*) e harmonia. Atributos relacionados com o timbre e variação temporal (variação do timbre ao longo do tempo), geralmente atribuídos como de baixo nível, são obtidos através de técnicas de processamento de sinais como as transformadas de Fourier e análises espectrais. Atributos de baixo nível também são entendidos como atributos sem contextualização, ou micro-estruturas, que são obtidas em intervalos pequenos de tempo. Este grupo de características tem sido usado de forma predominante em tarefas de classificação, pois são simples de serem obtidos e costumam apresentar bom desempenho. Entretanto, este grupo está mais distante de propriedades inerentes da música que fazem parte da experiência perceptual dos humanos. Atributos relacionados com o ritmo, altura de nota e harmonia

são conhecidos como de médio-nível, sendo mais próximos da experiência humana, porém a obtenção destas características através de dados de áudio ainda apresenta um desafio (5).

Diversos trabalhos podem ser encontrados na literatura envolvendo classificação automática de gêneros musicais com diferentes atributos e métodos de aprendizagem. A comparação de desempenho entre abordagens é uma tarefa custosa, uma vez que diferentes taxonomias de gêneros são empregadas nas tarefas de classificação. Na última década, entretanto, bancos de dados com arquivos de áudio têm sido disponibilizados para que as comparações de desempenho sejam mais viáveis, tais como o banco disponibilizado pela conferência ISMIR 2004 (9), e os bancos disponíveis por Homburg (110), e por Tzanetakis (10). Bancos de dados simbólicos para fins acadêmicos foram disponibilizados por Masataka (121), e por McKay (11), porém os autores deste trabalho desconhecem estudos acadêmicos que realizem comparações de desempenho em dados simbólicos como a versão encontrada para arquivos de áudio (5).

Uma descrição completa de abordagens recentes que empregam diferentes atributos de áudio e métodos de aprendizagem para a discriminação de gêneros musicais pode ser encontrada na literatura (5, 12). Comparações de desempenho entre abordagens através de um banco de dados comum também estão disponíveis. De acordo com os autores, a aprendizagem supervisionada é predominante em tarefas de classificação de gêneros, de forma que os métodos de aprendizagem mais utilizados são: máquina de vetores de suporte (SVM, do inglês *Support Vector Machines*), k -vizinhos mais próximos (k -NN, do inglês *k-Nearest Neighbor*), e modelos de mistura Gaussiana (GMM, do inglês *Gaussian Mixture Model*) (13). A taxa média de classificação correta está no intervalo de 60% a 90%. Em todas as abordagens analisadas pelos autores, taxas altas de classificações corretas são obtidas através da combinação de diferentes atributos. Neste trabalho adotou-se o classificador Gaussiano como método de aprendizagem supervisionada. Foi verificado que atributos simples do ritmo aplicados neste classificador podem permitir taxas de acerto comparáveis às anteriormente citadas. Esta averiguação está descrita no Capítulo 6.

Outras abordagens incluem o uso de aprendizado semi-supervisionado (14, 134), combinação de classificadores (130) e atributos extraídos de arquivos de áudio a partir de arquivos MIDI (8). Dentro das taxonomias adotadas em cada um destes casos, a taxa média de acerto está contida em um intervalo similar ao anteriormente indicado (entre 70% e 90%).

A classificação automática de gêneros musicais através de arquivos MIDI foi explorada por (103, 128), porém analisada de forma mais sistemática por (11). Os atributos estão relacionados com o timbre, instrumentação, ritmo, melodia e harmonia. O autor propôs um modelo para classificação hierárquica dos gêneros, na qual os atributos são selecionados conforme o nível da hierarquia. A tarefa de classificação foi realizada pela combinação de dois métodos: redes neurais e k -vizinhos mais próximos. Dentro de uma taxonomia de três gêneros raízes e 9 gêneros folhas, a acurácia de classificação foi, em média, de 98%

para o primeiro caso, e 90% para o segundo.

Entre as pesquisas apresentadas, abordagens que usam unicamente o ritmo como atributo na tarefa de classificação de gêneros musicais podem ser encontradas. Tzanetakis e Cook em (10), utilizou histogramas de batidas para a classificação hierárquica de nove gêneros. A taxa de acerto obtida pelos autores foi de, em média, 60%. Uma revisão sobre sistemas automáticos para descrição de ritmos é apresentada em (15, 107). Os autores relatam que, apesar do consenso de alguns conceitos rítmicos, não existe uma única representação do ritmo que seja adequada para diferentes aplicações, como, por exemplo, indução do tempo e compasso, reconhecimento de batidas e quantização do ritmo. Os autores também estudaram a relevância de três grupos de descritores do ritmo ao analisarem seus respectivos desempenhos em experimentos para classificação de oito estilos musicais. A combinação de descritores permitiu uma taxa de acerto de 90%. Também foi observado pelos autores e por Scaringella em (12) que, em muitas das abordagens existentes, não há uma adequada compreensão de como o ritmo contribui para a classificação, dado que ele está representado de forma limitada nestas abordagens. Karydis (115) propôs um sistema para classificar peças clássicas em cinco categorias utilizando atributos obtidos a partir dos histogramas das alturas e durações das notas, em que a acurácia de classificação obtida foi, em média, 70% através do algoritmo K-NN. No trabalho de Akhtaruzzaman (99), o ritmo é analisado em termos de suas propriedades matemáticas e geométricas e posteriormente utilizado em um sistema hierárquico de classificação de ritmos de diferentes regiões.

A classificação não supervisionada ou agrupamento de dados (do inglês, *clustering*), é relativamente pouco explorada para tarefas de classificação de gêneros musicais (12). Segundo McKay e Scaringella (12, 123) a razão para tal é que a noção de gênero pode desaparecer dado que agrupamentos baseados somente em padrões de similaridade provavelmente não irão refletir uma taxonomia de gêneros próxima à usada pelos humanos, uma vez que estas são ilógicas e inconsistentes. O uso do algoritmo k-médias (do inglês, *k-means*, agrupamentos hierárquicos, mapas auto-organizáveis (do inglês, SOM - *Self Organized Maps*, e modelos ocultos de Markov (do inglês, HMM - *Hidden Markov Models*) estão entre as técnicas mais utilizadas (16, 129). Nestas abordagens, há a combinação de diferentes atributos do ritmo, timbre e volume, e o desempenho do agrupamento em relação à taxonomia usada como referência é comparável aos métodos supervisionados.

Apesar das desvantagens relatadas, as técnicas de agrupamento de dados evitam a necessidade de lidar com as inconsistências e redundâncias das taxonomias reais, ainda mais evidentes para taxonomias mais realísticas, de forma que as categorias produzidas podem estar mais adequadas que as especificadas por especialistas em termos de uma função objetiva de similaridade. No contexto da análise rítmica como está sendo proposta neste trabalho, aplicações de busca por similaridade ou por aspectos emotivos podem se beneficiar deste aspecto, uma vez que emoção e humor em música estão altamente correlacionados com ritmo (5).

Com base nas pesquisas anteriores supracitadas, este trabalho objetiva verificar se mecanismos simples podem ser utilizados para representar e discriminar o ritmo de músicas de diferentes gêneros com desempenho comparável às propostas anteriores. Tais mecanismos são estabelecidos pelas probabilidades condicionais para indicar a dinâmica de transição das notas rítmicas. São utilizadas as abordagens de classificação supervisionada e não supervisionada. Uma das inovações incorporadas é o uso de grafos e redes complexas na modelagem do sistema.

As redes complexas têm sido referenciadas como importantes mecanismos para representar diversos aspectos da natureza, uma vez que suas características topológicas e estruturais contribuem para o entendimento das relações, e propriedades inerentes dos dados (17). No contexto de aspectos musicais e artísticos, redes de música têm sido construídas e suas propriedades topológicas ajudam na compreensão da dinâmica e relação entre os elementos envolvidos. Portanto, a motivação da tese está na possibilidade de construir redes de música e analisar as comunidades de gêneros encontradas nestas redes. A inclusão de medidas topológicas pode favorecer o entendimento das características rítmicas de um gênero musical e da relação entre gêneros.

A seção a seguir apresenta trabalhos anteriores que utilizam redes complexas à luz da modelagem da informação musical.

Redes complexas na representação de informação musical

Como mecanismos de representação, os grafos tem se mostrado ferramentas eficientes para caracterizar e complementar o entendimento de diversos sistemas reais complexos (59). Nestes modelos, os elementos de interesse são representados por nós, enquanto que a relação entre eles, por arestas. Um nó pode indicar, por exemplo, uma pessoa, uma cidade, um neurônio, e assim por diante. Uma conexão pode indicar, por exemplo, que duas pessoas se conhecem, que existe uma linha direta de ônibus entre as duas cidades, ou que dois neurônios utilizam a mesma sinapse para transmitir sinais. Se o grafo possui muitas conexões, é também referenciado como uma rede complexa. Dentro deste contexto, uma rede complexa pode informalmente ser entendida como um conjunto grande de nós que se conectam entre si. A Internet é um exemplo de rede complexa.

Um dos principais aspectos deste modelo é que ele é capaz de capturar características dos sistemas complexos como um todo. Sistemas complexos podem ser definidos como sistemas compostos por muitos elementos que interagem entre si de forma não trivial (60). Desta forma, a análise da estrutura formada pela modelagem destas inter-relações pode revelar informações sobre o sistema que não seriam acessíveis através da análise de suas partes separadas. Esta análise inclui a extração de medidas globais como a distribuição dos graus, que resalta como os vértices estão conectados e quais elementos tem papel importante dado que possuem quantidades maiores de conexões; e o coeficiente de aglomeração, que

reflete como os vértices estão conectados entre si, e se há uma tendência para a formação de grupos. Modelos de redes foram consolidados, nos quais medidas topológicas seguem um padrão específico, o que também pode revelar informações importantes sobre a origem e dinâmica de tais sistemas.

Em um estudo proposto por Silva e demais (18), músicas populares brasileiras são estudadas com o objetivo de representá-las em uma rede complexa, na qual os vértices são os compositores e as arestas representam a existência de pelo menos um cantor em comum entre dois compositores. Os autores constataram que a rede construída apresenta propriedades de uma rede de pequeno mundo, com um alto coeficiente de aglomeração e distribuição dos graus dada por uma lei de potência.

Na análise de Gleiser (19), duas redes de colaboração foram desenvolvidas: entre músicos de *jazz* e entre bandas de *jazz*. Na primeira, dois músicos estão conectados se já tocaram na mesma banda; na segunda, duas bandas estão conectadas se possuem um músico de *jazz* em comum. A rede de colaboração entre músicos apresenta alto coeficiente de aglomeração, assortatividade ^a e propriedades de rede de pequeno mundo. Nestas redes, uma pequena quantidade de músicos contém grande número de conexões. A rede de colaboração entre bandas também apresenta alto coeficiente de aglomeração, e sugere que as bandas estão interconectadas com números similares de *links*.

Uma rede social formada pela colaboração entre *rappers* foi proposta em (20), na qual dois *rappers* estão conectados se gravaram juntos. O autor utilizou dados extraídos de site *Original Hip-Hop Lyrics Archive* (21). A rede obtida possui propriedades de uma rede de pequeno mundo, com distribuição dos graus caracterizada pela lei de potência e alto coeficiente de aglomeração.

Os autores em (22) utilizaram o site *All Music* (23) para analisarem duas redes sociais de músicos contemporâneos. Uma das redes, nomeada de rede de colaboração, estabelece que dois músicos estão conectados se tocaram juntos ou geraram um álbum juntos. A outra rede, denominada de rede de similaridade, estabelece que dois músicos estão conectados se as músicas compostas por eles apresentam características similares, de acordo com a opinião de especialistas. As duas redes apresentam alto coeficiente de aglomeração e propriedades de rede de pequeno mundo. Entretanto, estas divergem em outras características como distribuição dos graus e assortatividade (a rede de similaridade é totalmente assortativa e a rede de colaboração é parcialmente assortativa).

Um estudo quanto a similaridade entre músicos e bandas através de redes complexas é apresentado em (24). Os músicos (ou bandas) estão conectados se já tocaram ou compuseram músicas em comum. Os bancos de dados utilizados foram: *All Music*, *MSN Entertainment*, *Amazon* e *Launch Yahoo*. Esses bancos possuem a característica de que, quando um usuário seleciona um artista, outros são sugeridos como próximos aos de

^a probabilidade alta de vértices com alta conectividade se conectarem com outros vértices de alta conectividade

interesse do usuário. A similaridade é estabelecida de duas formas diferentes: hábitos dos usuários em escolher tais artistas e a opinião de especialistas em música. Os autores verificaram que a rede de recomendação musical construída por colaboração (similaridade através das escolhas dos usuários) possui propriedades livres de escala (distribuição dos graus dada pela lei de potência); e a rede de recomendação musical construída pela supervisão de especialistas possui propriedades de pequeno mundo. Na prática isto pode indicar que os usuários tendem a seguir outros usuários que já possuem uma boa aceitação, enquanto especialistas tendem a formar grupos nos quais compartilham opiniões.

Músicas de *Bach*, *Mozart* e *Chopin* e variação de músicas chinesas também foram usadas para a construção de uma rede de músicas (25). Cada música é representada como uma sequência temporal de notas. Dessa forma, cada nota é representada por um vértice e as conexões indicam a ordem temporal que as notas ocorreram. As propriedades de redes são extraídas e comparadas entre redes de músicas de diferentes compositores. Observou-se que, propriedades como coeficiente de aglomeração e distribuição dos graus são similares em todas as redes.

As redes complexas foram utilizadas para modelar a colaboração entre músicos e compositores através de um amplo banco de dados de álbuns de música brasileira (108). O estudo das propriedades destas redes permitiu uma análise da estrutura musical brasileira, assim como um levantamento de músicos mais influentes ao longo das décadas.

A detecção de comunidades também foi estudada em redes de música. Duas redes sociais foram estudadas através de métricas de similaridade e colaboração entre músicos, de forma que padrões organizacionais e aspectos que influenciam o crescimento de tais redes foram identificados (27). Em outro estudo similar, os autores descreveram dificuldade em se definir uma taxonomia de gêneros a partir de análises de dados online (28).

De acordo com as pesquisas supracitadas, verifica-se que a utilização de grafos e redes complexas pode beneficiar estudos que buscam entender interações entre artistas e músicas. Entretanto, estas ferramentas ainda não foram sistematicamente exploradas para a representação de atributos musicais e para discriminação de gêneros ou análise de suas relações. Esse trabalho contribui para tais aplicações através do uso destas ferramentas para a representação dos atributos e discriminação dos gêneros musicais. Atributos topológicos e algoritmos de detecção de comunidades são explorados para discriminação dos gêneros. As comunidades são obtidas por diferentes algoritmos de agrupamento e medidas de distância, definindo uma abordagem não supervisionada e motivada pela inclusão de medidas topológicas na análise dos dados.

As comunidades obtidas são comparadas com os agrupamentos resultantes do agrupamento hierárquico aglomerativo. Observa-se que o uso de redes complexas pode beneficiar a análise e discriminação dos gêneros musicais. Estas análises estão descritas nos Capítulos 6 e 7.

1.1.2 Sistemas de recomendação

A disponibilização de recursos *online* com conteúdo musical também tem aumentado a importância de sistemas de recomendação, de forma que a geração de listas de reprodução de forma automática surge naturalmente como uma aplicação inerente deste contexto. De fato, a tarefa manual e demorada de selecionar músicas para uma lista de reprodução pode ser automatizada por algoritmos que criam listas de acordo com as preferências dos usuários ou através de algum critério de similaridade entre músicas subsequentes. Sistemas de recomendação têm se mostrado cada vez mais necessários com o avanço de bibliotecas e rádios *online* e com o aperfeiçoamento de dispositivos móveis.

Na literatura, é possível encontrar diversas abordagens para a geração automática de listas de reprodução (29–31, 100, 109, 120, 126). Estas abordagens exploram a similaridade entre músicas a partir de três principais formas: filtros colaborativos, análise de áudio, e uso de metadados. Os filtros colaborativos comparam padrões de escolha entre usuários. A análise de áudio estabelece comparações entre atributos como altura, ritmo e timbre. A similaridade por metadados é baseada em informações como gênero e estilos, na maioria das vezes rotuladas por usuários.

O algoritmo desenvolvido neste trabalho é baseado em caminhadas em árvores geradoras mínimas (AGM) de redes de músicas, possui baixo custo computacional e não requer prévia rotulação dos usuários ou análise de metadados, evitando a inserção de possíveis ruídos consequentes da subjetividade do processo de rotulação. Cada vértice da árvore representa uma música e o peso das arestas reflete a similaridade dos atributos rítmicos entre pares de música. Destaca-se que a abordagem desenvolvida apresenta as seguintes vantagens: a utilização das árvores geradoras mínimas, que oferecem uma estrutura de grafo na qual vértices vizinhos são os mais similares entre si; e a consulta dos pesos das arestas no processo de busca. Estes dois fatores maximizam a similaridade entre músicas subsequentes. Apesar de não ter sido realizada uma avaliação conclusiva dos usuários para a qualidade das listas geradas, busca-se discutir características gerais destas listas em relação às propriedades da AGM.

A motivação do método desenvolvido também está relacionada com a possibilidade de associar os gêneros com a presença de padrões temporais no ritmo das músicas como forma de expressar noções de emoção. Gabrielsson (106) sugere que padrões rítmicos regulares ou suaves indicam expressões como alegria, contentamento ou tranquilidade. Ritmos irregulares ou complexos indicam expressões como surpresa, tensão ou ansiedade. Desta forma, na lista de reprodução obtida pelo algoritmo desenvolvido, as músicas são similares entre si de acordo com aspectos rítmicos e, portanto, estão emocionalmente relacionadas. O algoritmo desenvolvido está apresentado no Capítulo 8.

1.1.3 Síntese automática

Síntese automática pode ser entendida como o processo de gerar sons de forma automática. Esta pode abranger aspectos de baixo nível através de técnicas de processamento de sinais, ou pode abranger estruturas mais abstratas como as notas musicais. Miranda (2) apresenta diversas técnicas para a composição de música por computadores, como o uso de gramáticas, autômatos, algoritmos iterativos, algoritmos genéticos e redes neurais artificiais.

O grande desafio da síntese automática está relacionado com a qualidade musical das peças compostas. Em um estudo proposto por Liu (25), músicas de *Bach*, *Mozart* and *Chopin* e variação de músicas chinesas são usadas para a construção de uma rede de músicas. Variações do método de caminhada aleatória foram utilizadas para a síntese de novas sequências de notas. Os autores verificaram que a presença de tais propriedades nem sempre garante que as novas músicas compostas sejam musicalmente interessantes.

Este trabalho também investigou a possibilidade da composição de novas sequências rítmicas com base em estatísticas coletadas dos atributos rítmicos das músicas de diferentes gêneros (Capítulo 5). Esta composição é realizada através de caminhadas aleatórias nas redes de notas rítmicas. Apesar do método de síntese não ser original, tem-se que a modelagem dos atributos como desempenhada neste trabalho permite que as novas sequências mantenham características inerentes da referência utilizada para execução do método, permitindo uma variabilidade “controlada”. Por exemplo, é possível gerar sequências rítmicas de rock, de reggae ou blues. Cada uma destas sequências expressam a assinatura do gênero utilizado como referência.

1.1.4 Evolução de gêneros musicais

A evolução da música é um processo complexo envolvendo fatores sociais e psicológicos que são muitas vezes difíceis de descrever e analisar (113). Formas musicais se mostram substancialmente diferentes de acordo com distintas regiões e culturas e estão em constante mudança ao longo do tempo. Esta complexidade contribui para que o estudo da evolução da música se torne uma tarefa difícil, que tem sido investigada principalmente dentro de uma perspectiva qualitativa por diversas disciplinas, tais como sociologia, musicologia e história da música (32–35).

A seção 1.1.1 ofereceu uma discussão sobre trabalhos que têm examinado atributos discriminativos de gêneros musicais (5, 12, 39), porém o estudo da sua evolução tem recebido menos atenção. Dentro deste contexto, este trabalho investigou como o ritmo de músicas pertencentes a um determinado gênero tem evoluído ao longo do tempo. Busca-se

evidências para respostas das seguintes perguntas: (i) Como a diversidade dentro de um gênero evolui? (em outras palavras, é possível manter constante inovação dentro de um gênero?); (ii) É possível explicar a evolução dos gêneros musicais a partir de mecanismos simples? Questões similares se aplicam para a evolução de gêneros literários (36).

Figura 1.1 – Esquema resumido dos fatores principais que influenciam a evolução dos gêneros musicais.

O esquema da Figura 1.1 resume alguns mecanismos assumidos neste trabalho como importantes para a evolução dos gêneros e para a inclusão de diversidade nas músicas. Por um lado, aspectos sócio-culturais, o desejo do compositor de inovar seu material artístico, influências provenientes da interação entre compositores, e interações entre gêneros mediadas pela sociedade favorecem a diversidade de material musical. Por outro lado, esta diversidade pode ser regulada pelo fato de que compositores precisam seguir determinadas regras de composição a fim de manter a identidade de cada gênero. O estudo realizado neste trabalho sugere que a inclusão de diversidade não se mantém indefinidamente. Para todos os gêneros analisados, observou-se que a inovação rítmica ocorre em períodos curtos de tempo (tipicamente anos), mas claramente existem momentos de ressurgimento das formas rítmicas após períodos longos de tempo (tipicamente décadas). Este comportamento é interpretado como um mecanismo de recuperação associado com a competição entre fatores que favorecem a inclusão de diversidade (por exemplo, o desejo dos compositores de inovar) e a obediência às regras de composição.

Neste trabalho, a evolução da música é explorada dentro de uma perspectiva quantitativa através de uma análise estatística da evolução do ritmo de quatro gêneros musicais, sendo eles: blues, country, reggae e rock. Mais especificamente, uma medida de dispersão dos atributos rítmicos é estabelecida e calculada para diversas músicas que foram gravadas ao longo de um intervalo de tempo e representam a história de cada gênero. A evolução da dispersão do ritmo é expressada em termos de séries temporais e da análise discriminante linear, permitindo que as principais características da dispersão do ritmo sejam rastreadas ao longo do tempo. Esta análise está descrita no capítulo 10.

1.1.5 Relação entre gênero e aspectos emotivos

Tão notável quanto as distinções gênero-estilo são os aspectos emotivos provenientes da música. Os fatores pelos quais emoções são estimuladas através da música são objetos de estudo há muitas décadas (114). Comparando com a classificação de gêneros, a classificação automática de emoção é uma tarefa mais complexa, uma vez que envolve conceitos subjetivos e a dificuldade em se obter anotações concisas e dicionários de termos não redundantes. Pesquisas relacionadas com a identificação automática de emoção em música ainda estão em estágios iniciais e têm recebido atenção nos últimos anos.

Desta forma, gênero e emoção compreendem elementos principais para a organização da informação musical (113), e têm sido amplamente aplicados para recuperação musical como processos independentes (5, 12, 39, 116). Entretanto, a relação intrínseca entre estes dois aspectos é pouco analisada atualmente. Por exemplo, uma música de *rock* é geralmente motivadora, enquanto que uma música de *blues* tende a ser mais sentimental.

Em Hu (112), as relações gênero-emoção, artista-emoção, e recomendação-emoção são exploradas através da correlação entre *tags* sociais dos sites *All Music*, *epinions.com* e *Last.fm*. Contudo, atributos musicais não são considerados. Os autores sugerem que gênero e emoção são independentes ao verificarem que rótulos de emoção correlacionados estão presentes em diferentes gêneros. Por outro lado, Lin (119) afirma que expressões de emoção que co-ocorrem em um gênero são relativamente diferentes para outros gêneros. Através de uma análise estatística, os autores demonstram que o desempenho da classificação de emoção é melhorado quando a informação do gênero está disponível antecipadamente.

Dada esta contextualização, observa-se que são poucos os estudos que relacionam gênero e emoção. A análise desempenhada neste trabalho tem como objetivo contribuir para um melhor entendimento da relação inerente entre gênero e emoção, tornando-os descritores complementares que podem ser utilizados para aprimorar aplicações existentes de sistemas de recuperação da informação musical. O uso de dados simbólicos para um estudo sistemático de padrões musicais temporais à luz da inter-relação gênero e emoção é original na literatura. Como consequência, também é original o algoritmo para a detecção automática de motivos musicais.

Neste estudo, a caracterização dos gêneros musicais é analisada à luz de um contexto musical estendido representado pelos motivos musicais, que são definidos como pequenas sequências de notas ou rítmicas que se repetem com certa frequência em uma composição musical. O propósito da investigação é averiguar se a presença destes padrões temporais rítmicos e tonais pode ser usada para estabelecer uma relação entre gêneros musicais e aspectos subjetivos como emoção/humor. Este estudo foi desempenhado em oito gêneros musicais e está descrito no capítulo 9.

1.2 Principais contribuições

A Figura 1.2 apresenta uma visão geral das direções deste trabalho. As contribuições estão descritas a seguir:

1. Os atributos rítmicos utilizados e sua modelagem são originais e trazem informações sobre a dinâmica rítmica das músicas;
2. Definição de uma metodologia para a discriminação dos gêneros musicais, através da classificação supervisionada, métodos tradicionais de agrupamento de dados, e algoritmos para detecção de comunidades em redes complexas. Esta metodologia define uma forma quantitativa para identificar as convergências/divergências da taxonomia usual dos gêneros, o que parece ser uma conjectura até o presente momento (36);
3. Inclusão de medidas topológicas de redes complexas na análise dos gêneros musicais;
4. Geração de novas composições contextualizadas pelos gêneros musicais;
5. Desenvolvimento de uma abordagem original para a geração automática de listas de reprodução contextualizadas por gêneros musicais e aspectos rítmicos das músicas;
6. Investigação quantitativa da evolução dos gêneros musicais através dos padrões rítmicos;
7. Caracterização dos gêneros através da presença de motivos musicais e associação destes motivos com noções subjetivas de emoção/humor.

Figura 1.2 – Visão geral das direções exploradas neste trabalho. A partir da identificação dos atributos rítmicos de músicas de diversos gêneros, tem-se as seguintes investigações: classificação, composição automática, geração de listas de reprodução para sistemas de recomendação, evolução de gêneros e associação dos mesmos com expressões subjetivas de emoção/humor.

1.3 Organização

Esta tese está organizada da seguinte forma:

- Capítulo 2: descreve de forma resumida conceitos musicais necessários para a compreensão do trabalho;
- Capítulo 3: apresenta a fundamentação teórica do trabalho e descrição dos algoritmos adotados;
- Capítulo 4 descreve os métodos desenvolvidos para obtenção dos atributos rítmicos;
- Capítulo 5: aborda a análise de síntese de novas sequências;
- Capítulo 6: expõe a classificação dos gêneros através de técnicas supervisionadas e de agrupamento de dados;
- Capítulos 7: apresenta resultados da detecção de comunidades em redes de música;
- Capítulo 8: descreve o sistema de recomendação musical;
- Capítulo 9: explora o estudo da identificação de padrões temporais à luz da relação com expressões de emoção/humor;
- Capítulo 10: apresenta estudo da evolução dos gêneros;
- Capítulo 11: descreve as conclusões principais do trabalho.

A seguir são apresentados conceitos iniciais de notação musical e do formato adotado para representação dos eventos musicais.

Capítulo 2

Fundamentação teórica I - teoria musical e o formato MIDI

Este capítulo tem como objetivo principal apresentar noções básicas de notação musical, assim como brevemente discutir aspectos inerentes da classificação automática de gêneros musicais, como, por exemplo, a dificuldade em se obter definições claras que acarretam o não estabelecimento de uma taxonomia concisa e objetiva. Também são apresentadas informações sobre o formato MIDI que será utilizado neste trabalho.

2.1 Notação musical

As notas musicais são escritas nas linhas e espaços do pentagrama e nomeadas conforme a clave musical atribuída (clave de Sol, de Fá e de Dó). A Figura 2.1 apresenta as sete notas musicais naturais para a clave de Sol. A clave de Sol, mais utilizada para instrumentos musicais de sons médios e agudos, determina que a nota localizada na segunda linha do pentagrama será a nota Sol. As demais notas são localizadas no pentagrama em função da diferença de frequência em relação à nota Sol, sendo as notas de frequência mais baixa nas linhas ou espaços inferiores e as de frequência mais alta nas linhas ou espaços superiores. Da mesma forma, as claves de Fá (mais utilizada para instrumentos de som graves) e Dó (mais utilizada para instrumentos de som médios) determinam a localização das notas que receberão os nomes Fá e Dó, respectivamente, no pentagrama.

Figura 2.1 – As sete notas musicais

As notas musicais também podem ser representadas de letras alfabéticas, conforme a Tabela 2.1.

Tabela 2.1 – Associação das notas musicais com as letras do alfabeto

Letras	A	B	C	D	E	F	G
Notas	Lá	Si	Dó	Ré	Mi	Fá	Sol

As notas musicais possuem atributos como altura, intensidade e duração. A altura está associada a uma frequência fundamental. A intensidade expressa a qualidade que caracteriza uma nota tocada como forte ou fraca. A duração está relacionada com o tempo em que uma nota é tocada ou o tempo de silêncio entre duas notas (pausas) (2). A duração é um dos elementos determinantes do ritmo, geralmente representada por notas de valores (Tabela 2.2)^a.

Tabela 2.2 – As cinco notas de valores mais comuns são: semibreve, mínima, semínima, colcheia e semicolcheia. A semibreve é a figura de maior valor e as demais são determinadas através de frações de tempo da semibreve. As pausas representam a duração de silêncio entre dois eventos e seguem os mesmos valores das respectivas notas. Duas notas ligadas são tocadas como se fossem uma única nota e a duração é determinada pela soma de seus valores. O ponto de aumento é um sinal colocado à direita da nota ou pausa, aumentando metade de seu valor original. Tercinas são alterações na duração padrão. No exemplo apresentado, a mínima e a semínima estão divididas em três partes iguais.

Nome	Símbolo	Pausa	Duração Relativa
Semibreve			1
Mínima			1/2
Semínima			1/4
Colcheia			1/8
Semicolcheia			1/16

ligadura	ponto de aumento		tercina

Outro aspecto importante é o conceito de intervalo musical. Define-se intervalo musical como sendo a diferença de altura entre duas notas. De forma geral, o semitom é considerado o menor intervalo entre duas notas. As notas musicais ilustradas na Figura 2.1 não representam devidamente todas as notas com o intervalo de semitom. Para tal, foram criadas as notações que alteram as notas. Na Figura 2.2 estão representadas, através de números inteiros, todas as notas com intervalo de semitom do Dó na terceira oitava (Dó3) ao Dó# na quinta oitava (Dó#5). As alterações são indicadas por sinais que antecedem

^a Notas de valores são representações gráficas das notas musicais com informação de duração. Outras denominações comuns são: figuras rítmicas, figuras de duração ou figuras de valores.

as notas escritas no pentagrama. Tais sinais são conhecidos como acidentes. Exemplos de acidentes são: o sustenido ($\sharp$), que aumenta a altura da nota em um semitom; e o bemol ($\flat$), que abaixa a altura da nota em um semitom.

Figura 2.2 – Representação das notas musicais em duas claves

Portanto, o intervalo entre duas notas s e t , representadas por valores inteiros, pode ser determinado da seguinte forma:

$$\text{int}(s, t) = t - s. \quad (2.1)$$

Vale destacar que os intervalos musicais são amplamente utilizados em aplicações de análise de conteúdo musical, pois são invariantes a tonalidade da sequência. Portanto, é uma forma de representação mais robusta, apesar da sua simplicidade.

Na música ocidental estabelece-se uma oitava^b como um intervalo que possui uma taxa de frequência 2:1, no qual as notas musicais atuam em ciclos proporcionais. Se a frequência da nota A em 220 Hz é duplicada para 440 Hz, ainda é possível perceber a nota A, entretanto, ela estará uma oitava acima da anterior. De maneira similar, se a frequência for diminuída pela metade, ou seja, se de 220 Hz for para 110 Hz, então essa nota A será percebida uma oitava abaixo. A indicação da oitava será feita nesse trabalho com um número na frente das letras. Por exemplo, C3 representa a nota C (Dó) na terceira oitava.

Mais informações sobre teoria musical podem ser encontradas nas referências (37, 38).

2.2 Motivos musicais

Motivos musicais podem ser definidos como padrões temporais de configuração de notas que se repetem ao longo de uma sequência musical. Basicamente os motivos musicais podem ser tonais ou rítmicos. No primeiro caso mantém-se a sequência das alturas das notas; e, no segundo, mantém-se a sequência das notas de valores.

A Figura 2.3 ilustra o conceito de motivo para as primeiras quatro notas da abertura da 5a. Sinfonia de Beethoven. O motivo tonal e rítmico equivalente ao fragmento apresentado na Figura 2.3a estão indicados nas Figuras 2.3b e 2.3c, respectivamente.

^b O nome oitava está relacionado com a sequência de oito notas sucessivas da escala natural: Dó Ré Mi Fá Sol Lá Si Dó, o segundo Dó é dito estar uma oitava acima do primeiro.

Figura 2.3 – Ilustração de (b) motivo tonal e (c) rítmico das primeiras quatro notas da 5a. Sinfonia de Beethoven em (a)

Em músicas populares é possível encontrar muitas repetições, nas quais motivos rítmicos, tonais, ou até mesmo motivos que mantêm ambas as sequências rítmicas e tonais ao mesmo tempo estão presentes.

