

Análise de Integração Audiovisual Para Imersão Em Jogos Digitais

Elyssa Dalva Ribeiro Bueno

Departamento de ciência da computação
IFSULDEMINAS câmpus Muzambinho
Muzambinho, Brasil
elyssadalva@gmail.com

Ricardo José Martins

Departamento de ciência da computação
IFSULDEMINAS câmpus Muzambinho
Muzambinho, Brasil
ricardo.martins@muz.ifsuldeminas.edu.br

Resumo—Considerando os aspectos artísticos de um jogo, e buscando uma melhora na imersão do jogador em métodos de geração de conteúdo, este trabalho propõe uma integração dinâmica entre as músicas de um jogo digital e seus elementos visuais e sonoros. Com o auxílio de ferramentas de reconhecimento de emoções em músicas (MER), e compreendendo sobre representações de emoções humanas, bem como a influência de cores e certos componentes musicais e visuais em tais emoções, um jogo de plataforma 2D foi desenvolvido de forma a se adaptar à emoção transmitida pela música escolhida pelos jogadores. Com os resultados obtidos através de um questionário respondido por um grupo de jogadores, percebe-se a eficácia da abordagem proposta.

Palavras-chave—Jogos digitais, Imersão, Reconhecimento de emoções em músicas (MER)

I. INTRODUÇÃO

Um jogo digital pode apresentar vários componentes que interagem sensorialmente com o jogador, como música, efeitos sonoros, efeitos visuais, entre outros. A forma como esta interação é feita tem um impacto direto na experiência de imersão do jogador [1].

Desenvolvedores de jogos podem escolher esses componentes de forma manual, ou utilizar de algoritmos procedurais, ou *Procedural Content Generation* (PCG) para acelerar o desenvolvimento [2]. Porém a praticidade de conteúdos procedurais pode afetar a qualidade do controle do conteúdo, pois um desenvolvedor humano tem um poder de avaliação melhor do que se encaixa ou não na temática do jogo, já que tal envolve parâmetros naturalmente subjetivos [3].

Um trabalho similar provou ser possível estabelecer uma ligação relevante entre música e PCG [4], não levando em conta porém os aspectos emocionais da música e do jogo produzido.

Considerando tais fatores, uma solução que dá importância a parâmetros subjetivos, que influenciam na imersão do jogador pode ser considerada, permitindo um desenvolvimento prático, sem comprometer a intenção artística de um jogo.

Baseado nisso, o jogo desenvolvido neste trabalho buscou identificar as emoções expressas em músicas previamente selecionadas, para atrelar sua intenção emocional com os elementos audiovisuais do jogo, criando assim um método procedural que leva em conta os conceitos subjetivos como emoções através da música como parâmetro principal.

II. REFERENCIAL TEÓRICO

Segundo Ermi e Mäyrä [5], o conceito de imersão é tornar-se parte de uma experiência, de forma física ou virtual. Este trabalho aborda principalmente a imersão sensorial [5], que trata-se das experiências decorridas com interações audiovisuais e táteis com o jogador, seja pelos gráficos, músicas e efeitos sonoros, ou ainda por forças exercidas no jogador através de *joy-sticks* ou periféricos.

É interessante apontar porém, que, como foi observado no trabalho de Ferreira e Oliveira [6], o desempenho de jogadores de “*Canabalt*” melhorou quando escolhiam uma música de seu gosto pessoal, portanto, é possível que o método proposto influencie outras dimensões da imersão. Isso não seria um fenômeno incoerente, considerando que Calleja [7] descreve a imersão como envolvimento internalizado que combina várias dimensões imersivas.

Um dos principais aspectos imersivos em um jogo, e o qual é um pilar dessa pesquisa, é a música. Como dito por Sanders [8], é provável que a música seja um fator capaz de estabelecer se a imersão será ou não capaz de ocorrer, daí o motivo de usar a música como parâmetro base. Correlacionado a isso, Tanskanen [1] aponta que a utilização de som e efeitos visuais de forma conjunta podem ter um maior impacto na experiência do jogador.

Identificar a intenção emocional de uma música porém, não é algo fácil segundo Laurier e Herrera [9], e existem diversas formas de se representar o espectro de emoções humanas. A representação utilizada nesse trabalho é a chamada representação dimensional de emoções de Russel [10], ilustrado na Fig. 1, onde emoções são dispostas sobre um plano cartesiano cujos eixos representam a excitação/energia e a intensidade/valência desses sentimentos. Cada ponto nessa representação localiza uma emoção em específico.

