

# Efeitos Sensoriais em Aplicações Multimídia

Eyre Brasil B. Montevecchi<sup>1</sup>, Débora C. M. Saade<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal Fluminense (UFF)  
Av. Gal. Milton Tavares de Souza, s/no São Domingos - 24.210–346 - Niterói - RJ - Brasil

eyre@montevecchi.com.br, debora@midia.com.uff.br

**Abstract.** *Multimedia applications are mostly composed by audiovisual content, that is, bisensorial. Whereas around 60% of human communication is non-verbal and we experience the world mainly through our five senses: vision, hearing, taste, touch and smell, MulSeMedia - Multiple Sensorial Media applications can provide a higher degree of immersion when stimulating other senses beyond sight and hearing, and can also be used for therapeutic purposes in patients in a range of learning difficulties to Autism, Dementia and Alzheimer's. This article investigates the state-of-the-art regarding the inclusion of multiple sensorial effects in multimedia content through tools, platforms and simulators that allow MulSeMedia execution and rendering in the web and preferentially using the MPEG-V standard.*

**Resumo.** *Aplicações multimídia tradicionais são em sua grande maioria compostas por conteúdos audiovisuais, ou seja, bissensoriais. Considerando que cerca de 60% da comunicação humana é não verbal e experimentamos o mundo utilizando principalmente cinco sentidos: visão, audição, paladar, tato e olfato, aplicações MulSeMedia - Multiple Sensorial Media podem proporcionar maior grau de imersão ao estimular outros sentidos além da visão e audição, podendo ainda serem utilizadas para fins terapêuticos em pacientes que possuem desde dificuldades de aprendizado a Autismo, demência e Alzheimer. Este artigo investiga o estado da arte referente à inclusão de múltiplos efeitos sensoriais em conteúdo multimídia, por meio de ferramentas, simuladores, plataformas e aparelhos que permitem execução e renderização MulSeMedia na web e utilizam preferencialmente o padrão MPEG-V.*

## 1. Introdução

Aplicações multimídia tradicionais são em sua grande maioria compostas por conteúdos audiovisuais, ou seja, bissensoriais. Considerando que cerca de 60% da comunicação humana é não verbal e experimentamos o mundo utilizando principalmente cinco sentidos: visão, audição, paladar, tato e olfato [Ghinea et al. 2014], essas aplicações podem proporcionar maior grau de imersão (i.e. melhorar a QoE - Qualidade de Experiência) ao estimular outros sentidos além da visão e audição, sendo então denominadas MulSeMedia - *Multiple Sensorial Media* [Ghinea et al. 2014]. Ainda, como apresentado em [Ghinea et al. 2014], aplicações MulSeMedia podem ser utilizadas para fins terapêuticos em pacientes que possuem desde dificuldades de aprendizado a Autismo, demência e Alzheimer, por meio de ambientes multissensoriais. Além do uso terapêutico, a adição de efeitos sensoriais tem sido bastante utilizada para entretenimento (e.g. simuladores, cinema 4D, jogos imersivos).

A Figura 1 de [Waltl et al. 2009] mostra de uma maneira geral o conceito de receber efeitos sensoriais simultaneamente ao conteúdo audiovisual. Uma mídia e os metadados SEM - *Sensory Effect Metadata* do padrão MPEG-V [Kim and Han 2014], que podem ser obtidos de um disco (e.g. DVD, Blu-Ray) ou da Internet são processados por um dispositivo intermediário, o *Media Processing Engine*. Esse é o responsável por interpretar e adaptar a mídia e o SEM à capacidade dos dispositivos que irão renderizar os efeitos sensoriais e as configurações de preferências do usuário.