2.3 Gêneros musicais

Gêneros musicais permitem que usuários agrupem e caracterizem peças musicais. A hierarquia ou taxonomia de gêneros é provavelmente a forma mais popular de classificar músicas (39). De forma geral, um gênero musical define um conjunto de músicas no qual aspectos musicais seguem características específicas, governadas por regras de composição que são aceitas por uma comunidade. Estas características podem estar relacionadas com instrumentação, estrutura rítmica e harmonia.

Gêneros são rótulos inerentes para classificação, pois é esperado que músicas de um mesmo gênero dividam similaridades entre si, tornando-as diferentes de músicas de outros gêneros (39).

Apesar do amplo uso, gêneros musicais não formam um conceito claro e definido e suas fronteiras são controversas (12, 124). Como consequência, o desenvolvimento de uma taxonomia se torna confuso e redundante, representando um desafio. Pachet e Cazaly (124) demonstraram que não há um acordo geral sobre taxonomias de gêneros musicais, as quais podem, inclusive, ser influenciadas por referências culturais. Até mesmo termos como rock, jazz, blues e pop não são totalmente claros e definidos, e apresentam divergências entre taxonomias existentes. Em muitas destas, propriedades intrínsecas das músicas não são adotadas, e as categorias podem refletir, por exemplo, interações entre artistas, fatores culturais, convenções históricas, e estratégias de *marketing* (11).

Apesar de ser esperado que uma peça musical seja a referência mais natural para a classificação dos gêneros, tem-se que, por muitas vezes, uma música contém aspectos que podem estar associados a mais de um gênero. Este fato ocorre de forma mais marcante com álbuns e artistas, uma vez que muitos álbuns contém material heterogêneo e que a maioria dos artistas tende a cobrir um amplo conjunto de gêneros durante suas carreiras. Desta forma, é difícil associar um álbum ou artista a um gênero específico. Esta confusão semântica existente entre as taxonomias pode causar redundâncias que, provavelmente são

mais simples de ser resolvidas por usuários, mas dificilmente serão tratadas por sistemas automáticos.

Dentro deste contexto, o problema de classificação automática de gêneros musicais se torna um desafio, e a escolha de formas de representação adequadas e discriminativas, assim como a escolha de métricas adequadas de similaridade configuram etapas essenciais na modelagem de um sistema de classificação. Argumentos que favorecem o contínuo aperfeiçoamento de sistemas de classificação automática de gêneros estão descritos a seguir e estão baseadas em (122).

2.3.1 Razões para uso de gêneros em sistemas de classificação

Pesquisas comprovam que gêneros musicais são categorias importantes e mais utilizadas por usuários no processo de busca de conteúdo musical. Rótulos como gêneros e emoção também são interessantes por fornecerem um vocabulário útil de definições e convenções que pode ser utilizado em discussões sobre categorias musicais. Tais convenções podem inclusive ser observadas diretamente pelo comportamento em sociedade. Por exemplo, fans de heavy metal e rap tendem a ter padrões característicos de vestimenta e fala.

A importância do gênero também é ressaltada através de pesquisas que mostram que estes podem influenciar no gosto de uma música, e que categorização de forma geral é um aspecto importante para a apreciação e cognição da música. Além disso, a classificação de gêneros pode contribuir para pesquisas nas áreas de teoria musical e musicologia, por exemplo, através do estudo da sua evolução ao longo do tempo, ou através da análise de correlação entre aspectos culturais e características musicais.

Observa-se, portanto, a necessidade e utilidade de se automaticamente classificar gêneros musicais. Enquanto que o tempo computacional para treinar e testar estes sistemas em larga escala e as inconsistências e ruídos as anotações de taxonomias existentes configuram desafios nestas tarefas, a categorização manual do grande volume de dados *online* se torna ainda mais inviável.

2.4 MIDI

De forma geral, MIDI (*Musical Instrument Digital Interface*) representa uma especificação que permite controlar e transferir informação musical entre instrumentos musicais e aparelhos eletrônicos (como computadores, sintetizadores, entre outros). Este formato se tornou um protocolo internacional de transmissão de dados em 1982 (37). Mensagens MIDI podem ser comparadas com uma notação de música conhecida como *piano roll*,

Figura 2.4 – As notas musicais, suas frequências e respectivos números MIDI (43). A nota C4 em destaque representa a localização do Dó central no piano.

uma vez que essas mensagens geralmente indicam instruções para controle dos eventos, ao invés de ondas sonoras propriamente ditas (40). Essas instruções podem determinar, por exemplo, o tempo de início e parada de uma nota, qual instrumento deve ser usado, nível de volume, assim por diante. Uma melodia é definida como uma sequência de mensagens MIDI contendo notas. Em (40, 41), encontram-se descrições detalhadas sobre a especificação MIDI e o funcionamento dos dois tipos principais de mensagens: mensagens de canais e mensagens de sistema. Informações adicionais também podem ser encontradas em *Midi Manufactures Association* (42).

Na literatura de recuperação de informação musical, MIDI é conhecido como um formato de representação simbólico (ou de alto nível) baseado em eventos (8). A visualização do formato MIDI pode ser, de uma forma geral, considerada como uma partitura musical digital na qual os instrumentos estão separados em vozes e cada nota é descrita em termos de altura, duração e intensidade. A altura está associada a uma frequência. A Figura 2.4 ilustra as frequências das notas e seus respectivos números MIDI, correspondentes às teclas do piano. Um arquivo MIDI pode ser visualizado em programas de edição de partituras

Figura 2.5 – Representação em pentagrama de um arquivo MIDI

Figura 2.6 – Ilustração de um pentagrama de percussão

que suportem este formato. A Figura 2.5 apresenta parte de um arquivo MIDI.

No arquivo MIDI da Figura 2.5 estão presentes as faixas de percussão, melodia, e vocais. A percussão é a faixa utilizada neste trabalho para representar o ritmo das músicas. Geralmente a percussão é indicada no arquivo MIDI como '*Percussion*' ou '*Drums*'.

Existem certas convenções para representar no pentagrama as notas tocadas por instrumentos de percussão. É possível, inclusive, que a cabeça da nota de valor seja representada por símbolos diferentes da forma tradicional, como ocorre com o chimbau indicado na Figura 2.6. A própria clave de percussão (primeiro símbolo do pentagrama na Figura 2.6) não estabelece uma nota musical de preferência, como acontece na clave de sol. Desta forma, a localização da nota no pentagrama não determina uma altura dentro da escala musical, mas sim qual o instrumento em questão (caixa, prato, bumbo) que está executando a nota^c.

Em percussão, a batida define a pulsação rítmica das notas de valores. As primeiras duas notas de valores circuladas em vermelho na Figura 2.6 (representando os instrumentos bumbo e chimbau) ocorrem simultaneamente na mesma batida.

^c Apesar que notas de instrumentos mais graves como o bumbo são geralmente localizadas na parte inferior do pentagrama; enquanto que notas de instrumentos mais agudos como os pratos são localizadas na parte superior do pentagrama.

2.4.1 Razões para o uso do formato MIDI

Em contrapartida à representação simbólica, a informação musical pode estar digitalmente armazenada através de dados de áudio. Neste caso, a informação musical é obtida pela amostragem da onda sonora. Exemplos de arquivos de áudio conhecidos são: MP3, WAV e AIFF. Atributos extraídos de ondas sonoras que não apresentam uma contextualização musical são geralmente referenciados como atributos de baixo nível, enquanto que aqueles extraídos de representações simbólicas que expressam sentido musical são conhecidos como de alto nível (8).

É conhecido que o padrão MIDI possui algumas desvantagens, geralmente relacionadas com limitações do sintetizador e com a dificuldade de mapear um sofisticado controle de parâmetros para muitos instrumentos (11). Portanto, uma gravação MIDI geralmente possuirá uma qualidade sonora inferior à respectiva gravação em arquivo de áudio.

Entretanto, é possível relatar vantagens que gravações MIDI possuem em relação a arquivos de áudio. MIDI é um formato muito mais compacto. Portanto, o armazenamento é menos custoso e o processamento e análises de arquivos se tornam mais ágeis (11). É simples editar arquivos MIDI, pois as instruções podem ser facilmente identificadas e alteradas.

Por ser um formato simbólico, é possível extrair informações de alto nível específicas. Em outras palavras, é possível conhecer claramente quais elementos musicais estão de fato contribuindo durante uma análise. No contexto de classificação de gêneros musicais, essa vantagem se torna ainda mais importante, pois busca-se identificar, por exemplo, padrões de notas, padrões de ritmos ou características harmônicas que influenciam na discriminação dos gêneros. Essas informações podem ser diretamente extraídas de arquivos MIDI. No entanto, extrair tais informações em arquivos de áudio é, atualmente, custoso e difícil (11).

2.5 Considerações finais

Este capítulo apresentou conceitos musicais e do formato MIDI que são necessários para o entendimento do restante da tese de doutorado. Também foram discutidos aspectos dos gêneros musicais e da dificuldade em se definir rótulos únicos e precisos para as músicas. Uma vez que a taxonomia existente é confusa e redundante, a classificação automática de gêneros musicais se torna um desafio, e pesquisas atuais buscam a generalização da taxonomia usual através de um estudo sistemático dos atributos musicais.

Para tal, é desenvolvido neste trabalho um estudo sistemático dos ritmos no tocante

à caracterização dos gêneros musicais que está fundamentado nos conceitos teóricos apresentados no próximo capítulo.

Capítulo 3

Fundamentação teórica II - métodos de análise

Este capítulo apresenta a fundamentação teórica dos métodos e algoritmos que serão utilizados neste trabalho. Descrições completas de cada um dos assuntos abordados podem ser encontradas nas respectivas referências indicadas.

Na seção 3.1 são descritas técnicas de extração de atributos, importantes em sistemas de classificação, especialmente quando o número de amostras disponíveis é reduzido. Uma vez que arquivos MIDIs são menos comuns que arquivos de áudio, a extração de atributos configura uma tarefa essencial no trabalho. Além disso, a identificação de atributos mais relevantes pode determinar quais sequências de notas são mais relevantes e discriminativas de cada gênero. Uma vez realizada a etapa de extração de atributos, qualquer sistema de classificação pode ser aplicado.

Em termos gerais, classificar significa associar amostras observadas às classes ou categorias de acordo com as propriedades que elas apresentam. As amostras, aqui representadas pelas músicas de diferentes gêneros, são representadas por vetores de atributos. Se exemplos conhecidos de cada classe estão disponíveis, a classificação é dita supervisionada. Na classificação supervisionada, busca-se formular regras de classificação que são primeiramente ajustadas através de exemplos musicais com rótulos conhecidos (conjunto de treinamento), e posteriormente aplicadas em novos exemplos (conjunto de teste). No contexto deste trabalho, a fase de “aprendizagem” consiste em testar os atributos rítmicos no conjunto de treinamento. Na fase de “generalização” ou “aplicação”, vetores de atributos de músicas pertencentes ao conjunto de teste são classificados em um dos gêneros através dos parâmetros estabelecidos durante o treinamento. Com isso, será possível verificar se tais atributos que configuram mecanismos simples do ritmo possuem potencial para a discriminação dos gêneros musicais.

Nesta tese, o classificador Gaussiano através de funções discriminantes foi aplicado como abordagem paramétrica de classificação supervisionada. O classificador Gaussiano

tem como base a Teoria de Decisão Bayesiana e utiliza as probabilidades condicionais de cada classe juntamente com as probabilidades *a priori* para atribuir cada amostra à classe com a máxima probabilidade *a posteriori*. Uma vez que as densidades das classes são adequadamente estimadas, esse método define uma regra de classificação ótima, no sentido que a probabilidade de erro de classificação é mínima (13). O classificador Gaussiano está descrito na seção 3.2.

Na classificação não supervisionada não há um conhecimento prévio sobre as classes, ou seja, o conjunto de treinamento é inexistente. Métodos de agrupamento de dados (do inglês, *clustering*) são amplamente utilizados na classificação não supervisionada, nos quais as amostras são agrupadas de acordo com algum critério de similaridade. O princípio básico é estabelecer os grupos de tal forma que a similaridade entre amostras de um mesmo grupo seja maximizada e a similaridade entre amostras de grupos distintos seja minimizada.

Tipos de agrupamento usualmente utilizados são o particional e o hierárquico. No primeiro caso, um número fixo de grupos é obtido através de uma única partição do espaço de atributos. Por outro lado, no agrupamento hierárquico uma série de partições é realizada de forma progressiva. O agrupamento hierárquico é dito aglomerativo (ou *bottom-up*) quando o algoritmo inicia com N amostras estabelecendo N grupos, e, sucessivamente, combina os grupos até que todas as amostras façam parte de um único grupo. Este método de agrupamento é adotado neste trabalho e está descrito na seção 3.3. De forma análoga, o agrupamento é dito divisivo (ou *top-down*) se, inicialmente, todas as amostras são agrupadas em um único grupo, e, progressivamente, se dividem em grupos menores.

Portanto, métodos de agrupamento de dados inferem de forma não supervisionada as características intrínsecas dos dados. Neste contexto, a detecção de comunidades em redes complexas podem contribuir para um melhor entendimento da relação entre as músicas envolvidas, uma vez que medidas topológicas podem ser inseridas na análise (seção 3.4). Os resultados dos métodos de agrupamento podem ser comparados com os obtidos pela classificação supervisionada, permitindo uma discussão das vantagens e desvantagens em cada uma das taxonomias obtidas.

3.1 Extração de atributos

Na maioria das aplicações de reconhecimento de padrões, a extração de atributos é uma etapa fundamental, uma vez que a relação quantidade de atributos e quantidade de amostras disponíveis influencia diretamente no desempenho do sistema de classificação (7, 13). A quantidade necessária de amostras de treinamento para a classificação supervisionada aumenta conforme a dimensionalidade do problema aumenta (44). Nesse sentido, métodos

de extração de atributos podem ser usados com o objetivo de potencializar esta relação, uma vez que eles ajudam na obtenção de atributos relevantes quando o conjunto de amostras é limitado.

As amostras (ou objetos) são representadas por vetores de atributos dentro de um espaço com d dimensões, sendo d o número total de atributos ou características de cada amostra. Geralmente, o objetivo principal de problemas de discriminação é maximizar a separabilidade entre as classes, isto é, vetores de atributos de diferentes classes devem ocupar regiões compactas e distintas no espaço de atributos. Portanto, atributos adequados e significantes devem ser escolhidos para permitir tal discriminação.

Métodos que criam novos atributos que são transformações ou combinações dos atributos originais são conhecidos como métodos de extração de atributos. O objetivo principal desses métodos é determinar um subespaço adequado de dimensionalidade m , de tal forma que $m < d$. Análise de Componentes Principais (PCA - *Principal Component Analysis*) e Análise Discriminante Linear (LDA - *Linear Discriminant Analysis*) são dois métodos amplamente utilizados para esta extração (7, 13, 45). Ambos aplicam transformações geométricas (rotações) ao espaço original de atributos, gerando novos atributos que são combinações lineares dos originais. PCA é um método não supervisionado que busca uma projeção na qual os novos atributos sejam ortogonais e descorrelacionados. A LDA, por outro lado, é um método supervisionado que busca uma projeção que melhor separe os dados.

A Figura 3.1 ilustra graficamente os princípios básicos dessas duas técnicas. A direção x' obtida pela técnica PCA permite que as duas classes sejam representadas com máxima dispersão. Entretanto, as densidades projetadas através da direção x' se sobrepõem, tornando essas duas classes dificilmente separáveis. Por outro lado, a direção y' obtida pela técnica LDA permite que as classes sejam separáveis. Costuma-se dizer que direções geralmente ideais para representar os atributos nem sempre são também ideais para melhor separá-los, refletindo assim os objetivos diferentes dessas duas abordagens.

Figura 3.1 – Ilustração de duas técnicas amplamente utilizadas na análise de atributos: Análise de componentes principais - PCA (otimizando a representação dos dados), e Análise discriminante linear - LDA (otimizando a separabilidade das classes), adaptado de (46).

Com o intuito de constatar o desempenho de estratégias supervisionadas e não supervi-

sionadas para a discriminação dos gêneros, ambos os métodos PCA e LDA serão adotados neste trabalho para a extração de atributos.

3.1.1 Análise de componentes principais

PCA é um método estatístico de segunda ordem e não supervisionado, relacionado com a transformação de Karhunen-Loève (13). Basicamente, o método computa a expansão dos vetores de atributos nos m autovetores da matriz de covariância associados aos m maiores autovalores. Uma vez que a nova transformação contém os atributos mais significativos, a representação dos dados no subespaço linear minimiza o critério de erro médio quadrático. Assumindo um conjunto de dados multivariado, essa técnica busca remover redundâncias presentes nos atributos originais, o que permite uma redução de dimensionalidade.

Considere um vetor $\mathbf{x}$ com n elementos, representando atributos ou medidas de uma amostra. O primeiro passo consiste em centralizar $\mathbf{x}$, subtraindo-o da respectiva média, de tal forma que $\mathbf{x} \leftarrow \mathbf{x} - E\{\mathbf{x}\}$. Na sequência, $\mathbf{x}$ é linearmente transformado em outro vetor $\mathbf{y}$. Geometricamente, tal procedimento é obtido através de uma rotação do sistema de coordenadas ortogonais, de forma que os elementos em $\mathbf{x}$ sejam não-correlacionados no novo sistema de coordenadas. Portanto, o método PCA maximiza as variâncias das projeções de $\mathbf{x}$ nos novos eixos (componentes). A primeira componente corresponde à maior variância; a segunda componente corresponde à maior variância na direção ortogonal da primeira, e assim por diante. Desta forma, as componentes associadas a pequenas dispersões (dadas pelos respectivos autovalores) podem ser descartadas sem que haja perda significativa de informação dos dados originais.

Informações adicionais sobre o método PCA e suas propriedades estatísticas podem ser encontradas na literatura de reconhecimento de padrões (6, 7, 13, 45–47).

3.1.2 Análise discriminante linear

A técnica LDA pode ser considerada uma generalização da Função Discriminante de Fisher para o caso multivariado (7, 13). É um método supervisionado no qual as projeções obtidas maximizam um critério de separabilidade dos dados. O critério de similaridade busca maximizar a distância média entre as classes e minimizar o espalhamento dentro da classe. Portanto, o objetivo principal desta técnica é que amostras de uma mesma classe sejam as mais similares possíveis, enquanto que amostras de classes distintas sejam as mais discriminativas possíveis quando representadas pelos novos atributos. Em outras palavras, o método LDA estabelece um novo espaço de atributos que maximiza a distância

inter-classes e minimiza a distância intra-classe. As projeções também permitem reduzir a dimensionalidade dos dados e, assim, podem ser utilizadas em problemas de classificação.

Pode ser mostrado que a maximização do critério de separabilidade entre classes conduz à definição de um problema de autovetor generalizado (7, 13). Portanto, é possível obter os autovalores e autovetores da matriz definida por $(S_{intra}^{-1} * S_{inter})$, em que S_{intra} é a matriz de espalhamento intra-classe, e S_{inter} é a matriz de espalhamento inter-classe. Os m autovetores associados aos m autovalores dessa matriz podem ser utilizados para projetar os dados. Porém, o $rank^a$ de $(S_{intra}^{-1} * S_{inter})$ é limitado a $C - 1$, onde C é o número de classes (gêneros). Como consequência, podem existir apenas $C - 1$ autovalores não nulos. Ou seja, o número de novos atributos está condicionado ao número de classes, $m \leq C - 1$. Outra questão importante desse método é que, para problemas de alta dimensionalidade, nos quais o número de amostras disponíveis é menor que a quantidade de atributos, S_{intra} tende a se tornar singular, não admitindo inversa.

Informações adicionais sobre LDA podem ser encontradas em (7, 13, 45, 46).

3.2 Funções discriminantes lineares e quadráticas

Ao assumir a distribuição Gaussiana multivariada sobre os dados, a função densidade de probabilidade é definida por:

$$p(\vec{x}_i | \omega_j) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_j|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\vec{x}_i - \vec{\mu}_j)^T \Sigma_j^{-1} (\vec{x}_i - \vec{\mu}_j) \right\}, \quad (3.1)$$

em que d é a dimensão dos dados, $\vec{x}_i$ é o vetor de padrões da amostra i , e $\vec{\theta}_j = \{\vec{\mu}_j, \Sigma_j\}$ representa o conjunto paramétrico da j -ésima classe ω_j , cujos parâmetros são o vetor média ($\vec{\mu}_j$) e matriz de covariância (Σ_j). Ao adotar os estimadores por máxima verossimilhança, tem-se que:

$$\hat{\vec{\mu}}_j = \frac{1}{N} \sum_{i=1}^N \vec{x}_i, \quad (3.2)$$

$$\hat{\Sigma}_j = \frac{1}{N} \sum_{i=1}^N (\vec{x}_i - \vec{\mu}_i) (\vec{x}_i - \vec{\mu}_i)^T, \quad (3.3)$$

Dentro desse contexto, a classificação pode ser realizada através de funções discriminantes, g_j , atribuindo cada vetor de padrões $\vec{x}_i$ à classe ω_j que possui o valor máximo da função discriminante. Considerando a regra de Bayes, não considerando os termos constantes e utilizando as estimativas dos parâmetros segundo as equações 3.2 e 3.3, uma

^a *Rank* de uma matriz é o número de linhas ou colunas que são linearmente independentes.

regra de decisão pode ser definida da seguinte forma: atribuir a amostra $\vec{x}_i$ à classe ω_j se $g_j > g_p$ para todo $p \neq j$. A função discriminante pode ser calculada como:

$$g_j(\vec{x}_i) = \log(p(\omega_j)) - \frac{1}{2} \log(|\hat{\Sigma}_j|) - \frac{1}{2} (\vec{x}_i - \hat{\mu}_j)^T \hat{\Sigma}_j^{-1} (\vec{x}_i - \hat{\mu}_j). \quad (3.4)$$

Classificar uma amostra i com vetor de padrões $\vec{x}_i$ com base nos valores de $g_j(\vec{x})$, $j = 1, \dots, C$ (C é o número de classes), define um classificador discriminante quadrático (também conhecido como classificador Gaussiano quadrático) (13).

A probabilidade *a priori*, $p(\omega_j)$, pode ser estimada por:

$$p(\omega_j) = \frac{n_j}{\sum_j n_j}, \quad (3.5)$$

em que n_j é o número de amostras da classe ω_j .

Na classificação multivariada, podem ocorrer problemas de singularidade em algumas matrizes de covariância. Este fato geralmente acontece quando a quantidade de amostras disponíveis é insuficiente para a estimação adequada da matriz de covariância amostral. Uma alternativa para minimizar esse problema consiste em estimar uma única matriz de covariância para todas as classes, $\hat{\Sigma} = \hat{\Sigma}_1 = \dots = \hat{\Sigma}_C$. Nesse caso, a função discriminante se torna linear em $\vec{x}_i$, e pode ser simplificada como se segue:

$$g_j(\vec{x}) = \log(p(\omega_j)) - \frac{1}{2} \hat{\mu}_j^T \hat{\Sigma}^{-1} \hat{\mu}_j + \vec{x}^T \hat{\Sigma}^{-1} \hat{\mu}_j \quad (3.6)$$

em que $\hat{\Sigma}$ é a matriz de covariância comum a todas as classes. A regra de decisão se mantém a mesma. Esse procedimento define um classificador discriminante linear (também conhecido como classificador Gaussiano linear) (13).

3.2.1 Coeficiente Kappa

Em tarefas de classificação supervisionada, é necessário definir um critério quantitativo para objetivamente avaliar o desempenho do processo. Os critérios mais utilizados são baseados na estimativa do erro de classificação e na acurácia obtida. Tal desempenho é geralmente medido na classificação das amostras do conjunto de teste através da matriz de confusão. A matriz de confusão é definida da seguinte forma:

$$C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1c} \\ c_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \\ c_{c1} & \dots & & c_{cc} \end{bmatrix}, \quad (3.7)$$

em que cada entrada da matriz c_{ij} representa o número de amostras da classe i classificadas como da classe j . Portanto, os elementos da diagonal indicam o número de classificações corretas. A matriz de confusão tem dimensão $C \times C$, em que C é o número de classes.

O coeficiente Kappa foi originalmente proposto por Cohen (48). No contexto da classificação supervisionada, este coeficiente quantifica a concordância entre as amostras previamente conhecidas (tais amostras são geralmente conhecidas como verdade terrestre, do inglês *ground truth*) e os resultados obtidos pelo classificador. Quanto melhor a acurácia de classificação, maior será o grau de concordância, e, conseqüentemente, maior o valor do coeficiente kappa.

O coeficiente Kappa é uma medida quantitativa bastante utilizada na literatura e simples de ser obtida. Essa medida possui propriedades estatísticas interessantes, tais como de ser assintoticamente normal com expressões bem definidas para estimação das suas variâncias, o que permite, por exemplo, a realização de testes de hipóteses estatísticos e a determinação de intervalos de confiança. Além disso, o coeficiente Kappa pode ser obtido diretamente da matriz de confusão como se segue (49):

$$\hat{k} = \frac{N \sum_{i=1}^C c_{ii} - \sum_{i=1}^C (x_{i+})(x_{+i})}{N^2 - \sum_{i=1}^C (x_{i+})(x_{+i})}, \quad (3.8)$$

em que x_{i+} é a soma dos valores da linha i , x_{+i} é a soma dos valores da coluna i , C é o número de classes, e N é o número total de amostras. A variância do coeficiente Kappa está relacionada com sua acurácia. Uma vez que o coeficiente Kappa é uma medida estatística, em geral, o uso de grandes bancos de amostras melhora seu desempenho ao minimizar sua variância. A variância do coeficiente kappa é estimada como se segue:

$$\hat{\sigma}_k^2 = \frac{1}{N} \left[\frac{\theta_1 (1 - \theta_1)}{(1 - \theta_2)^2} + \frac{2(1 - \theta_1)(2\theta_1\theta_2 - \theta_3)}{(1 - \theta_2)^3} + \frac{(1 - \theta_1)^2 (\theta_4 - 4\theta_2^2)}{(1 - \theta_2)^4} \right], \quad (3.9)$$

em que

$$\theta_1 = \frac{1}{N} \sum_{i=1}^C x_{ii}, \quad (3.10)$$

$$\theta_2 = \frac{1}{N^2} \sum_{i=1}^C (x_{i+})(x_{+i}), \quad (3.11)$$

$$\theta_3 = \frac{1}{N^2} \sum_{i=1}^C x_{ii} ((x_{i+}) + (x_{+i})), \quad (3.12)$$

$$\theta_4 = \frac{1}{N^3} \sum_{i=1}^C \sum_{j=1}^C x_{ij} ((x_{j+}) + (x_{+i}))^2. \quad (3.13)$$

Esse coeficiente indica que, quando $\hat{k} \leq 0$ não há qualquer concordância, e quando $\hat{k} = 1$ a concordância é total. Alguns autores sugerem interpretações relacionadas com os

Tabela 3.1 – Desempenho da classificação de acordo com o coeficiente Kappa

Kappa	Desempenho da classificação
$\hat{k} \leq 0$	Péssimo (<i>Poor</i>)
$0 < \hat{k} \leq 0.2$	Ruim (<i>Slight</i>)
$0.2 < \hat{k} \leq 0.4$	Razoável (<i>Fair</i>)
$0.4 < \hat{k} \leq 0.6$	Moderado (<i>Moderate</i>)
$0.6 < \hat{k} \leq 0.8$	Substancial (<i>Substantial</i>)
$0.8 < \hat{k} \leq 1.0$	Quase perfeito (<i>Almost Perfect</i>)

Figura 3.2 – Dendrograma para uma situação com 8 amostras, adaptado de (7)

valores obtidos do coeficiente Kappa. A Tabela 3.1 apresenta uma possível interpretação, proposta por (50).

3.3 Agrupamento hierárquico aglomerativo

O agrupamento hierárquico aglomerativo progressivamente agrupa as N amostras em C grupos, de acordo com um parâmetro definido, geralmente associado a uma medida de distância ou similaridade entre os vetores de atributos das amostras. No primeiro passo ($k = 1$), as N amostras são atribuídas como N grupos. No passo seguinte ($k = 2$), ocorre uma partição de $N - 1$ grupos, em seguida uma partição de $N - 2$ grupos e assim por diante. No último passo ($k = N$), todas as amostras fazem parte de um único grupo. Portanto, esta sequência progressivamente agrupa amostras que são mais similares entre si antes de considerar amostras menos similares. É possível dizer que, no k -ésimo passo, $C = N - k + 1$ (C é o número de grupos).

Para mostrar como as amostras são agrupadas, o agrupamento hierárquico pode ser representado por uma árvore correspondente, conhecida como dendrograma. A Figura 3.2

ilustra um dendrograma representando o resultado do agrupamento para um problema de 8 amostras. A medida de similaridade entre os grupos pode ser observada no eixo vertical. Diferentes números de grupos podem ser obtidos ao se cortar horizontalmente o dendrograma em diferentes valores de similaridade. O procedimento de agrupamento hierárquico requer três principais parâmetros (13, 51). O primeiro deles define uma medida para quantificar a similaridade (distância) ou dissimilaridade entre cada par de objetos ^b. Exemplos de medidas de distância são: Euclidiana, cosseno, *cityblock*, *mahalanobis*, entre outras (seção 3.3.1). O segundo parâmetro determina o critério de decisão para juntar e estabelecer a distância entre dois grupos de objetos (etapa conhecida no inglês como *linkage criteria*). O terceiro parâmetro define o número de grupos desejados, que está relacionado com o local em que o dendrograma será cortado, como ilustrado por C na Figura 3.2.

A seguir são apresentados métricas de distâncias e critérios de *linkage*.

3.3.1 Métricas de distâncias

Um conceito fundamental na utilização das técnicas de agrupamento de dados é na escolha de um critério que meça a distância entre dois objetos, ou que quantifique o quanto eles são parecidos. Esta medida é geralmente denominada coeficiente de parecença. Cabe observar que tecnicamente pode-se dividir o critério de medição da distância entre dois objetos em duas categorias: medidas de similaridades e de dissimilaridade (51). Em relação a segunda categoria, quanto maior a distância entre dois objetos, menos parecidos (mais dissimilares) eles serão. Coeficiente de correlação é um exemplo de medida de similaridade, enquanto que distância euclidiana é um exemplo de dissimilaridade (51).

A Tabela 3.2 apresenta algumas métricas de distâncias comuns em tarefas de agrupamento de dados (51). Para cada uma delas, considere $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_d)$ e $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_d)$ dois objetos em um espaço de dimensão d .

3.3.2 Critérios de *linkage*

É também necessário determinar uma medida de distância entre um objeto e um grupo, e entre dois grupos de objetos (etapa conhecida do inglês como *linkage criteria*). Alguns dos critérios mais utilizados estão apresentados na Tabela 3.3 (6, 51–54). Para os critérios apresentados, considere uma partição em que $C_1 = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r\}$ e $C_2 = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_s\}$ são dois grupos de objetos de tamanhos r e s , respectivamente.

^b Geralmente, similaridade e distância são referenciados como conceitos recíprocos (51)

Tabela 3.2 – Métricas de distância e suas respectivas definições

Métrica	Definição	Descrição
Distância Euclidiana	$D(\mathbf{x}, \mathbf{y}) = \left[\sum_{j=1}^d (x_j - y_j)^2 \right]^{\frac{1}{2}}$	Expressa a distância geométrica Euclidiana entre dois objetos em um espaço d -dimensional.
Distância de Manhattan ou <i>City-block</i>	$D(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^d x_j - y_j $	É definida como a distância entre dois objetos no espaço Euclidiano, com um sistema de coordenadas cartesianas fixo, como sendo a soma dos comprimentos das projeções dos segmentos de reta entre os pontos dos eixos das coordenadas. É estabelecida como a soma das distâncias para todos os atributos.
Distância máxima	$D(\mathbf{x}, \mathbf{y}) = \max_{i \leq k \leq d} x_j - y_j $	Também conhecida como distância “sup”, indica o valor máximo das distâncias dos atributos
Distância de Minkowski	$D(\mathbf{x}, \mathbf{y}) = \left(\sum_{j=1}^d x_j - y_j ^r \right)^{\frac{1}{r}}, r \geq 1$	É uma forma genérica para calcular a distância entre dois pontos no espaço d -dimensional (r indica a ordem da distância). As três primeiras distâncias apresentadas são casos particulares da distância de Minkowski. Ao considerar $r = 2, 1$ e ∞ , tem-se a distância Euclidiana, a distância de Manhattan, e a distância máxima, respectivamente.
Distância de Mahalanobis	$D(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y})^T \Sigma^{-1} (\mathbf{x} - \mathbf{y})}, \Sigma \text{ é a matriz de covariância dos dados}$	Essa distância incorpora a correlação entre objetos e os padroniza de modo que suas médias sejam 0 (zero) e desvio padrão 1(um).
Distância do Cosseno	$D_{ij} = 1 - \frac{\sum_{i=1}^d x_i y_j}{\sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_j^2}}$	Medida de semelhança entre dois vetores através do ângulo entre eles.

