

Formatação de documentos multimídia com múltiplos efeitos sensoriais

Marina Ivanov Pereira Josué¹

¹Instituto de Computação – Universidade Federal Fluminense (UFF)

marinaivanov@midia.com.uff.br

***Resumo.** Este trabalho apresenta algumas abordagens relacionados a conteúdo multimídia com múltiplos efeitos sensoriais, com o objetivo de obter-se o estado da arte em formatação desse tipo de conteúdo. A partir da análise dos trabalhos será feito um levantamento de requisitos desse componente responsável pela execução do conteúdo.*

1. Introdução

O formatador multimídia é um componente do sistema multimídia responsável por gerenciar a sincronização entre os objetos de mídia que compõem a aplicação, controlar os players de exibição e os eventos de interação, a partir de uma especificação do documento multimídia. Além disso, deve ser capaz de lidar com aplicações que permitem adaptação do conteúdo e apresentá-las aos usuário final com um certo nível de QoS (*Quality of Service*) e QoE (*Quality of Experience*).

A manutenção dos relacionamentos de sincronismo espaço-temporal definidos pelo autor da aplicação é um ponto importante para reprodução de uma aplicação multimídia com qualidade. Desse modo, o formatador deve tentar garantir que um objeto de mídia esteja disponível no receptor no momento em que deve ser exibido, quando o conteúdo multimídia é entregue através de alguma rede de comunicação.

O comportamento temporal de uma aplicação multimídia pode ser modelado através de uma estrutura de dados [Costa et al. 2008, Chung and Pereira 2005, TIEN and Roisin 2002], que fornece informações dos instantes em que os objetos de mídia são apresentados. A partir dessa estrutura, o formatador é capaz de determinar os tempos de escalonamento e as informações de sincronização dos objetos necessários. Essa estrutura pode também considerar eventos imprevisíveis, tais como interação do usuário e adaptações ao ambiente, e pode implementar mecanismos de reserva de recursos e mecanismos de adaptação.

Além de coordenar o sincronismo da aplicação e os eventos de interação, o formatador também deve gerenciar o carregamento dos players de mídia. Por exemplo, quando um vídeo é iniciado pela aplicação multimídia, uma instância do player de vídeo deve estar preparada para reproduzi-lo.

Uma aplicação multimídia pode prover pontos de interação que podem ser ou não dependentes do conteúdo apresentado. Exemplos de interações independentes do conteúdo apresentado, são as interações de controle de reprodução do conteúdo (*play*, *pause*, *stop*, *resume*, *fast-forward* e *rewind*). Já as interações dependentes do conteúdo

são definidas pelo autor do documento multimídia, e podem ocorrer de diferentes formas, como seleção de botões ou de objetos de mídia.

Atualmente, as aplicações multimídia têm sido empregadas em diferentes áreas como entretenimento, educação, serviços governamentais, saúde e comércio. Com o objetivo de aprimorar tais serviços e enriquecer a QoE dos usuários do conteúdo, alguns trabalhos propõem a adição de efeitos sensoriais à multimídia tradicional [Cha et al. 2009, Waltl et al. 2013, Murray et al. 2014, Nakamoto and Yoshikawa 2006, Luque et al. 2014]. Essa técnica permite que o autor de aplicações especifique a sincronização de atuadores e sensores com objetos de mídia, estimulando outros sentidos além da visão e audição, para transmitir uma informação.

Além de modificar a maneira como o autor transmite uma informação ao usuário, a apresentação de documentos multimídia com múltiplos efeitos sensoriais introduz alterações no formatador multimídia tradicional. Isso porque, além de coordenar a sincronização entre objetos de mídia, o formatador deverá também ser capaz de manter as relações de sincronismo entre mídia e efeito sensorial e entre dois ou mais efeitos sensoriais. E o formatador terá que manipular o carregamento de atuadores, além dos players de mídia tradicionais.

Nesse contexto, o objetivo deste artigo é estudar os trabalhos relacionados à formatação de aplicações multimídia com múltiplos efeitos sensoriais. A partir da análise dos trabalhos relacionados à formatação multimídia tradicional e multimídia multi-sensorial, busca-se levantar requisitos para um formatador de conteúdo multimídia com múltiplos efeitos sensoriais.

Este trabalho está estruturado da seguinte forma. A Seção 2 descreve as características de formatadores multimídia tradicionais e alguns de seus requisitos. A Seção 3 apresenta os trabalhos relacionados à formatação de aplicações multimídia com múltiplos efeitos sensoriais. A Seção 4 especifica requisitos de um formatador *multisemídia*, levantados com base nos trabalhos relacionados e no estudo de formatadores multimídia. Uma estrutura para modelar o comportamento temporal de aplicações multisemídia é apresentada na Seção 5. Por fim, a Seção 6 apresenta as conclusões do trabalho e os trabalhos futuros.

2. Formatação multimídia

A apresentação de um conteúdo multimídia conforme especificado por seu autor é coordenada pelo formatador, que pode desempenhar diferentes funcionalidades de acordo com a plataforma de execução, a linguagem de especificação e as características da aplicação.

A proposta de formatadores multimídia é apresentada em diferentes trabalhos [Buchanan and Zellweger 1993, Van Rossum et al. 1993, Abdelli and Badache 2011, Soares et al. 2000]. Em geral, esses formatadores controlam a apresentação guiados por um plano de execução preenchido com informações sobre as ações que devem ser disparadas pela ocorrência de mudança de estado de um evento. O plano de execução pode ser especificado, utilizando diferentes modelos, como modelo dirigido a eventos ou timeline, por exemplo.