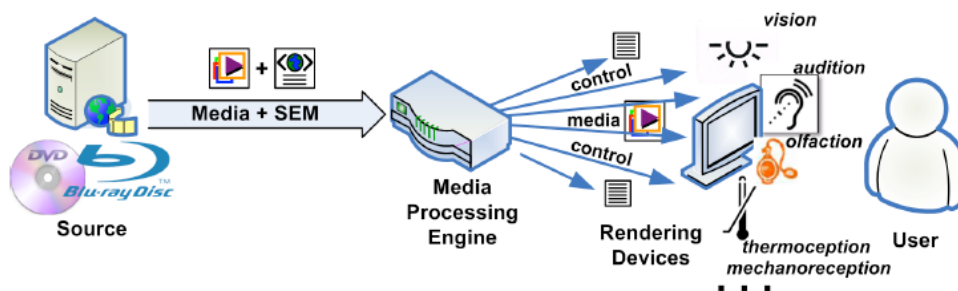


Figura 1. Conceito de efeitos sensoriais. [Waltl et al. 2009]

Este artigo investiga o estado da arte no que se refere a inclusão de múltiplos efeitos sensoriais em conteúdo multimídia, por meio de ferramentas, simuladores, plataformas e aparelhos que permitem execução e renderização MulSeMedia na web e utilizam preferencialmente o padrão MPEG-V. O restante do texto está organizado da seguinte maneira: a Seção 2 contém uma visão geral do padrão MPEG-V, a Seção 3 descreve o estado da arte em MulSeMedia para web, a Seção 4 contém a conclusão e trabalhos futuros.

## 2. MPEG-V

Como forma de estabelecer um padrão na comunicação entre o mundo real (e.g. sensores e atuadores) e o mundo virtual (e.g. aplicações, jogos etc), foi criado o padrão MPEG-V ou ISO/IEC 23005 [Ghinea et al. 2014, SALEME 2015]. Com foco em interoperabilidade, o padrão provê arquitetura e especifica a forma de representação da informação (i.e. metadados) para que essa comunicação ocorra, sendo dividido em 7 partes:

- **Parte 1 - Arquitetura:** trata da arquitetura geral, interfaces e questões de interoperabilidade.
- **Parte 2 - informações de controle:** fornece meios para descrever capacidades disponíveis dos dispositivos e informações de preferência do usuário. Aqui está especificada a linguagem de descrição de controle da informação (CIDL - *Control Information Description Language*) baseada em XML, assim como os vocabulários para descrição das capacidades dos dispositivos (DCDV - *Device Capability Description Vocabulary*) e sensores (SCDV - *Sensor Capability Description Vocabulary*), preferências de efeitos sensoriais do usuário (SEPV - *User's Sensory Effect Preference Vocabulary*) e para adaptação dos sensores (SAPV - *Sensor Adaptation Preference Vocabulary*).
- **Parte 3 - efeitos sensoriais:** provê meios para descrição de efeitos sensoriais, especificando uma linguagem para descrição de efeitos sensoriais (SEDL - *Sensory Effect Description Language*) baseada em XML. Os tipos de efeitos sensoriais

estão definidos em um vocabulário SEV - *Sensory Effect Vocabulary* e contendo essas descrições (SEDL e SEV), temos o elemento raiz SEM - *Sensory Effect Metadata*, que estará associado ao conteúdo multimídia.

- **Parte 4 - características de objetos do mundo virtual:** contém formatos para representação de dados que especificam objetos virtuais (semântica e sintaxe) para que exista compatibilidade com outros mundos virtuais ou seja possível controlá-los por meio de entradas do mundo real.
- **Parte 5 - formato de dados para dispositivos de interação:** fornece o formato de dados para dispositivos de interação (e.g. sensores e atuadores), ou seja, descrição de comandos para atuadores e de informações obtidas a partir de sensores. Aqui, a IIDL - *Interaction Interface Description Language*, baseada em XML, é a linguagem que permite definir interfaces para troca de mensagens entre os mundos real e virtual. Os vocabulários que auxiliam essa troca são o DCV - *Device Command Vocabulary* e SIV - *Sensed Information Vocabulary*.
- **Parte 6 - ferramentas e tipos:** contém especificação de sintaxe e semântica para ferramentas e tipos de dados comuns em outras partes do padrão MPEG-V, (e.g. tipos relacionados a cores, unidades etc), assim como esquemas de classificação para esses tipos.
- **Parte 7 - conformidade e software de referência:** possui a especificação de regras de conformidade e software de referência para validação e esclarecimento das especificações do próprio padrão MPEG-V.