Tabela 3.3 – Critérios de *linkage* e suas respectivas definições

Critério	Definição	Descrição
<i>Linkage</i> simples	$D(C_1, C_2) = \min_{\mathbf{x} \in C_1, \mathbf{y} \in C_2} d(\mathbf{x}, \mathbf{y})$	Também conhecido como “distância do vizinho mais próximo”, “distância mínima”, o critério de <i>Linkage</i> simples define a proximidade de dois grupos como a distância mínima entre dois objetos pertencentes a grupos distintos.
<i>Linkage</i> completo	$D(C_1, C_2) = \max_{\mathbf{x} \in C_1, \mathbf{y} \in C_2} d(\mathbf{x}, \mathbf{y})$	Define a distância de dois grupos como a distância máxima entre dois objetos de grupos distintos.
<i>Linkage</i> pela média	$D(C_1, C_2) = \frac{1}{rs} \sum_{\mathbf{x} \in C_1, \mathbf{y} \in C_2} d(\mathbf{x}, \mathbf{y}),$	Também conhecido como UPGMA (do inglês, <i>unweighted pair group method using arithmetic averages</i>), este critério define a distância entre dois grupos como a média das distâncias de todos os possíveis pares de objetos de grupos distintos.
<i>Linkage</i> por centróide	$D(C_1, C_2) = \ \mathbf{m}_1 - \mathbf{m}_2\ ^2$, em que $\mathbf{m}_1 = \frac{1}{r} \sum_{\mathbf{x} \in C_1} \mathbf{x}$	Também conhecido como UPGMC (do inglês, <i>unweighted pair group method using centroids</i>), este critério define a distância entre dois grupos como a distância no espaço euclidiano de seus respectivos centróides.
<i>Linkage</i> por Ward's	$D = \sum_{k=1}^K \sum_{\mathbf{x} \in C_k} \ \mathbf{x} - \mathbf{m}_k\ ^2$	Também conhecido como “método de variância mínima”, o método de <i>linkage</i> por <i>Ward's</i> tem como objetivo minimizar o critério de erro quadrático em termos da variância intra-classe. A cada passo, todas as possíveis uniões entre pares de grupos são consideradas e os grupos que apresentam a mínima dispersão de informação são selecionados para o agrupamento. Portanto, os objetos são agrupados de forma a garantir a menor dispersão intra-classe. Esse método é usualmente referenciado como um dos mais eficazes no agrupamento hierárquico (55–57).

3.4 Representação de sistemas por grafos

Esta seção apresenta conceitos iniciais sobre grafos, árvores geradoras mínimas e redes complexas, uma vez que tais representações são abordadas no trabalho. A seção 3.4 expõe os conceitos básicos de grafos. A seção 3.4.1 discute conceitos iniciais sobre árvores geradoras mínimas, enquanto que as seções 3.4.2 e 3.4.3 descrevem, respectivamente, modelos de redes complexas e métodos para detecção de comunidades.

Grafos direcionados e não-direcionados

Cada elemento de uma rede é representado por um vértice e as conexões entre os elementos são representadas pelas arestas. Basicamente, existem quatro tipos de redes: redes direcionadas (referenciadas como dígrafos) com pesos e sem pesos, e redes não-direcionadas (referenciadas como grafos) com pesos ou sem pesos. Para transformar um dígrafo em um grafo utiliza-se a operação de simetria; e transformação de grafos ou dígrafos com pesos para grafos ou dígrafos sem pesos é feita através da operação de limiarização.

Os elementos que definem um grafo direcionado com pesos, representado pela letra G , são (59, 60):

- Vértices (ou nós). Cada vértice é representado por um número inteiro $i = 1, 2, \dots, N$; $V(G)$ é o conjunto de vértices do grafo G e N indica o número de vértices ($|V(G)|$).
- Arestas (ou *links*). Cada aresta relaciona um par (i, j) que representa uma conexão do vértice i para o vértice j . Da mesma forma, $E(G)$ é o conjunto de arestas do grafo G e M é o número de arestas ($|E(G)|$).
- Mapeamento $\omega : E(G) \rightarrow R$. Cada aresta (i, j) possui um peso $\omega(i, j)$ associado a ela.

Os grafos não-direcionados (com pesos ou sem pesos) não possuem orientação em suas arestas, de forma que uma aresta (i, j) em $E(G)$ indica uma conexão do vértice i para o vértice j e do vértice j para o vértice i . Nesse caso, vértices i e j são considerados vértices adjacentes. Um grafo é dito ser completo se os N vértices são mutuamente adjacentes.

Grafos ou dígrafos com pesos são representados através de uma matriz de pesos W , de maneira que cada elemento da matriz, w_{ij} , associa um peso para uma conexão do vértice i para o vértice j ($w_{ij} = \omega(i, j)$).

Conceitos para grafos direcionados e não direcionados importantes para a compreensão do trabalho estão apresentados a seguir (59, 60):

Adjacência: A matriz de adjacência A de um grafo contém N linhas e N colunas e tem a propriedade de ser simétrica, ou seja, para todo vértice i e j , $A(i, j) = A(j, i)$. Portanto, dois vértices i e j são adjacentes ou vizinhos se $a_{ij} = 1$. Para dígrafos, utiliza-se conceitos de predecessores e sucessores. Se $a_{ij} \neq 0$, i é predecessor de j e j é sucessor de i . Predecessores e sucessores como vértices adjacentes.

Incidência: A matriz de incidência I de um grafo contém N linhas e M colunas, de forma que $I(i, j) = 1$ se a aresta e_j é incidente ao vértice v_i .

Vizinhança: Representada por $v(i)$, determina o conjunto de vértices vizinhos ao vértice i .

Grau de um vértice: Representado por k_i , determina o número de arestas conectadas ao vértice i , $k_i = \sum_j a_{ij} = \sum_j a_{ji}$. Para dígrafos, calcula-se o grau de entrada, k_i^{in} , como o número de arestas que chegam ao vértice i ; e grau de saída, k_i^{out} , como o número de arestas que partem do vértice i . Desta forma, $k_i^{in} = \sum_j a_{ji}$, e $k_i^{out} = \sum_j a_{ij}$. Portanto, o grau total é definido por $k_i = k_i^{in} + k_i^{out}$.

Grau médio: Média de k_i sobre todos os vértices do grafo: $\langle k \rangle = \frac{1}{N} \sum_i k_i = \frac{1}{N} \sum_{ij} a_{ij}$.

Força: Para dígrafos, a força de um vértice é definida como a soma total dos pesos associados a este vértice ($s_i^{in} + s_i^{out}$), ou seja, a soma dos pesos das respectivas arestas de entrada, $s_i^{in} = \sum_j w_{ji}$; e a soma dos pesos das respectivas arestas de saída, $s_i^{out} = \sum_j w_{ij}$.

Coefficiente de aglomeração: O coeficiente de aglomeração determina a presença de ciclos de ordem três. Essa medida traz informação de como os vizinhos de um vértice estão conectados uns aos outros^c. Portanto, tal coeficiente reflete a conectividade e estrutura cíclica dos dígrafos, e pode ser calculado como se segue: seja m_i o número de vizinhos do vértice i , e l_i o número de conexões entre os vizinhos do vértice i ; o coeficiente de aglomeração é obtido como $cc(i) = l_i / m_i(m_i - 1)$. Informações sobre essa medida e seus efeitos em diversas aplicações de grafos podem ser encontradas em (60).

Caminhada: Sequência de m arestas $(i, k_1), (k_1, k_2), \dots, (k_{m-1}, j)$ pelas quais dois vértices não adjacentes i e j estão conectados. O tamanho da caminhada é definido m . Dois vértices estão conectados se há pelo menos uma caminhada entre eles.

Caminhada aleatória ponderada: Caminhada com ponderação na escolha das arestas. Um vértice inicial é selecionado; o próximo vértice é randomicamente selecionado entre todos os vértices de saída do vértice inicial. Porém, essa escolha é ponderada pela probabilidade das arestas, ou seja, a probabilidade de escolha de um vértice é proporcional ao peso da aresta de saída do vértice anterior para ele. O procedimento continua até um número pré-determinado de passos ser atingido.

Loop ou ciclos: Ocorre quando a caminhada começa e termina com o mesmo vértice i , desde que cada vértice k_n seja visitado apenas uma vez.

Caminho: Ocorre quando todos os vértices da caminhada são diferentes. Para grafos não-direcionados e sem pesos, o número de vértices necessários para conectar o vértice i ao vértice j é conhecido como tamanho do caminho. O caminho mais curto define a quantidade mínima de vértices necessária para conectar os vértices i e j .

^c Para grafos não-direcionados e sem pesos, essa medida é conhecida como transitividade.

O tamanho desse caminho é representado por d_{ij} (para grafos com pesos, pode ser necessário levar em consideração os pesos das conexões). Quando não há caminhos disponíveis para conectar os vértices i e j , é dito que $d_{ij} = \infty$. Para grafos orientados, tem-se que, em geral, $d_{ij} \neq d_{ji}$, uma vez que o caminho para conectar o vértice i ao vértice j geralmente é diferente do caminho para conectar o vértice j ao vértice i .

3.4.1 Árvores geradoras mínimas

Dado um grafo $G = (V, E)$, em que cada aresta que liga dois vértices u e v possui um peso associado de forma que $(u, v) \in E = w(u, v)$, uma árvore geradora mínima (AGM) é um subconjunto de arestas $T \subseteq E$, que estabelece um subgrafo conexo e acíclico no qual o peso total $w(T) = \sum_{(u,v) \in T} w(u, v)$ seja minimizado.

Portanto, uma AGM é um grafo acíclico conectado de pesos mínimos no qual há apenas um único caminho do vértice u para qualquer vértice v . Além disso, de acordo com suas propriedades, se o vértice u está conectado aos vértices v e k através de uma medida de distância representada nos pesos das arestas, então ambos vértices v e k são os dois vértices mais similares ao vértice u (considerando a árvore e não o grafo original). Desta forma, visitar os vizinhos mais próximos numa AGM irá produzir uma sequência de vértices na qual o próximo elemento é o mais similar ao anterior na árvore. Este aspecto será explorado neste trabalho para a geração automática de listas de reprodução.

A Figura 3.3 apresenta um exemplo de um grafo conectado e uma AGM.

Figura 3.3 – Um exemplo de grafo conectado em (a), e uma árvore geradora mínima em (b), adaptado de (61)

3.4.2 Redes complexas

O estudo das redes complexas foi motivado por pesquisas iniciais envolvendo percolação e grafos randômicos (62–67). Verificou-se que os grafos randômicos dificilmente apresentam características necessárias para uma melhor representação de problemas complexos reais.

Modelos de redes complexas com propriedades importantes têm sido propostos na busca de um melhor entendimento para as redes reais. Exemplos destas redes estão apresentados na Figura 3.4. Trabalhos importantes como os de Erdős e Rényi (65–67) sobre redes complexas aleatórias (Figura 3.4a), de Watts e Strogatz (68), sobre redes do pequeno mundo (Figura 3.4b), de Barabasi e Albert (69) sobre redes livres de escala (Figura 3.4c), e de Girvan e Newman (70) sobre o estudo de estruturas comunitárias das redes complexas formam a base para os estudos atuais na área.

Figura 3.4 – Exemplo de (a) rede randômica (b) rede de pequeno mundo (c) rede livre de escala (59)

As redes complexas têm se mostrado úteis para o entendimento de muitos aspectos presentes na vida do ser humano. Elas conseguem representar de forma adequada estruturas de interesse ou fenômenos naturais como, por exemplo, a disposição das estradas de um país, o sistema aéreo mundial, participação de atores em um filme, as relações sociais de uma pessoa e até mesmo a Internet (71).

3.4.3 Detecção de comunidades em redes

A detecção de comunidades em redes complexas é um tema de pesquisa importante em muitas áreas como medicina, biologia, sociologia, economia, transporte (72), nas quais a interação entre elementos pode ser naturalmente representada por grafos. A obtenção de estruturas de alta-ordem pode estabelecer evidências sobre a organização e funcionamento dos elementos envolvidos, como por exemplo, a identificação funcional de proteínas ou grupos de pessoas com interesses em comuns. Em redes de música, as comunidades podem refletir a dinâmica da interação entre artistas ou compositores. De forma geral, comunidades indicam nós com grande conectividade intra-grupo, e pequena conectividade entre-grupo. O coeficiente de aglomeração apresentado anteriormente pode ser útil para determinar a tendência de grupos de nós em formar uma comunidade (58).

Muitas das técnicas tradicionais para agrupamento de dados podem ser utilizadas para a detecção de comunidades em redes (58). Algoritmos que desempenham agrupamentos

Figura 3.5 – Dendrograma ilustrando algoritmos que realizam agrupamento dos vértices para a detecção de comunidades. Os círculos na parte inferior da figura representam os vértices da rede. Subindo na árvore hierárquica, tem-se a formação de comunidades maiores. A linha tracejada indica as quatro comunidades encontradas para este nível do dendrograma.

hierárquicos dos vértices são clássicos na literatura de redes complexas e se baseiam em similaridades ou dissimilaridades entre pares de vértices na obtenção dos grupos (análogo ao dendrograma da Figura 3.5). Portanto, uma medida de similaridade entre vértices precisa ser previamente definida. Tal medida pode ser estrutural, levando em consideração a topologia da rede; ou pode ser baseada em atributos, levando em consideração as propriedades dos vértices. Uma medida de similaridade, s_{ij} , geralmente pode ser transformada em uma medida de dissimilaridade d_{ij} , da seguinte forma: $d_{ij} = k - s_{ij}$, para alguma constante k . Diferentes algoritmos e medidas de dissimilaridade resultam em diferentes estruturas de comunidades.

Particionamento através de medidas de modularidade

Ainda no contexto de partição hierárquica, os algoritmos propostos por Girvan e Newman (70, 73) utilizam índices de centralidade dos nós para detectar comunidades em redes a partir de uma medida conhecida como “vértices com centralidade de intermediação” (do inglês, *betweenness centrality*), que quantifica a influência de um vértice na rede e é definida como o número de caminhos curtos que passam por este vértice. Girvan e Newman generalizaram esta medida para o conceito de “arestas com centralidade de intermediação” (do inglês, *edge centrality*), a qual estabelece o número de caminhos curtos entre vértices que passam por uma determinada aresta.

É esperado que redes com estruturas de comunidades definidas contenham poucas arestas entre grupos distintos. Portanto, caminhos curtos entre nós de diferentes comunidades necessariamente precisam passar por estas arestas (Figura 3.6), as quais terão valores grandes de centralidade de intermediação e irão revelar a estrutura de comunidade da rede.

Figura 3.6 – Uma rede com três estruturas de comunidades (vértices em vermelho, azul e verde) de acordo com as abordagens de Girvan e Newman. As arestas destacadas na figura indicam as arestas com valores maiores de centralidade de intermediação, as quais oferecem caminhos entre as comunidades, adaptado de (73).

O algoritmo é executado através de uma abordagem divisiva: inicia-se com uma rede totalmente conectada e a cada passo remove-se uma aresta com valor alto de centralidade. Portanto, este algoritmo pode ser computacionalmente custoso para grandes redes. Newman and Girvan posteriormente propuseram medidas alternativas para quantificar a centralidade de intermediação das arestas (73).

Comunidades bem definidas devem surgir a medida que muitas arestas são excluídas. Entretanto, é esperado que os algoritmos descritos acima gerem progressivas partições da rede, que podem ser representadas por um dendrograma. Girvan and Newman (70) definiram uma medida nomeada “modularidade”, para atestar quais conjuntos de comunidades mais fielmente descrevem a estrutura da rede antes de estabelecer a partição final, de forma que a partição que proporciona o maior valor de modularidade é escolhida.

Dada a divisão de uma rede com um conjunto de vértices V em c comunidades, $V_1, V_2, \dots, V_c$, de forma que $\cup_{i=1}^c V_i = V$, o valor de modularidade Q desta partição é definido como:

$$Q = \sum_{i=1}^c \left[\frac{A(V_i, V_i)}{A(V, V)} - \left(\frac{A(V_i, V)}{A(V, V)} \right)^2 \right] \quad (3.14)$$

em que $A(S, T)$ corresponde ao número de arestas entre duas comunidades S e T , estabelecido como:

$$A(S, T) = \sum_{i \in S, j \in T} A_{ij} \quad (3.15)$$

O primeiro termo, $\frac{A(V_i, V_i)}{A(V, V)}$, estabelece a fração de arestas que conecta vértices dentro da comunidade V_i ; enquanto que o segundo termo, $\frac{A(V_i, V)}{A(V, V)}$, indica a fração de arestas que se conectam com vértices na comunidade V_i .

A equação 3.14 pode ser simplificada como se segue:

$$Q = \frac{1}{2m} \sum_{i=1}^n \sum_{j=1}^n \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \quad (3.16)$$

em que m é o número de arestas da rede, k_i é o grau do vértice i , c_i é a comunidade que o vértice i está inserido, e $\delta(c_i, c_j) = 1$ se os vértices i e j fazem parte da mesma comunidade.

A presença de comunidades bem definidas é evidenciada quando $Q \cong 1$, ao passo que $Q \leq 0$ indica que a estrutura obtida não é melhor que uma estrutura obtida ao acaso. Girvan e Newman utilizaram esta medida para determinar quando parar o particionamento da rede. Tal abordagem pode ser computacionalmente custosa, e otimizações foram posteriormente implementadas (74, 75). Clauset (76) introduz o conceito de “modularidade local”, para a identificação de estruturas locais de comunidade.

Particionamento espectral

Análises espectrais têm sido bastante exploradas para a caracterização de grafos e redes complexas, mostrando-se uma alternativa em relação às métricas baseadas na topologia dos grafos que, geralmente, são correlacionadas e dependentes (77). Em termos matemáticos, quando um grafo é analisado no domínio espectral, constitui-se uma representação em termos de componentes ortogonais, o que significa que a informação está, de alguma forma, não correlacionada. Neste contexto, uma análise dos autovetores e autovalores da matriz de adjacência ou da matriz Laplaciana pode identificar características que não estariam disponíveis em métricas definidas no domínio da topologia. Algoritmos para particionamento espectral em redes são geralmente baseados na decomposição de valores próprios da matriz Laplaciana.

Considere $\mathbf{A}$ e $\mathbf{B}$ a matriz de adjacência e incidência de um grafo $G = (V, E)$, em que V é o conjunto de vértices e E é o conjunto de arestas. A matriz Laplaciana, $\mathbf{Q}$, é dada por:

$$\mathbf{Q} = \mathbf{B}\mathbf{B}^T = \Delta - \mathbf{A} \quad (3.17)$$

em que Δ é a matriz diagonal dos graus de V . Na matriz diagonal dos graus, $v(i, i)$ é a soma dos graus do vértice i ; e $v(i, j) = 0$, para todo $i \neq j$.

O segundo menor autovalor da matriz Laplaciana representa a conectividade algébrica do grafo^d (78). Pode ser mostrado que o autovetor associado a esse autovalor, conhecido como vetor de Fiedler (26), está diretamente relacionado com a conectividade do grafo. Na prática, os sinais dos elementos do vetor de Fiedler podem ser usados para particionar o grafo em duas regiões. Esse processo pode ser visto como uma quantização para dígitos

^d A conectividade algébrica de um grafo G é um atributo espectral do mesmo, dado pelo segundo menor autovalor da matriz Laplaciana. O valor deste autovalor é maior que 0 se G é um grafo conexo. É utilizado, por exemplo, para medir o quão conectado um grafo é, e para sincronização em modelos de redes complexas.

binários (ou zero ou um). Entretanto, propõe-se quantizar os coeficientes do vetor de Fiedler em C valores, onde C representa o número desejado de grupos. Desta forma, o grafo é essencialmente particionado em C subgrafos ou comunidades, o que é equivalente a encontrar $C - 1$ vales no histograma que representa a distribuição de valores dos coeficientes do autovetor. Para o estudo realizado nesta tese, os limiares foram encontrados através de uma inspeção visual do histograma^e. Uma análise matemática mais detalhada sobre os autovetores da matriz Laplaciana e suas propriedades pode ser encontrada em (77).

3.5 Considerações finais

Este capítulo apresentou a fundamentação teórica de métodos de análise utilizados no trabalho. Os capítulos a seguir descrevem a forma com que estes métodos serão aplicados no contexto de síntese de novas sequências, classificação automática de gêneros musicais, e geração de listas de reprodução.

Para todas estas investigações os atributos do ritmo são comuns e refletem a dinâmica das sequências temporais das notas de valores. A metodologia para a obtenção destes atributos está descrita no próximo capítulo.

^e Futuramente, pretende-se usar métodos da literatura de processamento de imagens para a estimação automática destes limiares (79).

Capítulo 4

Materiais e métodos

Este capítulo apresenta os materiais e métodos adotados neste trabalho para identificação e modelagem dos padrões rítmicos de cada música do banco de dados. O formato MIDI foi utilizado para representação da informação musical. A edição e a análise dos arquivos MIDI foi realizada com auxílio do *software* para composição e notação musical *Sibelius*^a.

A partir de sites da Internet, arquivos MIDI de músicas de diferentes gêneros são selecionados. Para cada arquivo, a faixa da percussão é selecionada para análise e representada por um vetor sequencial de notas de valores (seção 4.1). A dinâmica das notas de valores presentes na percussão é modelada através de cadeias de Markov de primeira e segunda ordens (seção 4.2), e posteriormente representada através de dígrafos (seção 4.3). Também adotou-se a modelagem das notas através da concatenação das duas cadeias de Markov. Estas três perspectivas quanto à relação temporal das notas de valores estabelecem as características analisadas de cada música.

A Figura 4.1 apresenta um diagrama dos passos principais do desenvolvimento do trabalho, informando o capítulo da tese que descreve cada uma das estratégias de análise a partir dos padrões rítmicos. A notação matemática associada aos tipos de dados também está indicada, com o intuito de facilitar a localização das mesmas com cada etapa do trabalho no decorrer do texto.

4.1 Obtenção do vetor de notas de valores

A leitura das informações dos eventos musicais presentes na percussão foi realizada com o uso do MIDI *toolbox* para o Matlab (82). Este *toolbox* é gratuito e contém funções para analisar e visualizar arquivos MIDI no ambiente Matlab. Um arquivo MIDI é referenciado nesse *toolbox* através de uma matriz, na qual as linhas indicam os eventos individualmente;

^a A licença para esse *software* foi obtida através da reserva técnica do projeto temático FAPESP 2005/00587-5

Figura 4.1 – Visão geral dos passos principais para o desenvolvimento do trabalho

e colunas expressam informações deste evento, tais como a batida na qual o evento ocorreu, duração relativa, canal MIDI, altura MIDI (nota) e velocidade.

A Figura 4.2 apresenta a primeira parte da faixa de percussão do arquivo MIDI apresentado anteriormente (seção 2.4). A duração relativa é representada pelas notas de valores através da proporção de duração entre as notas (seção 2.1). A descrição das notas nas batidas correspondentes ao trecho circulado de vermelho estão indicadas na Tabela 4.1.

Como a instrumentação não está sendo adotada, se dois ou mais eventos ocorrem na mesma batida, a mediana das notas de valores é escolhida. Em termos musicais isto significa que, quando a quantidade de instrumentos na mesma batida for maior ou igual a três, o evento escolhido tenderá ser proveniente de instrumentos médios dentre os ocorridos^b. Se o número de eventos é par, a mediana pode resultar em um valor que

^b Por exemplo, eventos do chimbau (instrumento agudo) e do bumbo (instrumento grave) não serão escolhidos nestes casos

Figura 4.2 – Parte de uma faixa de percussão. A duração relativa e as batidas correspondentes aos eventos do trecho em destaque estão expressas na Tabela 4.1.

Tabela 4.1 – Disposição das notas de valores e batidas correspondentes ao trecho em destaque na faixa de percussão da Figura 4.2

Batida	8	9			9.5		10		10.25	10.5		11		11.5	11.75
Nota	1	0.5	0.5	0.5	0.5	0.5	0.25	0.5	0.25	0.5	0.5	0.5	1	0.25	0.25
Mediana	1	0.5			0.5		0.25		0.25	0.5		0.5		0.25	0.25

não corresponde a alguma das notas de valores utilizada. Nestes casos, a mediana será expressada pela nota de valor da mediana inferior. Por exemplo, a nota de valor selecionada para a batida 10 é 0.25, pois corresponde à mediana inferior das notas de valores $\{0.25, 0.5\}$ que ocorrem nesta batida. Considerando a mediana dos valores nos casos em que diferentes instrumentos ocorrem na mesma batida, o vetor final das notas de valores, $\mathbf{V}^{\text{dur}}$, deste trecho em destaque será: $[1 \ 0.5 \ 0.5 \ 0.25 \ 0.25 \ 0.5 \ 0.5 \ 0.25 \ 0.25]$. Para cada música este procedimento é realizado e um vetor $\mathbf{V}^{\text{dur}}$ é obtido a partir da faixa de percussão completa.

4.2 Obtenção dos atributos

Uma vez que o vetor $\mathbf{V}^{\text{dur}}$ das notas de valores está disponível, as cadeias de Markov são utilizadas para a modelagem do ritmo. Com o uso destas cadeias de Markov é possível determinar a probabilidade de ocorrência de eventos futuros dada a ocorrência de um ou mais eventos passados (2). A ordem da cadeia é indicada pela quantidade de eventos passados que são considerados. Uma cadeia de Markov de primeira ordem leva em análise apenas o predecessor de um evento. Cadeias de Markov de segunda ordem levam em consideração dois eventos passados e assim por diante. Geralmente, uma cadeia de Markov de ordem n é representada por uma matriz de transição de estados de $n + 1$ dimensões. Esta matriz fornece a probabilidade de ocorrência de um evento, dados os n estados anteriores. Para cada música, a sequência temporal na qual as notas de valores do vetor

$\mathbf{V}^{\text{dur}}$ ocorrem é observada e modelada através de uma matriz de transição de primeira ordem, $\mathbf{T}^{\text{P}}$, e de segunda ordem, $\mathbf{T}^{\text{s}}$.

A Figura 4.3a apresenta o vetor $\mathbf{V}^{\text{dur}}$ das notas de valores conforme estabelecido anteriormente na Tabela 4.1. A Figura 4.3c e Figura 4.3b apresentam, respectivamente, as tabelas de transição $\mathbf{T}^{\text{P}}$ e $\mathbf{T}^{\text{s}}$ correspondente à sequência das notas de valores da Figura 4.3a. Os eventos passados, $(n-1)$, estão listados verticalmente, enquanto que os eventos correntes, (n) , estão listados horizontalmente. Para encontrar, por exemplo a probabilidade de uma semicolcheia ocorrer após uma colcheia, identifica-se a colcheia na primeira coluna (neste caso a colcheia é o evento passado), e então move-se até o encontro da coluna correspondente à semicolcheia. Esta entrada está em destaque na Tabela 4.3c e tem valor 0.5, ou seja, de acordo com a sequência de notas observada na Figura 4.3a, em metade das vezes a colcheia antecedeu uma colcheia, e na outra metade das vezes, a colcheia antecedeu uma semicolcheia. De forma similar, a probabilidade de ocorrer uma semicolcheia após uma sequência de duas colcheias (entrada $\mathbf{T}^{\text{s}}(2, 3)$ em destaque na Tabela 4.3b) é 0,5, uma vez que a semicolcheia foi o único evento ocorrido após os eventos passados colcheia $(n-2)$ e colcheia $(n-1)$ de acordo com a sequência de notas da Tabela 4.3a.

Figura 4.3 – Matriz de transição de (c) primeira e (b) segunda ordem da sequência de notas em (a).

Cadeias de Markov de primeira ordem tendem a definir as relações subsequentes entre notas que são mais significativas. Por outro lado, cadeias de Markov de segunda ordem indicam a probabilidade de ocorrência de uma sequência de três notas consecutivas, capturando assim informações adicionais quanto às dinâmicas e dependências das notas de valores por contemplar um contexto estendido. Busca-se conferir a contribuição de cada uma destas cadeias para a discriminação dos gêneros quando utilizadas de forma separada, e quando utilizadas de forma conjunta através da concatenação das duas informações.

Uma vez que as matrizes de transição $\mathbf{T}^{\text{P}}$ e $\mathbf{T}^{\text{s}}$ estão disponíveis para a música m , os

vetores de atributos (ou padrões), $\mathbf{V}^m$ e $\mathbf{B}^m$, desta música são obtidos com a ordenação sequencial das colunas das matrizes de transição $\mathbf{T}_m^p$ ou $\mathbf{T}_m^s$, respectivamente.

Vale mencionar que, antes de estabelecer o número de eventos das cadeias de Markov dos vetores $\mathbf{V}^{dur}$ de todas as músicas do banco de dados, uma filtragem foi realizada com o objetivo de excluir notas que raramente ocorriam^c. Essa verificação foi necessária para garantir que os vetores de padrões tenham a mesma quantidade de atributos, e para permitir uma redução de dimensionalidade ao evitar notas que não contribuam significativamente para a análise. Após a filtragem, foram mantidas 18 notas de valores para a modelagem da cadeia de Markov de primeira ordem, resultando em uma matriz de transição, $\mathbf{T}^p$, com 18 linhas e 18 colunas. Para a modelagem da cadeia de Markov de segunda ordem são adotados 167 pares de notas de valores, resultando em uma matriz de transição $\mathbf{T}^s$ contém 167 linhas (pares de notas) e 18 colunas.

Desta forma, o vetor de atributos $\mathbf{V}^m$ da música m tem componentes $\{v_i^m, i = 1, 2, \dots, d\}$, em que $d = 324$ (18x18), enquanto que o vetor de atributos $\mathbf{B}^m$ da música m tem componentes $\{v_i^m, i = 1, 2, \dots, d\}$, em que $d = 3006$ (167x18).

Com a concatenação das duas cadeias, cada música também pode estar representada por um terceiro vetor de atributos, $\mathbf{VB}^m$, que contempla as probabilidades condicionais da cadeia de Markov de primeira ordem provenientes do vetor $\mathbf{V}^m$, e os 3006 de segunda ordem provenientes do vetor $\mathbf{B}^m$. Portanto, o vetor de padrões $\mathbf{VB}^m$ de cada música possui 3330 elementos.

4.3 Representação em dígrafos

As matrizes de transição $\mathbf{T}^p$ e $\mathbf{T}^s$ também podem ser representadas por dígrafos. Para a matriz de transição de primeira ordem, tem-se o que será referenciado neste trabalho como dígrafo de primeira ordem. Cada vértice do dígrafo representa uma nota das 18 notas de valores, como semínima, mínima, semicolcheia, e assim por diante. As arestas refletem a ocorrência temporal das notas. Por exemplo, a existência de uma aresta a partir de um vértice i , representado por uma semínima, para um vértice j , representado por uma semicolcheia, indica que uma semínima (evento passado) foi seguida por uma semicolcheia pelo menos uma vez. O peso das arestas expressa a frequência relativa com que as notas de valores seguem-se umas às outras na música e é definido pela entrada na matriz de transição $\mathbf{T}^p(i, j)$. Quanto mais espessa a aresta, maior é o peso da aresta entre esses dois vértices, ou seja, maior é a probabilidade desta sequência ocorrer.

De forma similar os pesos das arestas dos dígrafos de segunda ordem são manifestados

^c Foram descartadas as notas com probabilidade de ocorrência inferior a 0.001%, como notas de valor menor que a semicolcheia, e tercinas menos comuns.

pelas entradas da matriz de transição $\mathbf{T}^s$, em que $\mathbf{T}^s(i, j)$ associa um peso à conexão do vértice i (par de notas ocorrido anteriormente) para o vértice j (nota em questão). Neste caso, os pesos expressam a frequência com que as duas notas de valores anteriores foram seguidas pela nota de valor atual.

Exemplos de dígrafos de primeira ordem são apresentados na Figura 4.4. A Figura 4.4a mostra uma amostra de blues através da música *Looking Out The Window* por *Stevie Ray Vaughan*. A Figura 4.4b ilustra uma amostra de MPB através da música *Desalento* por Chico Buarque. Na Figura 4.4c uma amostra de reggae pode ser observada através da música *Three Little Birds* por *Bob Marley*. Finalmente, a Figura 4.4d apresenta uma amostra de rock, representada pela música *She Loves You* por *The Beatles*. Os vértices destes dígrafos estão atribuídos em posições fixas para permitir uma primeira inspeção visual quanto às diferenças e similaridades entre os ritmos de amostras de gêneros diferentes.

A relação temporal das notas de valores parece ser mais dinâmica para a música de blues, uma vez que o dígrafo da Figura 4.4a contém mais conexões quando comparado aos demais. Por outro lado, os padrões rítmicos estão expressados de forma menos dinâmica para a música de MPB (dígrafo da Figura 4.4b). De acordo com os exemplos adotados, as músicas de reggae e rock contém aspectos rítmicos parecidos. Porém, vale ressaltar que estes dígrafos retratam apenas uma música de cada gênero. Busca-se constatar se essa dinâmica se mantém discriminatória ao considerar todas as músicas de cada gênero. Pelos dígrafos da Figura 4.4, observa-se que algumas notas como a tercina de colcheia e a semicolcheia estão presentes nos quatro dígrafos, se mostrando notas expressivas na modelagem do ritmo conforme realizada no trabalho.

Os dígrafos de segunda ordem das mesmas músicas da Figura 4.4 podem ser observados na Figura 4.5. Neste caso, a visualização das 167 combinações de pares fica prejudicada. É por isso que os dígrafos da Figura 4.5 contém apenas os vértices que possuem ligações entre si.

A relação da dinâmica descrita anteriormente para os dígrafos de primeira ordem se mostrou similar aos de segunda ordem. O dígrafo de blues contém diversas conexões e vértices, enquanto que o de MPB contém apenas cinco vértices. Verifica-se que a modelagem pela cadeia de Markov de segunda ordem fornece mais informação quanto às dependências das sequências de notas.

4.4 Considerações finais

Este capítulo apresentou a metodologia para a obtenção e representação dos padrões rítmicos a partir de arquivos MIDI. A partir destes padrões, diferentes investigações são exploradas, conforme capítulos seguintes.

Figura 4.4 – Dígrafos de primeira ordem de quatro músicas. (a) *Looking Out the Window* (Stevie R Vaughan); (b) *Desalento* (Chico Buarque); (c) *Three Little Birds* (Bob Marley); (d) *She Loves You* (The Beatles)

O uso do formato MIDI favorece o entendimento das configurações de notas que contribuem para a análise, porém impõe a limitação de banco de dados menores para os experimentos.