Para automatizar o processo de reconhecer emoções em músicas, existem as chamadas ferramentas de reconhecimento de emoções em músicas ou *Music Emotion Recognition* (MER). Com o auxílio de uma ferramenta desse tipo, o jogo criado é capaz de detectar a emoção de uma música e se adaptar.

A geração procedural de conteúdo para jogos, ou *Procedural Content Generation for Games* (PCG-G), que segundo

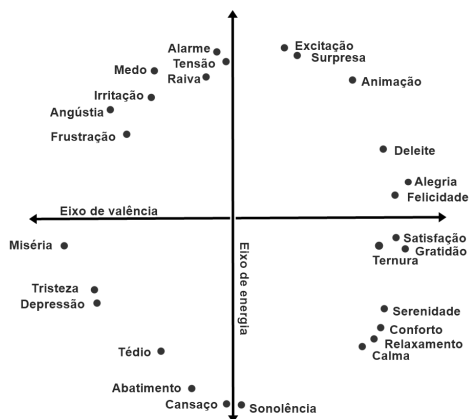


Fig. 1. Representação dimensional de emoções adaptado de Russel [10] e Ahn [11].

Hendrix et al [2] é a utilização de computadores para criação de conteúdo interessante para jogos. Sua utilização resulta em uma produção mais rápida e barata que a criação manual. Porém o PCG pode exigir de um computador tanto poder de processamento, quanto a habilidade de julgamento técnico e cultural de valores. Os autores descrevem 6 categorias de conteúdos que podem ser criados proceduralmente em um jogo, sendo o foco deste trabalho os chamados *Game Bits*, que consistem de componentes mais elementares de um jogo, como texturas, imagens e sons.

Um outro fator de extrema importância para a imersão de um jogador são os componentes visuais, que podem adicionar à experiência do jogador quando apresentam uma estética impactante [1].

Um importante elemento de uma composição visual são as cores. Nogueira [12] diz que a abordagem psicológica sobre uma possível linguagem universal das cores tem contribuído para maior compreensão sobre seus efeitos por meio de análise de respostas emocionais.

Levando em conta tal abordagem psicológica, Heller [13] propôs acordes cromáticos, combinações de cores que são relacionados a um determinado efeito, são associados à emoções específicas, buscando uma linguagem de cores universal. Acordes cromáticos utilizados especificamente neste trabalho estão descritos na próxima seção.

III. CONSTRUÇÃO E APLICAÇÃO DO JOGO

A. Desenvolvimento

O jogo desenvolvido trata-se de um jogo de plataforma 2D, utilizando scripts C# com o auxílio da Unity engine. O jogo focou-se em identificar e tratar 4 emoções humanas, cada uma pertencente a um quadrante diferente da classificação dimensional de Russel [10], sendo elas alegria, tristeza, medo e ternura.

Cada uma dessas foi relacionada à um acorde cromático [13], demonstrados na Fig. 2, utilizando de variações de seus tons, e à *sprites* e *backgrounds*, correspondentes à ideia que tais emoções e cores passam.

Por falta de acordes cromáticos específicos de tristeza e medo, foram escolhidas sensações relacionadas para substituí-los. Para a tristeza, foi selecionado o acorde cromático da frieza, baseado na definição da palavra "triste" e da palavra "frieza" pelo dicionário Michaelis [14] de língua portuguesa. O dicionário define "triste" como "Que não tem alegria; desgostoso" e frieza como "Ausência de entusiasmo, de ânimo, de envolvimento" [14], sendo portanto possível traçar um paralelo entre os dois conceitos.

O medo do desconhecido, ou FOTU (Fear Of The Unknown), muito abordado nas obras do gênero horror do autor H.P. Lovecraft, que o descrevia como o mais velho e mais forte tipo de medo [15], foi apontado como possivelmente o medo fundamental da humanidade por Carleton [16]. O mesmo, mais tarde, aprofundou sua pesquisa no assunto, onde reforçou sua afirmação do trabalho anterior com critérios revisados sobre define um medo fundamental [17]. Levando isso em consideração, o acorde cromático escolhido para representar o medo foi o do misterioso.



Fig. 2. Acordes cromáticos escolhidos.

Para identificar a emoção das músicas a serem selecionadas, a ferramenta MER escolhida foi o SentiFY, uma aplicação feita com o API do Spotify, que conta com os dados de quase todas as músicas disponíveis no serviço de streaming de música. A aplicação posiciona as músicas em uma representação dimensional de emoções baseado nos seus valores de energia e valência.