### 3. Estado da Arte

Dentre os trabalhos que fazem uso da inclusão de efeitos sensoriais, há vários deles que fazem uso principalmente da parte 3 do padrão MPEG-V [SALEME 2015]. Essa seção descreve pesquisas na área MulSeMedia com foco em utilização de tecnologias voltadas para a web, embora nem todas façam uso do padrão MPEG-V.

#### 3.1. Ferramentas e plataformas de efeitos sensoriais

[SALEME 2015] busca resolver problemas relacionados à interoperabilidade, alto acoplamento entre camada de renderização de efeitos sensoriais e a camada de apresentação (com *Media Players*), o que impede a reutilização e mobilidade dos dispositivos físicos em camadas de apresentação distintas. Além disso, esses problemas também dificultam a utilização de múltiplos efeitos sensoriais em áreas como medicina, jogos e educação, uma vez que essas áreas geralmente usam mídias baseadas em eventos e *Media Players* usam em sua maioria uma única linha temporal. Desta forma, o autor propõe uma plataforma denominada PlaySEM, composta por: um módulo renderizador SE Render (desacoplado do *Player*) de efeitos sensoriais que usa o protocolo de comunicação UPnP, tornando possível que dispositivos de diferentes fornecedores se comuniquem com o renderizador; um *Video Player* SE Video Player que lê metadados de efeitos sensoriais no formato MPEG-V (com extensão XML ou SEM) e se comunica com o renderizador; um *Device Control* vinculado ao renderizador, é um programa embarcado em microcontrolador que recebe os comandos para acionar os dispositivos de interação. Além disso, foi criada uma biblioteca Java MPEG *Metadata* para transformar XML em objeto Java reutilizável. Como resultado, o *Video Player* e o renderizador podem ser executados em sistemas operacionais diversos. A Figura 2 mostra o conjunto de hardwares no ambiente real do usuário

funcionando em conjunto com um mini-pc onde o SE Renderer está em execução e um monitor apresentando um vídeo no SE Video Player.

