

Capítulo

2

Ambientes Multissensoriais Aplicados à Saúde: Desenvolvimento de Aplicações e Tendências Futuras

Marina Ivanov, Eyre Montevecchi, Raphael Abreu, Fabio Barreto,
Joel dos Santos e Débora Muchaluat-Saade

Abstract

Incorporating sensory effects to traditional multimedia applications makes it possible to stimulate other human senses in addition to sight and hearing. Furthermore, this new type of application allows the construction of multisensory environments that can be used for healthcare applications, for example, stimulating multiple senses of patients through new types of therapy or cognitive training. In this context, this chapter presents and discusses the characteristics of applications with multiple sensory effects, highlighting possibilities of their use for healthcare. This chapter also describes some use cases of applications and multisensory ambients for helping the treatment of patients diagnosed with Alzheimer's, autism spectrum disorder, post-traumatic stress disorder, among others.

Resumo

A incorporação de efeitos sensoriais às aplicações multimídia tradicionais possibilita estimular outros sentidos humanos, além da visão e audição. Ainda, este novo tipo de aplicação permite a construção de ambientes multissensoriais que podem ser utilizados para aplicações em saúde, por exemplo, estimulando múltiplos sentidos de pacientes por meio de novos tipos de terapia ou treinamento cognitivo. Neste contexto, este capítulo apresenta e discute características de aplicações com múltiplos efeitos sensoriais, destacando possibilidades de sua utilização na área da saúde. Este capítulo também descreve alguns casos de uso de aplicações e ambientes multissensoriais, que podem servir como ferramenta auxiliar no tratamento de pacientes diagnosticados com Alzheimer, transtornos do espectro autista, estresse pós-traumático, dentre outros.

2.1. Introdução

Aplicações multimídia são amplamente utilizadas, a fim de transmitir informação utilizando diferentes mídias audiovisuais, como vídeo, texto, áudio ou imagem. Em geral, essas aplicações definem relacionamentos de sincronização entre esses tipos de mídia, e estimulam a audição e visão dos usuários que estão consumindo seu conteúdo. Os relacionamentos de sincronização entre os objetos de mídia os organizam tanto no espaço (tela do dispositivo de exibição) quanto no tempo. Por exemplo, um autor pode definir que um objeto de mídia do tipo áudio deve ser apresentado 5 segundos após o início da apresentação de um vídeo (organização no tempo); e que uma imagem deve ser apresentada no canto superior direito da tela (organização no espaço). Além disso, as aplicações interativas possibilitam a participação do usuário por meio de dispositivos de entrada, como controle remoto, teclado ou mouse, possibilitando assim um maior engajamento do usuário com o conteúdo oferecido.

Visando aumentar a qualidade de experiência do usuário em aplicações multimídia, alguns trabalhos propõem a incorporação de efeitos sensoriais ao conteúdo audiovisual, dando origem a um novo tipo de aplicação denominada aplicação mulsemídia (do inglês, *mulsemedia* – *Multiple Sensorial Media*) [Ghinea et al. 2014]. Em aplicações mulsemídia, é possível apresentar efeitos sensoriais sincronizados com o conteúdo audiovisual que está sendo apresentado [Josué et al. 2018, Rodrigues et al. 2019]. Por exemplo, em um vídeo que apresenta uma praia, o autor da aplicação poderia adicionar um efeito de aroma de mar e um efeito de vento, simulando a brisa da praia. Desse modo, as aplicações mulsemídia permitem aumentar a sensação de imersão do usuário, podendo ser utilizadas em diversas áreas como educação, entretenimento, turismo e saúde.

Diferente das aplicações multimídia tradicionais, que são exclusivamente bissensoresiais (i.e., sentidos de visão e audição), as aplicações mulsemídia são aquelas que envolvem três ou mais de nossos sentidos, e.g., visão, audição e olfato [Ghinea et al. 2014]. Para produzir tais aplicações, pode-se utilizar dispositivos sensores para identificar características do ambiente, tais como temperatura ou reação do usuário, e atuadores para renderizar¹ efeitos sensoriais, como vento, névoa ou calor.

Nesse contexto, este capítulo tem como objetivo apresentar o potencial de utilização de aplicações mulsemídia em saúde, por meio do estímulo multissensorial. Na literatura, algumas propostas estudam os benefícios do uso de ambientes multissensoriais para o tratamento de pacientes portadores de demência ou diagnosticados com alguma outra doença do envelhecimento [Ballard et al. 2002], de crianças com dislexia [Kast et al. 2007] e para transtorno de estresse pós-traumático (TEPT) [Rizzo et al. 2010]. Esses trabalhos se baseiam em indícios de que diferentes estímulos sensoriais promovem sensação de prazer e alívio da tensão e pressão, com consequente melhora no comportamento geral.

Em [Ghinea et al. 2014], os autores discutem o uso de efeitos sensoriais em terapias para pessoas com necessidades especiais como o autismo. Por exemplo, a utilização de efeitos sensoriais associados a jogos eletrônicos para testes e exercícios de estimula-

¹Renderizar é uma adaptação da palavra em inglês *render*. Que significa, em termos gerais, apresentar ou prover algo. No contexto deste capítulo, o termo renderizar é usado para significar o ato de apresentar efeitos sensoriais de forma física para o usuário.

ção cognitiva é uma abordagem que já vem sido empregada em algumas propostas na literatura [de Paula et al. 2019]. Jogos eletrônicos permitem a criação de exercícios que estimulam aspectos cognitivos como memória, raciocínio lógico, cálculo, criatividade, resolução de problemas e atenção. Em geral, esses jogos são de simples entendimento, apresentam níveis de dificuldade crescente, e em alguns casos podem reproduzir jogos tradicionais, como jogos de tabuleiro. Efeitos sensoriais podem melhorar a imersão do usuário no ambiente do jogo e aumentar a sua sensação de realidade, isto é, a percepção do usuário de estar fisicamente presente em um mundo virtual.

A integração de sensores às aplicações mulsemídia possibilita que a aplicação capte as reações do usuário que está consumindo o conteúdo e colete informações do ambiente de execução. A partir desta coleta de dados, a aplicação pode se adaptar, reagindo conforme o estado do usuário ou do ambiente. Suponha uma aplicação mulsemídia que apresenta efeitos sensoriais para um paciente com TEA (Transtorno do Espectro Autista). Nesta aplicação, pode-se especificar que o conteúdo sendo apresentado se modifique caso o sensor de batimento cardíaco do usuário alcance um valor alto, demonstrando agitação do paciente.

Visando aumentar a imersão do usuário e proporcionar uma experiência de presença mais realística, além dos efeitos sensoriais, algumas propostas na literatura têm utilizado realidade virtual [Mishkind et al. 2017, Rizzo et al. 2010]. A realidade virtual (do inglês, *Virtual Reality* - VR) é uma tecnologia que fornece uma experiência simulada ao usuário, através de um equipamento chamado óculos ou *headsets* de realidade virtual. Nessa experiência, o cenário apresentado para o usuário pode ser alterado, conforme ele realiza movimentos com a cabeça. Como discutido em [Mishkind et al. 2017], o uso de realidade virtual tem se revelado promissor para o tratamento de transtornos psiquiátricos. O TEPT e a ansiedade estão relacionados a padrões excessivos de estímulo e resposta baseados no medo e a um significado patológico na estrutura do medo, que mantém os sintomas e impede a recuperação [Mishkind et al. 2017]. Este capítulo discute também o uso de efeitos sensoriais sincronizados à realidade virtual para a simulação de situações causadoras de desconforto para certos pacientes com TEPT.

O restante do capítulo está organizado da seguinte forma. A Seção 2.2 apresenta o conceito de aplicações mulsemídia e os desafios de desenvolvimento e execução desse tipo de aplicação. Dispositivos utilizados para renderização de efeitos sensoriais e sensores para captura de informações do usuário e do ambiente são descritos na Seção 2.3. Na Seção 2.4, são descritos casos de uso de ambientes multissensoriais a fim de ilustrar sua aplicação no tratamento de doenças ou transtornos. A Seção 2.5 descreve desafios relacionados ao desenvolvimento de aplicações mulsemídia e apresenta novas tendências para a área. Por fim, a Seção 2.6 apresenta as considerações finais, resumindo os pontos descritos neste capítulo relacionados à utilização de aplicações mulsemídia aplicadas à saúde.

2.2. Aplicações Mulsemídia

Conforme comentado anteriormente, aplicações mulsemídia permitem a apresentação de estímulos a outros sentidos humanos, como tato, paladar e olfato, de forma sincronizada ao conteúdo audiovisual tradicional. Esta seção apresenta dois desafios principais relaci-

onados a tais aplicações. O primeiro remete à criação de aplicações mulsemídia, em que são necessárias formas de representação de efeitos sensoriais e sua sincronização, bem como um ferramental de suporte ao autor deste tipo de conteúdo. Em seguida, são apresentados os desafios na execução de tais aplicações, onde é necessária uma arquitetura para suporte à renderização de efeitos de forma sincronizada ao conteúdo audiovisual.

2.2.1. Autoria de Aplicações Mulsemídia

Assim como em multimídia, as aplicações mulsemídia focam em orquestrar a apresentação de conteúdo que estimule os sentidos do usuário. Sendo assim, o esforço de criar uma aplicação mulsemídia se assemelha a criação de uma aplicação multimídia. Portanto, ao longo desta seção, investiga-se como é feita a autoria multimídia para, em seguida, apresentar métodos e ferramentas para autoria mulsemídia.

Para criar uma aplicação multimídia é necessário que o autor da aplicação especifique tanto propriedades temporais quanto espaciais das mídias que participam da aplicação. Essas propriedades podem se apresentar de diversas maneiras tais como: o relacionamento entre duas ou mais mídias; o tempo de exibição de cada uma; a posição na tela em que ela é apresentada; as interações com o usuário final da aplicação; a adaptação de conteúdo e de leiaute, entre outras [Soares et al. 2000, Barreto et al. 2016].

Aplicações multimídia tradicionais são geralmente definidas com uma linguagem específica, chamada de *linguagem de autoria multimídia*. Tais linguagens concentram-se na definição das mídias que farão parte da aplicação, a forma como serão apresentadas e sua organização ao longo do tempo [Blakowski and Steinmetz 1996, Hardman 1998]. Um subconjunto dessas linguagens, geralmente baseado em XML, usa uma abordagem declarativa fornecendo construções de alto nível de abstração para definição de aplicações multimídia.

As linguagens de autoria declarativa visam proporcionar uma clara separação entre a descrição da aplicação e sua implementação, *i.e.*, sua execução [Hardman 1998]. Assim as evoluções futuras da forma de como uma aplicação é apresentada não exigem uma redefinição, ou alteração, de toda a base das aplicações previamente especificadas. Exemplos de linguagens de autoria multimídia declarativas são HTML5 (*HyperText Markup Language*)², SMIL (*Synchronized Multimedia Integration Language*)³ e NCL (*Nested Context Language*) [ITU-Recommendation 2009]. Esta última é a linguagem adotada como padrão para desenvolvimento de programas interativos no Sistema Brasileiro de TV Digital. De acordo com [Soares and Barbosa 2009], uma vantagem das linguagens de autoria multimídia declarativas, é que elas fornecem maior facilidade para criação de aplicações por autores que não são programadores. Isto é importante, uma vez que as aplicações multimídia podem ser utilizadas em diferentes áreas, tais como web, TV digital e IPTV (do inglês, *Internet Protocol Television*); e criadas por autores com diferentes conhecimentos e perfis, tais como artistas e produtores de conteúdo.

O crescimento do número de dispositivos inteligentes, em especial sensores e atuadores, tem o potencial de mudar a forma de interação entre pessoas e tecnologia. Isso representou uma oportunidade interessante para sistemas multimídia a fim de expandir

²<https://www.w3.org/TR/html5/>

³<http://www.w3c.org/TR/SMIL3>

possibilidades de interação com o usuário. Muitas linguagens multimídia declarativas permitem a integração com código procedural ou imperativo. Como exemplo, HTML com Javascript ou NCL com Lua. Através de rotinas desenvolvidos nestas linguagens imperativas, é possível comunicar-se com sensores e atuadores para construir uma aplicação mulsemídia. Usando código imperativo, as aplicações podem obter informações de sensores para adaptar seu comportamento ou ativar atuadores de efeitos sensoriais para estimular os sentidos do usuário. Porém, a utilização de módulos imperativos de programação requer profissionais específicos (programadores) que terão que enfrentar vários desafios, como por exemplo, a diversidade de tecnologias para comunicação com sensores e atuadores. Muita iniciativas presentes na literatura trabalham para mudança destes cenários, conforme é apresentado na próxima seção.

2.2.1.1. Modelos, Ferramentas e Padrões

Algumas propostas na literatura buscam estender linguagens de autoria multimídia permitindo definir efeitos sensoriais [Guedes et al. 2017], ou propõem modelos/ferramentas para autoria de tais aplicações [Choi et al. 2011, Kim and Han 2014, Walth et al. 2013, de Mattos et al. 2020]. Na proposta de Guedes et al. [Guedes et al. 2017], atuadores ou sensores são modelados individualmente. Essa abordagem possui como desvantagem o forte acoplamento da aplicação mulsemídia com dispositivos específicos. Dessa forma, qualquer alteração no modo como um efeito sensorial é implementado ou qualquer alteração na instalação física utilizada para executar um aplicativo leva a uma nova especificação da aplicação. Algumas ferramentas ou linguagens de autoria se baseiam em um modelo multimídia/mulsemídia para a definição de seus componentes e dos relacionamentos entre eles. Em [Mattos and Muchaluat-Saade 2018], é proposto um modelo denominado MultiSEM (do inglês, *Multimedia Sensory Effect Model*), empregado como base para a ferramenta de autoria STEVE [de Mattos and Muchaluat-Saade 2018]. O editor STEVE permite abstrair do autor detalhes inerentes de implementação dos efeitos sensoriais facilitando assim a autoria.

A modelagem de aplicações mulsemídia também é proposta por [Josué et al. 2018], que especificam a representação de efeitos sensoriais e características do ambiente e do usuário com alto nível de abstração, permitindo que aplicações sincronizem tais efeitos com conteúdo audiovisual. Um argumento, levantado em [Josué et al. 2018], é que um nó de mídia pode representar tanto fragmentos de informação multimídia tradicional ou efeitos sensoriais. Visto que informação multimídia tradicional e efeitos sensoriais tem em comum o estímulo a sentidos do usuário. Com isso, os instrumentos já existentes para sincronização de mídias podem ser utilizados para sincronizar efeitos sensoriais. Como tratado em [Josué et al. 2018], cada tipo de nó possui *players* específicos responsáveis por exibir o conteúdo de um nó. Esta extensão permite que *players* específicos (associados também a atuadores específicos) possam ser utilizados para renderizar diferentes tipos de efeito sensorial. Outro ponto positivo é que uma vez definindo efeitos sensoriais como nós, características gerais destes efeitos, como sua intensidade e duração, ou específicas, como cor da luz ou essência de cheiro, são determinadas por atributos destes nós. Em [Rodrigues et al. 2019] os autores apresentam uma implementação da proposta de [Josué et al. 2018] para a linguagem NCL.