A Tabela I apresenta as músicas escolhidas com a ajuda da ferramenta MER citada e seus respectivos valores de energia e valência.

TABELA I
MÚSICAS ESCOLHIDAS PARA O TRABALHO E SUAS RESPECTIVAS CARACTERÍSTICAS.

Nome	Compositor	Energia	Valência	Emoção
The big parade	Pascal Michael Stiefel	0,803	0,762	Alegria
The riverlands	Yasunori Nishiki	0,149	-0,668	Ternura
Ganado IV	Misao Senbongi	-0,611	0,105	Medo
Moonlight Sonata 1st movement	Ludwig van Beethoven	-0,0785	-0,123	Tristeza

O jogo possui scripts que foram desenvolvidos para fazer a leitura de um documento onde se encontram a valência e energia de cada música, essas obtidas com a ajuda da ferramenta SentiFY. Em seguida, localiza a emoção da música escolhida na representação dimensional de Russel, e faz a já citada substituição de *Game Bits*, de forma adequada. O jogo também possui um algoritmo que altera características de alguns objetos, como a escala, de acordo com a batida

da música, para gerar um efeito visual que conecta a música diretamente aos elementos do jogo. A Fig. 3 apresenta capturas de tela de cada uma das 4 opções disponíveis, sendo respectivamente alegria, ternura, medo e tristeza.

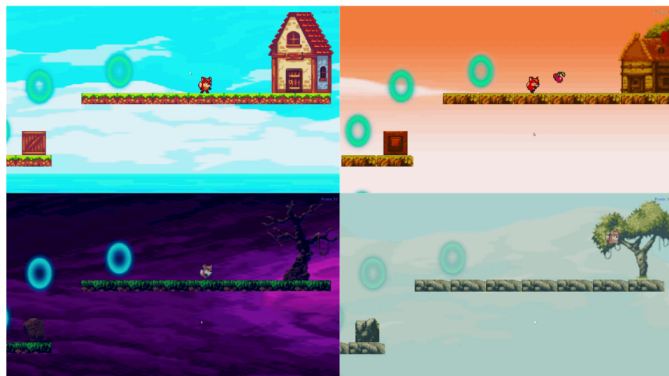


Fig. 3. Capturas de tela do jogo representando respectivamente as emoções de alegria, ternura, medo e tristeza.

Como controle, uma segunda versão do jogo foi feita, intitulada versão A, e essa possui um script exclusivo, que associa os componentes audiovisuais de maneira aleatória, para que após a aplicação, os jogadores possam relatar se há uma diferença significativa em comparação à ideia proposta, intitulada versão B.

B. Aplicação

Um grupo de 23 estudantes do curso de Ciência da Computação do Instituto Federal de Educação, Ciências e Tecnologia do Sul de Minas Gerais Campus Muzambinho, foi exposto às duas versões do jogo. O jogo foi executado em computadores de um laboratório de informática cedido pelo IFSULDEMINAS. Cada sessão de aplicação foi feita com 5 pessoas ou menos, dispostas de maneira a evitar que os participantes conversassem ou observassem os computadores uns dos outros.

Cada participante utilizou de fones de ouvido para melhor interação com a música e possíveis efeitos sonoros do jogo, além de jogarem ambas as versões do jogo em modo tela cheia.

Todas essas especificações são para proporcionar um ambiente adequado, pois o ambiente ao redor do jogador é fator importante para criar uma experiência mais imersiva [18]. Além disso, esse método impede que um jogador influencie outro quanto à experiência vivida no jogo.

O questionário era composto de:

- Cinco perguntas objetivas, com 5 alternativas, sendo a última opção em aberto caso nenhuma outra se aplicasse, questionando aos jogadores qual emoção eles sentiam relacionado à música, aos visuais da versão A e B e aos efeitos sonoros da versão A e B separadamente.
- Duas perguntas em escala de Lickert [19], questionando se as versões A e B adequam-se bem à música.
- Duas perguntas objetivas com 3 alternativas, perguntando em qual versão o jogador se sentia mais imerso e em qual

versão sentiu melhora no seu desempenho. As opções eram "Versão A", "Versão B" e "Indiferente" em ambas as perguntas

- Duas perguntas objetivas questionando a nota que o jogador daria de 1 a 5 para cada versão do jogo.
- Uma questão dissertativa opcional para os jogadores deixarem suas experiências e sugestões.