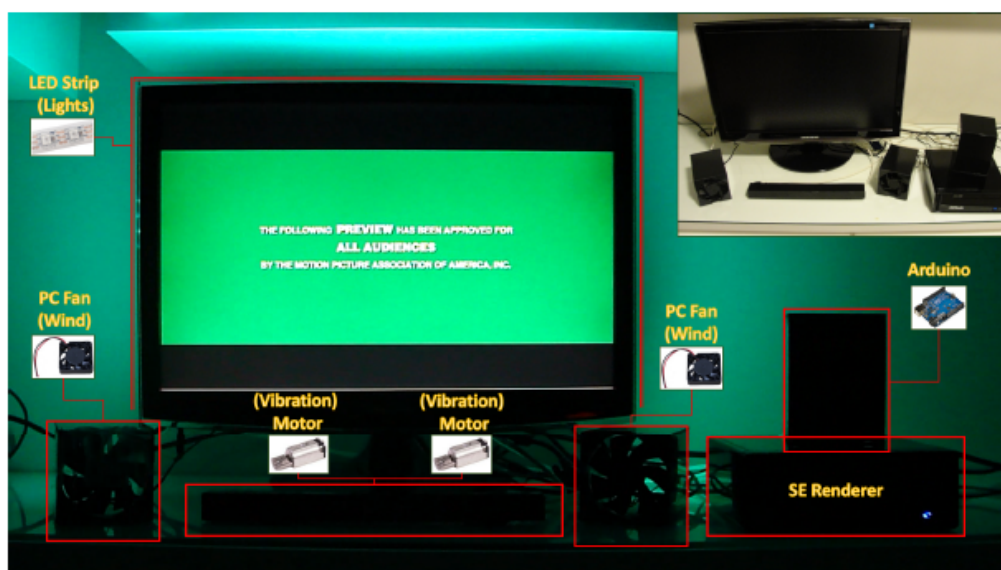


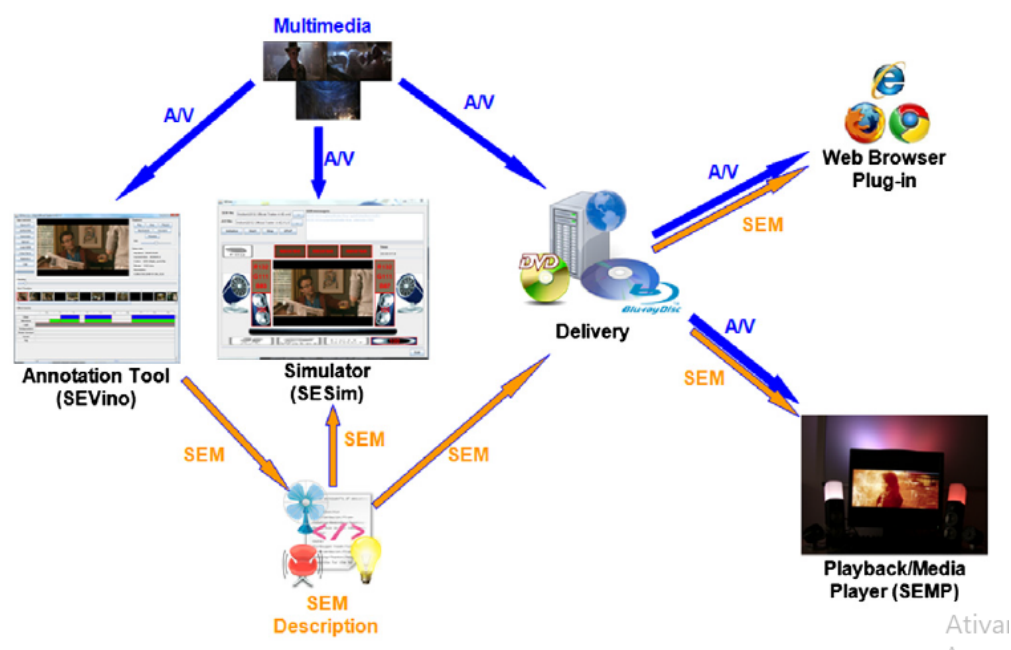
Figura 2. Hardwares no ambiente real do usuário [SALEME 2015]

Em [Waltl et al. 2013], com o objetivo de criar um *framework* para captura, medição, quantificação e avaliação da experiência do usuário de forma científica, os autores apresentam uma cadeia ponta-a-ponta de ferramentas que criam e usam metadados SEM por meio de vários canais de distribuição. A Figura 3 ilustra o *framework* desenvolvido, mostrando como as ferramentas se complementam em um contexto geral: conteúdo multimídia pode ter descrições de efeitos sensoriais adicionadas por meio do SEVino, gerando uma descrição SEM. Em seguida, o SEM e o conteúdo podem ser carregados para simulação no SESim ou podem ser entregues por meio da Internet, DVD ou Blu-Ray. Se o conteúdo está contido em um site, o plugin para Web Browser pode ser utilizado para execução da mídia e controle dos atuadores e sensores. Caso contrário, o *Player* SEMP será usado.

SEVino - *Sensory EffectVideoAnnotation* [Waltl et al. 2013] é uma ferramenta que permite realizar anotações em sequências de vídeo com efeitos sensoriais e gerar um SEM (em acordo com o MPEG-V), suportando os efeitos de vento, vibração, luz, temperatura, spray de água, aroma e névoa. Foi contruída em Java e incorpora, através do *framework* VLCj, o VLC *Player* para decodificação e renderização dos arquivos A/V.