A motivação para a representação em dígrafos está atribuída à possibilidade de utilizar medidas topológicas na análise. Além disso, caminhadas aleatórias nestes dígrafos podem ser diretamente aplicadas no contexto de síntese de novas sequências rítmicas. O próximo capítulo apresenta resultados desta análise.

Figura 4.5 – Dígrafos de segunda ordem de quatro músicas. (a) *Looking Out the Window* (Stevie R Vaughan); (b) *Desalento* (Chico Buarque); (c) *Three Little Birds* (Bob Marley); (d) *She Loves You* (The Beatles)

Capítulo 5

Síntese de sequências rítmicas

Este trabalho também investigou a possibilidade da composição de novas sequências rítmicas com base em estatísticas coletadas dos dígrafos de primeira ordem construídos a partir da percussão de músicas de diferentes gêneros. A motivação está na possibilidade de gerar novas sequências que mantenham características inerentes da referência utilizada para execução do método, permitindo uma variabilidade controlada.

A composição automática de sequências rítmicas, referenciada neste trabalho como síntese de ritmos, é realizada através de caminhadas aleatórias ponderadas em dígrafos de música ou em dígrafos de quatro gêneros, conforme a taxonomia apresentada na seção 5.1. Uma vez realizada, o resultado da síntese de ritmos também pode ser visualizado através de dígrafos, de forma a permitir comparações com as respectivas versões originais. A estratégia proposta para a síntese de novas sequências rítmicas a partir da percussão de músicas de diferentes gêneros está resumida na Figura 5.1 e detalhada na seção 5.2.

5.1 Taxonomia adotada

Tabela 5.1 – Taxonomia de gêneros para o sistema de composição de sequências rítmicas

Gêneros	Quantidade de Músicas
Blues	70
MPB	70
Reggae	70
Rock	70
Total: 280	

Para a investigação da síntese de ritmos foram utilizados os quatro gêneros musicais apresentados na Tabela 5.1.

Figura 5.1 – Etapas da abordagem proposta para a composição automática de sequências rítmicas contextualizadas por gêneros ou músicas

5.2 Estratégia

O algoritmo de síntese de novas sequências rítmicas é baseado no conceito de caminhada aleatória ponderada em dígrafos com pesos. De forma resumida, tal caminhada é realizada como se segue: primeiramente um vértice inicial que contém conexões de saída é escolhido; no segundo passo, um próximo vértice é randomicamente selecionado entre todos os vértices de saída do vértice inicial. Porém, esta escolha é ponderada pela probabilidade das arestas, ou seja, a probabilidade de escolha de um vértice é proporcional ao peso da aresta de saída do vértice anterior para ele (quanto maior o peso da aresta, maior a probabilidade). O segundo passo é repetido até que um número pré-determinado de passos seja atingido.

A estratégia descrita acima pode ser aplicada em dígrafos individuais de músicas, ou em dígrafos de gênero. No primeiro caso, o algoritmo de síntese é desempenhado sobre os dígrafos provenientes das matrizes de transição $\mathbf{T}^p$ de cada música, obtida conforme

apresentado no capítulo 4. Este procedimento estabelece um vetor $\mathbf{S}^{\text{dur}}$ que contém a sequência de notas de valores escolhidas pelo algoritmo. Com o cálculo das probabilidades de transição entre pares de notas do vetor $\mathbf{S}^{\text{dur}}$ através das cadeias de Markov de primeira ordem, uma nova matriz de transição $\mathbf{S}^{\text{p}}$ é determinada e utilizada para a construção do dígrafo resultante do processo de síntese.

Para a síntese em dígrafos de gêneros, tem-se uma abordagem semelhante. Primeiramente, um vetor de notas $\mathbf{G}^{\text{dur}}$ é obtido pela concatenação dos vetores $\mathbf{V}^{\text{dur}}$ de todas as músicas de um determinado gênero. A partir de $\mathbf{G}^{\text{dur}}$, uma matriz de transição $\mathbf{G}^{\text{p}}$ é calculada e representada como dígrafo, permitindo uma primeira inspeção visual das diferenças e semelhanças entre dígrafos de diferentes gêneros. Desta forma, são construídos um dígrafo para o gênero blues, um dígrafo para o gênero MPB, um dígrafo para gênero reggae e um dígrafo para o gênero rock. Na segunda etapa, tais dígrafos por gênero são usados como referência para a síntese de sequências rítmicas. Uma vez que o algoritmo de caminhada aleatória ponderada é desempenhado nestes dígrafos, tem-se o estabelecimento de um novo vetor $\mathbf{B}^{\text{dur}}$ que contém a sequência de notas de valores escolhidas pelo algoritmo a cada iteração. Calculando as probabilidades condicionais através do uso das cadeias de Markov de primeira ordem, obtém-se a matriz de transição resultante, $\mathbf{B}^{\text{p}}$, que também é representada como dígrafo para possibilitar uma comparação qualitativa em relação às respectivas versões originais.

5.3 Resultados e discussões

Alguns dos resultados de síntese estão indicados a seguir e divididos em dois grupos conforme a estratégia apresentada na seção anterior: síntese através de dígrafos de músicas; e síntese através de dígrafos de gênero.

5.3.1 Síntese por música

A Figura 5.2 (a) apresenta o dígrafo original da música “*Is This Love?*” por Bob Marley, enquanto que a Figura 5.2 (c) apresenta o dígrafo obtido pelo processo de síntese aplicado no dígrafo da Figura 5.2 (a). As primeiras notas de valores da música original podem ser observadas na Figura 5.2 (b); e as primeiras notas de valores que foram selecionadas no processo de composição podem ser observadas na Figura 5.2 (d).

Os dígrafos das Figuras 5.2 (a) e 5.2 (c) possuem quase todas as conexões em comum (exceto pela presença apenas no dígrafo original da aresta entre as notas colcheia e tercina de semicolcheia). Isto é esperado uma vez que o peso das arestas é elemento

Figura 5.2 – Dígrafos de música: (a) dígrafo de uma música de reggae; (b) parte inicial da seqüência rítmica original; (c) dígrafo gerado pelo processo de síntese da música em (a); (d) parte inicial da seqüência rítmica gerada a partir da caminhada aleatória no dígrafo em (c)

Figura 5.3 – Dígrafos de música: (a) dígrafo de uma música de mpb; (b) parte inicial da sequência rítmica original; (c) dígrafo gerado pelo processo de síntese da música em (a); (d) parte inicial da sequência rítmica gerada a partir da caminhada aleatória no dígrafo em (c)

determinante da escolha do próximo vértice e pelo fato de que há um conjunto pequeno de vértices participantes. Seguindo a mesma configuração do exemplo de reggae apresentado na Figura 5.2, a Figura 5.3 apresenta os resultados de síntese para a música de mpb “*Construção*” por Chico Buarque.

As sequências apresentadas nas Figuras 5.2 (b)-(c) e nas Figuras 5.3 (b)-(c) são simples de ser executadas e contêm passagens de notas em comuns. Em todas as sequências observa-se a prevalência da semicolcheia. Nas partituras de percussão analisadas neste trabalho, a semicolcheia representa principalmente o chimbau, que é o instrumento de percussão mais recorrente para os gêneros estudados. Portanto, é esperado que esta nota de valor se sobressaia em relação as demais. Apesar destas similaridades, as versões originais e compostas das sequências de cada música contêm passagens de notas incomuns, o que

permite a inclusão de uma variabilidade controlada nas composições.

Quanto à qualidade musical, de acordo com a opinião dos autores deste trabalho, as sequências podem soar um tanto monótonas. Uma explicação para este fato é que apenas um instrumento é considerado por batida, deixando a sequência um pouco sem contexto. A combinação de várias composições permitindo a simultaneidade de instrumentos em uma mesma batida pode ser favorável para que as novas sequências sejam dinâmicas e sonoramente agradáveis. Esta combinação pode se realizar, por exemplo, através de uma abordagem diferente para a escolha do instrumento representante de cada batida. Considere a situação em que, ao invés da mediana, o instrumento “mais grave” é escolhido. Combinando as composições desta abordagem com a estratégia atual, é possível que uma mesma batida tenha dois instrumentos diferentes, possibilitando a síntese de uma sequência rítmica mais rica.

5.3.2 Síntese por gênero

Os dígrafos originais de cada gênero estão indicados na Figura 5.4. Comparando de forma qualitativa os quatro dígrafos, é possível observar muitas conexões distintas. Por exemplo, a sequência de colcheia e semínima duplamente pontuada e a sequência de tercina de colcheia para mínima pontuada só ocorre no dígrafo de blues, enquanto que a transição da semicolcheia para a mínima pontuada só ocorre no dígrafo de rock. Algumas transições significativas (representadas por arestas com maiores pesos) são distintas entre os dígrafos, e algumas são comuns. Os pares de notas semicolcheia - semicolcheia, semicolcheia - colcheia pontuada, duas tercinas de colcheias, por exemplo, são sequências significativas em todos os gêneros.

Dígrafos de blues e rock contêm mais conexões que os dígrafos de MPB e reggae. Conforme já apresentado neste trabalho, o ritmo de blues tende a ser mais complexo entre os quatro adotados neste estudo. Muitos estilos são derivados do rock, de forma que a variabilidade deste gênero pode estar refletida nas diversas conexões do seu dígrafo.

Os dígrafos resultantes do processo de síntese estão ilustrados na Figura 5.5, enquanto que as novas sequências rítmicas geradas pelo algoritmo de síntese nos dígrafos de gêneros estão apresentadas na Figura 5.6.

As sequências da Figura 5.5 também são simples de ser executadas, contêm a prevalência da semicolcheia, e os autores têm a mesma opinião quanto à qualidade musical. A presença de alguns pontos de aumento da sequência de blues (Figura 5.5a), de uma certa forma, representa o aspecto desigual e “swingado” do ritmo deste gênero. A sequência de MPB (Figura 5.5b) é a mais repetitiva em termos da prevalência da semicolcheia. Vale notar que a tercina está presente na sequência de reggae (Figura 5.5c). A tercina é comum neste

Figura 5.4 – Dígrafos de gêneros: (a) dígrafo de blues, (b) dígrafo de mpb, (c) dígrafo de reggae, e (d) dígrafo de rock

gênero, principalmente em músicas com características do *One Drop Beat* (83). A presença da semínima acontece nas sequências de reggae e rock (Figura 5.5d). Este aspecto também é notável, pois a semínima geralmente representa o bumbo, um instrumento importante nestes dois gêneros pois contribui para a caracterização de batidas definidas.

Figura 5.5 – Dígrafos de gênero resultantes do algoritmo de caminhada aleatória sobre os dígrafos em 5.4: (a) dígrafo de blues, (b) dígrafo de de mpb, (c) dígrafo de reggae, e (d) dígrafo de rock

5.4 Considerações finais

Este capítulo apresentou a investigação da síntese de novas sequências rítmicas, aplicadas em dígrafos de música ou em dígrafos de gêneros.

As novas sequências são parecidas com as originais, mas contêm passagens diferentes que estão dentro do contexto da referência adotada. Este aspecto é interessante do ponto de vista artístico, uma vez que possibilita uma variabilidade controlada. Cada sequência pode ainda ser explorada de diferentes formas. Um exemplo seria inserir informações de altura em cada uma das notas de valores, tornando a sequência rítmico-melódica. Estas

Figura 5.6 – Primeiras 50 notas da nova sequência rítmica obtida pelo algoritmo de síntese nos dígrafos em 5.4. Em (a) sequência de blues, (b) sequência de de mpb, (c) sequência de reggae, e (d) sequência de rock

possibilidades podem servir de inspirações iniciais para composições musicais.

A síntese através dos dígrafos de segunda ordem, apesar de não explorada neste trabalho, pode incluir uma maior variabilidade rítmica nas sequências. Vale também mencionar que as estratégias de síntese desenvolvidas neste trabalho são diretamente aplicáveis para diferentes aspectos musicais, como a melodia e a harmonia. Isto pode ser realizado, por exemplo, pelo emprego de diferentes instrumentos na análise, como o baixo, a guitarra (sequência de acordes) e o vocal.

A combinação de diferentes instrumentos pode ser favorável para que as composições sejam mais completas e com significado musical. De fato, a inclusão de outras medidas do ritmo, como tempo e compasso, pode contribuir para o refinamento das sequências geradas e permitir uma análise mais sistemática em relação às características de cada gênero. Torna-se possível verificar, por exemplo, se o prolongamento da duração que favorece um aspecto “swingado” no blues ocorre em uma determinada batida. Também seria possível analisar, por exemplo, como uma melodia de blues se comportaria num ritmo de reggae.

Em relação aos dígrafos de gênero, observou-se que, enquanto muitas sequências importantes de pares de notas são comuns aos quatro gêneros, outras são peculiares de

cada um. Portanto, torna-se relevante verificar se os gêneros podem ser discriminados em termos destas sequências. O próximo capítulo apresenta esta investigação.

Capítulo 6

Classificação de gêneros musicais

Neste capítulo são apresentados os principais resultados da discriminação dos gêneros musicais através de métodos de aprendizagem supervisionada e não supervisionada.

Assim como diversos problemas envolvendo tarefas de reconhecimento de padrões, o processo de classificação automática de gêneros musicais pode ser dividido em três etapas principais: representação, extração de atributos, e *design* do classificador (6, 7). A representação das músicas se dá pelas probabilidades condicionais de sequências de pares e trios de notas de valores através das cadeias de Markov de primeira e segunda ordens (Capítulo 4). A extração de atributos é realizada neste estudo pelos métodos PCA e LDA, enquanto que a classificação é desempenhada através do classificador gaussiano e do agrupamento hierárquico de dados.

Tendo em vista a dificuldade de se definir uma taxonomia clara de gêneros musicais, é também investigado neste trabalho a possibilidade de classificar as músicas dentro de uma perspectiva de classificação multi-categórica. A motivação é complementar aos resultados de classificação, considerando as discussões sobre a redundância da taxonomia de gêneros. A possibilidade de uma classificação multi-gêneros pode ser promissora e está provavelmente mais próxima à experiência humana. Neste caso, uma música pode não estar associada somente a um único gênero, mas sim a um conjunto destes que a descreve.

A Figura 6.1 resume os passos principais das estratégias de classificação direcionadas neste capítulo. A taxonomia adotada está apresentada na seção 6.1. Detalhes das tarefas de classificação estão descritos na seção 6.2, enquanto que os resultados estão discutidos na seção 6.3. As considerações finais do capítulo estão expostas na seção 6.5.

6.1 Taxonomia adotada

Para a análise da classificação automática de gêneros musicais foi utilizada uma taxonomia de quatro gêneros musicais conforme apresentado na Tabela 6.1.

Figura 6.1 – Diagrama dos passos principais para classificação automática de gêneros musicais

6.2 Estratégia

Os métodos de aprendizagem são aplicados nas matrizes de dados $\mathbf{P}^p$ e $\mathbf{P}^s$, estabelecidas, respectivamente, através dos vetores de atributos $\mathbf{V}^m$ (cadeia de Markov de primeira ordem) e $\mathbf{B}^m$ (cadeia de Markov de segunda ordem) de cada música do banco de dados. A matriz de dados $\mathbf{P}^p$ tem dimensão 280 (quantidade de amostras) x 324 (quantidade de atributos); enquanto que a matriz de dados $\mathbf{P}^s$ tem dimensão 280 x 3006 (seção 4.2).

Perspectiva supervisionada

Como método de aprendizagem supervisionada, foi utilizado o classificador Gaussiano através de funções discriminantes e, como medida de desempenho, adotou-se o coeficiente kappa de Cohen e a porcentagem de acerto. Na estimação do erro das tarefas de classificação

Tabela 6.1 – Taxonomia de gêneros para o sistema de classificação. Cada amostra está associada a um único gênero, uma vez que a maioria dos bancos de dados MIDI ainda não é multi-categoria.

Gêneros	Quantidade de Músicas
Blues	70
MPB	70
Reggae	70
Rock	70
Total: 280	

a seguir, *Re-substitution* significa que todas as amostras (músicas) de cada gênero foram usadas como conjunto de treinamento (para estimar os parâmetros) e todas as amostras foram usadas como conjunto de teste. A abordagem *Re-Substitution* é conhecida como um estimador otimista, porém útil para verificar se o modelo está correto e para permitir a análise da contribuição dos atributos através de um custo computacional menor.

Para a validação cruzada *Hold-out* 70%-30%, 70% das amostras de cada classe foi utilizada como conjunto de treinamento e os outros 30% foram usados como conjunto de teste. Na validação cruzada *Hold-out* 50%-50%, as amostras foram separadas em dois grupos: 50% para treinamento e 50% para teste. Nestes dois casos de validação cruzada, cem iterações são realizadas de maneira que, em cada iteração, uma partição diferente é aleatoriamente selecionada. Os resultados a seguir apresentam, portanto, o coeficiente kappa médio, variância média e taxa de acerto média das iterações.

A classificação multi-categorias é desempenhada sobre a matriz de dados $\mathbf{P}^{\mathbf{P}}$. Neste caso, ao invés de atribuir uma amostra apenas à classe com máxima probabilidade *a posteriori*, a estratégia de classificação é estabelecida como se segue.

1. Primeiramente, as probabilidades *a posteriori* máximas da amostra m são normalizadas de forma que somem um, e ordenadas de forma ascendente em um vetor $\mathbf{P}_m^{\max}$.
2. Em seguida, dada uma amostra m , computa-se a distância entre seu vetor $\mathbf{P}_m^{\max}$ e cada um dos seguintes vetores: $\mathbf{R}_1 = [1, 0, 0, 0]$, representando a classificação em uma categoria; $\mathbf{R}_2 = [0.5, 0.5, 0, 0]$, representando a classificação em duas categorias; $\mathbf{R}_3 = [0.33, 0.33, 0.33, 0]$, para a classificação em três categorias; e $\mathbf{R}_4 = [0.25, 0.25, 0.25, 0.25]$, para a classificação em quatro categorias. A classificação final da amostra m é atribuída de acordo com a mínima distância entre seu vetor $\mathbf{P}_m^{\max}$ e cada um destes quatro vetores.

Para exemplificar esta idéia, considere a primeira amostra do gênero blues e seu respectivo vetor $\mathbf{P}_1^{\max}$ de probabilidades *a posteriori* máximas: 0.9521156 dado o gênero blues; 0.0108921 dado o gênero MPB; 0.0162640 dado o gênero reggae; e 0.0207283 dado o gênero rock. Ao computar a distância entre o vetor $\mathbf{P}_1^{\max}$, definido como $\mathbf{P}_1^{\max} =$

$[0.9521156, 0.0108921, 0.0162640, 0.0207283]$, e os vetores descritos no passo 2, certamente a menor distância será atribuída ao vetor $\mathbf{R}_1$, indicando que a amostra deve ser classificada somente como blues.

Agora considere a quinta amostra de blues e seu respectivo vetor de probabilidades *a posteriori* máximas: 0.4097450 dado o gênero blues; 0.0079218 para MPB; 0.5501617 para reggae; and 0.0321714 para rock. Após a ordenação, o vetor de probabilidades *a posteriori* máximas será $\mathbf{P}_5^{\max} = [0.5501617, 0.4097450, 0.0321714, 0.0079218]$, e a mínima distância é atribuída ao vetor $\mathbf{R}_2$, indicando que, de acordo com os padrões rítmicos e a abordagem adotada, esta música deve ser classificada em dois gêneros, blues e reggae.

Amostras (músicas) com valores similares de probabilidades *a posteriori* máximas tendem a refletir uma mistura de atributos de uma ou mais classes (gêneros). A motivação da abordagem multi-categorias é definir um critério no qual estas amostras podem ser adequadamente atribuídas a mais de uma classe, dependendo do nível de similaridade que apresentam.

Perspectiva não supervisionada

O agrupamento hierárquico aglomerativo foi adotado como um dos métodos de aprendizagem não supervisionada (seção 3.3). A distância Euclidiana foi estabelecida como medida de distância, enquanto que o método de *linkage* por *Ward* foi adotado como critério de decisão para juntar dois grupos de amostras(3.3.2). O número de grupos foi definido como quatro para permitir uma comparação com a abordagem supervisionada.

Para a classificação multi-categorias, a estratégia é definida como se segue. A partir do site LastFM (<http://www.lasfm.com.br>), uma rádio *online* interativa conhecida, verificou-se como os rótulos dos gêneros estavam associados com as músicas do banco de dados, com o intuito de estabelecer uma taxonomia multi-categorias das mesmas. O nome de cada uma das 280 músicas foi inserido no site e os gêneros mais comuns associados à ela (gêneros que aparecem junto com a música) foram identificados. Se entre estes gêneros encontrava-se algum dos gêneros desta investigação (blues, MPB, reggae e rock), esta música teria este gênero associado a ela também. Por exemplo, entre os gêneros mais comuns associados à música *A mess of blues* por “Elvis Presley” estão os gêneros blues (original) e rock. Esta música, portanto, foi rotulada como blues e rock. Este processo resultou no diagrama de Venn da Figura 6.2.

As músicas de MPB não tiveram outros rótulos (blues, reggae ou rock) associados a elas. Vinte amostras de blues tiveram o rótulo de “rock” associado a elas, 17 amostras de reggae tiveram o rótulo de “rock” associado a elas, e 3 amostras de rock tiveram o rótulo “blues” associado a elas.

Figura 6.2 – Diagrama de Venn para o banco de dados multi-categorias

6.3 Resultados e discussões

A seguir são apresentados os principais resultados da extração de atributos (seção 6.3.1), da classificação supervisionada (seção 6.3.2), do agrupamento de dados (seção 6.3.3), e da classificação multi-categorias (6.3.4).

6.3.1 Extração de atributos

Com o objetivo de reduzir a dimensionalidade dos dados, torna-se necessário identificar o número ideal de componentes principais que irão representar os novos atributos. É esperado que, em uma dimensão alta, as classes sejam separáveis. Entretanto, altas dimensões impõem alta complexidade ao sistema. Um procedimento para determinar a quantidade ideal de componentes principais é verificar o quanto da variância dos dados é preservada. Para tal, soma-se os valores dos primeiros l autovalores (l é a quantidade de componentes principais a ser verificada), e divide-se o resultado desta soma pela soma total dos valores de todos os autovalores. Se o resultado deste cálculo for igual ou maior que 75%, pode-se dizer que um número l de componentes (ou novos atributos) preserva pelo menos 75% da variância dos dados, um limiar geralmente aceitável em tarefas de classificação (7).

Apesar de reter apenas 36% da variância dos dados, a Figura 6.3 ilustra as duas componentes principais (novos atributos) obtidas com a aplicação do PCA. Na Figura 6.3a o PCA foi aplicado na matriz de dados $\mathbf{P}^{\mathbf{P}}$. Verificou-se que são necessárias 15 componentes principais para que 76% da variância dos dados seja retida. Em outras palavras, é possível reduzir a dimensionalidade dos dados de 324-D para 15-D sem uma perda significativa de informação. No entanto, será mostrado a seguir que a quantidade de componentes para se atingir melhores resultados depende da estratégia de classificação escolhida.

A Figura 6.3b apresenta a primeira e segunda componentes principais geradas pela

Figura 6.3 – Primeiro e segundo novos atributos obtidos pela aplicação da técnica PCA na matriz de dados (a) $\mathbf{P}^P$ e (b) $\mathbf{P}^S$

aplicação do PCA sobre a matriz de dados $\mathbf{P}^S$ (apenas 20% da variância dos dados foi capturada por esses dois eixos). Neste caso, são necessárias 30 componentes principais para preservar 75% da variância dos dados. Portanto, é possível reduzir a dimensionalidade dos dados de 3006-D para 30-D.

Nos dois casos, é possível observar que as classes estão substancialmente sobrepostas, o que torna este problema de classificação automática de gêneros uma tarefa não trivial. A Figura 6.4a ilustra a primeira e segunda componentes (novos atributos) obtidos pela aplicação da técnica LDA na matriz de dados $\mathbf{P}^P$, enquanto que a Figura 6.4b apresenta as duas primeiras componentes para a matriz de dados $\mathbf{P}^S$. Como mencionado anteriormente, a LDA tem a restrição de obter apenas $C - 1$ autovalores não-nulos, em que C é o número de classes (gêneros). Portanto, nesta investigação apenas três componentes podem ser obtidas. Por ser um método supervisionado, no qual o objetivo principal é maximizar a separabilidade entre as classes, os quatro gêneros estão mais claros que no espaço de atributos PCA (Figura 6.3), mas ainda assim é possível observar sobreposições substanciais, principalmente entre os gêneros rock e reggae.

Com a análise LDA, é possível quantificar a dispersão intra-classe e inter-classe da matriz de dados através do cálculo das matrizes de dispersão (6, 44). Esta análise foi desempenhada para a matriz de dados $\mathbf{P}^P$. A matriz de dispersão de cada classe, $S_{intraBlues}$, $S_{intraMPB}$, $S_{intraReggae}$ e $S_{intraRock}$, a matriz de dispersão inter-classe, S_{inter} , e um índice de separabilidade geral definido como o traço de $(S_{intra}^{-1} * S_{inter})$ foram calculados. Os valores deste último estão descritos a seguir:

$$\begin{aligned} trace(S_{intra}) &= 505.55 \\ trace(S_{intraBlues}) &= 154.96 \end{aligned}$$

Figura 6.4 – Primeiro e segundo atributos obtidos pela aplicação da técnica LDA na matriz de dados (a) $\mathbf{P}^P$ e (b) $\mathbf{P}^S$

$$\begin{aligned}
 \text{trace}(S_{\text{intraMPB}}) &= 127.41 \\
 \text{trace}(S_{\text{intraReggae}}) &= 91.76 \\
 \text{trace}(S_{\text{intraRock}}) &= 131.40 \\
 \text{trace}(S_{\text{inter}}) &= 22.08 \\
 \text{trace}(S_{\text{intra}}^{-1} * S_{\text{inter}}) &= 3.49
 \end{aligned}$$

Estes valores enfatizam a dificuldade deste problema de classificação: os traços das matrizes de dispersão intra-classe são grandes, enquanto que o traço da matriz de dispersão inter-classe e do índice de separabilidade geral são pequenos. Isto confirma que as classes se sobrepõem consideravelmente, uma vez que estes valores mostram que as classes não estão compactas no espaço de atributos, e, portanto, possuem pequenas distâncias entre seus elementos e respectivos centróides. O menor valor entre os traços das matrizes de dispersão intra-classe corresponde ao gênero reggae (é a classe mais compacta). Esta característica pode justificar o motivo pelo qual músicas pertencentes a este gênero foram, em grande parte, classificadas corretamente.

As técnicas de extração de atributos PCA e LDA ajudam a identificar quais atributos mais contribuem para a classificação. Uma análise interessante seria associar os valores de cada elemento dos primeiros autovetores com os atributos originais. Neste estudo, destacaram-se as dez sequências de pares de notas que mais contribuem para a representação e separação das classes (Figura 6.5) considerando a matriz de dados $\mathbf{P}^P$ e o primeiro e segundo autovetores obtidos pelo PCA e LDA. Os índices dos elementos mais expressivos (pesos maiores) destes autovetores que indicam atributos mais importantes são selecionados e associados às sequências de notas da matriz $\mathbf{P}^P$ original.

A Figura 6.5a e 6.5b apresentam as sequências mais importantes de acordo com dois primeiros autovetores do PCA. A espessura da aresta é definida pelo valor (peso) do índice

Figura 6.5 – As dez sequências de notas mais significantes de acordo com: (a) primeiro autovetor PCA (b) segundo autovetor PCA (c) primeiro autovetor LDA (d) segundo autovetor LDA

no autovetor. Vale mencionar que estas sequências são aquelas que mais frequentemente acontecem nos ritmos dos gêneros abordados. Em outras palavras, correspondem aos índices que possuem papel significativo na representação dos ritmos. Portanto, estabeleceu-se nesta investigação que são as “dez sequências mais representativas”, quando o primeiro e segundo autovetores da PCA são levados em consideração. Tercinas de colcheia e semicolcheia, por exemplo, são particularmente importantes nos gêneros blues e reggae.

De forma similar, as Figuras 6.5c e 6.5d apresentam as sequências mais importantes de acordo com o primeiro e segundo autovetores da LDA. Diferentemente das sequências obtidas pela PCA, estas sequências não são comuns para todos os ritmos, mas devem ocorrer com frequências distintas em cada gênero. Desta forma, elas são referidas neste trabalho como as “dez sequências mais discriminativas”, quando o primeiro e segundo autovetores da LDA são examinados.

Tabela 6.2 – Coeficiente kappa e desempenho de classificação dos novos atributos do PCA computados a partir da matriz de dados $\mathbf{P}^P$, que expressa a modelagem do ritmo através da cadeia de Markov de primeira ordem

PCA	Classificador Gaussiano Quadrático			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	0.60	0.0012	73.72%	Substancial
Hold-Out 70%-30%	0.34	0.0053	50.25%	Razoável
Hold-Out 50%-50%	0.33	0.0033	48.57%	Razoável
	Classificador Gaussiano Linear			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	0.64	0.0012	73.21%	Substancial
Hold-Out 70%-30%	0.33	0.0053	50%	Razoável
Hold-Out 50%-50%	0.31	0.0032	48%	Razoável

6.3.2 Classificador Gaussiano

Esta seção descreve os resultados de classificação de acordo com o classificador Gaussiano. Ambas as técnicas de extração de atributos PCA e LDA são aplicadas nas matrizes de dados, de forma que os resultados de classificação estão especificados de acordo com o método utilizado.

Atributos rítmicos das cadeias de Markov de primeira ordem: matriz de dados $\mathbf{P}^P$

A Tabela 6.2 apresenta os resultados de classificação obtidos pelo classificador Gaussiano sobre os atributos gerados pela aplicação do método PCA na matriz de dados $\mathbf{P}^P$. Os resultados estão descritos em termos de coeficiente kappa, variância, desempenho da classificação e interpretação do valor do coeficiente de acordo com (49).

A variância do kappa fornece evidências sobre a confiabilidade do seu desempenho. Quanto maior a variância, menor é a precisão do valor obtido pelo coeficiente. Uma vez que este coeficiente é uma medida estatística, em geral, o uso de banco de dados maiores melhora sua precisão, tornando menores os valores de variância. Observa-se pela Tabela 6.2 que a menor variância ocorreu para o cenário *Re-substitution*, no qual todas as amostras fazem parte do conjunto de teste. Este comportamento indica que a configuração do *Re-substitution* possibilitou a melhor precisão do kappa nestes experimentos. Por outro lado, a validação cruzada *Hold-Out 70%-30%* forneceu a maior variância do kappa, uma vez que apenas 30% das amostras formam o conjunto de teste.

A Tabela 6.2 indica que o desempenho para as partições *Hold-Out* foram em torno de 50%. Como PCA é uma técnica não supervisionada, o desempenho da estimação dos parâmetros (por exemplo, das matrizes de covariância) é prejudicado devido ao problema de haver relativamente poucas amostras para análise. Este comportamento deve ser

Tabela 6.3 – Coeficiente kappa e desempenho de classificação dos novos atributos do LDA computados a partir da matriz de dados $\mathbf{P}^{\mathbf{P}}$, que expressa a modelagem do ritmo através da cadeia de Markov de primeira ordem

LDA	Classificador Gaussiano Quadrático			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	0.62	0.0013	71.78%	Substancial
Hold-Out 70%-30%	0.57	0.0048	67.50%	Moderado
Hold-Out 50%-50%	0.6	0.0027	70%	Substancial
	Classificador Gaussiano Linear			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	0.63	0.0013	72.14%	Substancial
Hold-Out 70%-30%	0.58	0.0046	68.75%	Moderado
Hold-Out 50%-50%	0.66	0.0024	74.28%	Substancial

Tabela 6.4 – Matrizes de confusão dos resultados em destaque nas Tabelas 6.2 e 6.3

PCA kappa = 0.64					LDA kappa = 0.63				
	Blues	MPB	Reggae	Rock		Blues	MPB	Reggae	Rock
Blues	46	2	22	0	Blues	61	1	6	2
MPB	1	45	23	1	MPB	2	45	20	3
Reggae	0	0	70	0	Reggae	6	6	56	2
Rock	1	1	31	37	Rock	5	5	21	39

minimizado com a utilização de um banco de amostras maior.

A Tabela 6.3 apresenta os resultados de classificação obtidos pelo classificador Bayesiano sobre os atributos LDA através da matriz de dados $\mathbf{P}^{\mathbf{P}}$. Diferente do PCA, LDA é uma técnica que utiliza toda informação discriminativa disponível na matriz de dados. Pela Tabela 6.3, observa-se que os novos atributos LDA permitiram bons resultados nas situações *Hold-Outs*, o que reflete sua característica supervisionadas.

As matrizes de confusão da Tabela 6.4 são referentes à classificação da configuração *Re-Substitution* com atributos PCA e LDA e uso do classificador Gaussiano Quadrático. Apesar do valor e variância do coeficiente kappa serem parecidos nas duas classificações, as respectivas matrizes de transição são diferentes. Entretanto, nos dois casos, as classificações incorretas estão mais concentradas no gênero reggae: muitas músicas de outros gêneros foram classificadas como reggae. Com o PCA, todas as amostras de reggae foram classificadas corretamente.

Ainda a partir da matriz de dados $\mathbf{P}^{\mathbf{P}}$, a Figura 6.6 apresenta a variação do valor do coeficiente kappa de acordo com a quantidade de componentes principais PCA para o classificador Gaussiano quadrático e linear nas configurações *Re-Substitution*, *Hold-Out 70%-30%*, e *Hold-Out 50%-50%*.