Além da modelagem de aplicações, alguns trabalhos empregam padrões para representar aplicações com múltiplos efeitos sensoriais. Dentre esses padrões, destaca-se o padrão MPEG-V [Kim and Han 2014], que especifica um formato para a troca de dados entre o “mundo real” (*i.e.*, a instalação física onde uma aplicação é executada e percebida pelo usuário), e o “mundo virtual”, ou seja, a aplicação em si como ilustrado na Figura 2.1. Kim e Han [Kim and Han 2014], por meio do padrão MPEG-V, definem um conjunto de linguagens que permitem, entre outras coisas, a descrição de objetos virtuais, efeitos sensoriais e comando para sua implementação na instalação física. A linguagem *Sensory Effect Description Language* (SEDL) define um conjunto de elementos para a descrição de efeitos sensoriais. Uma descrição SEDL, em um arquivo *Sensory Effect Metadata* (SEM), associa efeitos a conteúdo multimídia. Esta associação é feita por meio dos tempos de início e fim de um dado efeito. Um adaptador (não especificado no padrão) transforma uma descrição SEM em comandos a atuadores. Comandos do atuador e informações capturadas por sensores são descritos com a linguagem *Interaction Interface Description Language* (IIDL). Para permitir que comandos aos atuadores considerem preferências do usuário, estas são descritas usando a linguagem *Control Information Description Language* (CIDL). Essa linguagem também pode ser utilizada para descrever a capacidade de atuadores e sensores. É importante notar que a sincronização temporal contemplada pelo padrão MPEG-V segue um modelo baseado em linha do tempo (*timeline*).

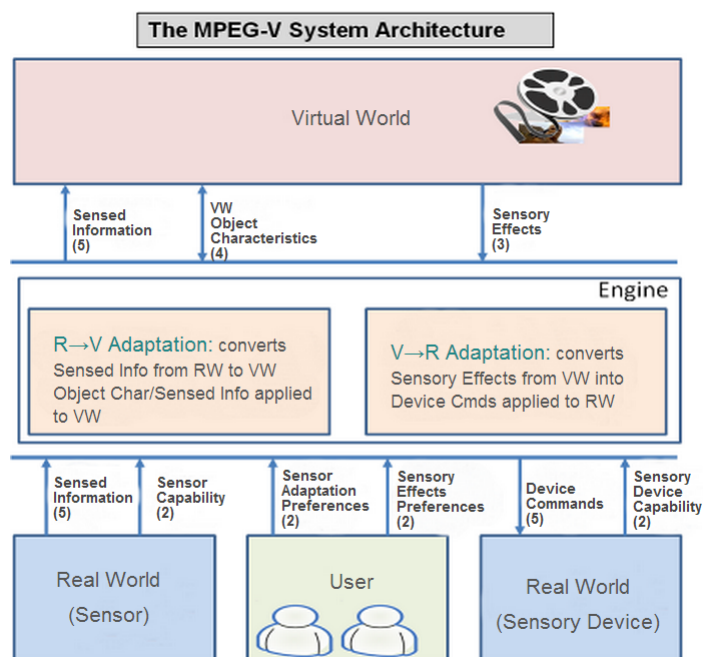


Figura 2.1. Arquitetura do padrão MPEG-V. ⁵

A especificação temporal através de *timeline* especifica sincronização entre os componentes de uma aplicação através do posicionamento dos objetos de mídia ao longo de um eixo temporal compartilhado. Como discutido em [Blakowski and Steinmetz 1996], essa abordagem de sincronização limita a expressividade do autor, pois não possibilita re-

⁵<http://wg11.sc29.org/mpeg-v/>

presentar eventos assíncronos como a interatividade do usuário. Em MPEG-V, especifica-se a sincronização entre um vídeo ou áudio e efeitos sensoriais. Uma outra limitação do padrão MPEG-V, é que, apesar de oferecer suporte à criação de metadados de efeitos sensoriais (arquivos SEM), ele não fornece mecanismos para especificar uma aplicação mulsemídia por completo contendo diversos objetos de mídia sincronizados (imagens, textos e outros vídeos e áudios).

O desenvolvimento de aplicações mulsemídia pode ser feita através da edição direta de um documento mulsemídia, utilizando uma linguagem específica, ou através de ferramentas gráficas. Walth et al. [Walth et al. 2013] propõem uma ferramenta denominada *SEVino* (*Sensory Effect Video Annotation*), que fornece uma linha do tempo dividida em canais, sendo um canal para cada tipo de efeito sensorial. Na linha do tempo, o autor define o tempo de início e duração do efeito. Para facilitar a definição destes tempos, a ferramenta permite a visualização do vídeo principal quadro a quadro. Após a fase de autoria, a ferramenta gera descrições compatíveis com o padrão MPEG-V indicando os tempos de início e fim dos efeitos criados. As descrições MPEG-V geradas pelo *SEVino* representam os efeitos sensoriais a serem executados em dispositivos físicos. A Figura 2.2 apresenta a interface do *SEVino*.

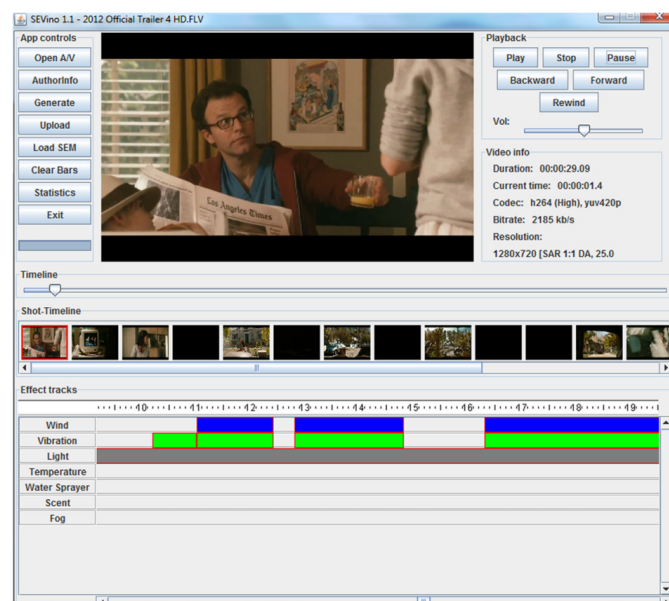


Figura 2.2. Interface gráfica do *SEVino* proposto por [Walth et al. 2013].

A criação de aplicações mulsemídia utilizando uma linha do tempo para definir a sincronização de efeitos sensoriais com conteúdos audiovisuais também é proposta por [Choi et al. 2011]. A ferramenta proposta por [Choi et al. 2011] é chamada *RoSE Studio* (do inglês, *Representation of Sensory Effects*) e exporta os efeitos sensoriais criados para os arquivos SEM. Outra característica interessante é que a ferramenta também oferece uma interface gráfica específica para definir efeitos de movimentos de poltrona. A Figura 2.3 apresenta a interface gráfica da ferramenta.

Outra ferramenta para criação de arquivos SEM através de uma linha de tempo é a ferramenta *SMURF* (*Sensible Media aUthoRing Factory*) [Kim 2013]. *SMURF* permite

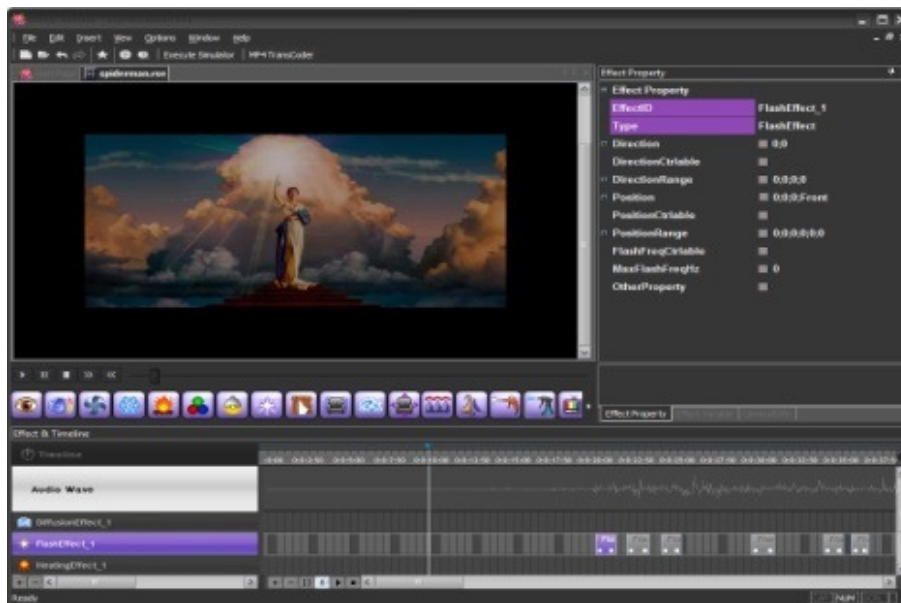


Figura 2.3. Interface gráfica do *RoSE Studio* proposto por [Choi et al. 2011].

ao autor criar uma composição de efeitos sensoriais facilitando assim a definição de novos efeitos. Além disso, os autores argumentam que a ferramenta também pode criar aplicações integradas a sensores do ambiente e do usuário. Informações coletadas do ambiente são por exemplo temperatura, umidade e iluminação. As informações vitais do usuário podem incluir sua pressão sanguínea ou frequência cardíaca. A Figura 2.4 apresenta a interface gráfica da ferramenta SMURF.

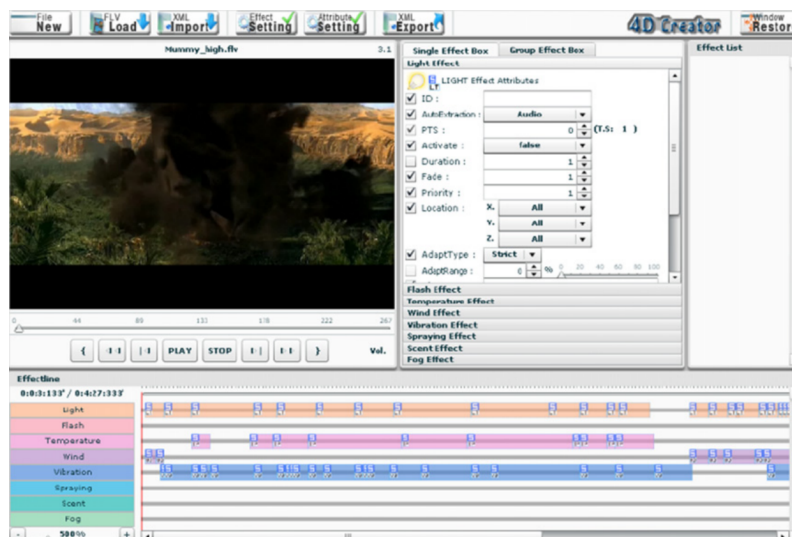


Figura 2.4. Interface gráfica do *SMURF* proposto por [Kim 2013].

Um conjunto de ferramentas para autoria de efeitos sensoriais também é proposto por [Shin et al. 2016]. Os autores se concentram em oferecer uma interface gráfica de fácil uso direcionada para construção de experiências imersivas para cinemas 4D. Duas ferramentas de autoria são propostas nesse trabalho. A ferramenta *Real4DEMaker* permite

a autoria de efeitos sensoriais individuais, enquanto a ferramenta *Real4DStudio* pode ser usada para autoria de diversos efeitos sensoriais em uma linha do tempo. O *Real4DStudio* pode ser usado para gerar um grupo de efeitos sensoriais e permite salvá-los para uso posterior. A Figura 2.5 apresenta a interface da ferramenta *Real4DStudio*. É importante notar que, como essas ferramentas são baseadas no padrão MPEG-V, elas herdam as limitações do paradigma de autoria de linha do tempo empregado pelo padrão. Uma dessas limitações é a falta de suporte para definição de eventos assíncronos tais como interatividade e testes condicionais.

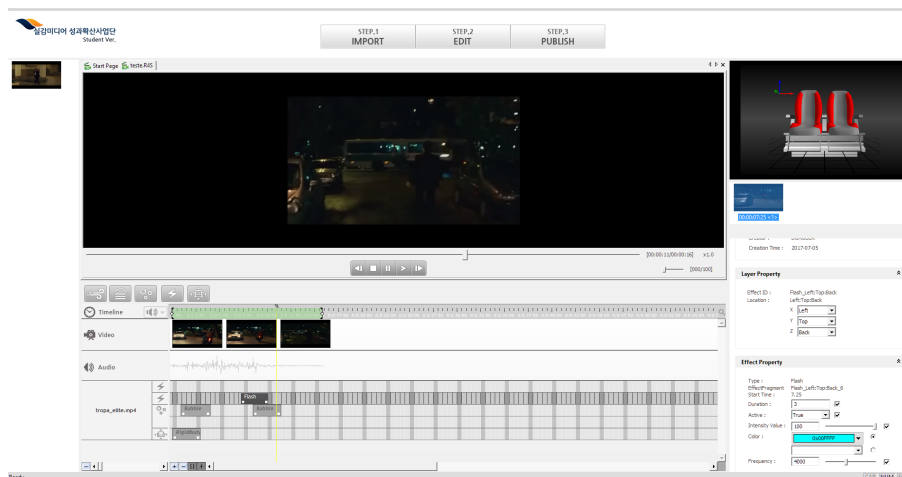


Figura 2.5. Interface gráfica do *Real4DStudio* - Student Version proposto por [Shin et al. 2016].

STEVE 2.0 [Mattos and Muchaluat-Saade 2018] é uma ferramenta de autoria multimídia que diverge das anteriores por não ser unicamente baseada no padrão MPEG-V. Com isso, *STEVE 2.0* oferece ao autor uma interface de linha do tempo, mas por outro lado, seu modelo temporal não é baseado em *timeline*, mas sim baseado em eventos. Utilizando *STEVE*, os efeitos sensoriais podem ser sincronizados com diversas mídias tradicionais (vídeos, imagens e textos) e não somente com um único vídeo. *STEVE 2.0* também permite sincronizar apenas efeitos sensoriais entre si sem a necessidade de conteúdo audiovisual. A interface gráfica do editor *STEVE 2.0* pode ser vista na Figura 2.6.

2.2.1.2. Linguagens

Sulema [Sulema 2017] propõe uma linguagem baseada em um sistema algébrico de conjuntos chamada ASAMPL para construção de aplicações que utilizem efeitos sensoriais. Segundo o trabalho de [Sulema 2017], é possível especificar não só variáveis de tipos primitivos mas também estruturas com conjunto de informações ordenadas, como encontrado em mídias contínuas. Um conceito interessante proposto pelo trabalho é poder agrupar estruturas compostas em contêineres chamados de agregados, além de definir toda a álgebra para manipulação destes agregados.

De acordo com [Sulema 2017], a utilização de estruturas de agregação é importante por permitir combinar mídias com informações que não são intrínsecas às mesmas.

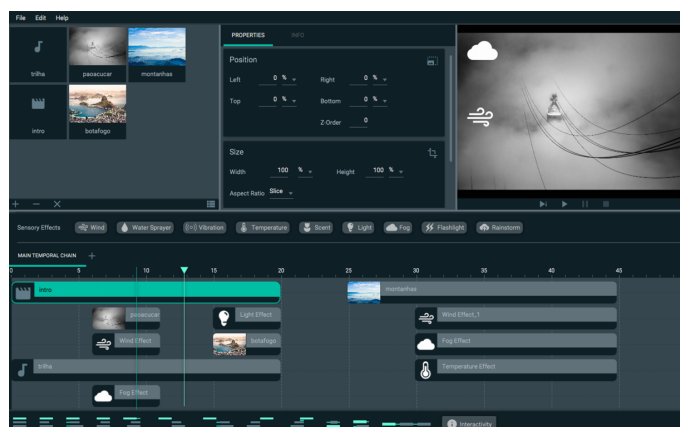


Figura 2.6. Interface gráfica do **STEVE 2.0** proposto por [Mattos and Muchaluat-Saade 2018].

Por exemplo, permite que informações, como temperatura em que a mídia deve ser exibida, ou a umidade do ambiente quando o vídeo foi gravado, sejam relacionadas ao conteúdo audiovisual. Outros trabalhos como [Soares and Barbosa 2009, Soares et al. 2010, W3C 2008] definem um conceito semelhante. O nó de composição permite o agrupamento de várias mídias. Além disso, cada mídia pode definir um conjunto de propriedades. Uma limitação do trabalho [Sulema 2017] é a estrutura de linha do tempo, impossibilitando contemplar eventos assíncronos e adaptação do conteúdo das aplicações.