SEMP - *Sensory EffectMediaPlayer* [Waltl et al. 2013] é um reproduzidor MulSe-Media que suporta os dispositivos amBX Gaming PC peripherals [PHILIPS 2015], Cyborg Gaming Lights for PC e Vortex Activ. A ferramenta possui um *parser* XML para leitura do SEM, usa a API DirectShow da Microsoft em conjunto com um módulo para cálculo e auto-extração de cor e usa um gerenciador de efeitos para controlar os dispositivos interativos.

A biblioteca AmbientLib [Waltl et al. 2013] executa o *parsing* dos arquivos SEM e controla os dispositivos de efeitos sensoriais (e.g. luzes, ventiladores, vibração). Integrada ao Web Browser plugin criado [Waltl et al. 2013], é possível renderizar Mul-

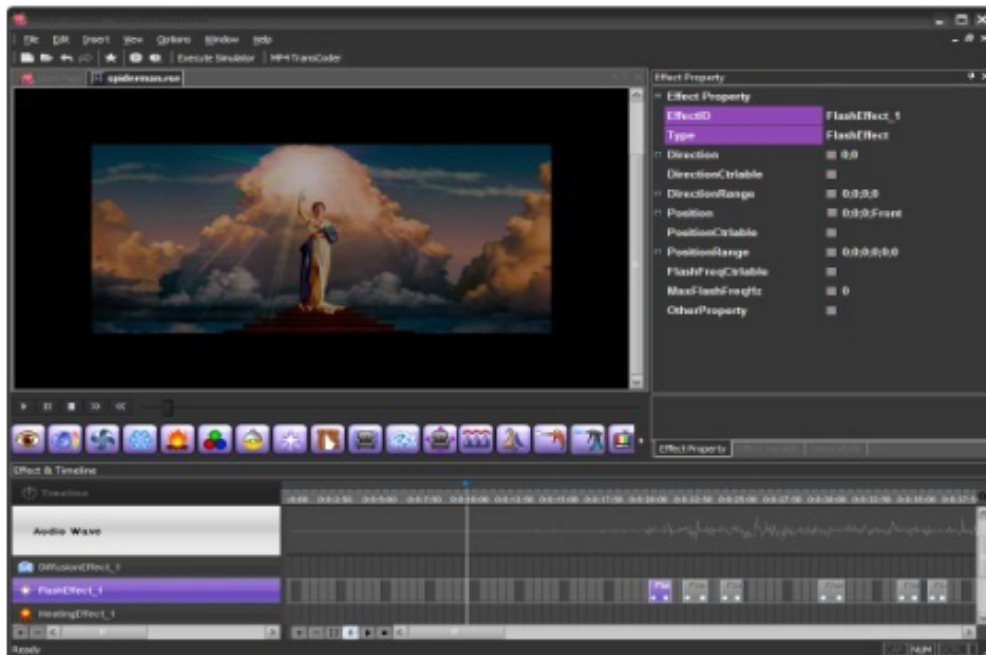


**Figura 3. Visão geral da cadeia de ferramentas ponta a ponta [Waltl et al. 2013]**

seMedia em vários Web browsers (e.g. Safari, Chrome, Firefox, Opera), sendo compatível com vários dispositivos (e.g. sistema amBX). A comunicação entre o plugin e as páginas Web é feita utilizando APIs dos browsers, por meio de chamadas a métodos JavaScript, definidos e exportados pelo plugin. Para que seja possível a compatibilidade com vários browsers, é utilizado o FireBreath, um *framework* para desenvolvimento de plugins. Uma vez que o HTML5 não permite manipulação de vídeos, é utilizado o elemento `< canvas >` em conjunto com o JavaScript.