O último valor dos gráficos da Figura 6.6 sugere que, a partir deste valor a classificação não pode ser efetuada devido aos problemas de singularidade envolvendo as inversões das matrizes de covariância. Pode ser observado pela Figura 6.6 que este limiar de singularidade é diferente em cada situação. Entretanto, para o classificador quadrático, este limiar está no intervalo de 38-48 componentes; enquanto que, para o classificador linear, está no

Tabela 6.5 – Coeficiente kappa e desempenho de classificação dos novos atributos do PCA e LDA computados a partir da matriz de dados $\mathbf{P}^s$, que expressa a modelagem do ritmo através da cadeia de Markov de segunda ordem

PCA	Classificador Gaussiano Linear			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	0.82	0.0011	87%	Quase perfeito
Hold-Out 70%-30%	0.35	0.0052	51%	Razoável
Hold-Out 50%-50%	0.34	0.0031	51%	Razoável
LDA	Classificador Gaussiano Linear			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	0.8	0.0012	85%	Substancial
Hold-Out 70%-30%	0.87	0.0047	90%	Quase perfeito
Hold-Out 50%-50%	0.84	0.0027	87.85%	Quase perfeito

intervalo de 79-105 componentes. O limiar menor para o classificador quadrático pode ser explicado pelo fato de que é necessária a estimação de quatro matrizes de covariância, cada uma estimada pelas amostras da classe respectiva. Como há 70 amostras para cada classe, problemas de singularidade irão ocorrer em um espaço de dimensão menor quando comparado ao classificador linear, o qual usa todas as 280 amostras para estimar uma única matriz de covariância. O número ideal de componentes principais permitindo o melhor valor do coeficiente kappa (destacado em vermelho em cada um dos gráficos) também segue este comportamento, e está em um intervalo de 5-37 componentes para o classificador quadrático; e 9-106 para o classificador linear.

Atributos rítmicos das cadeias de Markov de segunda ordem: matriz de dados $\mathbf{P}^s$

A Tabela 6.5 apresenta os resultados de classificação para os atributos PCA e LDA obtidas através da matriz de dados $\mathbf{P}^s$. A utilização de uma cadeia de Markov de segunda ordem para a modelagem do ritmo melhorou significativamente o desempenho de classificação. Isto indica que a análise de um contexto rítmico estendido (sequências de trios de notas ao invés de pares de notas) é sensível à discriminação dos gêneros.

A Tabela 6.6 apresenta a matriz de confusão da classificação *Re-substitution* com kappa = 0.82. Novamente, todas as amostras de reggae foram corretamente classificadas. Trinta amostras dos demais gêneros foram classificados como reggae, o que indica que estas amostras estão sobrepostas às de reggae no espaço de atributos deste gênero. Porém, a quantidade de amostras sobrepostas é menor que no caso apresentado anteriormente (Tabela 6.4 - kappa=0.61) para a matriz de dados $\mathbf{P}^p$, em que 76 amostras de outros gêneros estão sobrepostas às de reggae. Blues se manteve como o gênero com segunda maior quantidade de amostras classificadas corretamente, seguido de mpb e rock.

Figura 6.6 – Valores do kappa variando o número de componentes principais utilizadas nas tarefas de classificação. (a)Classificador Gaussiano Quadrático: Re-Substitution (b)Classificador Gaussiano Linear: Re-Substitution (c)Classificador Gaussiano Quadrático: Hold-Out (70%-30%) (d)Classificador Gaussiano Linear: Hold-Out (70%-30%) (e)Classificador Gaussiano Quadrático: Hold-Out (50%-50%) (f)Classificador Gaussiano Linear: Hold-Out (50%-50%)

Tabela 6.6 – Matriz de Confusão para a classificação com $\kappa = 0.82$

PCA $\kappa = 0.82$				
	Blues	MPB	Reggae	Rock
Blues	65	1	4	0
MPB	0	58	12	0
Reggae	0	0	70	0
Rock	1	5	14	50

Tabela 6.7 – Matriz de confusão para o agrupamento hierárquico aglomerativo

	Blues	MPB	Reggae	Rock
Grupo 1	25	32	46	30
Grupo 2	11	16	8	8
Grupo 3	12	14	6	19
Grupo 4	22	8	10	13

6.3.3 Agrupamento hierárquico aglomerativo

Os resultados do agrupamento hierárquico discutidos a seguir foram computados através da matriz de dados $\mathbf{P}^p$, ou seja, com a caracterização da dinâmica das notas de valores através da cadeia de Markov de primeira ordem.

Para facilitar as análises dos dendrogramas gerados, os gêneros foram representados em cores. As amostras de blues estão em vermelho, as amostras de MPB em verde, de reggae em azul e de rock em roxo.

A matriz de confusão da Tabela 6.7 tem como objetivo verificar a quantidade de amostras de cada gênero atribuídas a cada grupo. Por exemplo, é sabido que na matriz de dados $\mathbf{P}^p$, músicas de 1 a 70 pertencem ao gênero blues. Portanto, a primeira coluna dessa matriz de confusão indica que 25 amostras de blues foram atribuídas ao primeiro grupo, 11 foram atribuídas ao segundo grupo, 12 ao terceiro grupo e 22 ao quarto grupo. Também pode ser observado que no grupo 1, amostras de reggae são maioria; enquanto que no grupo 2, amostras de MPB são maioria, apesar da diferença pequena do número de amostras de blues; no grupo 3, amostras de rock são maioria; e no grupo 4, amostras de blues são maioria. Portanto, cada gênero está predominante em um grupo distinto, através das suas amostras mais discriminativas. Contudo, as amostras se sobrepõem substancialmente, refletindo novamente a complexidade do problema.

Comparando a matriz de confusão da Tabela 6.7 com a matriz de confusão da Tabela 6.4 para a classificação com os atributos PCA ($\kappa = 0.61$), vale notar que no primeiro caso, o primeiro grupo contém amostras dos quatro gêneros consideravelmente apresentadas (25 para o gênero blues, 32 para o gênero mpb, 46 para o gênero reggae, e 30 do gênero rock), em um total de 133 amostras neste grupo. Na classificação supervisionada com atributos do PCA, 22 de amostras de blues, 23 de mpb, e 31 de rock foram classificadas de forma incorreta como pertencentes ao gênero reggae, em um total de 146 amostras classificadas

Figura 6.7 – Dendrograma da formação dos quatro grupos (coloridos em azul - reggae, verde - MPB, roxo- rock e vermelho - blues)

como reggae (incluindo as de reggae propriamente ditas). Desta forma, há uma relação entre estas duas abordagens: enquanto que a representação PCA não foi eficiente para discriminar reggae dos demais gêneros, o grupo 1 foi o grupo que mais sobrepôs amostras de todos os gêneros.

A Figura 6.7 apresenta o dendrograma dos quatro grupos formados, coloridos de acordo com o gênero dominante de cada um. Os quatro grupos estão detalhados nas Figuras 6.8 - 6.11, nas quais cores das legendas reportam a classificação real de cada grupo. Diferentes cores foram adotadas para permitir uma análise visual. O primeiro grupo está colorido em azul, o grupo dois em verde, grupo três em roxo, e grupo quatro em vermelho. Estas cores foram escolhidas de acordo com o gênero dominante de cada grupo. Por exemplo, o grupo um está colorido em azul porque neste grupo amostras de reggae são maioria.

A classificação real está expressa na legenda da figura de cada dendrograma, em que cada amostra está colorida de acordo com as cores atribuídas aos gêneros. Observa-se a presença de sub-grupos formados por amostras de um mesmo gênero. Tais sub-grupos podem indicar a presença dos diversos estilos dentro de um gênero, o que deve ser investigado em estudos futuros.

A performance do agrupamento de dados foi computada seguindo o seguinte critério: n_i/N_i , onde n_i é o número de classificações corretas da classe i ; e N_i é o número total de amostras da classe i . Assume-se como classificação correta se a amostra de um determinado gênero foi atribuída ao grupo no qual este gênero é dominante (reggae no grupo 1, mpb no grupo 2, rock no grupo 3 e blues no grupo 4). Os desempenhos foram: $46/70 = 66\%$ para reggae; $16/70 = 23\%$ para mpb, $19/70 = 27\%$ para rock, e $22/70 = 31\%$ para blues. O desempenho geral de classificação segundo esta estratégia de avaliação foi de 36.75% . O agrupamento hierárquico com a mesma configuração de parâmetros mas a partir das matrizes de dados $\mathbf{P}^s$, e $\mathbf{P}^{ps}$ resultou em um desempenho geral de 39.32% e

Figura 6.8 — Dendrograma detalhado do primeiro grupo, em azul. Reggae é o gênero dominante desse grupo. As amostras estão especificadas da esquerda para a direita: 205 208 178 174 117 85 161 88 102 2 68 74 237 158 159 119 134 126 123 6 275 198 255 107 189 160 232 97 204 103 168 157 200 153 196 142 187 261 17 192 229 109 219 51 236 267 249 257 83 273 217 274 143 100 146 210 180 194 3 239 190 34 16 18 38 193 155 36 152 5 12 203 25 44 263 23 62 28 266 82 86 125 169 141 164 254 184 30 71 87 129 58 116 115 264 220 276 118 166 197 265 37 131 132 72 78 95 104 105 167 234 29 40 170 144 214 59 69 179 175 177 227 172 226 98 247 182 212 225 206 272 209 e 162

Figura 6.9 — Dendrograma detalhado do segundo grupo, em verde. MPB é o gênero dominante desse grupo. As amostras estão especificadas da esquerda para a direita: 8 9 32 26 79 10 260 21 147 80 56 150 202 218 76 230 139 120 90 133 13 112 122 185 96 46 111 173 279 70 149 154 278 188 128 84 43 259 94 113 135 235 e 224

Figura 6.10 — Dendrograma detalhado do terceiro grupo, em roxo. Rock é o gênero dominante. As amostras estão especificadas da esquerda para a direita: 4 19 269 55 238 136 110 251 199 280 50 61 245 244 52 54 63 89 262 106 181 250 137 186 81 248 228 240 252 242 233 22 215 165 60 101 288 73 121 130 108 148 271 195 277 124 127 140 216 39 e 66

Figura 6.11 — Dendrograma detalhado do quarto grupo, em vermelho. Blues é o gênero dominante. As amostras estão especificadas da esquerda para a direita: 1 77 14 67 65 20 27 24 49 151 211 222 33 171 35 48 191 91 183 92 270 223 7 42 114 176 207 201 163 243 246 256 241 253 11 213 41 145 47 75 99 258 156 15 138 53 93 231 31 45 57 e 64

39.4%, respectivamente.

Por trabalhar na dimensão original dos dados, o agrupamento hierárquico permitiu uma discriminação das classes próxima às abordagens PCA apresentadas anteriormente mesmo sem qualquer informação antecipada sobre as classes (compare com os cenários *Hold-Out* da Tabela 6.2). Reggae se manteve como a classe que mais se sobrepõe às demais, enquanto que blues manteve-se como o gênero mais discriminativo, seguido de MPB e rock.

6.3.4 Classificação multi-categorias

Esta seção apresenta os resultados da discriminação dos gêneros sobre a matriz de dados $\mathbf{P}^{\mathbf{P}}$ considerando uma classificação multi-categorias, dentro de uma perspectiva supervisionada e de uma perspectiva não supervisionada.

Perspectiva supervisionada

Apesar do fato de que os melhores resultados de classificação descritos na seção 6.3.2 foram obtidos com o classificador gaussiano linear sobre os atributos do PCA e do LDA (Tabelas 6.2 e 6.3, respectivamente), os resultados da classificação multi-categorias são apresentados através do classificador gaussiano quadrático sobre as duas primeiras componentes do LDA (com $\kappa = 0.48$), uma vez que se torna mais fácil de ilustrar graficamente.

As Tabelas 6.8 e 6.9 descrevem as categorias obtidas para cada amostra ou música. Com o objetivo de permitir uma visualização gráfica dos resultados, a Figura 6.12 mostra os contornos das densidades condicionais gaussianas para os dois atributos LDA de cada classe, juntamente com a dispersão dos atributos das amostras do banco de dados. Os contornos são obtidos através de sub-intervalos igualmente espaçados dentro da extensão $[p_{min}, p_{max}]$, em que p_{min} é o valor mínimo de probabilidade e p_{max} o valor máximo de probabilidade da distribuição. Para o caso da Figura 6.12, os contornos são estabelecidos pelas linhas de contorno (do inglês, *isolines*) centrais destes intervalos (aquelas que dividem a extensão $[p_{min}, p_{max}]$ em duas metades). É possível observar que os rótulos atribuídos às amostras das Tabelas 6.8 e 6.9 refletem sua posição no espaço de atributos.

Como um exemplo, considere a amostra de blues número 30 da Figura 6.12. Apesar de ter blues como rótulo original, seu vetor de atributos está mais próximo dos vetores de atributos das amostras de mpb, ou seja, está localizado mais próximo ao centro da densidade condicional de mpb do que ao centro da densidade condicional de blues. Portanto,

Figura 6.12 – Linhas de contorno das densidades condicionais gaussianas para os dois primeiros atributos LDA e o espaço de atributos das amostras: todos os pontos para um dado contorno são igualmente equiprováveis, uma vez que estão equidistantes do vetor média da classe correspondente, de acordo com a métrica de Mahalanobis

de acordo com os atributos rítmicos, é esperado que esta amostra seja classificada como pertencente ao gênero MPB. As amostras 56 e 265 foram classificadas como pertencentes aos quatro gêneros. Nenhuma das amostras de reggae foi classificada como reggae apenas, uma vez que as densidades condicionais de reggae e rock estão quase totalmente sobrepostas.

Perspectiva não supervisionada

O desempenho da classificação multi-categorias sobre a perspectiva não supervisionada foi obtido a partir do banco de dados apresentado na Figura 6.2 e os grupos obtidos pelo agrupamento de dados (Tabela 6.7). Por exemplo, ao invés de considerar apenas as 19 amostras de rock no grupo 3 como classificações corretas, agora é também verificado quantas das 17 amostras de reggae-rock e quantas das 23 amostras de blues-rock estão presentes no grupo 3, uma vez que neste caso elas também são referidas como amostras de rock. Neste caso, o total de amostras de rock para a classificação multi-categorias é 107 (67 + 23 + 17). Os desempenhos do gênero MPB se manteve o mesmo (23%), dado que nenhuma das amostras de blues, rock ou reggae foram também rotulados como MPB.

Tabela 6.8 – Resultados da classificação multi-categorias para as músicas de blues e mpb

Amos de blues	Amostras de mpb
1 - blues	71 - MPB
2 - reggae and rock	72 - MPB
3 - reggae and blues	73 - MPB
4 - blues	74 - MPB, rock and reggae
5 - blues	75 - MPB
6 - reggae and rock	76 - reggae, MPB, and rock
7 - blues	77 - MPB
8 - blues	78 - MPB
9 - blues	79 - MPB, rock and reggae
10 - blues	80 - rock
11 - blues	81 - MPB
12 - blues	82 - reggae, rock and MPB
13 - blues	83 - reggae and rock
14 - blues	84 - MPB
15 - blues	85 - MPB, reggae and rock
16 - blues	86 - reggae, rock and MPB
17 - reggae and rock	87 - MPB
18 - blues	88 - MPB
19 - blues	89 - reggae, rock and blues
20 - blues	90 - blues, reggae and rock
21 - blues	91 - MPB
22 - blues	92 - MPB
23 - blues	93 - MPB
24 - blues	94 - MPB
25 - blues	95 - MPB
26 - blues	96 - reggae, rock and MPB
27 - blues	97 - reggae and rock
28 - reggae, rock and MPB	98 - rock and reggae
29 - blues and reggae	99 - MPB, rock and reggae
30 - MPB	100 - reggae and rock
31 - blues	101 - MPB
32 - blues	102 - MPB
33 - blues	103 - reggae and rock
34 - blues	104 - MPB
35 - blues	105 - MPB
36 - blues and reggae	106 - reggae, rock and MPB
37 - rock, blues and MPB	107 - reggae and rock
38 - blues	108 - MPB
39 - blues	109 - reggae and rock
40 - blues	110 - MPB
41 - blues	111 - reggae and rock
42 - blues	112 - reggae, rock and MPB
43 - blues	113 - MPB
44 - blues	114 - MPB
45 - blues	115 - reggae, rock and MPB
46 - blues	116 - reggae, rock and blues
47 - blues	117 - MPB, reggae and rock
48 - blues	118 - reggae, rock and MPB
49 - blues	119 - MPB
50 - blues	120 - MPB
51 - reggae and rock	121 - MPB
52 - blues, reggae and rock	122 - reggae, rock and MPB
53 - blues	123 - MPB
54 - reggae, blues and rock	124 - MPB
55 - blues	125 - reggae, rock and MPB
56 - reggae, rock, blues and MPB	126 - MPB
57 - blues	127 - MPB
58 - blues and reggae	128 - MPB
59 - blues	129 - MPB
60 - blues	130 - MPB
61 - blues	131 - MPB
62 - blues	132 - MPB
63 - reggae, blues and rock	133 - MPB
64 - blues	134 - MPB
65 - blues	135 - MPB
66 - blues	136 - MPB
67 - blues	137 - MPB
68 - blues	138 - MPB
69 - rock and reggae	139 - MPB
70 - blues	140 - MPB

Tabela 6.9 – Resultados da classificação multi-categorias para as músicas de reggae e rock

Amostras de reggae	Amostras de rock
141 - reggae, rock and MPB	211 - rock
142 - rock and reggae	212 - rock and reggae
143 - reggae and rock	213 - rock
144 - blues	214 - reggae and rock
145 - rock and reggae	215 - rock
146 - reggae and rock	216 - rock
147 - reggae, rock and MPB	217 - reggae and rock
148 - reggae and rock	218 - reggae, rock and MPB
149 - reggae, rock and MPB	219 - reggae and rock
150 - reggae, MPB and rock	220 - reggae, rock and MPB
151 - rock and reggae	221 - rock
152 - blues and reggae	222 - rock
153 - reggae and rock	223 - rock
154 - reggae, blues and rock	224 - rock
155 - blues and reggae	225 - reggae and rock
156 - rock and reggae	226 - rock and reggae
157 - reggae and rock	227 - reggae and rock
158 - MPB, rock and reggae	228 - rock
159 - MPB, rock and reggae	229 - reggae and rock
160 - reggae and rock	230 - reggae, MPB and rock
161 - reggae and rock	231 - rock
162 - MPB	232 - reggae and rock
163 - rock and reggae	233 - rock
164 - reggae, rock and MPB	234 - MPB
165 - rock and reggae	235 - rock
166 - reggae, rock and MPB	236 - reggae and rock
167 - MPB	237 - MPB, rock and reggae
168 - reggae and rock	238 - blues
169 - reggae, rock and MPB	239 - blues and reggae
170 - blues	240 - rock
171 - rock and reggae	241 - rock
172 - rock and reggae	242 - rock
173 - rock and reggae	243 - rock
174 - MPB, reggae and rock	244 - rock
175 - rock and reggae	245 - blues and reggae
176 - reggae and rock	246 - MPB, reggae and rock
177 - rock and reggae	247 - rock and reggae
178 - MPB, reggae and rock	248 - rock
179 - rock and reggae	249 - reggae and rock
180 - reggae and rock	250 - rock
181 - reggae, rock and MPB	251 - rock
182 - rock and reggae	252 - rock
183 - rock and reggae	253 - rock
184 - reggae and rock	254 - reggae, rock and MPB
185 - reggae, rock and MPB	255 - reggae and rock
186 - rock and reggae	256 - reggae, rock and MPB
187 - rock and reggae	257 - reggae and rock
188 - blues and reggae	258 - reggae and rock
189 - reggae and rock	259 - rock
190 - reggae and blues	260 - MPB, reggae and rock
191 - rock and reggae	261 - rock and reggae
192 - reggae and rock	262 - reggae, rock and blues
193 - blues and reggae	263 - reggae and rock
194 - rock and reggae	264 - reggae, rock and MPB
195 - rock and reggae	265 - reggae, blues, rock and MPB
196 - reggae and rock	266 - reggae, rock and MPB
197 - reggae, rock and MPB	267 - reggae and rock
198 - reggae and rock	268 - MPB
199 - rock and reggae	269 - rock
200 - reggae and rock	270 - rock
201 - rock and reggae	271 - rock
202 - reggae, MPB and rock	272 - rock and reggae
203 - blues	273 - reggae and rock
204 - reggae and rock	274 - reggae and rock
205 - MPB, reggae and rock	275 - reggae and rock
206 - reggae and rock	276 - reggae, rock and MPB
207 - reggae and rock	277 - MPB and rock
208 - MPB, reggae and rock	278 - blues
209 - reggae and rock	279 - reggae, rock and MPB
210 - reggae and rock	280 - rock

Tabela 6.10 – Taxonomia de estilos para o sistema de classificação

Estilos	Quantidade de Músicas
Blues-Rock	15
Chicago-Blues	15
Funk	15
Rap-hardcore	15
Hard-Rock	15
Jazz-Soul	15
Metal	15
Rock-Psicodélico	15
Punk	15
Salsa	15
Total: 150	

O desempenho de reggae também se manteve o mesmo (66%), dado que nenhuma das amostras de blues, MPB ou rock foram também rotulados como reggae. Rock obteve desempenho de 26% $((19 + 2 + 7)/107)$, uma vez que duas amostras de reggae-rock estão no grupo 3, e 7 amostras de blues-rock estão no grupo 3. Por fim, blues obteve desempenho de 30% $((22/73))$, dado que nenhuma das 3 amostras de rock-blues estão presentes no grupo 4 (referenciado como o grupo do gênero blues).

6.4 Análises com uma taxonomia de estilos

Com o intuito de complementar as análises anteriores, verificou-se o comportamento dos métodos de aprendizagem para uma taxonomia de dez estilos musicais. Vale lembrar que neste trabalho o termo “estilo” refere-se a variações dentro de gênero, como, por exemplo, hard-rock ou chicago-blues. Os atributos referem-se aos vetores $\mathbf{V}^m$ de cada música, ou seja, indicam as probabilidades condicionais da cadeia de Markov de primeira ordem.

A taxonomia de estilos está exposta na Tabela 6.10. Estes arquivos foram, em sua maioria, obtidos através do banco de dados *Bodhidharma MIDI Database* (11), disponível para fins acadêmicos. Apesar da quantidade pequena de músicas por estilo, uma vez que nem todos os arquivos possuem a faixa da percussão, foi possível o levantamento de evidências sobre a caracterização e discriminação dos mesmos.

A Figura 6.13a e 6.13b apresentam, respectivamente, as duas primeiras componentes (novos atributos) PCA e LDA para a matriz de dados $\mathbf{P}^p$ das 150 músicas da Tabela 6.10. Apesar de ocorrer uma sobreposição considerável de músicas na plotagem PCA, observa-se que alguns grupos são mais visíveis que no caso da Figura 6.3a (veja, por exemplo, a aglomeração das músicas de chicago-blues, jazz-soul e punk). As duas componentes da

(a)

(b)

Figura 6.13 – Primeiro e segundo novos atributos obtidos pela aplicação das técnicas – (a) PCA e (b) LDA – na matriz de dados $\mathbf{P}^{\mathbf{P}}$ sobre taxonomia de 10 estilos

Figura 6.13a retêm 20% da variância dos dados. Na plotagem com os atributos LDA, fica ainda mais evidente que os dez grupos estão bem separados, em oposição aos quatro grupos da Figura 6.4a.

As Tabelas 6.11 e 6.12 apresentam os resultados de classificação. Mesmo com a presença de 6 classes a mais, observou-se que, em média, o desempenho foi superior que na situação dos quatro gêneros (Tabelas 6.2 e 6.3). As partições *Hold-Out* do PCA (Tabela 6.11)

Tabela 6.11 – Coeficiente kappa e desempenho de classificação dos novos atributos do PCA computados a partir da matriz de dados **PP** da taxonomia de 10 estilos da Tabela 6.10

PCA	Classificador Gaussiano Quadrático			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	1	0	100%	Ótimo
Hold-Out 70%-30%	0.23	0.0052	31%	Razoável
Hold-Out 50%-50%	0.21	0.0032	29%	Razoável
	Classificador Gaussiano Linear			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	0.96	0.0003	96%	Quase perfeito
Hold-Out 70%-30%	0.3	0.0055	37%	Razoável
Hold-Out 50%-50%	0.27	0.0033	34%	Razoável

Tabela 6.12 – Coeficiente kappa e desempenho de classificação dos novos atributos do LDA computados a partir da matriz de dados **PP** da taxonomia de 10 estilos da Tabela 6.10

LDA	Classificador Gaussiano Quadrático			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	1	0	100%	Ótimo
Hold-Out 70%-30%	0.98	0.0002	98%	Quase perfeito
Hold-Out 50%-50%	0.97	0.0004	97%	Quase perfeito
	Classificador Gaussiano Linear			
	Kappa	Variância	Desempenho	Interpretação
Re-Substitution	0.99	0.0005	99%	Quase perfeito
Hold-Out 70%-30%	0.99	0.0004	99%	Quase perfeito
Hold-Out 50%-50%	0.99	0.0005	99%	Quase perfeito

apresentaram desempenho inferior, pois são diretamente afetadas pelo tamanho limitado de amostras disponíveis. O mesmo não ocorreu para as partições *Hold-Out* do LDA (Tabela 6.12), refletindo a influência da sua característica supervisionada para a estimação das matrizes de covariância. A estimação de uma única matriz de covariância contribuiu para um suave aumento das tarefas *Hold-Out*, novamente pela consequência do número de amostras. Por outro lado, quando todas as amostras estão disponíveis (mesmo que apenas 15 de cada classe), as matrizes de covariância estimadas de cada classe permitem uma classificação de 100% em ambos os casos PCA e LDA. Isto não ocorreu para a situação de quatro gêneros. Através destes resultados, evidencia-se que os atributos do ritmo também podem ser utilizados para a caracterização dos estilos, e que tal caracterização ocorre de forma mais eficaz que para os gêneros.

A Tabela 6.13 mostra a matriz de confusão para o agrupamento hierárquico dos 10 estilos estudados. Verifica-se que em nove dos dez grupos, há a dominância de um estilo. Seguindo a mesma estratégia para obtenção do desempenho para o agrupamento de dados, tem-se que para a situação de 10 estilos, o desempenho foi de 33%.

Tabela 6.13 – Matriz de confusão para o agrupamento hierárquico aglomerativo dos 10 estilos

	Blues-Rock	Chicago-Blues	Jazz-Soul	Hard-Rock	Rock-Psic	Rap-HC	Metal	Punk	Funk	Salsa
Grupo 1	2	0	0	0	0	0	0	0	0	0
Grupo 2	2	8	2	0	0	0	0	0	1	0
Grupo 3	1	4	5	0	0	1	0	0	0	0
Grupo 4	2	2	1	3	0	0	1	0	0	0
Grupo 5	0	0	0	0	3	1	2	2	1	1
Grupo 6	2	0	2	4	3	7	3	4	5	1
Grupo 7	1	0	1	2	0	0	3	0	1	1
Grupo 8	3	0	2	1	6	3	2	7	4	1
Grupo 9	2	0	2	3	3	0	2	0	2	1
Grupo 10	0	1	0	2	0	3	2	2	1	10

6.5 Considerações finais

Este capítulo apresentou as estratégias de classificação automática de gêneros musicais de acordo com os atributos do ritmo. Ambas as abordagens supervisionadas e não supervisionadas foram investigadas.

Os resultados obtidos com a técnica LDA são particularmente significativos pois demonstram a natureza intrínseca dos gêneros e dos estilos. Apesar de altamente conhecidos, termos como rock, reggae ou pop permanecem não claramente definidos. Além disso, vale ressaltar que a intensidade da batida ou a informação do tempo, aspectos importantes na definição do ritmo não estão sendo considerados. Isso significa que analisar o ritmo apenas através de sequências de notas de valores, como atualmente desenvolvido, poderia impor dificuldades de classificação até mesmo para especialistas. As músicas classificadas de forma incorreta possuem propriedades similares descritas em termos das notas de valores e, como resultado, estabelecem matrizes de transição parecidas. Estas músicas podem contribuir para as redundâncias presentes nas taxonomias, pois fazem parte de um espaço de atributos comum a mais de um gênero.

Entretanto, a informação disponível do ritmo foi mais sensível para a representação dos estilos. Constatou-se que uma taxonomia de dez estilos permitiu resultados de classificação próximos ou superiores aos provenientes de uma taxonomia de quatro gêneros, ainda que no primeiro caso o número de amostras seja reduzido (15 para cada estilo). Vale mencionar que os estilos avaliados não estão significativamente representados no banco de dados usado para a análise dos gêneros.

Espera-se que as características musicais que definem um estilo sejam mais claras e definidas que para um gênero, por este ser mais amplo e, em muitos casos, abranger diversos estilos. Os resultados evidenciam, de fato, que a classificação automática de estilos musicais é menos complexa que a de gêneros. Uma análise mais minuciosa dos resultados da análise de gêneros poderia revelar subgrupos melhor definidos que poderiam corresponder a estes resultados encontrados com o banco de dados de estilos.

A despeito destas considerações, a metodologia proposta, apesar de requerer complementações relevantes, fornece uma contribuição pertinente na direção do desenvolvimento de uma abordagem alternativa para a classificação automática de gêneros musicais.

Tendo em vista que o problema de classificação automática de gêneros e estilos musicais não é uma tarefa trivial, os resultados indicam que mecanismos simples da modelagem do ritmo podem contribuir para a discriminação destes. Esta é uma das principais contribuições desta investigação. Apesar da dificuldade da comparação de desempenho entre trabalhos anteriores que exploram a classificação automática dos gêneros musicais, tem-se que a taxa de acerto média para as abordagens supervisionadas está no intervalo de 60%-90%. Geralmente estas taxas de desempenho são alcançadas através da combinação de diversos atributos ou combinação de classificadores.

As taxas de desempenho atingidas nas estratégias supervisionadas deste presente estudo estão no intervalo de 50%-90%. Apesar de serem taxas comparáveis às anteriores, tem-se uma percepção mais clara de quais aspectos influenciam a discriminação dos gêneros. Observou-se que as dependências entre as notas de valores da percussão podem ser úteis para a separação entre diferentes gêneros. Apenas através desta dependência, taxas comparáveis de desempenho foram atingidas.

Além disso é interessante a possibilidade de progressivamente investigar a contribuição de diferentes atributos para o entendimento dos gêneros. A presente análise viabiliza este estudo. Por exemplo, verificou-se que a dependência de três notas de valores fornece informações relevantes que atuam na discriminação e que não estão presentes nas dependências de pares de notas. Aspectos como tempo e intensidade da batida, por exemplo, podem ser inseridos gradativamente e a intervenção de cada um no contexto de caracterização dos gêneros pode ser estudada de forma mais sistemática, constituindo assim uma abordagem *bottom-up* para distinção dos gêneros.

O uso de uma classificação multi-categorias é particularmente apropriada, uma vez que reflete a natureza intrínseca do banco de dados. A classificação de gêneros musicais pode não ser uma tarefa simples até mesmo para especialistas, dado que uma música pode ser atribuída a mais de um gênero. Classificação automática de gêneros é uma tarefa mais complexa, pois muitas músicas podem conter vetores de atributos similares, definindo grandes áreas de sobreposição no espaço de atributos, como observado nos resultados deste capítulo. Dentro deste contexto, a classificação multi-categorias tem papel importante, tornando possível a generalização da taxonomia de gêneros, originalmente disponível no banco de dados.

Os resultados sugerem que, entre os gêneros estudados, blues é o mais discriminativo, seguido de MPB e rock. Reggae é o gênero no qual os vetores de atributos apresentam menor dispersão e, assim, são geralmente classificados corretamente. Porém, muitos vetores de atributos dos outros gêneros, principalmente de rock e MPB estão posicionados no espaço de atributos do reggae, o que acarreta em muitas classificações incorretas.

Um das principais motivações deste trabalho, é explorar a discriminação dos gêneros à luz da detecção de comunidades em redes de música. Em tais redes, os vértices representam as músicas do banco de dados, e as arestas retratam a relação de similaridade entre as

duas músicas conectadas. Algoritmos de detecção de comunidades utilizam propriedades topológicas das redes no processo de detecção de grupos. Estas informações adicionais podem favorecer a caracterização dos gêneros. Os resultados desta investigação estão abordados no próximo capítulo.

Capítulo 7

Detecção de comunidades em redes de música

Este capítulo apresenta os resultados principais para as tarefas de detecção de comunidades em redes de música para a discriminação de gêneros musicais. A motivação principal é promover uma comparação com os métodos do capítulo anterior e verificar se a inclusão de medidas topológicas destas redes podem favorecer a caracterização dos gêneros.

A estratégia desenvolvida para construir as redes de músicas, nas quais os vértices representam as músicas do banco de dados e as arestas a relação de similaridade entre as duas músicas conectadas está resumida na Figura 7.1 e descrita na seção 7.2. Tal estratégia foi aplicada sobre a taxonomia de gêneros da seção 7.1. Os resultados estão expressos na seção 7.3, enquanto que as considerações finais deste capítulo estão relatadas na seção 7.4.

Figura 7.1 – Diagrama dos passos principais para a detecção de comunidades em redes de música

Tabela 7.1 – Taxonomia de gêneros para a construção de redes de música e detecção de comunidades nestas redes

Gêneros	Quantidade de Músicas
Blues	70
MPB	70
Reggae	70
Rock	70
Total: 280	

7.1 Taxonomia adotada

Para a detecção de comunidades em redes de música foi utilizada uma taxonomia de quatro gêneros musicais conforme apresentado na Tabela 7.1.