Guedes et al. [Guedes et al. 2016] apresentam uma proposta de interação entre atuadores (luz, emissor de cheiro, ventilador) em sincronismo com conteúdo de TV seguindo o padrão brasileiro. Sua abordagem usa o ambiente declarativo Ginga-NCL, o *middleware* Ginga para enviar comandos para um dispositivo Intel Galileo⁶ com o kit Seed Studio Grove⁷. Com o objetivo de aumentar a sensação de imersão do usuário, a aplicação ativa atuadores presentes no ambiente da aplicação. A comunicação entre o Galileo e os atuadores é feita de forma cabeada.

A proposta de [Guedes et al. 2016] utiliza um *script* Lua de programação que recebe eventos da linguagem NCL através da API NCLua [Soares and Barbosa 2009]. Esses *scripts* são usados para acionar os dispositivos atuadores de acordo com a reprodução da mídia. Nessa arquitetura, os dispositivos devem ser previamente conhecidos e definidos na linguagem NCL, que especifica o documento mulsemídia, explorando um potencial já presente na linguagem. Por exemplo, a possibilidade de tratar eventos e interação com usuário ou compor vários objetos em um único contexto.

2.2.1.3. Ambiente Multiusuário e Interação Multimodal

Em aplicações mulsemídia, principalmente as que dão suporte a terapias, a participação de mais de um usuário é um ponto fundamental. Normalmente, essas terapias são conduzidas por um terapeuta que precisa controlar a execução da aplicação, e o paciente que

⁶<https://intel.com/content/www/us/en/do-it-yourself/galileo-maker-board.html>

⁷https://www.seedstudio.com/item_detail.html?p_id=1978

também pode interagir. Uma aplicação mulsemídia que dá suporte ao acesso simultâneo de múltiplos usuários, e consegue diferenciá-los, é considerada uma aplicação multiusuário. Como os diferentes usuários de uma aplicação multiusuário podem interagir com seu conteúdo, o trabalho [Guedes et al. 2017] propõe meios para especificar interações multiusuário. Na proposta de [Guedes et al. 2017], os usuários são especificados através do uso de um perfil, porém o trabalho não aborda como o sistema mulsemídia irá obter a descrição de tal perfil.

Algumas aplicações mulsemídia possibilitam que o usuário interaja com o conteúdo que está sendo apresentado. Quando uma aplicação mulsemídia dá suporte a diferentes tipos de interação (por controle remoto, por gesto ou por voz, por exemplo), elas são consideradas multimodais. Guedes et al. [Guedes et al. 2017] propõem um *framework* para apoiar interfaces de usuário multimodais em aplicações multimídia interativas. O *framework* integra diferentes tipos de modalidades de entrada e saída. O trabalho modela tipos de entrada multimodais que suportam novas modalidades de entrada, tais como gestos e reconhecimento de voz, e diferentes modalidades de saída, como os conteúdos audiovisuais tradicionais, sintetizadores de voz e atuadores. Na proposta de [Guedes et al. 2017], o mecanismo de reconhecimento de voz, utiliza arquivos SRGS (*Speech Recognition Grammar Specification*)⁸ para definir os trechos que devem ser reconhecidos durante a execução da aplicação. O *framework* é proposto para a linguagem NCL, entretanto não é implementado na prática.

A interação multimodal também é empregada no trabalho de [Santos et al. 2015], que usa uma abordagem orientada a eventos para construir ambientes interativos com efeitos sensoriais que são ativados pelo reconhecimento de gestos do usuário. O trabalho dividiu o processo de design em estágios bem definidos, mostrando o comportamento do sistema sob uma nova perspectiva e suas interações. Seguindo a abordagem proposta por [Santos et al. 2015], o ambiente foi descrito através de um conjunto de estímulos, ações e comportamentos e fornece meios para detalhar o comportamento desejado, de acordo com o contexto em que está inserido, com descrições de *frames* comportamentais e cadeias comportamentais. Segundo os autores, um *frame* comportamental é a unidade básica do comportamento desejado. Ela representa a associação entre eventos, relacionados às interações do usuário, e ações, relacionadas ao comportamento esperado do ambiente. As cadeias comportamentais são a definição de regras de comportamento que controlam o que é esperado em cada contexto de interação e as transições entre contextos.

O trabalho de [Santos et al. 2015] também apresenta a possibilidade de desenvolver e integrar conteúdos multimídia com interações naturais. Vale ressaltar que tanto os *frames* comportamentais quanto as cadeias comportamentais são representados pela descrição XML, sem detalhar a linguagem usada para descrever os gestos do usuário. Esse arquivo XML fornece a flexibilidade necessária para modificar dinamicamente, remover ou adicionar novos *frames* comportamentais às regras de comportamento. Os gestos foram reconhecidos utilizando linguagem C# em integração com o dispositivo Microsoft Kinect⁹.

Seguindo a mesma abordagem, o trabalho [Saleme et al. 2017] apresenta uma

⁸<http://www.w3.org/TR/speech-grammar/>

⁹<https://developer.microsoft.com/pt-br/windows/kinect/>

avaliação de tempo da integração entre uma plataforma mulsemídia distribuída chamada PlaySEM [Saleme et al. 2019] e uma aplicação interativa em que os usuários interagem por gestos, a fim de descobrir quanto tempo esse processo leva. Os resultados mostram valores na faixa de 27ms a 67ms, em média, gastos ao longo do processo, antes da efetiva ativação dos dispositivos de efeitos sensoriais em uma rede cabeada.

2.2.2. Execução de Aplicações Mulsemídia

Após a fase de autoria de uma aplicação mulsemídia, esta pode ser transmitida através de alguma rede de comunicação, ou ser distribuída através de alguma mídia de armazenamento. Nos casos em que a aplicação mulsemídia é transmitida através da rede, torna-se necessário algum mecanismo para gerenciar o carregamento e execução do conteúdo dos objetos de mídia e renderização de efeitos sensoriais que compõem a aplicação.

Em geral, essas tarefas são desempenhadas por um componente do ambiente de execução, denominado formatador mulsemídia ou máquina de apresentação. Dessa forma, durante a execução de uma aplicação mulsemídia, o formatador irá controlar a reprodução da aplicação, a comunicação com os dispositivos emissores de efeitos sensoriais (atuadores), e gerenciar os exibidores de conteúdo de mídia audiovisual, por exemplo decodificadores de áudio, vídeo e imagem.

A apresentação do conteúdo audiovisual de uma aplicação é realizada por *players* específicos, que em geral são acoplados ao sistema multimídia. Dessa forma, o acionamento dos *players* específicos para cada tipo de mídia pode ser feito de forma direta, sem a necessidade de um protocolo de comunicação. Já a comunicação com os atuadores, responsáveis pela renderização dos efeitos sensoriais, deve ser feita através de algum protocolo de comunicação, uma vez que esses dispositivos estão ligados ao sistema mulsemídia através de uma rede.

Como a comunicação do formatador com os atuadores ocorre através da rede, podem ocorrer atrasos no envio de comandos, caso esta rede sofra com algum tipo de congestionamento. Além disso, cada atuador possui um tempo específico para ser acionado e renderizar o efeito de fato, após o recebimento do comando de ativação. Por exemplo, um ventilador pode levar alguns segundos até conseguir atingir a velocidade desejada pelo autor da aplicação.

Outro ponto importante a ser considerado pelo formatador mulsemídia é o tempo que um efeito leva para ser percebido pelo usuário. Um exemplo disso é o efeito de aroma, que uma vez liberado pelo atuador, demora algum tempo até ser percebido pelo usuário. Esse tempo depende diretamente da distância em que o usuário se encontra do emissor de aroma. Caso essas questões não sejam consideradas no momento de execução da aplicação, podem ocorrer atrasos ou até mesmo falhas de sincronização do conteúdo audiovisual com os efeitos sensoriais, ou entre efeitos.

De acordo com [Josué et al. 2019], os atrasos na apresentação dos componentes de uma aplicação mulsemídia podem ser reduzidos, ou até mesmo evitados, caso seja empregado um mecanismo de preparação. A preparação proposta por [Josué et al. 2019], pode ser aplicada tanto a objetos de mídia tradicionais, como áudio e vídeo, quanto a efeitos sensoriais. Conforme apresentado em [Josué et al. 2019], a preparação de um

objeto de mídia consiste em instanciar o *player* específico para o tipo da mídia, e buscar antecipadamente parte do conteúdo conforme o espaço em *buffer* disponível no *player*. Já para os efeitos sensoriais, a preparação consiste no acionamento antecipado dos atuadores responsáveis pela renderização do efeito, considerando suas características.

A reprodução de aplicações mulsemídia é um tema abordado em diferentes propostas [Waltl et al. 2013, Cha et al. 2009, Murray et al. 2014, Luque et al. 2014], que visam aumentar a qualidade de experiência dos usuários que estão consumindo o conteúdo. [Waltl et al. 2013] apresentam um simulador que é capaz de exibir conteúdo multimídia e simular efeitos sensoriais sincronizados com a aplicação. A arquitetura do simulador é dividida em três camadas: camada de entrada de dados, camada núcleo e camada de interface com o usuário.

A camada de entrada de dados recebe as informações sobre os arquivos de áudio e vídeo a serem exibidos e os metadados de efeitos sensoriais. A camada núcleo contém os módulos responsáveis por processar as informações fornecidas pela camada de entrada. Nessa camada, um *parser* XML extrai as características dos efeitos sensoriais descritos nos metadados e encaminha para o módulo simulador.

Além de enviar os efeitos para o temporizador que irá escalonar a entrega dos efeitos, o simulador carrega os arquivos de áudio e vídeo e os envia para o decodificador. O temporizador então irá verificar se, no tempo de reprodução atual, um efeito deve ser disparado e ativar o dispositivo atuador correspondente. Por fim, a camada de interface com o usuário contém módulos responsáveis pela apresentação do conteúdo, como os *players* de áudio e vídeo, e os dispositivos atuadores. A Figura 2.7 apresenta o simulador proposto por [Waltl et al. 2013], que dá suporte a efeitos de luz, vento, umidade, vibração, temperatura, jatos de água e aromas.

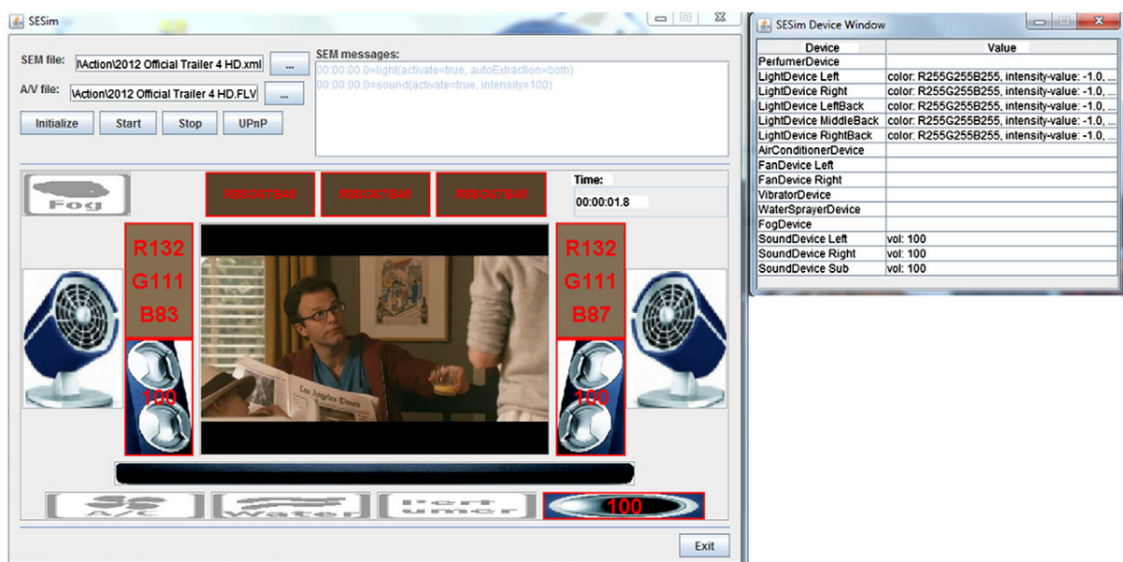


Figura 2.7. Simulador de aplicações multimídia com efeitos sensoriais proposto por [Waltl et al. 2013].

A simulação de efeitos sensoriais também é proposta por [Josué et al. 2018]. O

simulador desenvolvido como um plug-in para um software de edição 3D já existente, possibilita que o autor de uma aplicação mulsemídia organize objetos físicos e atuadores para construir um ambiente virtual 3D, como ilustrado na Figura 2.8. Durante a simulação dos efeitos sensoriais, o simulador monitora a rede, esperando a recepção de arquivos de metadados de efeitos para determinar quais atuadores devem ser ativados, e em qual momento. Quando um atuador é ativado, sua cor é alterada de acordo com o efeito que ele representa. Por exemplo, a cor branca é utilizada para representar efeitos de luz, o vermelho para efeitos de calor e verde para efeito de vento. A Figura 2.9 apresenta a interface do simulador em execução. Os atuadores ativos são destacados em vermelho. As intensidades de seus efeitos estão em 100%, ou seja, os atuadores estão com sua máxima cor.

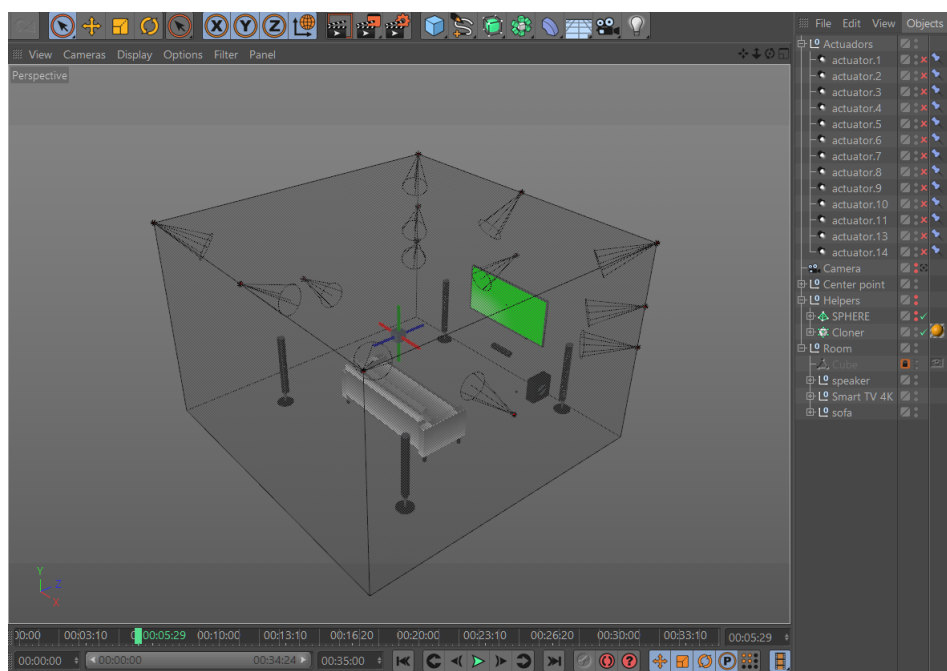


Figura 2.8. Interface do simulador de efeitos sensoriais proposto por [Josué et al. 2018].

Em [Cha et al. 2009], é proposto um *framework* para transmissão em *broadcast* de aplicações multimídia juntamente com estímulos e sensores táteis para prover maior sensação de imersão ao usuário do conteúdo. Na transmissão em *broadcast*, um mesmo conteúdo é transmitido para vários receptores ao mesmo tempo. Um exemplo de transmissão em *broadcast* é a que ocorre nos sistemas de televisão, onde uma emissora transmite o seu conteúdo para diferentes telespectadores.