Durante cada sessão, foi pedido aos jogadores para seguirem cada uma das seguintes instruções:

- Testarem seus fones de ouvido e verificarem se estavam funcionando devidamente.
- Alterarem o volume do áudio de forma que ficasse mais confortável à cada um.
- Jogarem ambas as versões do jogo, A e B, no modo tela cheia e utilizando os fones de ouvido disponibilizados. As duas versões do jogo poderiam ser jogadas sem ordem específica, mas era explicitamente obrigatório, que eles escolhessem o mesmo número de 1 a 4, representando cada uma das quatro músicas, em ambas as versões.
- Após jogarem ambas as versões, responderem o questionário que estava disponível através de um link em seus computadores.
- Não comentar sobre elementos do jogo ou do questionário com pessoas que não haviam sido expostas, para evitar influências na experiência de outros possíveis participantes.

É importante ressaltar que os participantes não sabiam qual versão era representada pela letra A e pela letra B, tampouco, quais músicas eram representadas pelos números de 1 a 4 apresentados no menu.

IV. RESULTADOS

O questionário focava em identificar as emoções do jogador em cada versão, bem como sua preferência entre as duas. Analisando individualmente as respostas das cinco primeiras perguntas, focadas em identificar a emoção quanto à música, aos visuais, e aos efeitos sonoros, foi possível construir a Tabela II.

TABELA II
ASSOCIAÇÃO ENTRE A EMOÇÃO DA MÚSICA E DOS VISUAIS E EFEITOS SONOROS.

	Versão A		Versão B	
	Sim	Não	Sim	Não
A emoção passada pela música e pelos visuais eram a mesma	30,40%	69,60%	69,60%	30,40%
A emoção passada pela música e pelos efeitos sonoros era a mesma	65,20%	34,80%	73,90%	26,10%

Utilizando o teste de McNemar [20] para validação estatística dos dados dessas cinco primeiras perguntas, chega-se a um $p = 0,0265$ em relação à música e aos visuais, sendo a versão B considerada mais condizente. Trata-se de uma diferença significativa considerando a taxa de significância de 5%. Quanto aos efeitos sonoros, a versão B continua mais condizente que a versão A, mas a diferença não é significativa.

Passando para as próximas duas questões, onde é perguntado diretamente se o jogador considera as versões A e B

adequadas à música, os resultados foram também favoráveis à versão B. Como pode ser observado na Tabela III, 47,8% concordam total ou parcialmente que a versão A é adequada, enquanto 86,9% concordam total ou parcialmente que a versão B é adequada.

TABELA III
OPINIÃO DOS JOGADORES QUANTO À ADEQUAÇÃO DAS VERSÕES A E B À MÚSICA.

Você concorda que a versão jogada se adéqua bem à música					
	Concordo totalmente	Concordo parcialmente	Indiferente	Discordo parcialmente	Discordo totalmente
Versão A	21,7%	26,1%	8,7%	17,40%	26,1%
Versão B	47,8%	39,1%	4,4%	8,7%	0%

A pergunta questionava diretamente em qual versão o jogador se sentia mais imerso. E o resultado apontou mais uma vez para a versão B, com 73,9% de preferência contra apenas 4,3% da versão A. O restante dos jogadores opinou sentir-se indiferente quanto às duas versões.

Quanto ao desempenho no jogo, 52,2% disseram terem se saído melhor na versão B, 30,4% na versão A e 17,4% disseram estar indiferentes.

Por fim, as últimas duas questões obrigatórias pediam que os jogadores avaliassem ambas as versões com notas de 1 a 5. A Tabela IV apresenta as notas dadas pelos jogadores à cada versão.

TABELA IV
NOTAS ATRIBUÍDAS PELOS JOGADORES À CADA VERSÃO DO JOGO.

Versão	Nota 1	Nota 2	Nota 3	Nota 4	Nota 5
A	0 (0%)	1 (4,3%)	5 (21,7%)	12 (52,2%)	5 (21,7%)
B	0 (0%)	0 (0%)	5 (21,7%)	5 (21,7%)	13 (56,5%)

V. CONCLUSÃO

Com base no que foi dito na seção de resultados, pode-se concluir que a ideia deste trabalho foi sim eficaz, e a integração dinâmica da música e dos efeitos visuais e sonoros beneficiou a experiência dos jogadores. Contudo, quanto aos efeitos visuais, esse benefício é mais expressivo do que quanto aos efeitos sonoros.

Em comparação com um método de PCG básico que foi utilizado na validação deste trabalho, a técnica proposta de integração com a música apresentou considerável melhoria na imersão e consequentemente na preferência dos jogadores.