7.2 Estratégia

As redes de música são construídas de forma que cada vértice representa uma música através do vetor de atributos $\mathbf{VB}^m$, resultante da concatenação dos vetores de atributos $\mathbf{V}^m$ (cadeia de Markov de primeira ordem) e $\mathbf{B}^m$ (cadeia de Markov de segunda ordem) ^a.

O peso da aresta que conecta dois vértices nas redes de música reflete a distância entre seus respectivos vetores de atributos, de acordo com alguma medida de distância. Se cada vértice está conectado a todos ou outros da rede, tem-se uma rede totalmente conectada. Entretanto, uma rede totalmente conectada dificulta a análise de estruturas intrínsecas (60). Propõe-se neste trabalho que apenas algumas conexões de cada vértice sejam mantidas. A abordagem utilizada para definir quais vértices serão conectados a um determinado vértice é descrita a seguir. Se o número mantido de conexões é igual para todos os vértices da rede, tem-se uma rede k -regular, em que k representa o número de conexões mantidas. Diferenças na construção das redes analisadas são consequências da escolha de distintas medidas de distância entre vértices, diferentes valores de K , e da execução ou não da etapa 3 a seguir.

1. A partir da matriz de dados ($\mathbf{P}^{ps}$), a distância entre pares de vetores de atributos (pares de músicas) é computada. Esse procedimento estabelece uma matriz simétrica de distâncias e uma rede totalmente conectada.

^a Vale mencionar que redes de gêneros construídas através do uso das cadeias de Markov de primeira e segunda ordens de forma separadas também foram analisadas. O o algoritmo de detecção de comunidades proposto por Clauset-Newman-Moore (75) identificou 5 grupos em cada caso, ao passo que com a concatenação das duas cadeias para formar o vetor de atributos, a aplicação do algoritmo de detecção de comunidades permitiu a formação de 4 grupos.

2. Cada vértice é conectado com apenas os seus K vizinhos mais próximos. O peso de cada aresta é definido por $1/d$, em que d é distância entre os respectivos vértices. Esta configuração estabelece uma rede k -regular.
3. Para cada vértice, calcula-se o peso médio das arestas, considerando os K vizinhos conectados. Mantém-se a aresta somente se o peso correspondente for maior que o peso médio.

Com as redes de música resultantes da etapa 2, algoritmos de detecção de comunidades serão testados com o objetivo de determinar grupos de músicas e constatar se tais grupos correspondem aos gêneros da taxonomia utilizada. Neste caso, adotou-se $k = 10$. Para a aplicação do particionamento espectral, adotou-se $k = 30$ e foi adicionado mais um passo para a construção das redes, indicado pelo passo 3. A etapa 3 estabelece redes de música não mais k -regulares. A motivação para o uso de redes não regulares está relacionada com o fato de que, se um grafo é k -regular, todos os autovalores da matriz Laplaciana, exceto pelo autovalor $av_N = 0$, são iguais a k . Nesta situação, a determinação do vetor de Fiedler (autovetor associado ao segundo menor autovalor da matriz Laplaciana), que reflete a conectividade do grafo, fica menos precisa (77).

7.3 Resultados e discussões

A rede de gêneros apresentada na Figura 7.2 foi obtida utilizando a distância do cosseno e com a aplicação do algoritmo de detecção de comunidades por Clauset-Newman-Moore (75). O algoritmo de Clauset-Newman-Moore é clássico na literatura de redes complexas e se baseia em similaridades ou dissimilaridades entre pares de objetos na obtenção dos grupos. Portanto, uma medida de similaridade entre nós precisa ser previamente definida. Tal medida pode ser estrutural, ou pode ser baseada em atributos através das propriedades dos nós (58). Este algoritmo desempenha um agrupamento hierárquico dos nós.

As músicas estão configuradas nas comunidades conforme a Tabela 7.2. Cada comunidade possui um gênero dominante. Amostras de blues e MPB estão mais concentradas nos grupos G3 e G1, respectivamente. Amostras de reggae estão proporcionalmente divididas nos quatro grupos. Aproximadamente metade das amostras de rock estão concentradas no grupo G1, sobrepondo-se com as amostras de MPB. O grupo G3 apresenta as músicas de blues que são mais discriminativas. O grupo G1 reflete principalmente a sobreposição entre as amostras de MPB, reggae e rock. Por fim, os grupos G2 e G4 refletem principalmente a sobreposição entre as amostras de reggae e rock.

Seguindo a estratégia do capítulo anterior (seção 6.3.3) para a obtenção do desempenho, tem-se que para a configuração da Tabela 7.2, a taxa de acerto foi de 44%. Este resultado

Figura 7.2 – As comunidades da rede 10-regular de gêneros musicais. Distância do cosseno entre vetores de atributos. Grupos obtidos pelo algoritmo de Clauset-Newman-Moore.

Tabela 7.2 – Configuração dos grupos da rede apresentada 10-regular com a distância do cosseno entre vetores de atributos e uso do algoritmo de Clauset-Newman-Moore.

	MPB	Rock	Blues	Reggae
G1	41	29	15	21
G2	4	13	1	11
G3	17	15	50	20
G4	8	13	4	18

foi superior aos 39.4% do resultado hierárquico apresentado no capítulo anterior.

Para a mesma rede descrita acima (10-regular com a distância do cosseno), a Figura 7.3 apresenta as nove comunidades obtidas de acordo com o algoritmo de Girvan e Newman (70, 73). Esse algoritmo utiliza os índices das arestas com centralidade de informação e medidas de modularidade para detectar as comunidades (seção 3.4.3). Vale comentar que muitas amostras de um mesmo gênero foram atribuídas no mesmo grupo. A concentração de amostras de um mesmo gênero nestas comunidades menores motiva estudos futuros para analisar a presença de estilos. Por exemplo, poderia ser interessante verificar se amostras de *blues-rock* ou *pop-rock* estão mais concentradas em uma comunidade.

A configuração dos grupos para a rede de nove comunidades está disposta na Tabela 7.3. Seguindo a estratégia de determinar as classificações corretas como sendo as classificações do gênero dominante de cada grupo, o desempenho deste algoritmo foi de 46%, com nove grupos.

Considerando a distância Euclidiana ao invés da distância do cosseno, as comunidades obtidas para a rede 10-regular com o uso do algoritmo Clauset-Newman-Moore estão

Figura 7.3 – A rede 10-regular de gêneros musicais. Distância do cosseno entre vetores de atributos. Grupos formados pelo algoritmo de Girvan e Newman

Tabela 7.3 – Configuração dos grupos da rede formada com a distância do cosseno e o emprego do algoritmo por Girvan e Newman para detecção de comunidades.

	Blues	MPB	Reggae	Rock
G1	12	18	6	17
G2	2	9	5	
G3	3	13	9	10
G4	13	6	8	4
G5	3	5	16	6
G6	0	5	10	14
G7	3	12	5	2
G8	4	6	3	5
G9	2	3	4	7

ilustradas na Figura 7.4. A Tabela 7.4 apresenta a configuração dos grupos. As principais diferenças são: amostras de reggae estão mais concentradas (31 no grupo G3); e o grupo G4 é menor que no primeiro caso com a distância do cosseno (Tabela 7.2), com apenas 12 amostras. Nesta configuração o desempenho de acerto foi de 41%.

Tabela 7.4 – Configuração dos grupos da rede apresentada 10-regular com a distância Euclidiana entre vetores de atributos e uso do algoritmo de Clauset-Newman-Moore.

	MPB	Blues	Reggae	Rock
G1	35	18	15	30
G2	19	41	23	10
G3	13	10	31	23
G4	3	1	1	7

Figura 7.4 – As comunidades da rede 10-regular de gêneros musicais. Distância Euclidiana entre vetores de atributos. Grupos obtidos pelo algoritmo de Clauset-Newman-Moore.

Comparando as comunidades da Tabela 7.4 com o agrupamento hierárquico apresentado anteriormente (os dois casos utilizam a distância Euclidiana entre vetores de atributos), tem-se que no primeiro caso as amostras de blues estão mais concentradas em um mesmo grupo, e que o grupo maior não concentra muitas amostras de todos os gêneros, o que não acontece na situação do agrupamento hierárquico. Esta verificação fornece outra evidência de que as informações topológicas inseridas no processo de detecção de comunidades nas redes de música, pode contribuir para melhorar o desempenho da classificação.

Resultados do particionamento espectral

A Figura 7.5 ilustra a rede de músicas resultante para a detecção de quatro comunidades.

O vetor de Fiedler para a rede da Figura 7.5 e o histograma correspondente da distribuição dos coeficientes estão apresentados nas Figuras 7.6 e 7.7, respectivamente. A distribuição dos valores dos coeficientes do autovetor associado ao segundo menor autovalor da matriz Laplaciana, sugere a presença de diferentes comunidades na rede de música (indicação de três vales no histograma).

A Tabela 7.5 apresenta a configuração dos grupos para o particionamento espectral. As principais diferenças para os casos anteriores são que as músicas de rock e MPB estão mais espalhadas entre os quatro grupos. Para o particionamento espectral, o desempenho foi de 41%.

Figura 7.5 – A rede de gêneros de acordo com o vetor Fiedler

Figura 7.6 – Vetor de Fiedler da rede apresentada na Figura 7.5

Tabela 7.5 – Configuração dos grupos da rede de gêneros de acordo com o vetor Fiedler.

	Reggae	Blues	Rock	MPB
G1	20	4	13	8
G2	21	46	6	15
G3	17	9	23	20
G4	12	11	18	27

Figura 7.7 – Distribuição dos valores dos coeficientes do vetor de Fiedler da rede apresentada na Figura 7.5

7.3.1 Inclusão de medidas topológicas

Para completar a análise do ritmo, diversas medidas de dígrafos foram extraídas e aplicadas nas tarefas de classificação. Para os dígrafos formados com os valores das cadeias de Markov 1D, as medidas utilizadas foram: grau total (grau de entrada + grau de saída), força total e coeficiente de aglomeração. Essas medidas foram computadas para cada vértice. Para obter os valores finais para cada dígrafo, considerou-se a entropia, valor médio, valor máximo, valor mínimo e desvio padrão de cada uma delas.

Dessa forma, a matriz de atributos resultante contém 280 linhas e 15 colunas. Verificou-se uma acurácia de 61.78% e $\kappa = 0.49$ através do classificador Bayesiano com a tarefa *Re-Substitution*. A combinação destes 15 atributos com os 324 do vetor $\mathbf{V}^m$ resultou numa acurácia de 85.72% e $\kappa = 0.81$, com a mesma estratégia de classificação.

Para os dígrafos da cadeia de Markov de segunda ordem, as medidas foram: grau de entrada e saída, e força de entrada e saída. Optou-se por não usar o coeficiente de aglomeração porque nestes dígrafos os vértices de notas individuais não possuem conexões de saída. Assim, o cálculo desta medida poderia ficar prejudicado. Para cada uma destas medidas utilizou-se valor médio, máximo, o desvio-padrão, e valor médio em conjunto (por exemplo, grau médio: considerando grau médio de entrada e grau médio de saída). A matriz de atributos resultante contém 280 linhas e 14 colunas. Desta forma, buscou-se garantir que os vetores de medidas dos dígrafos de primeira e segunda ordem tivessem quantidades próximas de atributos.

Para a classificação supervisionada, o desempenho de classificação foi de 60%, enquanto que o valor do κ foi $\kappa = 0.47$. A combinação destes 14 atributos com os 3006 do

vetor $\mathbf{B}^m$ resultou em um desempenho de 93% e $\text{kappa} = 0.9$, com a mesma estratégia de classificação.

Estes resultados sugerem que as medidas topológicas contribuem para melhorar o desempenho de tarefas de classificação automática dos gêneros musicais. De fato, tais medidas ressaltam as relações entre as notas, mais do que os vetores de atributos propriamente ditos. No capítulo 5 e nas análises da Figura 6.5 do capítulo 6, foi possível ressaltar pares de notas que tem papéis fundamentais na representação e na diferenciação dos gêneros. As medidas topológicas evidenciam informações globais sobre a relação entre as notas do ritmo que são sensíveis aos gêneros musicais.

7.4 Considerações finais

Através dos atributos rítmicos obtidos pelo uso das cadeias de Markov de primeira e segunda ordens, diferentes abordagens foram examinadas para a detecção de comunidades nas redes de gêneros musicais. As principais variações dessas abordagens se aplicam na escolha de diferentes medidas de distância entre vetores de atributos e diferentes algoritmos de agrupamento de dados em redes.

A classificação por detecção de comunidades em redes também se configura como uma abordagem não supervisionada, que apresentou desempenho superior em comparação aos resultados de classificação do agrupamento hierárquico. Além disso, informações adicionais podem ser levantadas ao examinar as redes de música. Por exemplo, amostras que estão nas fronteiras das comunidades podem representar as músicas que contribuem para redundâncias na taxonomia. Por outro lado, grupos definidos nas comunidades podem revelar músicas mais características de cada gênero, assim como artistas que mantêm tais características nas composições.

Analisando as comunidades, verifica-se que blues se mantém como a classe mais discriminativa, e que, na maioria dos casos, um grupo é sempre pequeno, em torno de 10 amostras. Esse aspecto sugere que, apesar da presença de quatro gêneros de acordo com a taxonomia usual, de acordo com os atributos rítmicos e amostras analisados, pode haver apenas três. Em outras palavras, os ritmos estão repetidos em gêneros diferentes. Esta sobreposição revela que alguns aspectos musicais podem coexistir entre os gêneros, o que não deixa de ser um aspecto interessante que pode ser mais cuidadosamente explorado.

Caminhadas nestas redes podem ser aplicadas para a geração automática de listas de reprodução, na qual músicas subsequentes são similares de acordo com o critério adotado para a construção da rede. Esta tese analisou esta possibilidade, de forma que um algoritmo para sistemas de recomendação musical foi desenvolvido. Esta investigação está apresentada no próximo capítulo.

Capítulo 8

Sistema de recomendação musical

Este trabalho investigou uma metodologia para a geração de listas automáticas de reprodução. Nestas listas, músicas subsequentes são similares de acordo com os atributos do ritmo gerados pela aplicação das cadeias de Markov conforme descrito no Capítulo 4.

Esta investigação foi aplicada em quatro gêneros musicais conforme a taxonomia da seção 8.1. A estratégia desenvolvida para a obtenção de uma lista sequencial de músicas consiste em caminhadas em uma árvore geradora mínima. Esta árvore foi construída a partir de uma rede de músicas em que cada vértice representa uma música do banco de dados através do vetor de atributos $\mathbf{VB}^m$ (seção 4.2). Esta estratégia está resumida na Figura 8.1 e detalhada na seção 8.2. A seção 8.3 apresenta uma possível árvore geradora mínima e discute características da lista de reprodução proveniente da mesma.

8.1 Taxonomia adotada

Para este estudo foram utilizados os quatro gêneros musicais da Tabela 8.1.

Tabela 8.1 – Taxonomia de gêneros para o sistema de recomendação musical

Gêneros	Quantidade de Músicas
Blues	70
MPB	70
Reggae	70
Rock	70
Total: 280

Figura 8.1 – Diagrama dos passos principais para a geração automática de listas de reprodução

8.2 Estratégia

A estratégia de busca desenvolvida é baseada em caminhadas em uma árvore geradora mínima (AGM) sem repetição de vértices. A AGM é obtida a partir da rede totalmente conectada na qual cada vértice representa uma música do banco de dados através de seu respectivo vetor de atributos $\mathbf{VB}^m$ (concatenação dos vetores $\mathbf{V}^m$ e $\mathbf{B}^m$ resultantes das probabilidades condicionais das cadeias de Markov de primeira e segunda ordem). O peso das arestas estabelece a distância dos vetores de atributos entre pares de vértices.

A distância do cosseno foi a medida escolhida para quantificar a similaridade entre os vetores de atributos (seção 3.3.1). Verificou-se em (84) que a distância do cosseno pode ser apropriada principalmente em situações em que a dimensionalidade dos vetores de padrões é alta.

A partir da rede de músicas, o algoritmo de Prim (85, 86) foi aplicado para a obtenção da AGM. Basicamente, o algoritmo realiza os seguintes passos: a partir de um vértice

escolhido aleatoriamente, escolhe-se a aresta com o peso mínimo que conecta um vértice v , do conjunto de vértices já visitados, a um outro vértice k , pertencente ao conjunto de vértices não visitados. O algoritmo termina quando $N - 1$ arestas são selecionadas, onde N é o número de vértices da rede.

Comparou-se a abordagem de caminhadas na AGM com uma abordagem simples dada pela busca em largura (BL). No caso da BL, o objetivo é formar uma lista de músicas visitando todos os vértices que estão a uma distância $d = 1$ do vértice inicial, na sequência visitar todos os vértices que estão a uma distância $d = 2$, e assim por diante até que todos os vértices sejam visitados ou até que uma profundidade máxima seja alcançada. Entretanto, na BL os pesos das arestas são completamente ignorados. Com isso, assume-se a hipótese de que todas as músicas são equidistantes no espaço de atributos, o que pode não ser interessante no contexto de listas de reprodução. Além disso, o conjunto de vértices que estão localizados a uma distância $d > 1$ do vértice inicial, em geral, não são conectados nem na AGM, nem no grafo original, o que significa que essas amostras podem não estar necessariamente próximas o suficiente no espaço de atributos. Esse aspecto se intensifica para grandes valores de d , em que as amostras equidistantes podem, inclusive, pertencer a diferentes gêneros.

Caminhadas na AGM, entretanto, podem atingir um vértice folha, o que significaria o fim de um caminho. A estratégia desenvolvida, nomeada neste trabalho como algoritmo *jumping walk* (JW), lida com essa situação ao pular para o vizinho mais próximo ainda não visitado e iniciando uma nova caminhada.

8.2.1 Algoritmo *Jumping-walk*

O algoritmo *jumping walk* consiste em sequências de caminhadas em uma AGM de uma rede de músicas. O primeiro passo do algoritmo é escolher randomicamente um vértice inicial de um determinado gênero, que pode ser dado como entrada pelo usuário. Após isso, uma sequência de vértices é definida pela contínua escolha do vizinho mais próximo (aresta com o peso mínimo) ainda não visitado para ser a próxima música da lista. Se a música selecionada é uma folha, então os vértices que estão a uma distância $d = 2, 3, 4, \dots$ são verificados, até que um vértice ainda não visitado seja encontrado. Se houver mais de uma opção, o vértice que oferece menor custo é escolhido, o que é equivalente a encontrar o caminho mais curto entre a folha e um vértice ainda não selecionado de acordo com o algoritmo Dijkstra (60). Mesmo ocorrendo um pulo, e, como consequência, uma descontinuidade na sequência, pelas propriedades da AGM é garantido que a próxima música da sequência é a mais similar ainda não tocada (considerando a árvore).

Um problema potencial com esse método é a geração de listas de reprodução com L

Figura 8.2 – Uma ilustração do algoritmo *jumping walk*

músicas, onde $L \approx N$. Com a regra de caminhar sempre para o próximo vizinho, alguns ramos da AGM podem ser negligenciados, o que pode causar pulos grandes conforme o número de vértices não visitados se aproxima de zero. Duas formas de evitar esse problema são: 1) ter uma quantidade grande suficiente de músicas; 2) interromper as caminhadas sempre que o número de vértices visitados atinja um limiar superior previamente definido (por exemplo, $0.75 * N$), e estabelecer um novo começo para a lista.

A Figura 8.2 ilustra o processo do salto. Considere A como o vértice inicial. A sequência de vértices visitados antes de atingir a folha **D** é A, B e C. Após isso, é efetuado um salto para o vértice mais próximo não visitado. Nota-se que as arestas AB e BC estão no caminho mínimo do vértice C para o D e também no caminho mínimo do vértice C para o vértice E. Portanto, a próxima música selecionada (D, nesse caso) será a mais similar tanto de C (posição i na lista) quanto de A (posição $i - 2$ na lista).

8.3 Resultados e discussões

A Figura 8.3 apresenta a AGM para a rede de músicas com a distância do cosseno entre os vetores de atributos. Nota-se que muitas músicas de um mesmo gênero estão localizadas em ramos de uma mesma sub-árvore, como as músicas de blues na ramificação inferior esquerda da árvore, ou as músicas de MPB, mais aglomeradas na parte inferior direita. Músicas de reggae e rock estão mais espalhadas, apesar de haver uma aglomeração de músicas de reggae na parte superior esquerda da árvore.

Figura 8.3 – Árvore geradora mínima para a rede de gêneros considerando o algoritmo de Prim (85) e distância do cosseno entre os vértices

Ao aplicar o algoritmo JW, escolhendo a música 39 como o vértice inicial, as primeiras 25 músicas da lista de reprodução estão indicadas a seguir. O gênero de cada música também está especificado na lista.

- 39 - (*blues*) Johnny Winter - Good morning little schoolgirl
- 50 - (*blues*) Stevie Ray Vaughan - Cold shot
- 62 - (*blues*) Stevie Ray Vaughan - Tell me
- 23 - (*blues*) Delmore Brothers - Blues stay away from me
- 203 - (*reggae*) Third World - Now that we've found love
- 25 - (*blues*) ElvisPresley - A mess of blues
- 44 - (*blues*) Ray Charles - Born to the Blue
- 12 - (*blues*) Barbra Streisand - Am I Blue?
- 38 - (*blues*) John Lee Hooker - One bourbon one scotch one beer
- 16 - (*blues*) Boy Williamson - Dont start me talking
- 18 - (*blues*) Boy Williamson - Keep it to yourself
- 152 - (*reggae*) Bob Marley - I shot the sheriff
- 36 - (*blues*) John Lee Hooker - Boom boom boom
- 3 - (*blues*) AlbertKing - Stormy monday
- 34 - (*blues*) Jimmie Cox - Before you accuse me
- 190 - (*reggae*) Natiruts - Liberdade pra dentro da cabeca
- 5 - (*blues*) BB King - Dont answer the door
- 193 - (*reggae*) Nazarite Skank
- 263 - (*rock*) Rolling Stones - Angie
- 239 - (*rock*) Metallica - Fuel
- 245 - (*rock*) Metallica - Sad but true
- 29 - (*blues*) Freddie King - Hide away
- 170 - (*reggae*) Cidade Negra - Eu fui eu fui
- 40 - (*blues*) Koko Taylor - Hey bartender
- 59 - (*blues*) Stevie Ray Vaughan - Manic depression

Com o objetivo de estabelecer comparações entre listas de reprodução, a busca em largura foi desempenhada na AGM, e no dígrafo na Figura 7.2 do Capítulo 7. Também é realizada uma comparação através de uma versão ponderada da busca em largura (BLP). Neste caso, as músicas disponíveis para uma distância d são ordenadas conforme o custo em relação ao vértice inicial (equivalente a aplicar o algoritmo Dijkstra entre o vértice inicial e os vértices para uma distância d e ordenar estes vértices de acordo com o custo mínimo que eles oferecem em relação ao vértice inicial).

Para as primeiras 100 músicas das listas de reprodução (usando a música 39 como vértice inicial), a distância e a alternância de gêneros entre músicas subsequentes na lista são analisadas. Para estes dois aspectos, o algoritmo JW é comparado com a BL e BLP na AGM, e com a BL e BLP no dígrafo 10-regular.

Pela Figura 8.4 é possível observar diferentes comportamentos de similaridade dependendo da estratégia adotada. Por exemplo, a maioria dos valores de distância é menor que 1 para o algoritmo JW (Figura 8.4a), enquanto que o contrário ocorre para a busca em largura no dígrafo regular (Figura 8.4c). A distância média entre músicas subsequentes também é apresentada para cada caso. Nota-se que o algoritmo JW possibilitou a distância média mínima entre as músicas. Isso significa que, em média, as músicas selecionadas estão mais ritmicamente relacionadas que nos outros casos. Eventualmente, a distância entre um par de vértices em um dígrafo pode ser não definida (como acontece na Figura 8.4c).

Figura 8.4 – Distâncias entre músicas subsequentes para as primeiras 100 músicas da lista de reprodução obtida pelo algoritmo JW e das listas de estabelecidas por diferentes variações da busca em largura

A Figura 8.5 mostra o comportamento das transições de gêneros entre músicas subse-

Figura 8.5 – Alternância de gêneros entre músicas subsequentes para as primeiras 100 músicas da lista de reprodução obtida pelo algoritmo JW e das listas de estabelecidas por diferentes variações da busca em largura

quentes conforme as estratégias de busca adotadas para comparação. De acordo com esta análise, a quantidade de transições entre gêneros com o algoritmo JW foi 57, enquanto que com a BL na AGM, BL no dígrafo 10-regular, BLP na AGM, e BLP no dígrafo 10-regular a quantidade de transições foi, respectivamente 60, 73, 70, e 72. Estes números indicam que a lista de reprodução gerada pelo algoritmo JW permitiu uma redução no número de transições entre os gêneros.

8.4 Considerações finais

Este capítulo apresentou a estratégia investigada para o desenvolvimento de um sistema de recomendação musical. O algoritmo desenvolvido é baseado em caminhadas nas árvores geradoras mínimas, e pode ser estendido para o emprego de outros atributos musicais. Músicas com configuração semelhante de notas de valores estão proximamente localizadas nas listas de reprodução. Dessa forma, o algoritmo não irá necessariamente gerar uma sequência de músicas pertencentes ao mesmo gênero, mas sim uma sequência na qual as músicas são mais similares entre si de acordo com os padrões rítmicos analisados.

É válido mencionar que diferentes AGMs (obtidas através de diferentes redes de música, por exemplo, com o uso de outras medidas de similaridade) resultam em diferentes listas de reprodução. De fato, múltiplas execuções do algoritmo de PRIM na mesma rede não garantem que a AGM obtida será a mesma. No contexto de geração de listas de reprodução, esses aspectos permitem uma variabilidade controlada.

Uma vez que músicas em posições próximas da lista possuem padrões rítmicos similares, uma relação pode ser estabelecida entre gênero e aspectos subjetivos como emoção/humor em música. Gabrielsson (106) descreve que ritmos regulares sugerem, em geral, expressões de emoção tais como “alegria” e “contentamento”. Por outro lado, ritmos irregulares sugerem, em geral, expressões de emoção/humor tais como “tensão” e “desconforto”. Desta forma, as músicas das listas de reprodução obtidas nesta investigação são ritmicamente relacionadas e podem estar emocionalmente ligadas.

Dentro deste prisma, este trabalho propôs uma análise inicial da relação entre gêneros musicais e expressões de emoção/humor através da presença de padrões temporais de notas nas músicas. Esta averiguação está descrita no próximo capítulo.

Capítulo 9

Detecção de motivos e relação com expressões de emoção/humor

Padrões temporais são elementos comuns em composições e usualmente referenciados como motivos musicais. Motivos musicais rítmicos são definidos como repetições de uma sequência de notas de valores. Motivos musicais melódicos são definidos como repetições de uma sequência de notas musicais (mais detalhes na seção 2.2). Este trabalho explorou de forma investigativa se músicas de diferentes gêneros podem ser discriminadas pela frequência de padrões temporais que se repetem ao longo da música.

Além do uso de um contexto musical estendido quando comparado com as cadeias de Markov de primeira e segunda ordens, verificou-se também a possibilidade de associar a presença destes padrões temporais com expressões de emoção/humor. Esta associação é motivada na relação entre fatores musicais e rótulos de emoção proposto por Gabriellson (106). A tabela 9.1 sumariza estas associações para três fatores musicais: ritmo, melodia e forma musical.

Tabela 9.1 – Expressões de emoção/humor de acordo com diferentes fatores musicais: ritmo, melodia e forma musical. Adaptado de (106)

Aspecto Musical	Característica	Rótulos de emoção mais comuns
Ritmo	Regular/previsível	felicidade, dignidade, glória, tranquilidade
	Irregular/complexo	divertido, desconforto, irritação
	Fluente	alegria, sereno, agradável.
Melodia	Longa extensão melódica	alegria, fantasia, desconforto
	Curta extensão melódica	tristeza, sentimental, delicado
	Movimento gradativo	tédio
Forma Musical	Muita complexidade	tensão, tristeza
	Pouca complexidade	alegria, paz, relaxamento

A presença de motivos melódicos e rítmicos foi verificada na melodia de músicas de sete diferentes gêneros, conforme a taxonomia apresentada na seção 9.1. A motivação para analisar a melodia está relacionada com o fato de que a melodia é um dos primeiros

aspectos musicais que permitem o reconhecimento e a memorização de uma música (132). Desta forma, parte-se da hipótese de que a melodia é um aspecto musical influente na percepção de emoção/humor em música.

A estratégia proposta para este estudo está resumida na Figura 9.1 e detalhada na seção 9.2. Resultados preliminares são apresentados na seção 9.3.

Figura 9.1 – Etapas da estratégia proposta para o estudo da ocorrência de motivos musicais na melodia de diferentes gêneros. A partir de arquivos MIDI, a faixa referente à melodia é extraída e representada por um vetor de notas, indicadas conforme os números MIDI respectivos (por exemplo, C5 é 72 e C4 é 60); e por um vetor de notas de valores conforme a duração relativa (por exemplo, a colcheia é representada por 0.5 enquanto que a semínima por 1)

9.1 Taxonomia adotada

Para esta investigação foram utilizados os sete gêneros musicais apresentados na Tabela 9.2.

Tabela 9.3 – Detecção de motivos melódicos da sequência de notas da Figura 9.2. A cada combinação de par de notas, combinação é verdadeira (V) quando as notas são iguais, ou falsa (F) caso contrário

vetor NA	60	62	64	62	60	62	62	57	60	62	64	62	60	72	72	71
deslocamento 1	-	60	62	64	62	60	62	62	57	60	62	64	62	60	72	72
combinações	-	F	F	F	F	F	V	F	F	F	F	F	F	F	V	F
vetor NA	60	62	64	62	60	62	62	57	60	62	64	62	60	72	72	71
deslocamento 2	-	-	60	62	64	62	60	62	62	57	60	62	64	62	60	72
combinações	-	-	F	V	F	V	F	F	F	F	F	V	F	F	F	F
...																
vetor NA	60	62	64	62	60	62	62	57	60	62	64	62	60	72	72	71
deslocamento 8	-	-	-	-	-	-	-	-	60	62	64	62	60	62	62	57
combinações	-	-	-	-	-	-	-	-	V	V	V	V	V	F	F	F

primeiro passo consiste em iterativamente comparar este vetor com versões deslocadas dele mesmo. Esta comparação é feita nota por nota, como ilustrado na Tabela 9.3. O tamanho do motivo é determinado pelo número de notas no vetor original que coincidem com as notas do vetor deslocado. A cada iteração, é computada a quantidade de motivos de diferentes tamanhos que ocorreram para este deslocamento. Por exemplo, a Tabela 9.3 mostra que ao deslocar o vetor NA em duas posições, tem-se como resultado três motivos de tamanho um, enquanto o deslocamento em oito posições resulta em um motivo de tamanho quatro, o qual representa o principal padrão melódico do exemplo adotado.

Este processo estabelece uma matriz $\mathbf{M}$ para cada música em que as linhas representam as iterações e as colunas expressam a quantidade de motivos de diferentes tamanhos que ocorreram para o deslocamento da respectiva iteração. Por exemplo, $\mathbf{M}(2, 1)$ indica a quantidade de motivos de tamanho um ocorridos quando o vetor NA foi deslocado duas vezes. Seguindo o exemplo da Tabela 9.3, $\mathbf{M}(2, 1) = 3$. O valor médio de cada coluna j da matriz $\mathbf{M}$ indica a frequência média com que motivos de tamanho j ocorreram na música que está sendo analisada. Por exemplo, o valor médio da coluna $j = 3$ estabelece a frequência média que motivos de tamanho três ocorreram considerando todas as iterações de deslocamento. Motivos menores que três não são atribuídos na análise, uma vez que estes não necessariamente representam um padrão de notas significativo, mas são esperados expressarem um caráter randômico. Este limiar foi escolhido de forma arbitrária. Para cada gênero, estabelece-se uma matriz $\mathbf{P}_j^m$ em que as colunas da linha i são os valores médios de cada coluna da matriz $\mathbf{M}$ para a música i pertencente ao gênero j . Ao computar a média das colunas da matriz $\mathbf{P}_j^m$, tem-se um vetor final $\mathbf{T}_j^m$. Os elementos deste vetor indicam a frequência média que motivos de tamanhos diferentes ocorrem no gênero j .