O *framework* proposto por [Cha et al. 2009] possibilita a reprodução de aplicações com informações que estimulam receptores nervosos na pele (estímulos táteis) e informações sentidas através de força e movimento aplicado nos músculos (estímulos cinestésicos). A aplicação mulsemídia háptica apresentada em [Cha et al. 2009] é especificada utilizando o padrão MPEG-4 BIFS [Signes et al. 2000] para sintetizar e sincronizar os objetos de mídia audiovisuais com os efeitos relacionados ao sentido do tato.

⁹www.maxon.net/products/cinema-4d

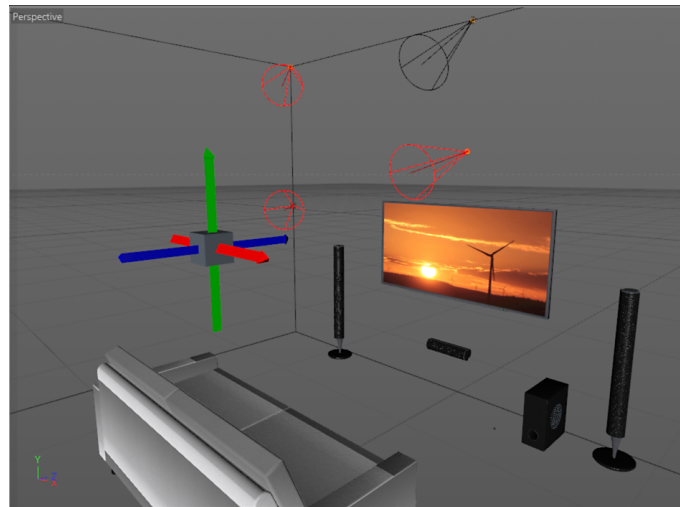


Figura 2.9. Ambiente de simulação em execução com atuadores ativos. [Josué et al. 2018]

Durante a reprodução do conteúdo, o formatador proposto por [Cha et al. 2009] obtém e analisa o grafo de cena BIFS para acionar os *players* responsáveis por reproduzir o conteúdo audiovisual e também os renderizadores de efeitos. Nos casos onde a aplicação possibilita interação do usuário por meio do tato, os sensores enviam a informação obtida do toque do usuário (força, posição, direção do movimento, por exemplo) para o formatador, como exibido na Figura 2.10. O formatador então processa a informação de interação do usuário e realiza alguma ação na apresentação do conteúdo.



Figura 2.10. Sensor tátil para interação do usuário empregado no *framework* proposto por [Cha et al. 2009].

A sincronização entre o fluxo audiovisual e fluxos de aromas é apresentada em [Murray et al. 2014], com o objetivo de analisar o uso desse tipo de efeito e o impacto da perda de sincronização entre os fluxos na QoE (do inglês, *Quality of Experience*) do usuário do conteúdo. O trabalho de [Murray et al. 2014] não especifica o ambiente de formatação do conteúdo, mas aponta alguns desafios para implementação do mesmo.

Na proposta de [Murray et al. 2014], os aromas são distribuídos em um eixo tem-

poral que segue uma recomendação definida em [Nakamoto and Yoshikawa 2006]. De acordo com [Nakamoto and Yoshikawa 2006], o tempo mínimo de apresentação entre dois aromas consecutivos deve ser de 5 segundos. Esse requisito pode ser validado na fase de autoria, e deve ser respeitado pelo formatador do conteúdo. Outra questão levantada pelo trabalho de [Nakamoto and Yoshikawa 2006] é o tempo que o usuário demora para detectar a presença de odores, após serem emitidos pelos dispositivos.

A utilização de estímulos multissensoriais para complementar o conteúdo audiovisual principal de um sistema de televisão é apresentada em [Luque et al. 2014]. O trabalho de [Luque et al. 2014] se baseia no fato de que a imersão completa do usuário é obtida quando todos os sentidos do corpo humano são acionados pela experiência multimídia multissensorial. [Luque et al. 2014] estuda a viabilidade da inserção de efeitos sensoriais e interações multimodais no sistemas de TV atuais. Para o estudo, [Luque et al. 2014] implementam a transmissão de uma partida de futebol em que diferentes dispositivos atuadores são disparados de maneira síncrona para entregar uma combinação de estímulos sensoriais, como ilustrado na Figura 2.11.

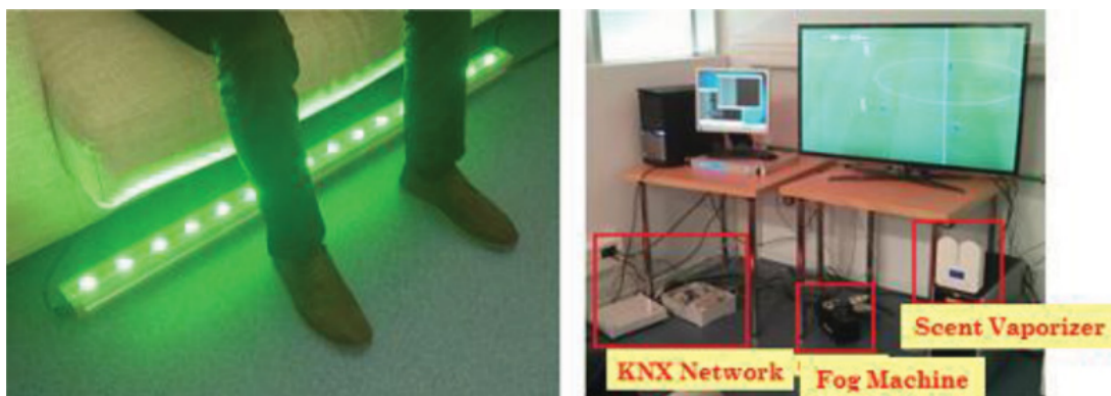


Figura 2.11. Ambiente de execução da aplicação proposta por [Luque et al. 2014].

A solução proposta por [Luque et al. 2014] dá suporte a utilização de alguns mecanismos definidos pelo MPEG-V para descrição dos efeitos sensoriais e a sincronização entre efeitos e objetos de mídia. O gerenciamento e a entrega dos diferentes tipos de dados do conteúdo multimídia multissensorial é realizado por um componente denominado “*Receiver Gateway*”, que pode utilizar os metadados de efeitos sensoriais propostos pelo MPEG-V para gerar os comandos para os dispositivos atuadores. Aplicações multimídia podem ser reproduzidas em diferentes plataformas, como em TV Digital, aparelhos celulares ou computadores. Nessas diferentes plataformas, a máquina de apresentação, responsável por coordenar a exibição do conteúdo pode ser implementada através de softwares dedicados ou até mesmo como um *plugin* para o navegador Web. A incorporação de efeitos sensoriais em aplicações Web é apresentada em [Waltl et al. 2014], que sincroniza os elementos da aplicação através da análise do tempo de reprodução do conteúdo audiovisual.

Na proposta de [Waltl et al. 2014], durante a execução da aplicação multimedial, quando a marcação de tempo atual de reprodução da aplicação é igual ao instante de apre-

sentação do efeito sensorial especificado pelos metadados, o *plugin* dispara a renderização do efeito. O instante de apresentação de efeito é especificado através de um intervalo de tempo, pois o *plugin* verifica se existe algum efeito a ser disparado a cada 30 ms.

Os trabalhos mencionados anteriormente apresentam soluções que permitem simular ou apresentar efeitos sensoriais acoplados a um ambiente de execução específico. Com o objetivo de permitir que o mesmo renderizador de efeitos seja reutilizado por diferentes interfaces de apresentação como Sistemas de TV Digital ou *videogames*, o trabalho de [Saleme et al. 2019] propõe uma plataforma denominada PlaySEM. A plataforma PlaySEM utiliza o protocolo MPEG-V apresentado anteriormente e possui três componentes: o *player* de vídeo SE (do inglês, *Sensory Effect*), o renderizador, e o controlador de dispositivo. O *player* é responsável pela execução do vídeo e leitura do arquivo MPEG-V que contém os metadados de efeitos sensoriais (arquivo SEM). A conversão dos metadados em comandos é realizada pelo renderizador, e o controle dos dispositivos físicos responsáveis pela renderização do efeito é feito pelo controlador. A utilização do PlaySEM, onde o renderizador está executando em um computador conectado a um monitor que exibe o conteúdo de vídeo, e se comunica com alguns atuadores responsáveis por efeitos de vento, luz e vibração, é apresentada na Figura 2.12.

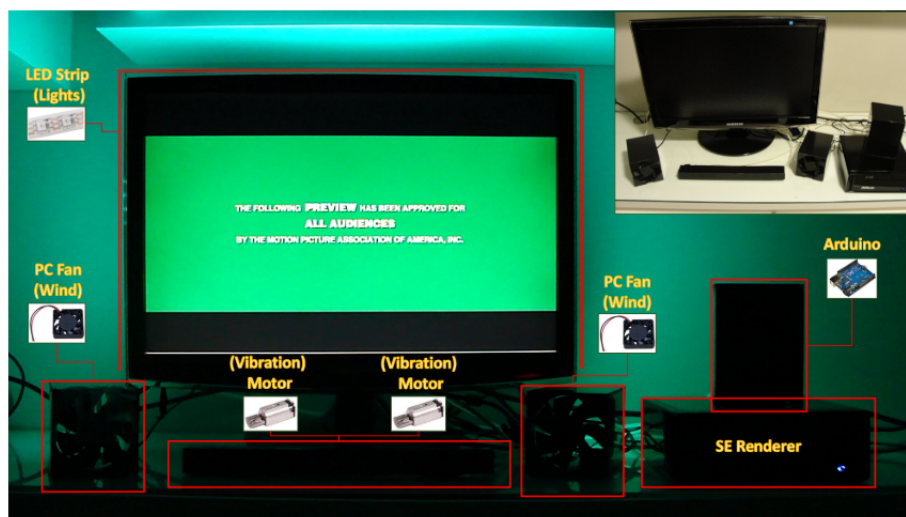


Figura 2.12. Ambiente de execução da aplicação mulsemídia proposta por [Saleme et al. 2019].

O Sistema Brasileiro de televisão digital possibilita que aplicações multimídia sejam transmitidas juntamente com o conteúdo do canal das emissoras de TV. Essas aplicações podem estar relacionadas ao conteúdo que está sendo exibido no canal ou não. Por exemplo, durante exibição de um filme, a emissora pode transmitir uma aplicação multimídia que permite ao telespectador comprar algum item que está sendo apresentado no filme.

Visando a incorporação de efeitos sensoriais em aplicações para o ambiente de TV Digital, [Rodrigues et al. 2019] propõem uma extensão ao *middleware* Ginga-NCL [ITU-Recommendation 2009]. O *middleware* Ginga-NCL é o componente do sistema de TV digital responsável pela apresentação e controle de aplicações multimídia especificadas na linguagem NCL. O *player* desenvolvido por [Rodrigues et al. 2019], processa o

documento mulsemídia especificado em NCL e gera metadados de efeitos sensoriais de acordo com o padrão MPEG-V para controlar os atuadores. Outra característica do *player* de efeitos sensoriais proposto por [Rodrigues et al. 2019] é que ele dá suporte à animação de efeitos. Isto é, permite que o posicionamento de um certo efeito sensorial seja alterado ao longo de sua renderização, fornecendo a sensação de que o efeito está “se movimentando” pelo ambiente de execução. Essa funcionalidade pode ser vista na Figura 2.13, que apresenta um efeito de luz sincronizado com a movimentação do sol sendo exibido no vídeo.

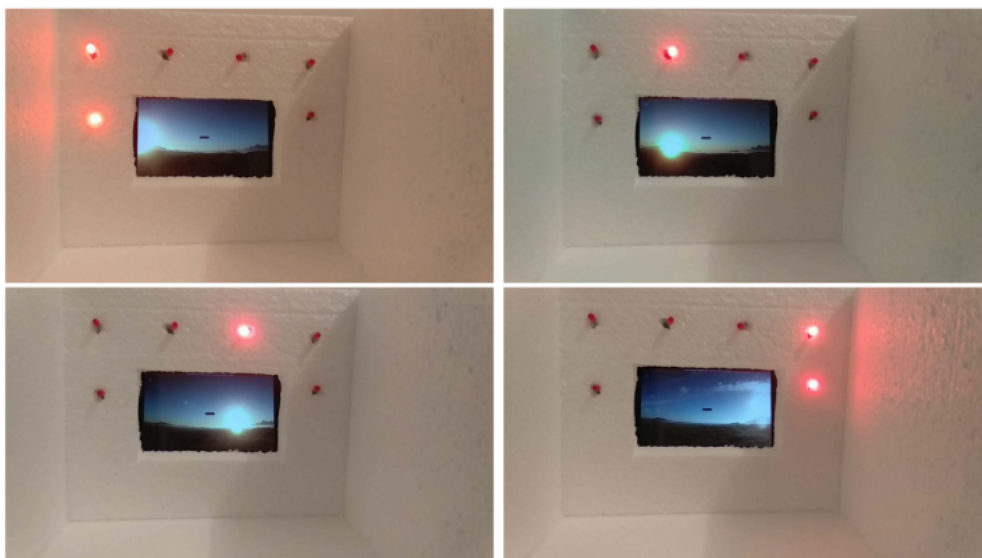


Figura 2.13. Exemplo de aplicação mulsemídia executando no *player* proposto em [Rodrigues et al. 2019].

2.3. Dispositivos para Ambientes Multissensoriais

A execução de aplicações mulsemídia demanda a incorporação de dispositivos ao ambiente de reprodução para renderização de efeitos sensoriais. Além dos dispositivos atuadores de efeitos, as aplicações também podem utilizar dispositivos sensores para permitir a coleta de informações do usuário ou do ambiente multissensorial. As seções a seguir descrevem alguns dispositivos disponíveis no mercado atualmente, para possibilitar renderização de efeitos e interação multimodal.

2.3.1. Dispositivos para Renderização de Efeitos Sensoriais

A execução de aplicações mulsemídia demanda a incorporação de dispositivos atuadores ao ambiente de reprodução para renderização de efeitos sensoriais. Esses atuadores podem renderizar efeitos olfativos, táteis (e.g vibração) e gustativos de forma sincronizada com a aplicação, proporcionando uma experiência mais imersiva ao usuário.

O difusor de aroma Moodo¹⁰, mostrado na Figura 2.14, é um dispositivo desenvolvido para casas inteligentes e funciona com um sistema de cápsulas com cristais de aromas: até quatro cápsulas podem ser acopladas e acionadas por meio de seus respecti-

¹⁰<https://moodo.co/>

vos *coolers* com intensidades reguláveis. O Moodo pode ser controlado por meio do aplicativo proprietário instalado em um *Smart Phone*, usando as teclas de controle do próprio difusor ou por meio de assistentes de voz e plataformas compatíveis. Para desenvolvedores, a Moodo também disponibiliza uma Interface de Programação de Aplicações (API) RESTful¹¹ que usa o formato JSON - *JavaScript Object Notation* para entrada e saída de informações. Uma das vantagens desse modelo é ser menos intrusivo que a *Feelreal Sensory Mask*, e a quantidade de formas para acesso e controle permitidas, incluindo conectividade Wi-Fi. A API RESTful também possui documentação interativa, que mostra como deve ser o JSON recebido e enviado e permite executar comandos. Como desvantagens, somente podem ser usadas as cápsulas fabricadas pela Moodo.