Em [Choi et al. 2011], a ferramenta para autoria MulSeMedia RoSE Studio é apresentada como parte do *framework* para serviços de *streaming* com efeitos sensoriais, em uma proposta de trazer o cinema 4D para esse tipo de serviço, seguindo o padrão MPEG-V. A Figura 4 mostra a principal interface da RoSE Studio, cujas funcionalidades estão distribuídas em três janelas: no canto superior esquerdo está a janela de reprodução; canto superior direito para modificar ou inserir valores para informações gerais da mídia e propriedades de efeitos. Na janela inferior, há uma linha do tempo que mostra os efeitos adicionados aos fragmentos de efeitos (menor unidade possível para identificar um controle de efeito sensorial). Logo abaixo da janela de reprodução, há uma fileira de ícones, representando os efeitos que podem ser adicionados à mídia. RoSE Studio pode gerar dois tipos de arquivo de saída: arquivo XML independente contendo o SEM ou esses metadados podem ser multiplexados com o conteúdo do vídeo para então ser transmitido usando o padrão MPEG-2 TS (*Transport Stream*).

SMURF - *Sensible Media aUthoRing Factory* [Kim 2013] é uma GUI - *Graphics User Interface* de autoria MulSeMedia com funcionalidades mais eficientes (segundo o autor) em comparação com as ferramentas RoSE Studio [Choi et al. 2011] e SEVino [Waltl et al. 2013], pois estende as funcionalidades do SEVino, incluindo novo suporte a mais três elementos do MPEG-V: *GroupOfEffects*, *Declaration* e *ReferenceEffect*. Possui validação automática de instâncias XML além da possibilidade de ser renderizada em



**Figura 4. RoSE Studio [Choi et al. 2011]**

Web browsers, uma vez que a GUI foi implementada em Flex, que pode ser executado em Adobe Flash Player. Na Figura 5 é possível visualizar a GUI SMURF, que possui cinco partes: componente para controle do vídeo, composição dos efeitos e registro de efeitos, menu de ferramentas, linha de tempo de efeitos.

### 3.2. Simuladores

SESim - *Sensory Effect Simulator* é um simulador para MulSeMedia que faz parte do *framework* apresentado em [Waltl et al. 2013] e descrito na Seção 3.1. Recebe um arquivo com metadados SEM, sendo desenvolvido em Java e com o *Player* VLC incorporado à ferramenta através do VLCj, similarmente ao SEVino. Como aceita SEM em conformidade com o MPEG-V, o usuário possui a flexibilidade de carregar no SESim, SEMs advindos de outras ferramentas, desde que atendam ao MPEG-V também.

Em [Kim et al. 2013] é apresentado o *Sensible Media Simulation System*, um simulador de efeitos sensoriais para automóveis com interface ilustrada na Figura 6. Na construção da interface foi utilizado o Adobe Flex para permitir sua utilização na web e Java, assim como o SMURF [Kim 2013], citado pelos autores do artigo em questão. O *core* do sistema é um componente denominado SMCS - *Sensible Media Control System*, formado por uma ECU - Unidade de Controle Eletrônica (para controlar os dispositivos do carro), um motor de renderização (para controle do processamento dos efeitos), um módulo de geração e interpretação de metadados XML e uma unidade de controle integrado para controle geral do sistema. O conector BlazeDS foi usado para troca de informações entre a interface (Adobe Flex) e o Java. A Figura 7 mostra os atuadores de efeitos sensoriais e seus respectivos módulos de controle. A Figura 7(a) é um ventilador, Figura 7(b) é para efeito de calor, Figura 7(d) é um LED, Figura 7(e) é para efeito de vibração, Figuras 7(c) e (f) são os módulos controladores.