Os motivos rítmicos são obtidos seguindo a mesma abordagem através dos vetores NV. A Tabela 9.4 apresenta as comparações das notas de valores para o exemplo da Figura 9.2. Com nove deslocamentos do vetor NV, é possível determinar a principal sequência rítmica do exemplo. Comparando a Tabela 9.4 com a Tabela 9.3, observa-se que a informação

Tabela 9.4 – Detecção de motivos rítmicos da sequência de notas da Figura 9.2. A cada combinação de par de notas de valores, combinação é verdadeira (V) quando as durações são iguais, ou falsa (F) caso contrário

vetor NV	0.5	0.5	0.5	0.5	1	0.5	1	0.5	3	0.5	0.5	0.5	0.5	1	0.5	1	0.5	3
deslocamento 1	-	0.5	0.5	0.5	0.5	1	0.5	1	0.5	3	0.5	0.5	0.5	0.5	1	0.5	1	0.5
combinações	-	V	V	V	F	F	F	F	F	F	V	V	V	F	F	F	F	F
vetor NV	0.5	0.5	0.5	0.5	1	0.5	1	0.5	3	0.5	0.5	0.5	0.5	1	0.5	1	0.5	3
deslocamento 2	-	-	0.5	0.5	0.5	0.5	1	0.5	1	0.5	3	0.5	0.5	0.5	0.5	1	0.5	1
combinações	-	-	V	V	F	V	V	V	F	V	F	V	V	F	V	V	V	F
...																		
vetor NV	0.5	0.5	0.5	0.5	1	0.5	1	0.5	3	0.5	0.5	0.5	0.5	1	0.5	1	0.5	3
deslocamento 9	-	-	-	-	-	-	-	-	-	0.5	0.5	0.5	0.5	1	0.5	1	0.5	3
combinações	-	-	-	-	-	-	-	-	-	V	V	V	V	V	V	V	V	V

Figura 9.3 – O vetor NA de duas melodias. (a) Melodia de blues através da música *Sweet Sixteen* por *BB King*, e (b) melodia de rap através da música *B-Boys e Flygirls* por *Bomfunk MCs*

proveniente dos motivos rítmicos é diferente da informação trazida pela análise dos motivos melódicos, sendo estes últimos menos comuns no exemplo supracitado.

9.3 Resultados e discussões

Conforme descrito anteriormente, calcula-se os vetores $\mathbf{T}^m$ e $\mathbf{T}^r$, que indicam, respectivamente, a quantidade média que motivos melódicos e rítmicos de distintos tamanhos ocorrem em músicas de diferentes gêneros. Os motivos melódicos e rítmicos trazem informações sobre o contorno melódico e rítmico da melodia. Melodias mais previsíveis geralmente têm contornos constantes, com muitos trechos que se repetem. Para ilustrar este aspecto, a Figura 9.3 apresenta o contorno melódico através do vetor NA de duas melodias que representam os gêneros blues e rap. Enquanto que o contorno da melodia de rap é mais constante, o contorno da melodia de blues é mais dinâmico.

A Figura 9.4 apresenta os vetores NV das mesmas músicas da Figura 9.3. O contorno rítmico da melodia de blues também é mais dinâmico. A melodia de rap apresenta um padrão rítmico mais modular que a de blues, mas vale notar que seu contorno

Figura 9.4 – O vetor NV das melodias da Figura 9.3

rítmico apresenta mais variações quando comparado com o contorno melódico ilustrado na Figura 9.3b. Assim, ambos os aspectos rítmicos e melódicos são examinados para esta investigação preliminar da relação entre gêneros musicais e expressões de emoção/humor.

As Figuras 9.5 (a) e 9.5 (b) apresentam, respectivamente, os vetores $\mathbf{T}^m$ e $\mathbf{T}^r$ dos sete gêneros. Em ambos os casos, é indicado tamanho do motivo e a quantidade média de vezes que este tamanho ocorreu em cada gênero. As anotações de emoção/humor para os gêneros da taxonomia adotada foram obtidos principalmente pelo site *All Music Guide* (23).

Entre os gêneros estudados, rap é o gênero que apresenta a maior quantidade de motivos melódicos, e o gênero com segunda maior quantidade de motivos rítmicos. De acordo com (87), rap é conhecido como um gênero de discurso rítmico com rimas e de fluxo regular. É também conhecido que melodias de rap são bem simples, uma vez que, em geral, o texto da canção é mais importante do que a melodia em si. Artistas de rap recebem anotações de emoção/humor tais como “alegre”, “motivante” e “relaxante”. Quando comparado com a Tabela 9.1, as anotações estão relacionadas com expressões associadas ao ritmo regular, fluente e forma musical de complexidade baixa. Rap recebeu influências do reggae (87), compartilhando características como o uso de ritmos sincopados e regulares. Melodias de reggae são geralmente simples e com senso de fraseado (23). De acordo com os resultados, reggae é o segundo gênero com maior quantidade de motivos melódicos e o terceiro gênero em relação à quantidade de motivos rítmicos. Anotações comuns para álbuns de reggae são “reflexivo”, “relaxante” e “calmo”, em que os dois últimos parecem concordar com expressões ligadas à forma musical de complexidade baixa da Tabela 9.1.

Conforme os resultados da Figura 9.5, dance é o gênero com maior quantidade de motivos rítmicos. Este gênero tem uma batida rítmica definida e regular (23). Anotações comuns para álbuns de dance são “alegre”, “glória” e “animado”, e são similares às expressões da Tabela 9.1 para ritmo regular. A constância e modularidade em músicas de dance determinam uma característica intrínseca do gênero: de apresentar um ritmo constante que estimula os movimentos do corpo.

Figura 9.5 – A configuração dos motivos (a) melódicos e (b) rítmicos para a melodia dos gêneros musicais. Cada gênero está representado por uma cor. Enquanto que rap é o gênero com maior quantidade de motivos melódicos, dance tem maior quantidade de motivos rítmicos.

Os resultados sugerem que blues e country não agregam muitos motivos repetidos em suas melodias, tanto em aspectos melódicos ou rítmicos. Blues é uma origem estilística de country. As melodias destes dois gêneros são caracterizadas pela seleção de notas específicas (por exemplo, terças, quintas e sétimas alteradas), e o uso de extensões melódicas curtas (88). Ambos os gêneros recebem anotações de emoção/humor tais como “complexo”, “sofisticado”, “sentimental”, “melancólico” e “dramático”. Estes termos possuem significados similares aos encontrados na Tabela 9.1 para formas musicais de alta complexidade e melodia de curta extensão.

Músicas de rock e punk apresentaram estrutura rítmica e melódica simples, mas não tão constantes quanto as observadas para rap, dance ou reggae. Anotações comuns para álbuns de rock e punk são: “energético”, “alegre” e “emocionante”. Punk é conhecido como um gênero com progressões básicas de acordes e melodias simples (estimulada de uma maneira mais intensa e rápida (23)). De acordo com os resultados, principalmente em relação à análise rítmica da melodia, o gênero punk parece ter uma repetição de motivos rítmicos que se assemelha ao reggae, talvez pelo fato do andamento não está sendo considerado.

9.4 Considerações finais

Neste capítulo foi descrita uma investigação para a ligação entre gênero e expressões de emoção/humor através da presença de motivos melódicos e rítmicos nas melodias de diversas músicas. A melodia (ou faixa vocal) é extraída de arquivos MIDI e representada por um vetor de notas ou por um vetor de notas de valores. Foi desenvolvido um método para identificar motivos rítmicos e melódicos em cada melodia e relacioná-los com a ocorrência de expressões de emoção/humor. Para efeitos de validação, anotações de emoção/humor para artistas do banco de dados adotado foram coletadas do site *All Music Guide* (23).

Gêneros como rap, reggae e rock, conhecidos por apresentarem padrões de ritmos constantes, expressaram grande quantidade de motivos nas melodias. Blues e country confirmaram a “fama” de gêneros “mais sofisticados”, uma vez que apresentaram menor quantidade de motivos.

Esta abordagem reflete os passos iniciais para o estudo da relação gênero e emoção/humor. A inclusão de outras características das músicas como instrumentação ou harmonia se torna interessante para que os gêneros sejam examinados de forma mais sistemática.

Os atributos do ritmo também foram aplicados na investigação da evolução dos gêneros musicais. Este estudo está apresentado no capítulo a seguir.

Capítulo 10

Evolução de gêneros musicais

O estudo da evolução da música se torna uma tarefa desafiadora dado o caráter complexo em que a música se insere na sociedade. A história da música fornece evidências sobre os aspectos evolutivos, porém estudos que buscam uma compreensão objetiva dos principais mecanismos que fazem parte da evolução da música ainda são relativamente pouco explorados na literatura.

Neste trabalho a evolução de gêneros musicais foi investigada dentro de uma perspectiva quantitativa através de uma análise estatística da evolução do ritmo de quatro gêneros musicais, conforme a taxonomia apresentada na seção 10.1.

A partir dos vetores de características $\mathbf{V}^m$ (probabilidades condicionais provenientes da cadeia de Markov de primeira ordem) uma medida de dispersão dos atributos rítmicos é estabelecida e calculada para diversas músicas, organizadas cronologicamente de acordo com o ano que foram gravadas. Esta estratégia está resumida na Figura 10.1 e detalhada na seção 10.2. Resultados preliminares que evidenciam aspectos evolutivos de cada gênero estão apresentados na seção 10.3.

10.1 Taxonomia adotada

Para esta análise foram utilizados os quatro gêneros musicais da Tabela 10.1.

Tabela 10.1 – Taxonomia adotada na investigação da evolução dos gêneros

Gêneros	Quantidade de Músicas
Blues	75
Country	79
Rock	99
Reggae	71
Total: 324

Figura 10.1 – Diagrama da metodologia proposta para a investigação da evolução dos gêneros musicais. A evolução do ritmo de cada gênero é descrita em termos de séries temporais, $X(t)$, e em termos da dispersão das principais características do ritmo obtidas pela análise discriminante linear (LDA)

10.2 Estratégia

Assim como apresentado no Capítulo 4, o ritmo da música m é representado por um vetor $\mathbf{V}^m$, com componentes $\{v_i^m; i = 1, 2, \dots, 324\}$, obtidas pela ordenação sequencial da matriz de transição $\mathbf{T}_m^p$. As N_m músicas de cada gênero são ordenadas cronologicamente, e o estudo da evolução do ritmo é desempenhado em termos de séries temporais (seção 10.2.1), e em termos da evolução dos principais atributos através da análise discriminante linear (seção 10.2.2).

10.2.1 Análise da dispersão global do ritmo através de séries temporais

Para definir uma série temporal a partir dos atributos do ritmo, as músicas de cada gênero são ordenadas em ordem cronológica conforme o ano, t_m , que elas foram originalmente gravadas. Os vetores de atributos $\mathbf{V}^m$ são ordenados de acordo com o índice da música, m , em ordem cronológica ascendente de acordo com os anos de gravação $t_{m+1} \geq t_m$. Com o intuito de reduzir o tamanho das flutuações na série temporal, os valores da mesma não estão associados a uma única música, mas cada valor resume as características de um grupo de M músicas consecutivas. As propriedades das músicas no intervalo $[m, m + M - 1]$ são descritas em termos do valor médio, e_i , e dispersão, d_i , dos elementos dos vetores de atributos, $\{v_i^m\}$, sobre as músicas, estimados como se segue:

$$e_i(t) = \frac{1}{M} \sum_{q=m}^{m+M-1} v_i^q \quad (10.1)$$

$$d_i(t) = \frac{1}{M-1} \sum_{q=m}^{m+M-1} (v_i^q - e_i(t))^2, \quad (10.2)$$

para todas as probabilidades de pares de notas de valores, $i = 1, 2, \dots, 324$.

Na equação (10.2), a dispersão é dada pela variância de v_i^q . Optou-se pela divisão por $M - 1$ ao invés de M uma vez que é esperado fornecer uma estimacão mais adequada da dispersão para valores pequenos de M (89). As músicas em cada conjunto $[m, m + M - 1]$ foram tipicamente gravadas em diferentes anos, de forma que a relação entre o ano e as características estatísticas fornecidas pelas equações (10.1) e (10.2) não é unicamente definida. O ano representativo será considerado como inteiro mais próximo da média dos anos para as músicas do intervalo $[m, m + M - 1]$, ou seja,

$$t = \text{nint} \left(\frac{1}{M} \sum_{q=m}^{m+M-1} t_q \right). \quad (10.3)$$

A série temporal para o valor médio e dispersão do atributo i é obtida calculando $e_i(t)$, $d_i(t)$, e t para os $N_m - M + 1$ possíveis intervalos, $[m, m + M - 1]$, das M músicas consecutivas do conjunto total de N_m músicas disponíveis para cada gênero. Cada uma destas séries temporais consiste de $N = N_m - M + 1$ valores. Para casos nos quais t é idêntico para dois conjuntos consecutivos de músicas, calcula-se o valor médio dos valores correspondentes com o intuito de estabelecer um único valor para cada ano. Para extrair conclusões a partir da análise da série temporal de cada atributo individualmente seria necessário analisar $N_t = 324$ séries temporais. A estratégia adotada utiliza uma descrição estatística baseada na série temporal da dispersão global para cada conjunto de músicas,

$[m, m + M - 1]$, definida como a soma de $d_i(t)$ sobre todos os atributos.

$$D(t) = \sum_{i=1}^{N_t} d_i(t) . \quad (10.4)$$

Por fim, utiliza-se a versão normalizada de $D(t)$ como se segue:

$$X(t) = \frac{D(t) - \langle D \rangle}{\sqrt{\text{Var}[D]}}, \quad (10.5)$$

em que $\langle D \rangle$ e $\text{Var}[D] = \langle D^2 \rangle - \langle D \rangle^2$ representam o valor médio e variância de D correspondentes aos N valores da série temporal.

Uma vez que a série temporal $X(t)$ foi obtida para todos os gêneros, a evolução da dispersão é analisada em termos da função de autocorrelação de $X(t)$. Em termos matemáticos, a autocorrelação é definida por (90):

$$a_k = \frac{\sum_{t=1}^{N-k} (X(t) - \langle X \rangle) (X(t+k) - \langle X \rangle)}{\sum_{t=1}^N (X(t) - \langle X \rangle)^2}, \quad (10.6)$$

A autocorrelação indica o quanto uma série temporal assemelha-se a ela mesma considerando um intervalo (*lag*) de k unidades de tempo. Em outras palavras, a_k quantifica a auto-similaridade de $X(t)$ ao longo do tempo. O conceito da autocorrelação está apresentado na Figura 10.2 através da ilustração de duas séries temporais. Considere uma série temporal $X(t)$ dada pela onda quadrada de período τ . Neste caso, o correlograma a_k exibe um comportamento oscilatório com o mesmo período, τ , da série temporal $X(t)$ (Figura 10.2(b)). Os valores absolutos de a_k nos pontos máximos e mínimos decaem conforme k aumenta, devido ao tamanho finito da série temporal. O decaimento inicial de a_k corresponde às correlações de curto prazo (do inglês, *short-term correlations*) dado que $X(t)$ tipicamente se auto assemelha para valores pequenos de k . Valores negativos de a_k correspondem aos intervalos de tempo nos quais a série temporal exibe comportamentos contrastantes. No contexto desta investigação, estes intervalos podem ser associados com períodos de inovação do processo $X(t)$ que representa a dispersão do ritmo de cada gênero. O contraste temporal máximo é obtido para $t = n\tau/2$ ($n = 1, 3, 5, \dots$) em que a_k assume valores mínimos.

Por outro lado, valores positivos de a_k estão associados aos intervalos de tempo nos quais a série temporal se auto assemelha. Estes intervalos de tempo referem-se, nesta análise, aos períodos de recuperação de um gênero, nos quais os mesmos retomam as características principais do ritmo após o período de inovação. O valor máximo de concordância neste exemplo ocorre para $t = n\tau$ ($n = 1, 2, 3, \dots$), correspondendo ao valores máximos de a_k .

Em geral, séries temporais reais possuem um comportamento randômico que torna a interpretação dos correlogramas mais desafiadora do que no exemplo apresentado na

Figura 10.2. Para uma série temporal completamente randômica, na qual $X(t)$ assume um valor aleatório independente a cada intervalo de tempo, a autocorrelação assume valores randômicos em função de k que são aproximadamente normalmente distribuídos com média zero e variância $1/N$ (isto é, $a_k \sim \mathcal{N}(0, 1/N)$). De acordo com o esperado, 95% dos valores de a_k caem no intervalo $[-2/\sqrt{N}, 2/\sqrt{N}]$ (90). Os valores de a_k para séries não randômicas também podem cair dentro deste intervalo de confiança se o tamanho da série temporal, N , é pequeno (compare o correlograma para a autocorrelação da onda quadrada com o intervalo de confiança na Figura 10.2(b)). Entretanto, isso não necessariamente significa que tal série não pode ser distinguida de uma série randômica. De fato, o correlograma de uma série temporal com um componente não randômico significativo irá tipicamente exibir certas regularidades (isto é, oscilações não randômicas) que não existem em séries randômicas (90).

Figura 10.2 – Correlogramas de duas séries temporais. (a) •: série temporal definida por uma onda quadrada de período $\tau = 10$; x: série temporal definida por uma sequência randômica que segue uma distribuição normal, contém média zero e variância 0.25. Ambas as séries consistem de $N = 35$ valores. (b) Autocorrelação para a onda quadrada (•) e para a série randômica (x) apresentadas em (a). A região em cinza indica o intervalo de confiança $[-2/\sqrt{N}, 2/\sqrt{N}]$ contendo 95% dos pontos em a_k para a série randômica

10.2.2 Evolução dos componentes principais do ritmo

Como uma segunda abordagem para o estudo da evolução dos gêneros, utilizou-se a análise discriminante linear (LDA) (13, 44) sobre os vetores de atributos $\{v_i^m\}$ com o objetivo de extrair seus atributos principais e analisar sua evolução temporal. De forma similar à estratégia anterior, adotam-se grupos de M músicas consecutivas. Desta forma, busca-se a extração de $L \leq N_t$ componentes principais, $\{\xi_l, l = 1, 2, \dots, L\}$. Como resultado, tem-se a descrição do espaço de atributos em termos de um conjunto ortogonal

de atributos não correlacionados, em que o objetivo principal é capturar a informação discriminativa da matriz de dados original, $\mathbf{P}^p$, em termos de um número relativamente pequeno de componentes principais (seção 3.1.2).

Nesta estratégia, foi considerado um espaço reduzido de atributos através das duas componentes principais (ξ_1 e ξ_2). Desta forma, cada ponto no espaço de atributos LDA corresponde ao valor médio de ξ_1 e ξ_2 para cada conjunto de M músicas consecutivas. A sequência de valores ξ_1 e ξ_2 para conjuntos consecutivos de música fornece uma descrição estatística da evolução temporal dos atributos rítmicos.

10.3 Resultados e discussões

As séries temporais para cada gênero apresentadas nesta seção foram obtidas através do uso de grupos de $M = 10$ músicas consecutivas. Primeiramente, é discutida a evolução de cada gênero em termos dos respectivos correlogramas (seção 10.3.1). Em seguida é apresentada a evolução dos principais atributos rítmicos, obtidos com LDA (seção 10.3.2).

10.3.1 Correlação temporal dos atributos do ritmo

A Figura 10.3 (a) mostra que a dispersão do gênero blues exhibe flutuações (veja, por exemplo, o pico na década de 60 e a elevação durante a década de 80). As oscilações regulares do correlograma na Figura 10.3 (b) sugerem que a série temporal $X(t)$ não é completamente randômica (compare com o correlograma para a série randômica da Figura 10.2) apesar do fato que muitos valores de a_k estão dentro do intervalo de confiança para uma série temporal randômica de mesmo tamanho (região sombreada da Figura 10.3 (b)). A autocorrelação cai para zero em um intervalo de tempo de $k \simeq 5$, indicando a existência de correlações de curto prazo, o que implica que $X(t)$ se altera de forma gradual (Figura 10.3(b)). O valor mínimo atingindo em $k \simeq 11$ anos reflete o período de tempo necessário para blues incorporar novos atributos rítmicos significantes. Em contraste, a correlação positiva observada para $k \simeq 22$ anos indica uma certa periodicidade na série temporal. Isto sugere que não é possível constantemente inovar sem necessariamente violar as características básicas que definem o gênero.

As oscilações da dispersão de rock e country podem ser observadas nas Figuras 10.4 e 10.5, respectivamente. O comportamento dos correlogramas é qualitativamente similar para os dois gêneros. De fato, ambos exibem correlações de curto prazo, como demonstrado para o decaimento inicial de a_k em zero (para $k \simeq 3$ e $k \simeq 7$ anos para rock e country, respectivamente). Além disso, a sequência de um mínimo em a_k seguido por um máximo

Figura 10.3 – Evolução do ritmo de blues. (a) Dispersão do ritmo, $X(t)$, e (b) a autocorrelação, a_k . O valor mínimo de a_k em $k \simeq 13$ anos, e o valor máximo em $k \simeq 22$ anos indicam, respectivamente, um período de inovação seguido por um período de recuperação. A região em cinza expressa o intervalo de confiança, $[-2/\sqrt{N}, +2/\sqrt{N}]$, para uma série randômica com $N = 60$ valores (90)

Figura 10.4 – Representação similar como na Figura 10.3 para a evolução do ritmo de rock. O comportamento das oscilações de a_k indica a presença de períodos de recuperação e inovação os quais ocorrem em intervalos de aproximadamente 5 anos

pode ser observado para ambos os gêneros, o que novamente sugere um mecanismo de recuperação associado com o compromisso entre inovação e aderências às regras de cada gênero.

A série temporal do gênero reggae também exibe correlações de curto prazo, conforme indicado pelo decaimento inicial de a_k para $k \simeq 7$ (Figura 10.6(b)). As oscilações do correlograma estão menos claras que nos outros casos, de forma que interpretações das mesmas se tornam mais difíceis para este gênero. As oscilações aparentam suportar

Figura 10.5 – Representação similar como na Figura 10.3 para a evolução do ritmo de country. As oscilações de a_k sugerem que as mudanças no ritmo ocorrem sobre períodos de $\simeq 11$ anos.

a existência de períodos de aproximadamente 12 anos nos quais fases de inovação e recuperação alternam (veja o mínimo em $k \simeq 12$ anos e o máximo em $k \simeq 24$ anos na Fig. 10.6(b)). Entretanto, o período de 12 anos representa uma fração significativa do período total compreendido na análise de reggae, dificultando a extração de conclusões sobre a evolução deste gênero através dos dados utilizados.

Figura 10.6 – Representação similar como na Figura 10.3 para a evolução do ritmo de reggae

10.3.2 Evolução temporal dos atributos principais do ritmo

Figura 10.7 – Evolução temporal das características principais do ritmo no espaço de atributos reduzido, (ξ_1, ξ_2) , através da LDA. Símbolos diferentes correspondem à gêneros diferentes: $\square$ - blues; $*$ - rock; $\bullet$ - country; $\circ$ - reggae

A Figura 10.7 apresenta as trajetórias dos quatro gêneros musicais no espaço de atributos definido pelas principais componentes LDA (ξ_1, ξ_2) dos vetores de atributos $\mathbf{V}^m$ de cada música. Cada ponto nas trajetórias corresponde a um conjunto distinto de $M = 10$ músicas consecutivas com uma sobreposição de três músicas. Estes conjuntos consecutivos de músicas ordenadas cronologicamente estabelecem uma trajetória no espaço de atributos que pode ser associada com a evolução temporal das características do ritmo de cada gênero. Pela Figura 10.7 observa-se que as trajetórias seguem características de dispersão diferentes para cada gênero.

Uma inspeção mais detalhada destas trajetórias (Figura 10.8) sugere a existência de

mecanismos de inovação e recuperação, em acordo com as discussões levantadas na seção anterior a partir da análise das séries temporais. Considere, por exemplo, a trajetória do gênero reggae (Fig. 10.8c). Inicialmente, a trajetória sai da origem, indicando uma certa inovação (cf. pontos 1, 2 e 3 na Figura 10.8c). Entretanto, a trajetória retorna para as proximidades do ponto 1 depois de dois passos subsequentes (ponto 5), sugerindo o período de recuperação. A trajetória parte novamente nos passos subsequentes, mas eventualmente retorna novamente para as proximidades da origem nos passos 8 e 9. Padrões similares de partida e retorno podem ser observados nos demais gêneros.

Figura 10.8 – Ampliação ao redor da origem para as trajetórias da Figura 10.7: (a) blues, (b) country, (c) reggae, e (d) rock

De acordo com o site *All Music* (23), blues manteve suas características fundamentais

durante os primeiros anos de sua existência. Isto está evidenciado na análise uma vez que os primeiros sete pontos na trajetória de blues da Figura 10.8a ($\square$) estão próximos uns aos outros no espaço de atributos. A trajetória exibe um desvio da origem entre os pontos 8 e 10, os quais contêm principalmente músicas das décadas de 70 e 80. Este fato parece estar associado com a emergência do sub-gênero conhecido como ‘Texas rock-blues’ (91) no começo da década de 70. De fato, muitas das músicas de blues utilizadas são do artista “Stevie Ray Vaughan”, que é um músico proeminente do estilo ‘Texas’.

Ainda a partir da Figura 10.8a, observa-se que o ponto de número 10 está próximo da região da origem, indicando alguma recuperação. Este resultado está de acordo com (91), que sugere que músicas de blues dos anos 80 são caracterizadas por um renascimento do blues tradicional, de forma que estão presentes músicas de formas mais tradicionais e músicas com formas mais inovativas. Além disso, blues tem como uma de suas formas derivativas o ‘*rhythm and blues*’ (R&B), que originou da combinação do *jazz* e do ‘jump blues’. De acordo com Ripani (92), o ritmo de R&B foi se alterando com o passar dos anos, sendo que mudanças mais significativas ocorreram durante as décadas de 70 e 80. Este fato também pode estar contribuindo para os desvios dos pontos 8 e 9 na trajetória de blues apresentada na Figura 10.8a.

A trajetória do country indicada nas Figuras 10.7 se mostra mais estável, sugerindo que os atributos rítmicos deste gênero não apresentaram mudanças significativas com o passar dos anos. Tal estabilidade está em acordo com o fato que não existem muitos estilos e formas derivativas do country (23). Os pontos 6 e 7 da trajetória do country apresentada na Figura 10.8b ($\bullet$) correspondem a músicas gravadas entre os anos de 70 e 80, as quais apresentam atributos rítmicos que mais se distanciam da origem. Uma possível associação deste desvio pode estar relacionada com a influência de um movimento marcante deste gênero, conhecido como “outlaw movement”, que se tornou popular durante as décadas de 60 e 70 (93). De fato, o banco de dados utilizado nesta investigação contém músicas de alguns artistas que tiveram presença significativa dentro deste movimento, tais como “Johnny Cash”, “Willie Nelson” e “Hank Williams”. A década de 70 também é caracterizada pela emergência do estilo “*country pop*”, enquanto que na década de 80 houve a emergência do “*neotraditional country*” (93, 94). O ponto 8 consiste principalmente de músicas da década de 80, que aparentam configurar um retorno em direção à origem. Esta recuperação das características do ritmo de músicas anteriores pode, portanto, ser atribuída ao movimento neotradicionalista que ocorreu na década de 80. Neste movimento, artistas que não estavam de acordo com a tendência *country pop* começaram a compor música country de uma forma mais tradicional. Músicas country da década de 90, principalmente representadas pelos pontos 9 e 10, são geralmente referenciadas como “country moderno”. Influenciadas pelo rock, principalmente pela emergência do rock alternativo, as músicas de country deste período tendem a exibir uma mistura entre formas tradicionais e formas mais melódicas.

Similar ao country, não há muitos estilos derivados do reggae (23). Este aspecto pode ser indicador do comportamento estável da trajetória para este gênero (Fig. 10.7). O ponto de número 3 na Figura 10.8c (o) representa as músicas de reggae da metade da década de 70, quando o estilo conhecido como “reggae de raiz” (do inglês “*roots reggae*”) emergiu e se tornou popular (95). O banco de dados contém músicas dos artistas “Bob Marley” e “Peter Tosh”, que são os artistas mais influentes deste estilo durante este período. São características deste estilo as mensagens das letras das músicas, que geralmente abordam assuntos como opressão racial e pobreza, e a inclusão de padrões mais complexos de percussão, como o uso do “*one drop*”.^a Em geral, os pontos de número 4 até 10 mostram desvios menores em relação à origem. Esta circunstância pode ser reflexo da co-existência de características das músicas de reggae posteriores à década de 70: enquanto que algumas composições incorporam as particularidades do reggae de raiz, outras se mantêm mais próximas às formas mais tradicionais de reggae, aspecto similar ao anteriormente observado para o country.

A Figura 10.8d mostra que o ritmo de rock tem se alterado consideravelmente com o tempo. Isto está de acordo com o fato de que rock é um gênero amplo que engloba uma grande variedade de estilos e origens estilísticas, incluindo R&B, blues, rock ‘n’ roll e country (23, 96). Isto torna difícil a tarefa de obter um conjunto adequado de atributos para descrever as características principais do gênero, assim como seus aspectos evolutivos (97).

10.4 Considerações finais

É esperado que a evolução do ritmo em gêneros musicais tenha influências de muitos fatores que são difíceis de identificar e/ou quantificar. Apesar desta complexidade, a análise estatística apresentada neste capítulo revela certas evidências mais genéricas que poderiam ser explicadas em termos de mecanismos relativamente simples. A principal descoberta é que a complexidade do ritmo, quantificada aqui por medidas de dispersão, é estocástica mas também exhibe certos níveis de periodicidade.

O fato que evidências de periodicidade foram observadas para todos os gêneros sugere a existência de mecanismos responsáveis por tal comportamento. Nesta investigação, propõe-se que a competição entre criatividade e obediência às regras de composição de cada gênero são provavelmente os fatores principais responsáveis para as oscilações na evolução do ritmo.

Para visualizar a interação entre estes fatores, pode ser interessante descrever o ritmo

^a De forma simplificada, “*one drop*” tem como característica a batida mais forte no terceiro tempo do compasso (considerando um compasso 4/4), ao invés de ocorrer no primeiro tempo, como usualmente esperado.

de um gênero como um sistema dinâmico randômico (98) que se move em um espaço de estados definidos pelos atributos rítmicos. Definir este espaço em termos precisos é uma tarefa desafiadora e a análise LDA desempenhada neste estudo pode ser vista como uma tentativa inicial para explorar a evolução dos gêneros musicais em uma projeção do espaço rítmico.

Regras de composição impõem restrições na evolução, o que significa que cada gênero pode explorar apenas uma área limitada do espaço rítmico. O fato da região permitida de exploração ser limitada explica o período de recuperação. A trajetória dos gêneros tende a visitar pontos do espaço rítmico que são próximos a outros que já foram visitados no passado. Dentro deste paradigma, a recuperação ocorre independente de regras particulares que governam a evolução do ritmo, uma vez que elas apenas podem explorar uma região limitada do espaço rítmico. Períodos de recuperação foram observados para os quatro gêneros, o que suporta esta conjectura. Entretanto, uma análise mais completa é necessária para validar as suposições levantadas neste capítulo.

Este capítulo finaliza os estudos desenvolvidos neste trabalho. A seguir são apresentadas as principais conclusões envolvendo as análises do ritmo para a caracterização de gêneros musicais, assim como as perspectivas de continuidade do trabalho.

Capítulo 11

Conclusões

Com o avanço da Internet, bancos de dados musicais online têm crescido em tamanho e quantidade, motivando pesquisas envolvendo análises de conteúdo e recuperação da informação musical. Extensivamente usados para organizar coleções musicais, gêneros musicais constituem descritores particularmente importantes, uma vez que sumarizam características comuns (padrões) em peças musicais, o que é uma propriedade importante em aplicações de reconhecimento automático de padrões. Entretanto, gêneros musicais não são um conceito bem definido, com confusas e redundantes taxonomias, tornando um desafio a definição destas de forma clara e objetiva. Deste modo, o estudo dos gêneros musicais tem se tornado cada vez mais importante em pesquisas de recuperação de informação musical.

Classificação de gêneros e estilos musicais

As tarefas de classificação foram exploradas de forma supervisionada e não supervisionada. No primeiro caso, constatou-se que os padrões rítmicos capturados são sensíveis à discriminação dos gêneros, oferecendo taxas de desempenho comparáveis aos estudos anteriores. A vantagem sobre estes estudos está relacionada com o entendimento claro de fatores que contribuem para separação dos gêneros e da possibilidade de um estudo sistemático que gradualmente insere novos aspectos musicais e estabelece a contribuição particular de cada um.

A partir dos experimentos realizados, nota-se que a dimensionalidade do problema é alta, ou seja, os ritmos destes quatro gêneros são complexos. Muitas músicas de gêneros distintos estão sobrepostas, uma vez que estas apresentam vetores de padrões similares. Consequentemente, muitas dimensões (atributos) são necessárias para separá-los. Mesmo considerando um contexto estendido (como no caso da cadeia de Markov de segunda ordem) a cada tempo para capturar as dinâmicas e dependências dos padrões rítmicos, a dimensionalidade do problema continua alta.

Nesta circunstância, o uso de uma classificação multi-categorias mostrou-se particularmente apropriado, pois revela a natureza intrínseca do banco de dados e está mais próxima da experiência humana. Com a metodologia proposta de classificação multi-gêneros, novos subgêneros como *rock-blues* foram criados através de um critério objetivo a partir dos gêneros originais.

O agrupamento hierárquico aglomerativo opera nas dimensões originais dos dados, promovendo uma classificação de em média 38% sem qualquer informação prévia das classes. Este aspecto ressalta a complexidade do problema. Por outro lado, o uso de representações em grafos (ao invés de métodos de agrupamentos em um espaço vetorial) é promissor, uma vez que permite de forma não supervisionada uma combinação de atributos topológicos juntamente com métricas de similaridade para inferirem estruturas presentes nos dados. O desempenho de classificação pelas comunidades encontradas nas redes de música foi superior ao agrupamento hierárquico.

Além disso, a possibilidade de estudar comunidades menores e relacioná-las com a presença de estilos dentro de um gênero é bastante promissora. De fato, constatou-se que a classificação automática de estilos se mostra menos desafiadora que a classificação de gêneros. Isto é de certa forma esperado, uma vez que dentro de um estilo as regras de composição são mais bem-definidas que dentro de um gênero, por este ser mais abrangente e incorporar diversos estilos.