Figura 2.14. Difusor de aroma *Moodo* com cápsulas e imagem do aplicativo de controle do dispositivo¹³

O sistema *Scent Palette SDS100*¹⁴ mostrado na Figura 2.15 possui um dispersor de aroma USB com oito câmaras pressurizadas, onde cartuchos individuais de cheiro são inseridos. Ar comprimido é utilizado para projetar o aroma por uma duração pré-definida e em seguida uma rajada de ar inodora é usada para dispersá-lo. O SDS100 tem alcance de 3 a 6 metros na frente da unidade, dependendo de quantos *coolers* são usados e é possível controlar os momentos de acionamento e duração do efeito por meio dos softwares incluídos no sistema: uma aplicação para controle manual e o *Vizard Lite*¹⁵ para utilização em ambiente VR. A *BIOPAC Systems Inc*¹⁶ também disponibiliza códigos-fonte de exemplos, escritos em linguagem Python, para criação de aplicações customizadas para controle da *Scent Palette*. Apesar de oferecer um menor número de interfaces compatíveis que o Modoo, e necessitar da conexão USB, o SDS100 possui menor latência de renderização do efeito devido ao seu mecanismo de funcionamento. Outra vantagem é não

¹¹<https://rest.moodo.co/>

¹³<https://moodo.co/>

¹⁴<https://www.biopac.com/product/scent-delivery-system>

¹⁵<https://www.worldviz.com/virtual-reality-software-features>

¹⁶<https://www.biopac.com/>

ser obrigatório o uso das fragrâncias vendidas pela BIOPAC Systems Inc. Mesmo oferecendo aromas mais realísticos, provavelmente por ser especializada em produtos para educação e pesquisa, foram encontrados trabalhos [Rizzo et al. 2010] utilizando aromas não vendidos pela empresa ou que usaram paletas de cerâmica embebidas em fragrância [Munyan 2018].



Figura 2.15. Sistema difusor de aroma *Scent Palette*¹⁸

A *Feelreal Sensory Mask*¹⁹, mostrada na Figura 2.16 é uma máscara que permite renderizar efeitos de aroma, vento (frio e quente), calor, vibração, névoa e água. Pode ser usada acoplada em um *headset* VR ou sozinha como difusor de aroma. Possui um cartucho substituível com capacidade para 9 cápsulas de aroma individuais e possui modelos compatíveis com vários *headsets* VR: Samsung Gear VR, Oculus Rift, Oculus Go, HTC Vive e PlayStation VR. A comunicação pode ser feita via Bluetooth ou Wi-Fi e a empresa proprietária também disponibiliza um aplicativo mobile para controle do *Feelreal Sensory Mask*.

Outros aplicativos também disponibilizados para usuários são um editor (*Feelreal Video Editor*) e um *player* (*Feelreal Player*) compatíveis com os sistemas operacionais Windows e MacOS, Neles é possível customizar a sincronização dos efeitos sensoriais em vídeos 360 e customizar conjuntos de fragâncias. O dispositivo já está disponível para ser usado em alguns jogos (por meio da instalação de *mod*), filme e experiências (por exemplo, meditação guiada). Para a criação de jogos compatíveis, existem SDKs criados pela empresa para as *engines* Unity e Unreal.

Transdutores táteis e *bass shakers* permitem que usuários sintam as vibrações do som, transformando sons em estímulos táteis. Podem ser colocados sob o chão e assentos de móveis (cadeiras gamer, sofás e carros, por exemplo). Esse tipo de dispositivo precisa ser conectado a um amplificador ou amplificador subwoofer para o funcionamento

¹⁸<https://www.biopac.com/product/scent-delivery-system>

¹⁹<https://www.feelreal.com>

²¹<https://feelreal.com/product/feelreal-mask-for-htc-vive/>



Figura 2.16. Dispositivo *Feelreal Sensory Mask*²¹

esperado²². Na Figura 2.17 é mostrado um exemplo de transdutor tátil²³ e de um *bass shaker*²⁴.



Figura 2.17. Dispositivo transdutor tátil *TST429 Platinum Transducer* e *bass shaker ButtKicker LFE* respectivamente

Uma outra forma de estimular o tato é através da utilização de uma luva que suporta reprodução passiva na mão do espectador [Cha et al. 2009]. Essa luva é composta por 76 motores vibratórios, controlados por um microcontrolador, conforme mostra a Figura 2.18. Para a execução do efeito, o microcontrolador recebe os metadados referente ao estímulo, que contém informações sobre a ativação dos motores e a intensidade a ser aplicada.

2.3.2. Dispositivos Sensores

Dispositivos sensores permitem que a aplicação mulsemídia receba dados coletados do usuário e do ambiente multissensorial. Desta forma, podem ser utilizados para controle da aplicação, como por exemplo para iniciar ou parar um atuador ou modificar parâmetros

²²<http://www.bassshakers.com/how-to-install-bass-shakers/>

²³<https://clarksynthesis.com/tst429-platinum-transducer/>

²⁴<https://thebuttkicker.com/buttkicker-lfe/>



Figura 2.18. Luva para estímulo tátil. [Cha et al. 2009]

da aplicação, para prover diferentes formas de interação, como por exemplo para capturar movimento dos olhos, voz ou gestos, ou para realização de parametrização e calibração da aplicação.

Dispositivos rastreadores oculares i.e. *eye trackers*, permitem capturar para onde o usuário está olhando (i.e. *gaze point* ou *point of gaze*), detectando atenção, foco, presença do usuário [Tobii 2020] e em alguns casos, medições de tamanho das pupilas (i.e. pupilometria). A Figura 2.19 possui o esquema de funcionamento desses sensores de rastreamento ocular baseado em vídeo. Projetores de luz infravermelha formam um padrão dos olhos (i.e. reflexão córnea) e câmeras de alta resolução capturam imagens desse padrão e de movimentos do centro da pupila. Algoritmos matemáticos de processamento de imagem e de aprendizado de máquina são então utilizados nessas imagens para determinar a posição dos olhos e o *gaze point* [Tobii 2020].

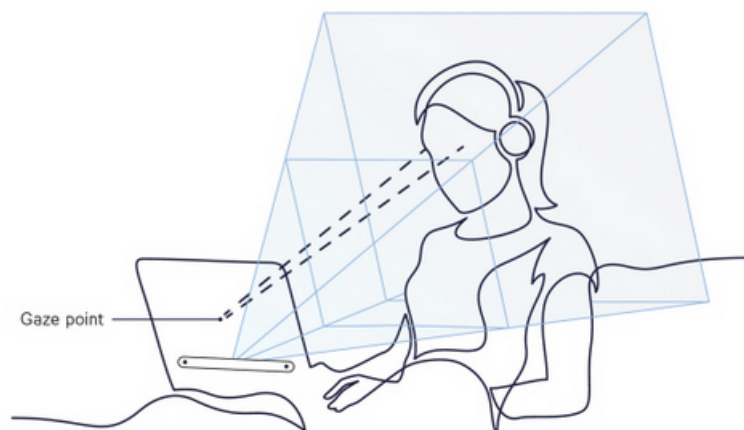


Figura 2.19. Ilustração do esquema de funcionamento do *eye tracker* [Tobii 2020]

O *eye tracking* proporciona diferentes tipos de interação com os olhos a pessoas que não conseguem ou não querem interagir por meio de outros dispositivos tradicionalmente utilizados, como mouse, teclado ou controle remoto. Pode ainda permitir o controle de uma aplicação multimedial seguindo a movimentação dos olhos do usuário [Silva et al. 2019]. Pode proporcionar mapas que mostram em que os usuários mantêm

sua atenção, de que forma e com o que se distraem. Combinado com outros dispositivos de entrada, pode criar novas experiências de interação em jogos. Além disso, *eye trackers* podem ajudar em pesquisas que utilizam a pupilometria para determinar situações de estresse, por exemplo.

A Figura 2.20 mostra um exemplo de sensor *eye tracker*: *Tobii Eye Tracker 4C*²⁵, que pode ser fixado na região inferior da tela onde será feita a interação. Esse modelo de sensor é compatível com o sistema operacional Windows, onde pode ser utilizado para interação, navegação e em aplicativos compatíveis (jogos). A *Tobii* também disponibiliza kits de desenvolvimento de software (*Software development kit - SDK*), como o *Core SDK* para desenvolvimento de aplicações e como base para outros kits, *Tobii Unity SDK* para desenvolvimento de jogos na *engine Unity* e *Stream Engine SDK* para criação de aplicações compatíveis por usuários avançados. A *Tobii Interaction Library*, parte do *Tobii EyeX Software Development Kit*, permite a criação de aplicações e jogos, provendo recursos em alto nível, tais como: *eye-gaze point*, posição dos olhos, fixação (foco), saber quando um *gaze point* sai ou entra em uma região, regiões clicáveis, focáveis ou realocáveis com o *gaze point* e detecção de presença do usuário. Os SDK são utilizáveis nas linguagens de programação C, C++ e C#.

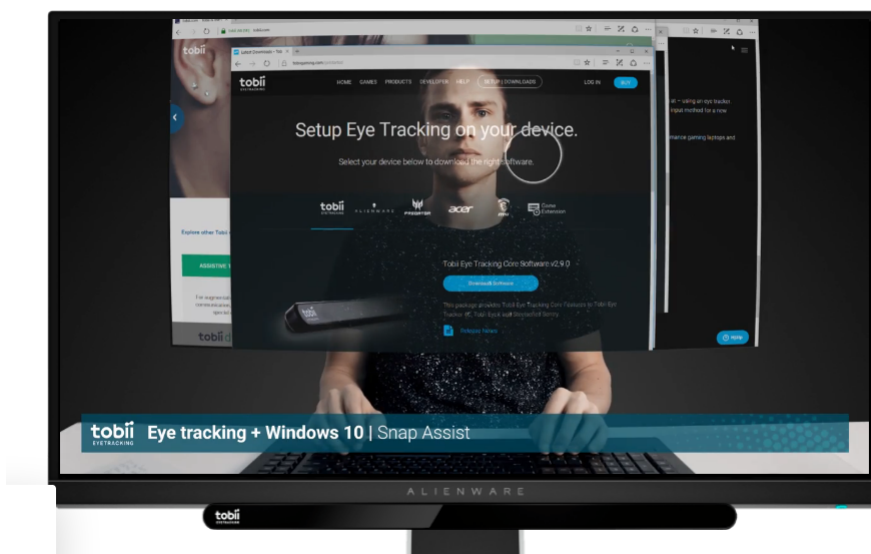


Figura 2.20. Sensor de *eye tracking* *Tobii Eye Tracker 4C* acoplado em um monitor [Tobii 2020]

O monitoramento pelo próprio indivíduo têm sido difundido pela indústria com a utilização de dispositivos vestíveis, como pulseiras inteligentes que monitoram nossas atividades, tais como *fitbit*²⁶ e *apple watch*²⁷, permitem o sensoriamento do batimento cardíaco em tempo real e geram relatórios que podem ser utilizados pela própria pessoa para avaliar sua atividade física ou levado a profissionais de saúde para auxiliar em diagnósticos [Piwek et al. 2016].

²⁵<https://gaming.tobii.com/tobii-eye-tracker-4c/>

²⁶<https://www.fitbit.com>

²⁷<https://www.apple.com>

Assistentes por voz são sensores que provêm diferentes formas de interação com a aplicação. O autor da aplicação pode, por exemplo, definir comandos por voz que alteram a sua execução e quais usuários podem vocalizá-los. As empresas desenvolvedoras de assistentes como Alexa²⁸, Google Assistant²⁹, Siri³⁰, Cortana³¹ e Bixby³² fornecem meios para que desenvolvedores possam integrá-los às suas aplicações e suportam a criação de rotinas (ou *skills*) que permitem, por meio de interface *drag-and-drop*, que usuários realizem várias ações por meio de um só comando.

2.4. Casos de Uso em Saúde

Como discutido anteriormente, aplicações mulsemídia podem ser utilizadas com foco em saúde, como em diferentes cenários para tratamento de doenças e transtornos, estímulo cognitivo e recuperação de lesões cerebrais. Como forma de ilustrar essas possibilidades de utilização, as seções abaixo apresentam alguns casos de uso de aplicações mulsemídia nesses cenários em questão.

2.4.1. Integração Sensorial em Indivíduos com TEA

Crianças com TEA são caracterizadas clinicamente por déficits de interação social e comunicação, bem como por interesses em atividades repetitivas [Posar and Visconti 2018]. As alterações sensoriais também são uma característica muito frequente, mas não são específicas do autismo, sendo uma característica frequentemente descrita em indivíduos com deficiência intelectual. Alguns padrões sensoriais são representados no transtorno do espectro do autismo em relação a estímulos sensoriais: hiporreatividade, hiperreatividade, busca sensorial e percepção aprimorada. As alterações sensoriais podem afetar negativamente a vida desses indivíduos e de suas famílias.

O trabalho de [Posar and Visconti 2018] discute uma deficiência não apenas das modalidades não sensoriais, mas também da integração multissensorial, isto é, problemas em processar diferentes modalidades sensoriais em conjunto. Dessa maneira, prover tipos específicos de estímulos sensoriais em intensidades e ambiente controlados pode melhorar a capacidade do sistema nervoso de lidar com processos de estímulos sensoriais. Com essa melhora, resultados apontam para redução de problemas de comportamentos e uma maior eficiência no aprendizado [Lang et al. 2012].

"Entender quais entradas sensoriais específicas causam desconforto em determinado indivíduo é o pré-requisito para reorganizar o ambiente em que ele vive e sua rotina diária para reduzir o máximo possível esse desconforto; nesse ponto de vista, um programa de dessensibilização pode ser útil. Em alguns casos, o uso de salas sensoriais, nas quais os indivíduos são submetidos a experiências sensoriais agradáveis de vários tipos, pode ser muito útil." [Posar and Visconti 2018]

²⁸<https://developer.amazon.com/pt-BR/alexa>

²⁹<https://assistant.google.com/>

³⁰<https://www.apple.com/br/siri/>

³¹<https://www.microsoft.com/en-us/cortana>

³²<https://www.samsung.com/br/apps/bixby/>

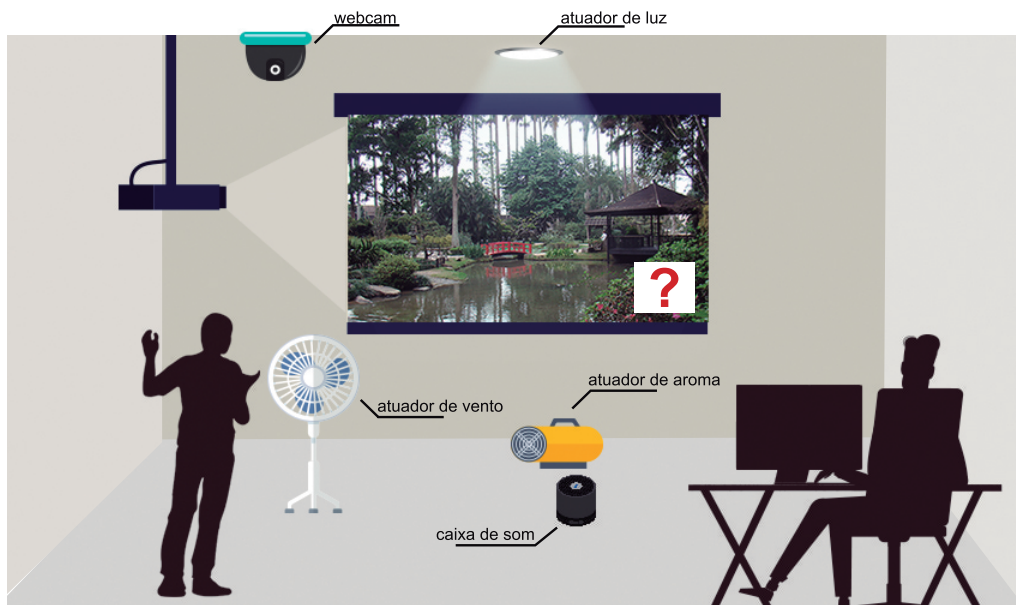


Figura 2.21. Exemplo de sala de terapia multiusuário

O caso de uso a seguir é motivado pelo trabalho feito por [Garzotto et al. 2019], em que é apresentado o *Magic Room*. O sistema tem objetivo de prover integração sensorial para crianças com TEA, por meio de efeitos sensoriais. A aplicação pode oferecer uma gama muito maior de oportunidades de interações sensoriais do que as terapias tradicionais. O trabalho obteve sucesso em integrar conteúdo audiovisual e efeitos sensoriais para desenvolver terapias de integração sensorial. Sendo assim, o *Magic Room* pode abrir caminho para novas intervenções terapêuticas para crianças com TEA.