O método proposto mostra-se relevante por tratar-se de uma forma procedural de geração de conteúdo, que considera aspectos emocionais e artísticos frequentemente apresentados em jogos digitais, permitindo que o desenvolvedor crie seu jogo de forma prática, e obtenha resultados satisfatórios com relação à experiência de seus jogadores.

Como trabalhos futuros podem ser considerados os seguintes tópicos:

- Verificar a eficácia quanto à outro tipo de público alvo.

- Verificar a eficácia em outros gêneros de jogos, bem como em jogos em 3D.
- Comparar sua eficácia com formas mais complexas de PCG.
- Fazer uma integração com o API do *Spotify* para permitir que qualquer música do aplicativo seja disponibilizada e automaticamente classificada.
- Criar uma inteligência artificial que consiga classificar automaticamente a emoção passada pelos *Game Bits*
- Expandir a integração com a música em outros elementos do jogo, como dificuldade e a jogabilidade.
- Fortalecer integração com a área da psicologia para ajudar a identificar ou manipular as emoções dos usuários.

REFERÊNCIAS

- [1] S. Tanskanen, "Player immersion in video game: Designing an immersive game project," 2018.
- [2] M. Hendrikx, S. Meijer, J. Velden, and A. Iosup, "Procedural content generation for games: A survey," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, vol. 9, 02 2013.
- [3] D. Bengtsson and J. Melin, "Constrained procedural floor plan generation for game environments," 2016.
- [4] P. M. R. V. de Castro, "Music-based procedural content generation for games," Master's thesis, Faculdade de Engenharia Universidade do Porto, 2017.
- [5] L. Ermi and F. Mäyrä, "Fundamental components of the gameplay experience: Analysing immersion," *Worlds in play: International perspectives on digital games research*, vol. 37, no. 2, pp. 37–53, 2005.
- [6] E. Ferreira and T. Oliveira, "Som, imersão e jogos eletrônicos: um estudo empírico," in *Anais do X Simpósio Brasileiro de Games e Entretenimento Digital*, 2011.
- [7] G. Calleja, *From Immersion to Incorporation*. The MIT Press, 2011.
- [8] T. Sanders and P. Cairns, "Time perception, immersion and music in videogames," in *Proceedings of the 24th BCS interaction specialist group conference*. British Computer Society, 2010, pp. 160–167.
- [9] C. Laurier and P. Herrera, "Automatic detection of emotion in music: Interaction with emotionally sensitive machines," in *Machine Learning: Concepts, Methodologies, Tools and Applications*. IGI Global, 2012, pp. 1330–1354.
- [10] J. A. Russell, "A circumplex model of affect," *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.
- [11] J. Ahn, S. Gobron, Q. Silvestre, and D. Thalmann, "Asymmetrical facial expressions based on an advanced interpretation of two-dimensional russells emotional model," 01 2010.
- [12] K. A. Nogueira, "Estudo de respostas emocionais às cores no contexto de cartazes de cinema," *Design e Tecnologia*, vol. 8, no. 15, pp. 1–11, 2018.
- [13] E. Heller, *A psicologia das cores: como as cores afetam a emoção e a razão*. Editora Gustavo Gili, 2012.
- [14] D. Michaelis, "Disponível em; <http://michaelis.uol.com.br/>," *Acesso em*, vol. 11, no. 04, 2011.
- [15] H. Lovecraft, J. Turner, R. Bloch, R. Campbell, and B. Lumley, "The oldest and strongest emotion of mankind is fear, and the oldest and strongest kind of fear is fear of the unknown," *HP LOVECRAFT," Supernatural Horror in Literature*.
- [16] R. N. Carleton, "Into the unknown: A review and synthesis of contemporary models involving uncertainty," *Journal of anxiety disorders*, vol. 39, pp. 30–43, 2016.
- [17] —, "Fear of the unknown: One fear to rule them all?" *Journal of anxiety disorders*, vol. 41, pp. 5–21, 2016.
- [18] A. Nordin, "Immersion and players' time perception in digital games," Ph.D. dissertation, University of York, 2014.
- [19] R. Likert, "A technique for the measurement of attitudes," *Archives of Psychology*, vol. 22, no. 140, pp. 44–53, 1932.
- [20] Q. McNemar, "Note on the sampling error of the difference between correlated proportions or percentages," *Psychometrika*, vol. 12, no. 2, pp. 153–157, 1947.