As medidas topológicas extraídas diretamente dos dígrafos de música (tais como grau, força e coeficiente de aglomeração) levaram a resultados razoáveis em uma dimensão bem menor de atributos. Diferentes medidas podem ainda ser incluídas. A combinação dessas medidas com as medidas obtidas pelas cadeias de Markov possibilitou uma melhora no desempenho da classificação supervisionada, sugerindo que a caracterização dos gêneros musicais pode ser favorecida com o uso de tais medidas topológicas. Outra motivação para o uso destas medidas se deve ao fato de que com elas é possível identificar notas-chaves ou pares e trios de notas que tem papéis importantes na análise dos gêneros.

Por fim, verifica-se que, dentre os gêneros explorados para classificação, blues é mais discriminativo. Representando o gênero mais antigo, e por possuir características específicas, blues pode ter influenciado os gêneros subsequentes, o que contribuiu para a mistura de características entre eles ao longo dos anos. Reggae, rock e MPB são gêneros mais similares, com muitas amostras que se sobrepõem. De acordo com os resultados, músicas de MPB tendem a ser mais discriminativas que as de rock, que por sua vez tendem a ser mais discriminativas que as de reggae. Músicas de MPB passaram a incluir diferentes ritmos como rock e música latina. Reggae possui origens em jazz, “rock and blues”, *rocksteady*, entre outros. Essas tendências são relevantes e estão de alguma forma refletidas nos resultados.

Síntese de novas sequências rítmicas

A metodologia desenvolvida para síntese de novas sequências musicais pode ser útil como inspirações para composições musicais. Além disso, por serem simples, tais sequências podem ser convenientes como exercícios para alunos iniciantes. Quando comparadas com as respectivas versões originais, as frases rítmicas possuem variações de passagens de notas, mas estão dentro do esperado para cada contexto. Com os dígrafos formados por todas as músicas de um gênero, verificou-se que alguns pares de notas estão presentes em todos os casos, enquanto que outros são específicos de cada um. Esta análise foi posteriormente complementada com a verificação dos pares de notas mais significantes de acordo com os autovetores das técnicas PCA e LDA sobre a matriz de dados.

As novas sequências rítmicas refletem a dinâmica das notas de valores nas músicas do banco de dados. Apesar de soarem um pouco repetitivas, representam um passo inicial para a implementação de um sistema mais completo de composição, que pode ser diretamente beneficiado da metodologia utilizada.

Geração automática de listas de reprodução

Também foi desenvolvido um algoritmo para a geração automática de listas de reprodução nas quais as músicas são similares de acordo com padrões rítmicos. O algoritmo é baseado em caminhos em árvores geradoras mínimas, não requer metadados de usuários e pode ser adaptado para outros atributos musicais. Em resumo, as principais contribuições do algoritmo são: 1) a distância entre músicas subsequentes na lista é minimizada quando comparada com o uso da busca em largura; 2) o algoritmo reduz o número de transições entre os gêneros; 3) músicas de um mesmo gênero estão proximamente localizadas na lista se possuem padrões rítmicos similares, o que permite uma variedade controlada; e, 4) músicas com padrões rítmicos semelhantes tendem a possuir aspectos emotivos parecidos, de acordo com o estudo levantado neste trabalho da relação entre gêneros e expressões de emoção/humor.

Por outro lado, não é possível garantir a preferência dos usuários pela lista de reprodução gerada pelo algoritmo desenvolvido. Uma avaliação realizada por diversos usuários é necessária para a validação do algoritmo. Apenas é possível sugerir que, tendo em vista os aspectos temporais apresentados na Figura 8.5, a lista de reprodução obtida, em média, maximiza a similaridade entre músicas subsequentes de acordo com uma medida de distância e um aspecto musical.

Relação de gêneros com expressões de emoção/humor

A relação entre gêneros e expressões de emoção/humor foi analisada, ainda que de forma exploratória. A motivação para esta investigação se deve ao fato de que estudos

anteriores sugerem que fatores musicais podem influenciar em rótulos de emoção/humor atribuídos a diferentes músicas. Por exemplo, músicas com ritmos regulares tendem a receber rótulos como “felicidade”, enquanto que músicas com ritmos mais complexos tendem a receber rótulos como “desconforto”.

O aspecto do ritmo ser regular ou complexo é atribuído neste trabalho à presença de motivos musicais em melodias do banco de dados. Verificou-se a relação da quantidade média de motivos encontrados em cada gênero com rótulos de emoção/humor obtidos de sites com o *All Music*. Gêneros como blues e country mantiveram a reputação de ser gêneros mais sofisticados, que recebem rótulos de emoção como “complexo”, “melancólico” e “dramático”; enquanto que rock, dance, reggae e rap são caracterizados como gêneros com ritmos mais regulares que usualmente recebem rótulos como “alegre”, “motivante” e “animado”.

Apesar dos resultados evidenciarem alguma relação entre gêneros e rótulos correntes de emoção/humor, constatou-se que alguns rótulos podem ser comuns a músicas com ritmos de diferentes aspectos. Os autores acreditam que esta análise possa ser mais sistematicamente validada partindo dos estilos (e não gêneros) como referência, uma vez que é esperado que estes possuam aspectos musicais mais definidos e menos heterogêneos.

Evolução de gêneros musicais

A análise da evolução dos gêneros foi direcionada pelos aspectos rítmicos de músicas de diferentes gêneros. A principal descoberta é que tendências de inovação estão presentes em todos os gêneros estudados mas a diversidade do ritmo não se mantém indefinidamente. De fato, observou-se que períodos de inovação são seguidos por períodos de recuperação que ocorrem após intervalos de tempo que são distintos para cada gênero. A existência de ambas as etapas de inovação e recuperação pode ser explicada em termos do compromisso entre renovação que promove evolução e a obediência às regras de composição que previne de inovações excessivas.

Com base apenas no estudo realizado, não é possível inferir detalhes precisos sobre os fatores que governam a evolução do ritmo. Entretanto, é interessante observar que a evolução do ritmo parece estar correlacionada com índices de popularidade, com a abundância de diferentes movimentos estilísticos (principalmente nas décadas de 70 e 80) dentro de cada gênero, e de influências mútuas entre gêneros. A comparação dos resultados de períodos de recuperação de cada gênero com os índices de popularidade da rádio *online* Lastfm suporta a hipótese de que alta popularidade acarreta em gêneros mais dinâmicos e que exibem períodos curtos de recuperação. De fato, de acordo com o site da Lastfm os gêneros contemplados neste estudo podem ser ordenados em ordem decrescente de popularidade como se segue: rock, blues, reggae e country (os índices para reggae e country são similares e esta ordenação proposta não é definitiva). Rock é o gênero mais

popular e, de fato, ambas as análises da série temporal e da evolução dos atributos LDA mostraram que este gênero tem o período mais curto de renovação. A influência significativa de estilos presentes no rock (é o gênero que abrange maior quantidade de estilos dentre os examinados) também favoreceu seu caráter mais dinâmico. Em contraste, gêneros menos populares como reggae e country, que também possuem menos estilos derivativos que rock e blues, são caracterizados por grandes períodos de inovação e recuperação.

11.1 Trabalhos futuros

O uso de um banco de dados maior deve ser considerado para garantir uma melhor validação dos resultados. Aspectos do ritmo, como por exemplo, a batida e o tempo, podem ser analisados. Vale a pena verificar a contribuição de cada um destes aspectos à luz das aplicações compreendidas neste trabalho. A metodologia desenvolvida facilita este estudo.

Outras dimensões como a melodia e a instrumentação também podem ser averiguadas, de forma a verificar se os comportamentos observados se mantêm para outras características musicais. Além disso, é promissor o estudo dos estilos musicais no contexto das aplicações investigadas. A classificação multi-categorias pode ser mais sistematicamente explorada com o uso de outros algoritmos, como, por exemplo, algoritmos de agrupamento de dados baseados em sistemas *fuzzy*.

Nos últimos anos, tem-se tornado mais comum a disponibilização de banco de dados MIDI para fins acadêmicos. Uma das contribuições desta tese é viabilizar de forma *online* o banco de dados construído para que pesquisas posteriores possam ser realizadas.

Referências

- 1 PRIOLLI, M. L. de M. *Princípios básicos da música para a juventude*. Rio de Janeiro: Casa Oliveira de Músicas LTDA, v.1, 2007.
- 2 MIRANDA, E. R. *Composing music with computers*. Oxford: Focal Press, 2002.
- 3 HOLT, F. *Genre in popular music*. Chicago: University of Chicago Press, 2007.
- 4 MOORE, A. F. Categorical conventions in music discourse: style and genre. *Music & Letters*, v. 82, n. 3, p. 432–442, 2001.
- 5 FU, Z.; LU, G.; TING, K. M.; ZHANG, D. A survey of audio-based music classification and annotation. *IEEE Transactions on Multimedia*, v. 13, n. 2, p. 303–319, 2011.
- 6 COSTA, L. D. F.; CESAR JR, R. M. *Shape analysis and classification: theory and practice*. Boca Raton, FL, USA: CRC Press, 2001.
- 7 DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern classification*. New York: John Wiley & Sons, Inc., 2001.
- 8 CATALTEPE, Z.; YASIAN, Y.; SONMEZ, A. Music genre classification using MIDI and audio features. *Journal on Advances of Signal Processing*, v. 2007, n. 1, p. 1–8, 2007.
- 9 ISMIR 2004 genre contest. Disponível em: <<http://ismir2004.ismir.net>>. Acesso em: 10 Out. 2012.
- 10 TZANETAKIS, G.; COOK, P. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, v. 10, n. 5, p. 293 – 302, 2002.

-
- 11 MCKAY, C. *Automatic classification of MIDI recordings*. 2004. (Master Thesis) - Faculty of Music McGill University, Montreal, 2004.
- 12 SCARINGELLA, N.; ZOIA, G.; MLYNEK, D. Automatic genre classification of music content: a survey. *IEEE Signal Processing Magazine*, v. 23, n. 2, p. 133 –141, 2006.
- 13 WEBB, A. R. *Statistical pattern recognition*. New York: John Wiley & Sons Ltd, 2002.
- 14 SONG, Y.; ZHANG, C. Content-based information fusion for semi-supervised music genre classification. *IEEE Transactions on Multimedia*, v. 10, n. 1, p. 145 –152, 2008.
- 15 GOUYON, F.; DIXON, S. A review of automatic rhythm description system. *Computer Music Journal*, v. 29, n. 1, p. 34–54, 2005.
- 16 MOSTAFA, M. M.; BILLOR, N. Recognition of western style musical genres using machine learning techniques. *Expert Systems with Applications*, v. 36, n. 8, p. 11378–11389, 2009.
- 17 NEWMAN, M. E. J. The structure and function of complex networks. *Society for Industrial and Applied Mathematics Review*, v. 45, n. 2, p. 167–256, 2003.
- 18 SILVA, D. D. L.; SOARES, M. M.; HENRIQUES, M. V. C.; ALVES, M. T. S.; AGUIAR, S. G.; CARVALHO, T. P.; CORSO, G.; LUCENA, L. S. The complex network of the brazilian popular music. *Physica A: statistical mechanics and its applications*, v. 332, n. 1, p. 559–565, 2004.
- 19 GLEISER, P. M.; DANON, L. Community structure in jazz. *Advances in Complex Systems*, v. 6, n. 4, p. 565–573, 2003.
- 20 SMITH, R. D. The network of collaboration among rappers and its community structure. *Journal of Statistical Mechanics: theory and experiment*, v. 1, n. 2, p. 1–15, 2006.
- 21 THE ORIGINAL hip-hop (rap) lyrics archive. Disponível em: <<http://www.ohhla.com>>. Acesso em: 04 Nov. 2012.

-
- 22 PARK, J.; CELMA, O.; KOPPENBERGER, M.; CANO, P.; BULD, J. M. The social network of contemporary popular musicians. *International Journal of Bifurcation and Chaos*, v. 17, n. 7, p. 2281–2288, 2007.
- 23 ALL music: music search, recommendations, videos and reviews. Disponível em: <www.allmusic.com>. Acesso em: 12 Out. 2012.
- 24 CANO, P.; CELMA, O.; KOPPENBERGER, M.; BULD, J. M. Topology of music recommendation networks. *Chaos*, v. 16, n. 1, p. 1 – 6, 2006.
- 25 LIU, X.; TSE, C. K.; SMALL, M. Complex network structure of musical compositions: algorithmic generation of appealing music. *Physica A*, v. 389, n. 1, p. 126 – 132, 2010.
- 26 Czechoslo, M., F. Algebraic connectivity of graphs. *Mathematical Journal*, v. 23, n. 98, p. 298–305, 1973.
- 27 TEITELBAUM, T.; BALENZUELA, P.; CANO, P.; BULD, J. M. Community structures and role detection in music networks. *Chaos*, v. 18, n. 043105, p. 043105:1–043105:7, 2008.
- 28 LAMBIOTTE, R.; AUSLOOS, M. On the genre-fication of music: a percolation approach. *The European Physical Journal B - condensed matter and complex systems*, v. 50, p. 183–188, 2006.
- 29 TATA, S.; EUGENIO, B. D. Song recommend: from summarization to recommendation. *Natural Language Engineeting*, v. 1, n. 1, p. 1–39, 2012.
- 30 ANDRIC, A.; HAUS, G. Automatic playlist generation based on tracking user's listening habits. *Multimedia Tools Applications*, v. 29, n. 2, p. 127–151, 2006.
- 31 PAUWS, S.; VERHAEGH, W.; VOSSEN, M. Music playlist generation by adapted simulated annealing. *Information Science*, v. 178, n. 3, p. 647–662, 2008.
- 32 TEMPERLEY, D. Communicative pressure and the evolution of musical styles. *Music Perception: an interdisciplinary journal*, v. 21, n. 3, p. 313–337, 2004.
- 33 LENA, J. C.; PETERSON, R. A. Classification as culture: types and trajectories of music genres. *American Sociological Review*, v. 73, n. 5, p. 697–718, 2008.

-
- 34 CAPUZZO, G. C. Neo-riemannian theory and the analysis of pop-rock music. *Music Theory Spectrum*, v. 26, n. 2, p. 177–200, 2004.
- 35 RENTFROW, P. J.; GOSLING, S. D. Message in a ballad: the role of music preferences in interpersonal perception. *Psychological Science*, v. 17, n. 3, p. 236–242, 2006.
- 36 MORETTI, F. *Graphs, maps, trees: abstract models for literary history*. London: Verso Books, 2005.
- 37 BENSON, D. J. *Music: a mathematical offering*. New York: Cambridge University Press, 2007.
- 38 CARDOSO, B.; MASCARENHAS, M. *Curso completo de teoria musical e solfejo*. Sao Paulo: Vitale, 1996.
- 39 AUCOUTURIER, J.; PACHET, F. Representing musical genre: a state of the art. *Journal of New Music Research*, v. 32, p. 83–93, 2003.
- 40 ROADS, C. *The computer music tutorial*. Massachusetts: MIT Press, 1996.
- 41 ROTHSTEIN, J. *MIDI: a comprehensive introduction*. Madison: A-R Editions, 1995.
- 42 MIDI manufactures association. Disponível em: <www.midi.org>. Acesso em: 06 Ago. 2012.
- 43 NOTE names, MIDI numbers and frequencies. Disponível em: <<http://www.phys.unsw.edu.au/jw/notes.html>>. Acesso em: 10 Ago. 2012.
- 44 FUKUNAGA, K. *Introduction to statistical pattern recognition*. Boston: Academic Press, 1990.
- 45 DEVIJVER, P. A.; KITTLER, J. *Pattern recognition: a statistical approach*. Englewood Cliffs NJ: Prentice Hall International, 1981.
- 46 THERRIEN, C. W. *Decision estimation and classification: an introduction to pattern recognition and related topics*. New York: John Wiley & Sons, 1989.

-
- 47 HYVARINEN, A.; KARHUNEN, J.; OJA, E. *Independent component analysis*. New York: John Wiley & Son, Inc., 2001.
- 48 COHEN, J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, v. 20, n. 1, p. 37–46, 1960.
- 49 CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Enviroment*, v. 37, n. 1, p. 35–46, 1991.
- 50 LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *International Biometric Societyrics*, v. 33, n. 1, p. 159–174, 1977.
- 51 GAN, G.; MA, C.; WU, J. *Data clustering: theory, algorithms, and applications*. Virginia: Society for Industrial and Applied Mathematics, 2007.
- 52 JAIN, A. K.; DUBES, R. C. *Algorithms for clustering data*. Upper Saddle River, NJ, USA: Prentice Hall, 1988.
- 53 ANDERBERG, M. R. *Cluster analysis for applications*. New York: Academic Press, 1973.
- 54 ROMESBURG, H. C. *Cluster analysis for resarches*. Morrisville, NC: LULU Press, 1990.
- 55 KUIPER, F. K.; FISHER, L. A. Monte Carlo comparison of six clustering procedures. *Biometrics*, v. 31, n. 3, p. 777–783, 1975.
- 56 BLASHFILED, R. K. Mixture model tests of cluster analysis: accuracy of four agglomerative hierarchical methods. *Psychological Bulletin*, v. 83, p. 377–388, 1976.
- 57 MOJENA, R. Hierarchical grouping methods and stopping rules: an evaluation. *The Computer Journal*, v. 20, n. 4, p. 359–363, 1977.
- 58 WEBB, A. R.; COPSEY, K. D. *Statistical pattern recognition*. 3. ed. Chichester: John Wiley, 2011.

-
- 59 COSTA, L. D. F.; RODRIGUES, F. A.; TRAVIESO, G.; BOAS, P. R. V. Characterization of complex networks: a survey of measurements. *Advances in Physics*, v. 56, n. 1, p. 167–242, 2007.
- 60 STEEN, M. V. *Graph theory and complex networks: an introduction*. Lexington: Maarten van Steen, 2010.
- 61 XU, R.; WUNSCH II, D. C. *Clustering*. New Jersey: John Wiley & Son, Inc., 2009.
- 62 RAPOPORT, A. Nets with distance bias. *Bulletin of Mathematical Biophysics*, v. 13, n. 2, p. 85–91, 1951.
- 63 RAPOPORT, A. Spread of information through a population with sociostructural bias I: assumption of transitivity. *Bulletin of Mathematical Biophysics*, v. 15, n. 1, p. 523–533, 1953.
- 64 RAPOPORT, A. Contribution to the theory of random and biased nets. *Bulletin of Mathematical Biophysics*, v. 19, n. 4, p. 257–277, 1957.
- 65 ERDŐS, P.; RENYI, A. On random graphs. *Publicationes Mathematicae*, v. 6, n. 2, p. 290–297, 1959.
- 66 ERDŐS, P.; RENYI, A. On the evolution of random graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, v. 5, n. 1, p. 17–61, 1960.
- 67 ERDŐS, P.; RENYI, A. On the strenght of connectedness of a random graph. *Acta Mathematica Scientia Hungary*, v. 12, n. 1, p. 261–267, 1961.
- 68 WATTS, D. J.; STROGATZ, S. H. Collective dynamics of small-world networks. *Nature*, v. 393, p. 440–442, 1998.
- 69 BARABASI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *Science*, v. 286, p. 509–512, 1999.
- 70 GIRVAN, M.; NEWMAN, E. Community structure in social and biological networks. *Proceedings of The National Academy of Sciences of the United States of America - PNAS*, v. 99, n. 12, p. 7821–7826, 2002.

- 71 COSTA, L. D. F. Redes complexas: modelagem simples da natureza. *Ciencia Hoje*, v. 36, n. 213, p. 34–39, 2005.
- 72 COSTA, L. D. F.; OLIVEIRA JR, O. N.; TRAVIESO, G.; RODRIGUES, F. A.; BOAS, P. R. V.; ANTIQUEIRA, L.; VIANA, M. P.; ROCHA, L. E. C. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. *Advances in Physics*, v. 60, n. 3, p. 329–412, 2011.
- 73 NEWMAN, M. E. J.; GIRVAN, M. Finding and evaluating community structure in networks. *Physical Review E*, v. 69, n. 5, p. 026113–1 – 026113–15, 2004.
- 74 NEWMAN, M. E. J. Fast algorithm for detecting community structure in networks. *Physical Review E*, v. 69, n. 066113, p. 066133–1:066133:5, 2004.
- 75 CLAUSET, A.; NEWMAN, M. E. J.; MOORE, C. Finding community structure in very large networks. *Physical Review E*, v. 70, n. 6, p. 1–6, 2004.
- 76 CLAUSET, A. Finding local community structure in networks. *Physical Review E*, v. 72, n. 026132, p. 1–6, 2005.
- 77 MIEGHEM, P. V. *Graph spectra for complex networks*. New York:Cambridge University Press, 2011.
- 78 CVETKOVIC, D. M.; BOOB, M.; SACHS, H. *Spectra of graphs, theory and applications*. 3. ed. New York: Academic Press, 1995.
- 79 LIAO, P. S.; CHEN, T. S.; CHUNG, P. C. A fast algorithm for multilevel thresholding. *Journal of Information Science and Engineering*, v. 17, n. 1, p. 713–727, 2001.
- 80 DUNN, J. C. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, v. 3, n. 3, p. 32–57, 1973.
- 81 DAVIES, D. L.; BOULDIN, D. W. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 1, n. 2, p. 224–227, 1979.
- 82 EEROLA, T.; TOIVIAINEN, P. *MIDI toolbox: MATLAB tools for music research*. Jyväskylä, Finland: University of Jyväskylä, 2004. Disponível em: <www.jyu.fi/musica/miditoolbox/>. Acesso em: 04 Nov. 2012.

-
- 83 JOHNSTON, R. *How to play rhythm guitar: the basics and beyond*. Milwaukee: Backbeat Books, 2004.
- 84 RIJSBERGEN, C. J. *The geometry of information retrieval*. Cambridge: Cambridge University Press, 2004.
- 85 PRIM, R. C. Shortest connection networks and some generalizations. *Bell System Technical Journal*, v. 36, n. 2, p. 1389–1401, 1957.
- 86 CORMEM, T. H.; LEISERSON, C. E.; RIVEST, R. L.; STEIN, C. *Introduction to algorithms*. 3. ed. Massachusetts: The MIT Press, 2009.
- 87 EDWARDS, P. *How to rap: the art & science of the hip hop MC*. Chicago: Chicago Revire Press, 2009.
- 88 HOME page in oxford music online. Disponível em: <www.oxfordmusiconline.com>. Acesso em: 05 Jul. 2012.
- 89 JOHNSON, R. A.; WICHERN, D. W. *Applied multivariate statistical analysis*. 5. ed. New Jersey: Prentice Hall, 2002.
- 90 CHATFIELD, C. *The analysis of time series: an introduction*. 6. ed. New York: Chapman & Hall/CRC, 2004.
- 91 KOMARA, E. *Encyclopedia of the blues 2 - volume set*. New York: Routledge, 2005.
- 92 RIPANI, R. J. *The new blue music: changes in rhythm*. Jackson: University Press of Mississippi, 2006.
- 93 KINGSBURY, P. *The encyclopedia of country music: the ultimate guide to the music*. Oxford: Oxford University Press, 2004.
- 94 RUSSELL, T. *Country music originals: the legends and the lost*. New York: Oxford University Press, 2010.
- 95 THOMPSON, D. *Reggae and caribbean music*. San Francisco: Backbeat Books, 2002.

-
- 96 SCARUFFI, P. *A history of rock music: 1951:2000*. Lincoln, NE: iUniverse, Inc., 2003.
- 97 WICKE, P.; FOQQ, R. *Rock music: culture, aesthetics and sociology*. Cambridge: Cambridge University Press, 1990.
- 98 BHATTACHARYA, R.; MAJUMDAR, M. *Random dynamical systems: theory and applications*. New York: Cambridge University Press, 2007.
- 99 AKHTARUZZAMAN, M. Representation of musical rhythm and its classification system based on mathematical and geometrical analysis. In: INTERNATIONAL CONFERENCE ON COMPUTER AND COMMUNICATION ENGINEERING, 2008, Kuala Lumpur, Malaysia. **Proceedings...** Malaysia:IEEE, 2008. p.466 –471.
- 100 ALGHONIEMY, M.; TEWFIK, A. H. A network flow model for playlist generation. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, 2001, Tokyo, Japan. **Proceedings...** Tokyo:IEEE, 2001. p.329–332.
- 101 BRECHEISEN, S.; KRIEGEL, H.-P.; KUNATH, P.; PRYAKHIN, A. Hierarchical genre classification for large music collections. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, 6., 2006, Toronto, Canada. **Proceedings...** Toronto, Canada:IEEE, 2006. p.1385–1388.
- 102 BURRED, J. J.; LERCH, A. A hierarchical approach to automatic musical genre classification. In: INTERNATIONAL CONFERENCE ON DIGITAL AUDIO EFFECTS, 6., 2003, London, UK. **Proceedings...** London, UK:DAFx, 2003. p.8–11.
- 103 CHAI, W.; VERCOE, B. Folk music classification using hidden Markov models. In: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE, 2., 2001, Las Vegas, NV. **Proceedings...** Las Vegas, NV:ICAI, 2001. p.1–6.
- 104 DECORO, C.; BARUTCUOGLU, Z.; FIEBRINK, R. Bayesian aggregation for music hierarchical genre classification. In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL, 8., 2007, Vienna, Austria. **Proceedings...** Vienna, Austria:ISMIR, 2007. p.1–4.
- 105 FABBRI, F. Browsing music spaces: categories and the musical mind. In: INTERNATIONAL ASSOCIATION FOR THE STUDY OF POPULAR MUSIC, 1999, Sydney, Australia. **Proceedings...** Sydney, Australia:IASPM, 1999. p.1–14.

- 106 GABRIELSSON, A. The relationship between musical structure and perceived expression. In: HALLAM, S.; CROSS, I.; THAUT, M. (Ed.). *The oxford handbook of music psychology*. Oxford:Oxford University Press, 2009. p.141–150.
- 107 GOUYON, F.; DIXON, S.; PAMPALK, E.; WIDMER, G. Evaluating rhythmic descriptors for musical genre classification. In: INTERNATIONAL AUDIO ENGINEERING SOCIETY CONFERENCE, 25., 2004, London, UK. **Proceedings...** London, UK:AES, 2004. p.6–2.
- 108 GUNARATNA, C.; STONER, E.; MENEZES, R. Using network sciences to rank musicians and composers in brazilian popular music. In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL CONFERENCE, 12., 2011, Miami, FL. **Proceedings...** Miami, FL:ISMIR, 2011. p.441–446.
- 109 HAUVER, D.; FRENCH, J. Flycasting: using collaborative filtering to generate a playlist for online radio. In: INTERNATIONAL CONFERENCE ON WEB DELIVERING OF MUSIC, 2001, Washington, DC. **Proceedings...** Washington, DC:IEEE, 2001. p.123–.
- 110 HOMBURG, H.; MIERSWA, I.; MOLLER, B.; MORIK, K.; WURST, M. A benchmark dataset for audio classification and clustering. In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL, 6., 2005, London, UK. **Proceedings...** London, UK:ISMIR, 2005. p.528–531.
- 111 HONG, J.; DENG, H.; YAN, Q. Tag-based artist similarity and genre classification. In: KNOWLEDGE ACQUISITION AND MODELING WORKSHOP, 2008, Wuhan, China. **Proceedings...** Wuhan, China:IEEE, 2008. p.628 –631.
- 112 HU, X.; DOWNIE, J. S. Exploring mood and metadata: relationships with genre, artist and usage metadata. In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL, 8., 2007, Vienna, Austria. **Proceedings...** Vienna, Austria:ISMIR, 2007. p.1–7.
- 113 HURON, D. Perceptual and cognitive applications in music information retrieval. In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL, Plymouth, USA. **Proceedings...** Plymouth, USA:ISMIR, 2000. p.1–6.
- 114 JUSLIN, P. N. Emotional responses to music. In: HALLAM, S.; CROSS, I.; THAUT, M.(Ed.). *The oxford handbook of music psychology*. Oxford: Oxford University Press, 2009. p.131–140.

- 115 KARYDIS, I. Symbolic music genre classification based on note pitch and duration. In: EAST EUROPEAN CONFERENCE ON ADVANCES IN DATABASES AND INFORMATION SYSTEMS, 10., 2006, Berlin, Heidelberg. **Proceedings...** Berlin, Heidelberg:Springer-Verlag, 2006. p.329–338.
- 116 KIM, Y. E.; SCHMIDT, E. M.; MIGNECO, R.; MORTON, B.; RICHARDSON, P.; SCOTT, J.; SPECK, J. A.; TURNBULL, D. Music emotion recognition: a state of the art review. In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL, 11., 2010, Utrecht, Netherlands. **Proceedings...** Utrecht, Netherlands:ISMIR, 2010. p.255–266.
- 117 LI, T.; OGIHARA, M. Music genre classification with taxonomy. In: INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 30., 2005, Philadelphia, USA. **Proceedings...** Philadelphia, USA:IEEE, 2005. p.197–200.
- 118 LI, T.; OGIHARA, M.; SHAO, B.; WANG, D. Machine learning approaches for music information retrieval. In: LI, T.; OGIHARA, M.; SHAO, B.; WANG, D.(Ed.). *Theory and novel applications of machine learning*. Vienna, Austria:I-Tech., 2009. p.259–278.
- 119 LIN, Y.-C.; YANG, Y.-H.; CHEN, H. H.; LIAO, I.-B.; HO, Y.-C. Exploiting genre for music emotion classification. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, 9., 2009, New York, NY. **Proceedings...** New York, NY:IEEE, 2009. p.618–621.
- 120 LOGAN, B. Content-based playlist generation: exploratory experiments. In: INTERNATIONAL SOCIETY ON MUSIC INFORMATION RETRIEVAL, 3., 2002, Paris, France. **Proceedings...** Paris, France:ISMIR, 2002. p.1–2.
- 121 MASATAKA, G.; HASHIGUCHI, H.; NISHIMURA, T.; OKA, R. RWC music database: popular, classical, and jazz music databases. In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL, 3., 2002, Paris, France. **Proceedings...** Paris, France:ISMIR, 2002. p.1–2.
- 122 MCKAY, C.; FUGINAGA, I. Musical genre classification: is it pursuing and how can it be improved? In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL, 7., 2006, Victoria, Canada. **Proceedings...** Victoria, Canada:ISMIR, 2006. p.101–106.
- 123 MCKAY, C.; FUGINAGA, I. Automatic genre classification using large high-level music feature sets. In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION

RETRIEVAL CONFERENCE, 5., 2004, Barcelona, Spain. **Proceedings...** Barcelona, Spain:ISMIR, 2004. p.525–530.

124 PACHET, F.; CAZALY, D. A taxonomy of musical genres. In: CONTENT-BASED MULTIMEDIA INFORMATION ACCESS, 6., 2000, Paris, France. **Proceedings...** Paris, France:IEEE, 2000. p.1–8.

125 PANAGAKIS, I.; BENETOS, E.; KOTROPOULOS, C. Music genre classification: a multilinear approach. In: INTERNATIONAL SOCIETY ON MUSIC INFORMATION RETRIEVAL, 9., 2008, Pennsylvania,USA. **Proceedings...** Pennsylvania, USA:ISMIR, 2008. p.583–588.

126 PAUWS, S.; EGGEN, B. PATS: realization and user evaluation of an automatic playlist generator. In: INTERNATIONAL SOCIETY ON MUSIC INFORMATION RETRIEVAL, 3., 2002, Paris, France. **Proceedings...** Paris, France:ISMIR, 2002. p.222–230.

127 SCARINGELLA, N.; ZOIA, G. On the modeling of time information for automatic genre recognition systems in audio signals. In: INTERNATIONAL SOCIETY ON MUSIC INFORMATION RETRIEVAL, 6., 2005, London ,UK. **Proceedings...** London ,UK:ISMIR, 2005. p.666–671.

128 SHAN, M.-K.; KUO, F.-F.; CHEN, M.-F. Music style mining and classification by melody. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, 2., 2002, Lausanne,Switzerland. **Proceedings...** Lausanne, Switzerland:IEEE, 2002. v.1, p.97–100.

129 SHAO, X.; XU, C.; KANKANHALLI, M. Unsupervised classification of music genre using hidden Markov model. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, 4., 2004, **Proceedings...** Taipu, Taiwan:IEEE, 2004. v.3, p.2023 – 2026.

130 SILLA, C.; KAESTNER, C.; KOERICH, A. Automatic music genre classification using ensemble of classifiers. In: INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS, 4., 2007, Singapore. **Proceedings...** Singapore:IEEE, 2007. p.1687 –1692.

131 SILLA JR, C. N.; FREITAS, A. A. Novel top-down approaches for hierarquical classification and their application to automatic music genre classification. In: INTERNATIONAL CONFERENCE ON SYSTEMS, MAN, AND CYBERNETICS, 2009, San Antonio, Texas. **Proceedings...** San Antonio, Texas:IEEE, 2009. p.3499–3504.

-
- 132 SNYDER, B. Memory for music In.: HALLAM, S.; CROSS, I.; THAUT, M.(Ed.). *The oxford handbook of music psychology*. Oxford:Oxford University Press, 2009. p.107–117.
- 133 TZANETAKIS, G.; ESSL, E.; COOK, P. Automatic musical genre classification of audio signals. In: INTERNATIONAL SOCIETY FOR MUSIC INFORMATION RETRIEVAL, 2., 2001, Indiana, USA. **Proceedings...** Indiana, USA:ISMIR, 2001. p.293–302.
- 134 WANG, L.; HUANG, S.; WANG, S.; LIANG, J.; XU, B. Music genre classification based on multiple classifier fusion. In: INTERNATIONAL CONFERENCE ON NATURAL COMPUTATION, 4., 2008, Jinan, China. **Proceedings...** Jinan, China:IEEE, 2008. p.580 –583.