A aplicação idealizada para esse caso de uso tem como objetivo ajudar na realização de terapias de integração sensorial em indivíduos com TEA. Esta consiste em orquestrar um ambiente imersivo onde estaria inserido o indivíduo e o seu terapeuta. São proporcionadas várias experiências sensoriais controladas pelo terapeuta. Desta forma, os efeitos sensoriais podem ser interrompidos a qualquer momento através do comando de voz “STOP” dito pelo terapeuta, quando o paciente demonstrar insatisfação através de seu semblante ou ainda quando seu batimento cardíaco atingir um patamar maior que um limiar pré-estabelecido. Os sinais biológicos como batimento cardíaco serão lidos por sensores conectados ao paciente.

A Figura 2.21 apresenta o layout da sala de terapia³³. Na sala, há um vídeo projetado e dois usuários, um é o terapeuta e outro é o usuário autista. Nessa sala, existem quatro atuadores de efeito sensorial: vento, luz, aroma e som. Além disso, uma *webcam* é usada para capturar os movimentos das mãos do usuário autista e um microfone é utilizado para gravar comandos de voz.

Neste caso de uso, há necessidade de diferentes tipos de usuários na sessão (terapeuta e paciente). Coletando dados do paciente e ambiente, o terapeuta pode verificar o andamento da terapia. Portanto é necessário especificar o perfil do paciente e do terapeuta

³³Imagens e gráficos obtidos de www.freepik.com

que participam da aplicação mulsemídia.

Uma vez definido o perfil, é possível armazenar informações associadas a cada paciente. Por exemplo, um conhecimento importante a ser extraído da interação do paciente é identificar limiares sensoriais do paciente no ambiente imersivo. Considere por exemplo a luminosidade. Qual seria o limite de luminosidade que um determinado paciente autista suporta? Neste caso, o terapeuta ao observar que o paciente se desorganiza diante de alguma característica do ambiente, pode determinar seu limite. Desta forma, a intensidade de cada um dos efeitos naquele momento pode ser armazenada como o limiar do paciente.

2.4.2. Relaxamento para Alzheimer

O Alzheimer é uma doença neurodegenerativa crônica e a forma mais comum de demência. A cada ano, 4,6 milhões de casos de Alzheimer são previstos e espera-se que esse número quase dobre em 2030 [Khan et al. 2018]. O paciente com Alzheimer necessita de cuidados especiais, precisando de supervisão constante na maioria do tempo. Esta tarefa é em geral de membros da família ou de profissionais contratados. Em especial, o cuidador familiar está mais suscetível a diversas patologias a decorrer do impacto que esta tarefa tem na privação de tempo e de suas atividades [Radziszewski et al. 2017].

Fraturas e quedas são principais motivos de internação hospitalar de idosos com Alzheimer [Van Doorn et al. 2003]. Um dos objetivos de supervisão de um paciente com Alzheimer é justamente evitar tais acidentes. Uma forma é amparar o idoso caso ele apresente comportamentos típicos da doença, como agitação, agressividade ou desorientação [Ballard et al. 2009]. Tais comportamentos, se não forem amenizados, podem eventualmente levar a problemas tanto para a pessoa com Alzheimer como para as que estão à sua volta [Hanford and Figueiro 2013].

Neste contexto, a fim de diminuir riscos a saúde, é desejável tratar a agitação e agressividade de um paciente com Alzheimer. Para amenizar tais comportamentos, pode-se optar por terapias medicamentosas. Porém os medicamentos causam efeitos colaterais e riscos para a saúde do paciente. Os efeitos colaterais variam desde desorientação, risco de quedas e até aumento no risco de morte [Ballard et al. 2009]. Em consequência disto, uma variedade de terapias alternativas são propostas para tratar agitação e agressividade em pacientes com Alzheimer.

Dentre estas, a terapia com música acumula maior base de evidências que indicam eficácia na redução de agitação em idosos com Alzheimer [O'Neil et al. 2011]. A terapia consiste da reprodução de sons que desencadeiem algum estado emocional no paciente. Na terapia com música, os sons podem ser genéricos, como uma música clássica de relaxamento ou individualizados, isto é, de acordo com a preferência musical do paciente. O trabalho [Gerdner 2000] indica uma redução significativa na agitação em tratamento com música individualizada em comparação com a música clássica. Além disso, vídeo em conjunto com música também tem sido utilizado como forma de aprimorar a qualidade de vida de pacientes com Alzheimer [Rubbi et al. 2016].

A terapia com aromas também é apresentada como uma intervenção possível para reduzir a agitação em formas de demência. Há indicações de que certos aromas podem

ser eficazes para ajudar as pessoas com demência a relaxar [O'Neil et al. 2011]. Dentre os aromas, destacam-se terapias com óleo de lavanda e erva-cidreira com possível efeito de redução da agitação em indivíduos com Alzheimer [Ballard et al. 2009].

Uma forma de terapia é o ambiente *Snoezelen* [Sarkar et al. 2018] para prover relaxamento em idosos com Alzheimer ao envolvê-los em atividades prazerosas e estimulantes. Tais ambientes podem dispor de recursos computacionais para oferecer às pessoas múltiplas oportunidades de interação e estímulo sensorial. Como apresentado no trabalho de [Sarkar et al. 2018], um dos recursos possíveis é a integração de uma aplicação interativa com sensores e atuadores de efeitos sensoriais.

Dessa maneira, a aplicação idealizada para este caso de uso tem como objetivo ajudar na realização de terapias de relaxamento em salas multissensoriais para doentes com Alzheimer. Esta consiste em possibilitar o terapeuta manipular efeitos sensoriais em ambiente imersivo onde estaria inserido seu paciente. Após o início da aplicação, o terapeuta pode iniciar ou parar a execução de efeitos sensoriais a qualquer momento por meio de comandos de voz, como "INICIA AROMA" ou "PARA LUZ".

A Figura 2.22 apresenta o layout da sala de terapia para relaxamento. Na sala, há um vídeo sendo apresentado na televisão e o usuário. Nessa sala, existem três atuadores de efeito sensorial: luz, aroma e som. Neste caso de uso, a aplicação multimedial executa na televisão. O vídeo apresenta um campo de lavandas e o aroma de lavanda é renderizado em sincronismo com o vídeo. Uma luz em tonalidade lilás é exibida em conjunto com uma música personalizada. Para conteúdo personalizado, o usuário precisa de um perfil definindo qual música prefere escutar para relaxar.



Figura 2.22. Exemplo de sala de terapia de relaxamento multissensorial

2.4.3. Exercícios Cognitivos

Conforme apresentado em [Antunes et al. 2006], o sistema funcional cognitivo consiste nas fases do processo de informação, como percepção, aprendizagem, memória, atenção, vigilância, raciocínio e solução de problemas. Esse sistema pode ser estimulado, através de diferentes atividades, que abordam diretamente a função cognitiva e/ ou o comprometimento cognitivo, e podem ser agrupadas de forma padronizada, criando um treinamento cognitivo [Martin et al. 2011].

A demência é um distúrbio cerebral que leva à redução do desempenho cognitivo e perda de memória. Desse modo, pessoas que possuem esse tipo de distúrbio podem sofrer com problemas na identificação de tempo, espaço e pessoas. A fim de auxiliar pacientes com demência, [Taher et al. 2015] propõem utilizar estímulos multissensoriais, através de efeitos de luz, cores, músicas, como técnica de terapia focando no relaxamento, estimulação e respostas individualizadas para estímulos diretos. Nos testes realizados por [Taher et al. 2015], durante oito semanas de intervenção com estímulos multissensoriais, foi possível verificar que este tipo de tratamento possui efeitos significativos na função cognitiva e na qualidade de vida de idosos com demência, além de diminuir a ansiedade e a depressão.

O estímulo multissensorial também é apresentado em [de Paula et al. 2019], porém aplicado ao contexto de jogos para exercício cognitivo. O jogo digital proposto por [de Paula et al. 2019] se baseia no teste do *Stroop*, criado por John Ridley Stroop em 1935 [Stroop 1935]. O teste do *Stroop* tem como objetivo verificar a reação do cérebro ao ser exposto a situações de discordância entre a leitura e a visualização de cores. No teste do *Stroop*, é exibido ao participante o nome de uma cor impresso em cor diferente (ex.: a palavra “azul” impressa em cor vermelha), e este deve dizer a cor em que a palavra está impressa, ao invés da palavra em si, como ilustrado na Figura 2.23.



Figura 2.23. Tela do jogo do *Stroop* proposto por [de Paula et al. 2019]

O jogo do *Stroop* proposto por [de Paula et al. 2019] incorpora o efeito sensorial de luz ao teste do *Stroop*, através de uma fita de LED acoplada à tela onde o texto com as cores é apresentado. Nessa versão do jogo do *Stroop*, a fita de LED era acessa com a cor correspondente à resposta correta, como apresentado na Figura 2.24. De acordo com os experimentos realizados, o uso do efeito de luz foi visto como um ponto positivo pelos usuários que participaram dos testes de usabilidade. Além disso, observou-se que houve

uma melhora no tempo de finalização da tarefa e na quantidade de acertos, ao comparar os testes realizados com o efeito de luz e sem o efeito.

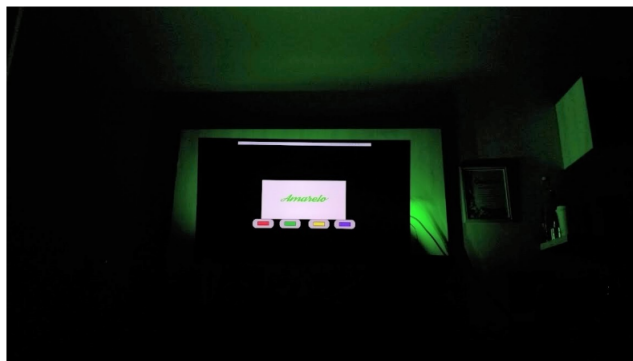


Figura 2.24. Jogo do Stroop com efeito sensorial de luz proposto por [de Paula et al. 2019]

Atividades ou jogos que fazem com que a pessoa memorize palavras ou imagens também são exemplos de exercícios cognitivos, e se enquadram na categoria de técnica de estimulação cognitiva. No trabalho de [Tortell et al. 2007], foi criado um jogo para testar a capacidade dos usuários em memorizar objetos de um ambiente virtual. O teste de cada usuário foi realizado em duas partes. Na primeira, foi apresentado ao usuário um ambiente virtual simulando uma região subterrânea como mostrado na Figura 2.25, onde o usuário poderia navegar através dela. Após 4 minutos percorrendo o ambiente virtual, o usuário ia para a segunda parte, onde tinha que responder a um questionário, marcando quais itens não foram observados no ambiente.

O trabalho de [Tortell et al. 2007] teve como objetivo avaliar o efeito do aroma na memorização de um ambiente virtual. Dessa forma, para um grupo de usuários, durante a navegação pelo ambiente, foi liberado um aroma composto de fragrâncias à base de óleo relacionado ao ambiente simulado. Esse aroma era liberado através de uma espécie de colar com 4 dispersores, conforme mostra a Figura 2.26. Além disso, outro grupo de usuários navegou pelo ambiente virtual, sem que nenhum aroma fosse liberado. Após a realização dos testes, [Tortell et al. 2007] verificou que os usuários que experimentaram o jogo com emissão de aroma tiveram, em média, 3,03 mais acertos do que aqueles que realizaram o experimento sem o aroma.

Uma outra possibilidade de estimulação cognitiva é utilizar um ambiente que forneça mecanismos para reabilitação cognitiva, conforme proposto por [Figuerola 2017]. Nesse trabalho, é proposto um ambiente personalizado multissensorial para estimular diretamente o sistema nervoso central, favorecendo o desenvolvimento de processos como atenção, memória e linguagem. Esse ambiente personalizado é definido como um espaço físico, onde o paciente pode receber vários estímulos auditivos, visuais, olfativos e táteis de forma clara.

O protótipo do ambiente personalizado desenvolvido por [Figuerola 2017] teve como objetivo a reabilitação de pacientes que sofreram um derrame. Com base neste protótipo, foram realizados alguns testes para verificar se o ambiente proposto possibilita que o paciente identifique cenários e recupere memórias. Além disso, os testes realiza-



Figura 2.25. Ambiente virtual utilizado na proposta de [Tortell et al. 2007]

dos buscaram analisar se um sentido pode ser considerado melhor do que os outros para estimular a memória.

Como resultado dos testes, foi possível verificar que os pacientes conseguiram identificar cenários, com base nos estímulos do ambiente proposto. Além disso, o percentual de pessoas que identificou corretamente o cenário, onde havia apenas estímulo auditivo foi de 1,9% nos testes com estímulo olfativo apenas. Entretanto, nos testes em que foram utilizados estímulos auditivos, olfativos e visuais, ao mesmo tempo, 59,4% das pessoas identificaram corretamente o cenário apresentado.

Os testes realizados por [Figueroa 2017] também buscaram verificar se a pessoa era capaz de se recordar de uma memória, com base no cenário visto. Nesses casos, foi possível observar que quando um único estímulo (auditivo ou olfativo) estava presente, 36 dos 74 indivíduos afirmaram recordar uma memória, quando havia 2 estímulos presentes (auditivo e olfativo), 23 dos 74 indivíduos recordaram uma memória e, quando os três estímulos foram apresentados ao mesmo tempo, 15 dos 74 indivíduos recordaram uma memória.

2.4.4. Terapia de Exposição em Realidade Virtual para Tratamento de TEPT

Como discutido nos trabalhos de Mishkind et al. [Mishkind et al. 2017] e Rizzo et al. [Rizzo et al. 2010], o uso de realidade virtual tem se revelado promissor para o tratamento de condições psiquiátricas, dentre essas o transtorno de estresse pós-traumático (TEPT). O TEPT está relacionado a padrões excessivos de estímulo e resposta baseados no medo e a um significado patológico na estrutura do medo, que mantém os sintomas e



Figura 2.26. Colar emissor de aroma utilizado na proposta de [Tortell et al. 2007]

impede a recuperação [Mishkind et al. 2017]. A terapia de exposição imaginária é uma das terapias recomendadas para serem utilizadas como primeira linha de tratamento para TEPT e requer que o paciente conte o trauma várias vezes [Mishkind et al. 2017]. Porém, devido ao trauma, muitos não querem participar do tratamento ou não conseguem falar sobre ele de forma vívida, com envolvimento emocional. Combinado com a ativação do medo em ambiente controlado e seguro, o envolvimento emocional é crítico para a terapia de exposição, fazendo com que a estrutura do medo mude. Nesse contexto, a realidade virtual pode servir como ferramenta para facilitar esse engajamento emocional por meio do sentimento de imersão e sentimento de presença que propicia. Além disso, como relatado no trabalho de Mishkind et al. [Mishkind et al. 2017], pacientes parecem mais dispostos a considerar a terapia de exposição com realidade virtual a outras formas de terapia de exposição. Outra vantagem da VR é o ambiente totalmente controlado, o que nem sempre é possível na exposição *in vivo*, e a possibilidade de reconstruir situações mais complexas de reproduzir e controlar, como um assalto, um ataque terrorista ou operações militares.