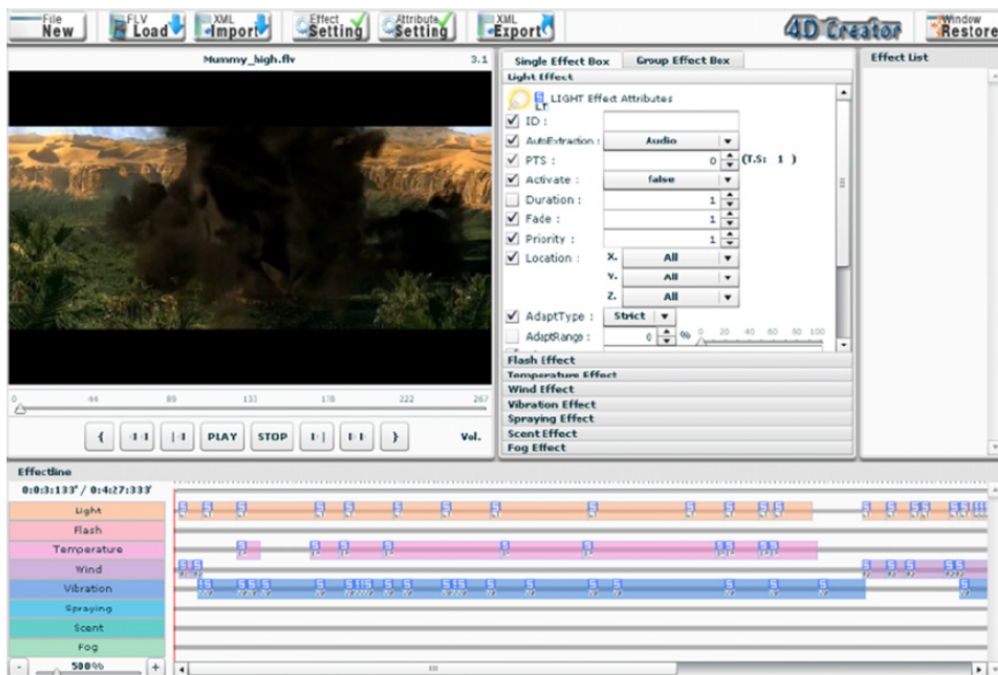
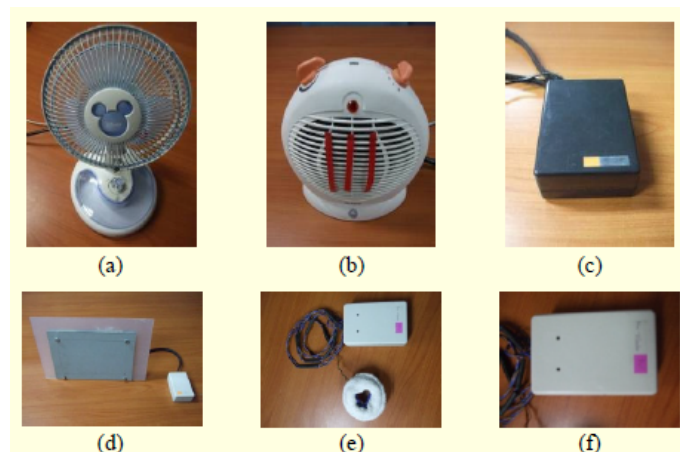


Figura 5. Interface gráfica do usuário SMURF [Kim 2013]



Figura 6. Interface gráfica do Sensible Media Simulator renderizando calor, vibração e luz [Kim et al. 2013]

Em [de Sousa et al. 2017] é apresentada a ferramenta MDD - *Model-Driven Development* denominada MulSeMaker, para desenvolvimento de aplicações MulSeMedia na Web baseada no padrão MPEG-V. O suporte declarativo do HTML5 foi estendido de forma a evitar o uso de JavaScript (diminuindo a complexidade de desenvolvimento) no suporte à renderização de efeitos sensoriais. O conjunto de tecnologias chamado *Web*



**Figura 7. Atuadores (a), (b), (d), (e) e seus respectivos módulos de controle (c), (f) [Kim et al. 2013]**

*Components*, que funcionam como *widgets* de interface de usuário reutilizáveis, foi usado para promover reuso e encapsulamento na criação de novas *tags* sensoriais. A interface é baseada em *Wizard*, onde os autores propõem como trabalho futuro, a modificação para *drag and drop*.