O trabalho retratado neste caso de uso é uma aplicação de Terapia de Exposição com Realidade Virtual (VRET) denominada *Virtual Iraq/Afghanistan*, com desenvolvimento apresentado em Rizzo et al. [Rizzo et al. 2010]. *Virtual Iraq* foi criado para tratamento de TEPT relacionado à exposição a combates nas operações militares no Iraque e Afeganistão (*Operation Enduring Freedom - OEF* e *Operation Iraqi Freedom - OIF*) em militares de serviço ativo e veteranos. Testes clínicos foram realizados com 20 militares (19 homens e 1 mulher, com idades que variam de 21 a 51 anos) em serviço ativo recém transferidos do Iraque, que já haviam tentado tratamentos para TEPT sem melhoria. Os resultados iniciais relatados são promissores: foram feitos 20 tratamentos completos, dos quais 16 militares não mais se enquadravam nos parâmetros para TEPT.

Virtual Iraq consiste em um conjunto de ambientes VRET com cenários de cidade e rodovia desértica customizáveis e ambientados no Oriente Médio, mostrados na Figura 2.27. As configurações do cenário “cidade” permitem inclusão de elementos como

estradas desoladas, um mercado, postos de controle, armazéns, mesquitas, lojas, edifícios em ruínas, e lotes de terra com lixo espalhado. O terapeuta que controla a sessão pode colocar o usuário em pontos específicos do cenário, incluir ou retirar veículos ativos nas vias assim como incluir ou retirar pedestres militares e civis animados, de acordo com as necessidades do paciente. É permitido que o usuário acesse o interior de edifícios e telhados e navegue pelos 24 bairros da cidade desértica a pé.



Figura 2.27. Cenários de cidade, estrada no deserto (utilizando o veículo HUMVEE) e com visão noturna ativada em *Virtual Iraq* [Rizzo et al. 2010]

O cenário “rodovia deserta” é formado por uma rodovia em uma extensa área desértica e contém elementos como duna de areia, algumas pequenas áreas de vegetação, pontes, destroços de batalhas, estruturas quebradas e intactas, postos de controle, detritos e figuras humanas. Nesse ambiente, o usuário está no interior de um veículo automóvel multifunção de alta mobilidade (HUMVEE), e pode ser posicionado como motorista, passageiro ou nas torres acima do teto do veículo. O HUMVEE suporta também a percepção de viagem em comboio ou solo. É possível selecionar o número de soldados na cabine e o quão machucados eles podem ficar durante ataques ao telhado ou com Dispositivos Explosivos Improvisados - IED (do inglês, *Improvised Explosive Device*), por exemplo.

O paciente navega pelos cenários usando controles com vibração (*force feedback*) ou uma réplica da arma M4, em ambos os casos sem a possibilidade de atirar. O clínico em controle da sessão utiliza a interface apresentada na Figura 2.28 para ajustar e controlar elementos como clima, período do dia, iluminação, visão noturna e sons ambientes (vento, barulho da cidade, alguém rezando). A interface também possibilita a ativação em tempo real de efeitos como vibração, áudio direcional 3D e estímulos olfativos. Assim, o clínico responsável customiza a experiência incluindo e retirando os estímulos de acordo com as

necessidades individuais do paciente.



Figura 2.28. Versão *wireless* da interface utilizada pelos clínicos para controle da aplicação *Virtual Iraq* [Rizzo et al. 2010]

Como parte dos estímulos auditivos, é possível ativar gatilhos auditivos como motores, disparos de armas e vento. A interface de controle também permite acionar eventos audiovisuais dinâmicos como sobrevoos de helicópteros, ataques a pontes, veículos explodindo e IEDs. Para os estímulos de vibração, são usados controles com *force feedback* e transdutores de som áudio-tátil, estes últimos localizados sob o chão da plataforma onde o paciente está e abaixo do assento desse.

Os estímulos olfativos são renderizados por meio de uma *Scent Palette*. É possível disparar aromas de fumaça (por exemplo, enquanto o usuário caminha próximo a um veículo em chamas), comidas típicas sendo cozinhadas, borracha queimada, lixo, explosivo sem fumaça, odor corporal, diesel, combustível, especiarias alimentares iraquianas e pólvora.

A Figura 2.29 mostra parte da estrutura do *Virtual Iraq*³⁴: o clínico controla os estímulos recebidos pelo paciente usando um interface, enquanto acompanha o que o paciente experiencia a partir de um monitor que espelha o conteúdo do *headset*. A Figura 2.30 mostra um protótipo inicial que foi utilizado em testes centrados no usuário com apoio da *Army Combat Stress Control Team* no Iraque.

2.4.5. Livro Multissensorial para Crianças com Dislexia

A dislexia é um transtorno definido pela incapacidade do cérebro em executar tarefas que requerem processamento de estímulos breves em rápida sucessão temporal, impedindo parcialmente que a pessoa disléxica consiga ler e compreender o que está lendo e dificultando seu aprendizado [Kast et al. 2007].

Treinamentos usando mídias audiovisuais para tratamento de crianças disléxicas e não-disléxicas apresentado em Kast et al. [Kast et al. 2007] obtiveram melhorias significativas na habilidade de escrita desses dois grupos de crianças. Em Silva et al. [Silva et al. 2019], são discutidos trabalhos que usam rastreamento ocular e estímulo de

³⁴<https://youtu.be/OkTHba0XM2c>



Figura 2.29. Estrutura do *Virtual Iraq* mostrada em vídeo de demo feita por Skip Rizzo para seção de documentários (Frontline) da emissora PBS [Rizzo et al. 2010]



Figura 2.30. Testes centrados no usuário feitos com protótipos iniciais do *Virtual Iraq* e com *feedback* da *Army Combat Stress Control Team* no Iraque [Rizzo et al. 2010]

outros sentidos, além da visão e audição, com resultados positivos para melhoria de aprendizado.

O caso de uso retratado nesta seção é apresentado em [Silva et al. 2019] e consiste em uma ferramenta multimídia para leitura de livros infantis. A Figura 2.31 mostra a ferramenta em funcionamento: estímulos auditivos, táteis (vento), visuais (variação de intensidade e cor da luz), e olfativos são acionados de forma automática à medida que o usuário lê o conteúdo do livro. Um *eye-tracker* é usado para a sincronização dos efeitos sensoriais com a leitura e para realizar ações como avançar e voltar pelas páginas do livro.

Após o usuário escolher o livro, cada página é carregada com seguintes elementos: linhas com o texto a ser lido, botões para avançar e voltar, título do livro, contagem de páginas e botão para encerrar a leitura. Para acionar os botões, o usuário deve fixar o olhar por 3 segundos no botão da ação desejada. À medida que o usuário lê as linhas de texto,

os efeitos vão sendo automaticamente renderizados.

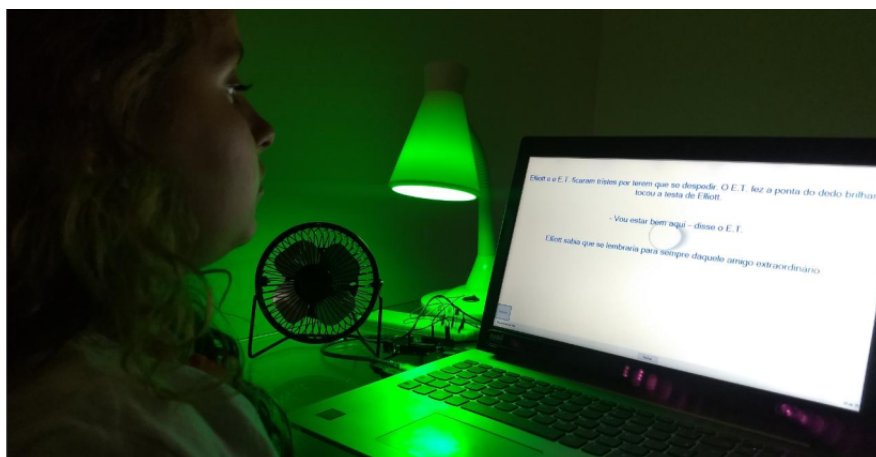


Figura 2.31. Livro multissensorial em uso [Silva et al. 2019]

2.5. Desafios e Tendências Futuras para Aplicações Mulsemídia

Diferente do conteúdo audiovisual que pode ser representado através de unidades de informação (por exemplo, os *pixels* de uma imagem ou os *frames* de um vídeo), a representação de um efeito sensorial através de unidades de informação é algo ainda pouco explorado na literatura. Desse modo, esta seção irá levantar alguns dos desafios relacionados ao tema deste capítulo, e também apresentar algumas propostas relacionadas à codificação digital de efeitos sensoriais.

O desenvolvimento e utilização de aplicações mulsemídia impõem alguns desafios tanto para os autores de aplicação, quanto para os fabricantes de dispositivos responsáveis pela renderização de efeitos sensoriais. Em relação ao ambiente de reprodução, os principais desafios estão na construção de dispositivos para a apresentação dos efeitos para estímulo do olfato e paladar.

Ambientes mulsemídia são por definição multimodais [Ghinea et al. 2014]. Os efeitos sensoriais (associados a conteúdo audiovisual) precisam ser entregues por seus atuadores apropriados ou devem ser modificados por informações provenientes de sensores. Para suportar esse tipo de relação, os dispositivos distribuídos devem ser capazes de se comunicar uns com os outros e também com seu controlador. Esta comunicação deve ser adaptável para uma série de dispositivos heterogêneos. Como visto neste capítulo, a comunicação de eventos entre sensores e atuadores pode permitir que a aplicação mulsemídia gere efeitos adaptados ao ambiente ou ao usuário. Adicionalmente, é de grande benefício adicionar também a comunicação destes dispositivos com servidores na nuvem [Alvi et al. 2015], de forma a fazer uso de um poder computacional que seria inviável para usuários convencionais. Este novo sistema integrado de sensores, atuadores, aplicação e Internet constitui a base de um novo ambiente mulsemídia inteligente.

Alguns desafios também são introduzidos ao utilizar desta abordagem, tal como a garantia da Qualidade do Serviço (QoS) em aplicações mulsemídia. A resposta dos atuadores deve ser entregue próximo ao tempo real para preservar a experiência do usuário,

no entanto a latência de processamento de algumas aplicações em nuvem (e.g reconhecimento de vídeo) pode ser proibitiva [Alvi et al. 2015]. Uma potencial solução para este problema é o avanço da implementação da futura geração de telecomunicação móvel, o 5G. A latência de comunicação ponto-a-ponto pela rede poderá ser diminuída significativamente. Com isso, a conectividade de rede ultra-confiável e ultra-responsiva permitirá fornecer remotamente experiências em tempo real que serão percebidas como tão rápidas quando a sensação tátil [Simsek et al. 2016]. Esta mudança de paradigma permite que sensores e atuadores sejam interfaces de comunicação de longa distância para as sensações humanas. Exemplos de experiências mulsemídia deste tipo são ambientes virtuais colaborativos multissensoriais. Um usuário não só irá visualizar todo o ambiente como também poderá interagir com outros participantes em tempo real. Sensores e atuadores hápticos poderão transmitir a sensação de toque entre os participantes. Essa nova rede que dar á suporte a essas aplicações, permitindo atrasos menores do que 1ms, vem sendo chamada de Internet Tátil (*Tactile Internet*).

Apesar dos avanços em ferramentas e modelos para facilitar o esforço de autoria, o processo de criação de uma aplicação mulsemídia ainda é um trabalho muito custoso em termos de esforço e tempo. Tal situação é agravada quando é necessário especificar a sincronização entre um grande número de objetos de mídia audiovisuais e efeitos sensoriais, bem como diferentes tempos de execução desses efeitos. Essa tarefa é muito custosa e pode induzir a erros [Abreu et al. 2018]. Neste contexto, [Waltl et al. 2013] indicam que uma forma de gerar tais marcações de maneira semiautomática incentivaria a adoção de tais aplicações pela comunidade. Diversas ferramentas possuem uma certa geração automática, porém ela se limita apenas a analisar quadros ou som do vídeo para emitir efeitos sensoriais. Um efeito de aroma, por exemplo, necessita de mais informação contextual da cena para ser automatizado, por exemplo indicando qual objeto está emitindo o aroma, ou qual é a distância e direção deste objeto em relação ao usuário. Portanto recentemente uma nova metodologia para resolver esse problema tem surgido ao empregar algoritmos de aprendizado de máquina para compreender cenas audiovisuais e correlacioná-las com a execução de certos efeitos sensoriais [Abreu et al. 2018]. Assim sendo, valiosos esforços nessa área serão a construção de *datasets* para treinamento de redes neurais para reconhecimento de efeitos sensoriais em diversos contextos. Outro valioso trabalho futuro é integrar em ferramentas de autoria mulsemídia a aprendizagem por reforço (*reinforcement learning*), que permite ao autor adaptar (i.e., treinar em tempo real) o reconhecimento da rede neural de acordo com seus objetivos.

2.6. Considerações Finais

A apresentação de diferentes efeitos sensoriais de forma sincronizada possibilita a criação de ambientes multissensoriais e aplicações, que podem ser empregadas em diferentes áreas, dentre elas a saúde. Ao estimular diferentes sentidos, essas aplicações fornecem uma maior sensação de imersão do usuário [Luque et al. 2014] no ambiente virtual. A sensação de imersão em ambientes virtuais é um ponto importante para a criação de simulações, e até mesmo para tornar o conteúdo mais atrativo para o usuário.

Uma outra forma de construir ambientes imersivos é através do uso da tecnologia de realidade virtual. Na área da saúde, essa tecnologia pode ser empregada para tratamento de condições psiquiátricas, como transtorno de estresse pós-traumático e a

ansiedade [Mishkind et al. 2017], por exemplo. Através da realidade virtual é possível apresentar ao usuário um cenário que simula alguma situação do dia-a-dia que lhe causa desconforto ou ansiedade, utilizando um óculos de realidade virtual. Dessa forma, é possível criar um ambiente controlado e seguro para realizar terapias de exposição em usuários com esse tipo de doença.

Os efeitos sensoriais também podem ser empregados para auxiliar em terapias de relaxamento, como a aromaterapia [O’Neil et al. 2011, Ballard et al. 2009] e terapia com música [O’Neil et al. 2011]. Esse tipo de terapia pode ser aplicada a pacientes com Alzheimer, que em alguns casos apresentam agitação e agressividade. Dessa forma, aplicações multimedíadas podem ser utilizadas para criar um ambiente multissensorial onde o terapeuta manipula efeitos sensoriais a fim de provocar o relaxamento no paciente.

Outros sintomas comuns em pacientes com Alzheimer são a redução do desempenho cognitivo, e perda de memória. De acordo com alguns estudos [Taher et al. 2015, Martin et al. 2011], os exercícios de estimulação cognitiva podem ser utilizados para preservar ou melhorar o desempenho das funções cognitivas. Nesse contexto, os efeitos sensoriais podem ser empregados em jogos digitais, de forma a oferecer estímulos multissensoriais.

Atualmente no mercado, existem diferentes dispositivos que permitem a renderização de efeitos sensoriais. Esses dispositivos permitem a emissão de aromas, de estímulos táteis e vibração, de efeitos luz, vento e temperatura. Além dos renderizadores de efeitos, existem dispositivos capazes de capturar informações do ambiente e do usuário, chamados sensores.

Algumas aplicações multimedíadas podem dar suporte à interação do usuário e à adaptação de conteúdo. Nessas aplicações, os sensores podem ser empregados tanto para capturar interações do usuário (por exemplo, interação por voz ou por movimentação dos olhos), quanto obter informações sobre as condições do usuário, como batimento cardíaco. Além disso, os sensores podem ser empregados para monitorar o estado do ambiente, como temperatura, luminosidade e pressão, por exemplo.

Este capítulo apresentou o conceito de aplicações multimedíadas aplicado à área da saúde. Foram apresentados alguns casos de uso, relacionados à integração sensorial em indivíduos com Transtorno do Espectro do Autismo, relaxamento de pacientes com Alzheimer, exercícios cognitivos, terapia de exposição para pacientes com transtorno de estresse pós-traumático e dislexia.

O presente capítulo também apresentou as principais linguagens e ferramentas para criação e execução de aplicações multimedíadas, propostos na literatura. Além disso, foram levantados alguns desafios e tendências futuras relacionados à construção de ambientes multissensoriais, que é uma área bastante fértil para pesquisas e trabalhos aplicados à saúde.

Referências

[Abreu et al. 2018] Abreu, R., dos Santos, J., and Bezerra, E. (2018). A bimodal learning approach to assist multi-sensory effects synchronization. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

- [Alvi et al. 2015] Alvi, S. A., Afzal, B., Shah, G. A., Atzori, L., and Mahmood, W. (2015). Internet of multimedia things: Vision and challenges. *Ad Hoc Networks*, 33:87–111.
- [Antunes et al. 2006] Antunes, H. K., Santos, R. F., Cassilhas, R., Santos, R. V., Bueno, O. F., and Mello, M. T. d. (2006). Exercício físico e função cognitiva: uma revisão. *Revista Brasileira de medicina do esporte*, 12(2):108–114.
- [Ballard et al. 2009] Ballard, C. G., Gauthier, S., Cummings, J. L., Brodaty, H., Grossberg, G. T., Robert, P., and Lyketsos, C. G. (2009). Management of agitation and aggression associated with alzheimer disease. *Nature Reviews Neurology*, 5(5):245.
- [Ballard et al. 2002] Ballard, C. G., O’Brien, J. T., Reichelt, K., and Perry, E. K. (2002). Aromatherapy as a safe and effective treatment for the management of agitation in severe dementia: The results of a double-blind, placebo-controlled trial with melissa. *The Journal of clinical psychiatry*.
- [Barreto et al. 2016] Barreto, F., Batista, D. T., dos Santos, J. A., and Muchaluat-Saade, D. C. (2016). Ncl-tester: Graphic application for ncl documents temporal test creation. In *Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web*, pages 91–94. ACM.
- [Blakowski and Steinmetz 1996] Blakowski, G. and Steinmetz, R. (1996). A media synchronization survey: Reference model, specification and case studies. *Journal on Selected Areas in Communications*, 14(1):5–35.
- [Cha et al. 2009] Cha, J., Ho, Y.-S., Kim, Y., Ryu, J., and Oakley, I. (2009). A framework for haptic broadcasting. *IEEE MultiMedia*, 16(3):16–27.
- [Choi et al. 2011] Choi, B., Lee, E.-S., and Yoon, K. (2011). Streaming media with sensory effect. In *Information Science and Applications (ICISA), 2011 International Conference on*, pages 1–6. IEEE.
- [de Mattos and Muchaluat-Saade 2018] de Mattos, D. P. and Muchaluat-Saade, D. C. (2018). Steve: a hypermedia authoring tool based on the simple interactive multimedia model. In *Proceedings of the ACM Symposium on Document Engineering 2018*.
- [de Mattos et al. 2020] de Mattos, D. P., Muchaluat-Saade, D. C., and Ghinea, G. (2020). An approach for authoring mulsemmedia documents based on events. In *2020 International Conference on Computing, Networking and Communications (ICNC)*, pages 273–277. IEEE.
- [de Paula et al. 2019] de Paula, G., de Melo, L. F., dos Santos, M. H., Seixas, F., Santana, R., and Muchaluat-Saade, D. (2019). Efeitos sensoriais em jogos cognitivos para idosos: Jogo do stroop. In *Anais do XIX SBCAS*, pages 187–198, Porto Alegre, RS, Brasil. SBC.
- [Figuroa 2017] Figuroa, R. C. (2017). Personal spaces for multisensory stimulation as support to rehabilitate patients with cognitive disabilities. In *Proceedings of the XVIII International Conference on Human Computer Interaction*, pages 1–2.

- [Garzotto et al. 2019] Garzotto, F., Gelsomini, M., Gianotti, M., and Riccardi, F. (2019). Engaging children with neurodevelopmental disorder through multisensory interactive experiences in a smart space. In *Social Internet of Things*, pages 167–184. Springer.
- [Gerdner 2000] Gerdner, L. A. (2000). Effects of individualized versus classical “relaxation” music on the frequency of agitation in elderly persons with alzheimer’s disease and related disorders. *International Psychogeriatrics*, 12(1):49–65.
- [Ghinea et al. 2014] Ghinea, G., Timmerer, C., Lin, W., and Gulliver, S. R. (2014). Mulsemmedia: State of the art, perspectives, and challenges. *ACM TOMM*, 11(1s):17.
- [Guedes et al. 2016] Guedes, Á. L., Cunha, M., Fuks, H., Colcher, S., and Barbosa, S. D. (2016). Using ncl to synchronize media objects, sensors and actuators. In *Anais Extendidos do XXII Simpósio Brasileiro de Sistemas Multimídia e Web*, pages 184–189. SBC.
- [Guedes et al. 2017] Guedes, Á. L. V., de Albuquerque Azevedo, R. G., and Barbosa, S. D. J. (2017). Extending multimedia languages to support multimodal user interactions. *Multimedia Tools and Applications*, 76(4):5691–5720.
- [Hanford and Figueiro 2013] Hanford, N. and Figueiro, M. (2013). Light therapy and alzheimer’s disease and related dementia: past, present, and future. *Journal of Alzheimer’s Disease*, 33(4):913–922.
- [Hardman 1998] Hardman, H. L. (1998). *Modeling and Authoring Hypermedia Documents*. PhD thesis, Universität Amsterdam.
- [ITU-Recommendation 2009] ITU-Recommendation (2009). Nested context language (ncl) and ginga-ncl for iptv services.
- [Josué et al. 2018] Josué, M., Abreu, R., Barreto, F., Mattos, D., Amorim, G., dos Santos, J., and Muchaluat-Saade, D. (2018). Modeling sensory effects as first-class entities in multimedia applications. In *Proceedings of the 9th ACM MMSys*, pages 225–236.
- [Josué et al. 2019] Josué, M., Moreno, M., and Muchaluat-Saade, D. (2019). Mulsemmedia preparation: a new event type for preparing media object presentation and sensory effect rendering. In *Proceedings of the 10th ACM Multimedia Systems Conference*, pages 110–120.
- [Kast et al. 2007] Kast, M., Meyer, M., Vögeli, C., Gross, M., and Jäncke, L. (2007). Computer-based multisensory learning in children with developmental dyslexia. *Restorative Neurology and Neuroscience*, 25(3-4):355–369.
- [Khan et al. 2018] Khan, S. S., Ye, B., Taati, B., and Mihailidis, A. (2018). Detecting agitation and aggression in people with dementia using sensors—a systematic review. *Alzheimer’s & Dementia*, 14(6):824–832.
- [Kim and Han 2014] Kim, S. and Han, J. (2014). Text of white paper on mpeg-v. In *San Jose, USA. MPEG Group Meeting, ISO/IEC JTC*, volume 1.

- [Kim 2013] Kim, S.-K. (2013). Authoring multisensorial content. *Signal Processing: Image Communication*.
- [Lang et al. 2012] Lang, R., O'Reilly, M., Healy, O., Rispoli, M., Lydon, H., Streusand, W., Davis, T., Kang, S., Sigafos, J., Lancioni, G., et al. (2012). Sensory integration therapy for autism spectrum disorders: A systematic review. *Research in Autism Spectrum Disorders*, 6(3):1004–1018.
- [Luque et al. 2014] Luque, F. P., Galloso, I., Feijoo, C., Martín, C. A., and Cisneros, G. (2014). Integration of multisensorial stimuli and multimodal interaction in a hybrid 3dtv system. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 11(1s):16.
- [Martin et al. 2011] Martin, M., Clare, L., Altgassen, A. M., Cameron, M. H., and Zehnder, F. (2011). Cognition-based interventions for healthy older people and people with mild cognitive impairment. *Cochrane database of systematic reviews*, (1).
- [Mattos and Muchaluat-Saade 2018] Mattos, D. P. and Muchaluat-Saade, D. C. (2018). Multisem: A mulsemmedia model for supporting the development of authoring tools. In *Proceedings of the 24th Brazilian Symposium on Multimedia and the Web, WebMedia '18*, pages 109–116, New York, NY, USA. ACM.
- [Mishkind et al. 2017] Mishkind, M. C., Norr, A. M., Katz, A. C., and Reger, G. M. (2017). Review of virtual reality treatment in psychiatry: Evidence versus current diffusion and use. *Current Psychiatry Reports*, 19(11):80.
- [Munyan 2018] Munyan, B. G. (2018). *Odorants, Memory, and Presence in Warfighters: Do the Scents of War Matter?* PhD thesis, University of Central Florida.
- [Murray et al. 2014] Murray, N., Lee, B., Qiao, Y., and Muntean, G.-M. (2014). Multiple-scent enhanced multimedia synchronization. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 11(1s):12.
- [Nakamoto and Yoshikawa 2006] Nakamoto, T. and Yoshikawa, K. (2006). Movie with scents generated by olfactory display using solenoid valves. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 89(11):3327–3332.
- [O'Neil et al. 2011] O'Neil, M. E., Freeman, M., Christensen, V., Telerant, R., Adleman, A., Kansagara, D., et al. (2011). *A systematic evidence review of non-pharmacological interventions for behavioral symptoms of dementia*. Department of Veterans Affairs Washington, DC.
- [Piwek et al. 2016] Piwek, L., Ellis, D. A., Andrews, S., and Joinson, A. (2016). The rise of consumer health wearables: promises and barriers. *PLoS medicine*, 13(2).
- [Posar and Visconti 2018] Posar, A. and Visconti, P. (2018). Sensory abnormalities in children with autism spectrum disorder. *Jornal de Pediatria*, 94:342 – 350.

- [Radziszewski et al. 2017] Radziszewski, R., Ngankam, H. K., Grégoire, V., Lorrain, D., Pigot, H., and Giroux, S. (2017). Designing calm and non-intrusive ambient assisted living system for monitoring nighttime wanderings. *International Journal of Pervasive Computing and Communications*, 13(2):114–129.
- [Rizzo et al. 2010] Rizzo, A. S., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., and Mclay, R. (2010). Development and early evaluation of the virtual iraq/afghanistan exposure therapy system for combat-related ptsd. *Annals of the New York Academy of Sciences*, 1208(1):114–125.
- [Rodrigues et al. 2019] Rodrigues, R. O., Josué, M. I., Abreu, R. S., Amorim, G. F., Muchaluat-Saade, D. C., and Santos, J. A. d. (2019). A proposal for supporting sensory effect rendering in ginga-ncl. In *Proceedings of the 25th Brazillian Symposium on Multimedia and the Web*.
- [Rubbi et al. 2016] Rubbi, I., Magnani, D., Naldoni, G., Di Lorenzo, R., Cremonini, V., Capucci, P., Artioli, G., and Ferri, P. (2016). Efficacy of video-music therapy on quality of life improvement in a group of patients with alzheimer’s disease: a pre-post study. *ACTA BIO-MEDICA DE L’ATENEO PARMENSE*, 87(4-S):30–37.
- [Saleme et al. 2017] Saleme, E. B., Celestrini, J. R., and Santos, C. A. S. (2017). Time evaluation for the integration of a gestural interactive application with a distributed mulsemedia platform. In *Proceedings of the 8th ACM on Multimedia Systems Conference*, pages 308–314. ACM.
- [Saleme et al. 2019] Saleme, E. B., Santos, C. A., and Ghinea, G. (2019). A mulsemedia framework for delivering sensory effects to heterogeneous systems. *Multimedia Systems*, 25(4):421–447.
- [Santos et al. 2015] Santos, C. A. S., Neto, A. N. R., and Saleme, E. B. (2015). An event driven approach for integrating multi-sensory effects to interactive environments. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pages 981–986. IEEE.
- [Sarkar et al. 2018] Sarkar, S., Karthikeyan, B., Ajai, S. S., Kumar, G. D., and Sharath, C. M. (2018). Relaxation aid for intellectual disabilities. In *2018 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, pages 60–64.
- [Shin et al. 2016] Shin, S. H., Ha, K. S., Yun, H. O., and Nam, Y. S. (2016). Realistic media authoring tool based on MPEG-V international standard. *International Conference on Ubiquitous and Future Networks, ICUFN*, 2016-Augus:730–732.
- [Signes et al. 2000] Signes, J., Fisher, Y., and Eleftheriadis, A. (2000). Mpeg-4’s binary format for scene description. *Signal Processing: Image Communication*, 15(4):321–345.
- [Silva et al. 2019] Silva, E. P., Amorim, G. F., and Santos, J. A. F. d. (2019). Adding temporal semantic to textual media objects with eye tracking technology. In *Proceedings of the 25th Brazillian Symposium on Multimedia and the Web*, WebMedia ’19, page 217–220, New York, NY, USA. Association for Computing Machinery.

- [Simsek et al. 2016] Simsek, M., Aijaz, A., Dohler, M., Sachs, J., and Fettweis, G. (2016). 5g-enabled tactile internet. *IEEE Journal on Selected Areas in Communications*, 34(3):460–473.
- [Soares and Barbosa 2009] Soares, L. F. G. and Barbosa, S. D. (2009). Programando em ncl 3.0: Desenvolvimento de aplicações para o middleware ginga. *Campus, Rio de Janeiro, RJ*.
- [Soares et al. 2010] Soares, L. F. G., Rodrigues, R. F., Cerqueira, R., and Barbosa, S. D. J. (2010). Variable and state handling in ncl. *Multimedia Tools and Applications*, 50(3):465–489.
- [Soares et al. 2000] Soares, L. F. G., Rodrigues, R. F., and Saade, D. C. M. (2000). Modeling, authoring and formatting hypermedia documents in the hyperprop system. *Multimedia Systems*, 8(2):118–134.
- [Stroop 1935] Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6):643.
- [Sulema 2017] Sulema, Y. (2017). Asampl: Programming language for mulsemedia data processing based on algebraic system of aggregates. In *Interactive Mobile Communication, Technologies and Learning*, pages 431–442. Springer.
- [Taher et al. 2015] Taher, A. V., Ahmadi, M. K., and Zamir, F. P. (2015). Effects of multi-sensory stimulation on cognition function, depression, anxiety and quality of life in elderly persons with dementia. *International Journal of Sport Studies*, 5(3):355–360.
- [Tobii 2020] Tobii (2020). This is eye tracking. Acessado em: 21 de maio de 2020.
- [Tortell et al. 2007] Tortell, R., Luigi, D., Dozois, A., Bouchard, S., Morie, J. F., and Ilan, D. (2007). The effects of scent and game play experience on memory of a virtual environment. *Virtual Reality*, 11(1):61–68.
- [Van Doorn et al. 2003] Van Doorn, C., Gruber-Baldini, A. L., Zimmerman, S., Richard Hebel, J., Port, C. L., Baumgarten, M., Quinn, C. C., Taler, G., May, C., Magaziner, J., et al. (2003). Dementia as a risk factor for falls and fall injuries among nursing home residents. *Journal of the American Geriatrics Society*, 51(9):1213–1218.
- [W3C 2008] W3C (2008). Synchronized multimedia integration language - smil 3.0 specification. <http://www.w3c.org/TR/SMIL3>. World-Wide Web Consortium Recommendation.
- [Waltl et al. 2013] Waltl, M., Rainer, B., Timmerer, C., and Hellwagner, H. (2013). An end-to-end tool chain for sensory experience based on mpeg-v. *Signal Processing: Image Communication*.
- [Waltl et al. 2014] Waltl, M., Timmerer, C., Rainer, B., and Hellwagner, H. (2014). Sensory effects for ambient experiences in the world wide web. *Multimedia tools and applications*, 70(2):1141–1160.