### 3.3. Atuadores e sensores

O amBX Gaming PC peripherals [PHILIPS 2015] ilustrado na Figura 8 é um produto da Philips, que foi utilizado nos trabalhos [Waltl et al. 2013, Kim et al. 2013]. Como avaliado em [SALEME 2015], permite aos autores de conteúdo MulSeMedia a adição de efeitos sensoriais de luz, vibração e vento a aplicações executadas a partir do computador. É compatível com o Sistema Operacional Windows (XP ou superior), sendo necessária a instalação de *drivers*.



**Figura 8. amBX Gaming PC peripherals [PHILIPS 2015]**

## 4. Conclusão

A maior parte das ferramentas de autoria MulSeMedia encontradas gera como resultado ou recebe um arquivo SEM com os metadados de efeitos sensoriais em conformidade com o MPEG-V, permitindo maior leque de opções para autoria e simulação desse tipo de aplicação. Tendo em vista que os usuários de autoria podem ter dificuldade com a descrição SEM, que requer conhecimento em XML, os trabalhos encontrados propõem ferramentas ou plataformas com interface gráfica, que possuem menor complexidade e esforço na criação dos metadados de efeitos sensoriais.



A maior parte das plataformas e ferramentas encontradas fazem uso de conjuntos de dispositivos montados pelos próprios autores. Alguns trabalhos oferecem compatibilidade com produtos que foram descontinuados (Cyborg Gaming Lights para luz e Vortex Activ para aroma). A solução comercial mais citada foi o sistema amBX, que possui um conjunto limitado de efeitos que podem ser executados: luz, vibração e vento.

Os trabalhos [Kim 2013, Kim et al. 2013], proporcionam a execução das ferramentas desenvolvidas em Web browsers por meio do Adobe Flex e Flash Player. Porém, a Adobe já comunicou a descontinuação desse plugin de mídia até 2020<sup>1</sup>, orientando os usuários a migrarem o conteúdo para HTML5, WebGL e formatos WebAssembly.

Não foi encontrada ou não foi detalhado nos trabalhos encontrados uma biblioteca específica para a renderização dos efeitos e controle dos atuadores, ou seja, a execução de efeitos de forma reutilizável e encapsulada, fazendo desse um problema em aberto que pode ser explorado em futuros trabalhos.

## Referências

- 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 Sousa, M. F., Ferraz, C. A. G., Kulesza, R., Ayres, I., and Lima, M. (2017). Mulsemaker: An mdd tool for mulsemimedia web application development. In *Proceedings of the 23rd Brazillian Symposium on Multimedia and the Web*, pages 317–324. ACM.
- Ghinea, G., Timmerer, C., Lin, W., and Gulliver, S. R. (2014). Mulsemimedia: State of the art, perspectives, and challenges. *ACM Trans. Multimedia Comput. Commun. Appl.*, 11(1s):17:1–17:23.
- 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, S.-K. (2013). Authoring multisensorial content. *Signal Processing: Image Communication*, 28(2):162 – 167. MPEG-V.
- Kim, S.-K., Joo, Y.-S., and Lee, Y. (2013). Sensible media simulation in an automobile application and human responses to sensory effects. *ETRI Journal*, 35(6):1001–1010.
- PHILIPS (2015). Philips ambx gaming pc peripherals sgc5103bd premium kit. Acessado em: 18 de novembro de 2018.
- SALEME, E. B. (2015). Playsem: uma plataforma para renderização de efeitos sensoriais compatível com o mpeg-v. Master's thesis, Universidade Federal do Espírito Santo.
- 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*, 28(2):136 – 150. MPEG-V.
- Waltl, M., Timmerer, C., and Hellwagner, H. (2009). A test-bed for quality of multimedia experience evaluation of sensory effects. In *Quality of Multimedia Experience, 2009. QoMEX 2009. International Workshop on*, pages 145–150. IEEE.

---

<sup>1</sup><https://www.techtudo.com.br/noticias/2017/07/adobe-vai-parar-de-distribuir-e-de-atualizar-o-flash-player-em-2020.ghtml>. Acessado em: 10 de dezembro de 2018