Firefly [Buchanan and Zellweger 1993] utiliza uma estrutura baseada em grafos temporais para controlar a execução do formatador, e concatena cadeias temporais par-

ciais para evitar problemas durante a apresentação dos objetos de mídia relacionados a eventos imprevisíveis. No CMIF [Van Rossum et al. 1993] o controle da execução se dá através de uma estrutura de grafo direcionado de dependência de tempo, similar a uma rede de Petri.

Durante a apresentação de um documento multimídia pode ocorrer incompatibilidade no tempo esperado para ocorrência de um evento que não pode ser atendido pelo player. Isso ocorre, por exemplo, quando o objeto de mídia transmitido pela rede, não é entregue a tempo do seu instante de execução. Desse modo, o formatador hipermídia pode implementar a pré-busca [Rodrigues and Soares 2002, Kim et al. 2001] de objetos de mídia, com base no plano de execução, para melhorar a qualidade da apresentação e diminuir a probabilidade de inconsistências temporais e espaciais. Essa técnica se aplica somente quando o conteúdo multimídia é obtido em modo pull, ou seja, os dados da aplicação são enviados a partir de uma solicitação do receptor.

A especificação temporal do documento pode ser feita de maneira flexível, com os instantes e durações de apresentação dos objetos podendo variar para preservar as relações entre mídias, ou de maneira fixa, com valores estáticos de duração e instante de apresentação. Quando a especificação é flexível, é possível fazer ajustes na apresentação do conteúdo “diminuindo” ou “esticando” o tempo de duração de alguns objetos de mídia. Esta funcionalidade, que altera a duração dos objetos de mídia é denominada computação de tempo elástico [Kim and Song 1995], e pode ser implementada pelo formatador.

[Candan et al. 1998] propõem técnicas para construção de escalonadores de recuperação e de apresentação de objetos de mídia considerando especificação temporal flexível. O trabalho divide o processo de recuperação em duas fases: (i) fase que define o plano de apresentação que satisfaz a especificação temporal flexível contida no documento, e (ii) fase que gera um plano de recuperação que especifica os instantes em que o cliente deve fazer uma requisição ao servidor pelos objetos que compõem o documento multimídia.

A pré-busca pode se basear em informações obtidas através da análise do canal de comunicação e da plataforma de transmissão, ou então por estimativas, como por heurísticas que envolvem o tamanho do dado a ser buscado. Essas estimativas de pré-busca podem ser armazenadas e utilizadas posteriormente para outras apresentações do conteúdo, a fim de otimizar a técnica. Entretanto, em aplicações multimídia em que a ocorrência de eventos não-determinísticos é muito grande, a pré-busca torna-se ineficiente, devido a dificuldade em se prever quando o evento irá ocorrer, e se irá ocorrer.

Ao especificar uma aplicação multimídia, o autor pode definir pontos de adaptação da aplicação através de objetos de mídia alternativos ou alternativas de exibição de um mesmo conteúdo. Os mecanismos de adaptação são interessantes tanto para adequar o conteúdo a um certo perfil de usuário, quanto para adequar a aplicação às condições do ambiente de execução.

Aplicações multimídia que especificam adaptações relacionadas às condições do ambiente exigem que o formatador implemente um componente responsável pelo gerenciamento do contexto. A gerência do contexto consiste na manutenção dos valores de parâmetros do usuário e do seu ambiente.

[Soares et al. 2000] definem que o formatador multimídia é responsável por con-

trolar a apresentação do documento baseado na sua especificação e na descrição da plataforma ou ambiente de reprodução. As informações que descrevem a plataforma de reprodução podem ser fornecidas por um administrador do sistema ou da rede, ou até mesmo pelo autor de um documento específico, e serem utilizadas para adaptação da aplicação.

Em [Abdelli and Badache 2011], o processo de adaptação da qualidade da aplicação ocorre em tempo de compilação, baseado em um escalonador que determina como cada objeto utilizado na apresentação multimídia deve ser obtido - se seus dados devem ser pré-buscados, entregues em tempo real ou substituídos. Para isso, o formatador considera os parâmetros de largura de banda e memória, que são monitorados durante a reprodução da aplicação.

Em geral, a adaptação de uma aplicação em tempo de execução pode demandar uma quantidade grande de recursos, nem sempre disponível no dispositivo de reprodução. [Abdelli and Badache 2011] propõem uma solução para economizar os recursos do receptor, inserindo uma instância do formatador em um servidor de proxy que obtém informações do perfil do usuário e contexto do sistema de reprodução para realizar alterações no documento multimídia especificado em SMIL (*Synchronized Multimedia Integration Language*) [Consortium et al. 2008].

Além das funcionalidades gerais que devem estar presentes em formatadores multimídia, existem outras funcionalidades que são características da linguagem de especificação do documento interpretada pelo formatador e da aplicação alvo. O formatador multimídia pode, por exemplo, implementar técnicas para verificar a consistência do documento multimídia em tempo de execução, mecanismos de segurança para interface com sistemas de pagamento para aplicações de comércio eletrônico, e interface com os players de mídia para obter informações sobre a reprodução da mídia.

Um exemplo de funcionalidade dependente da linguagem de especificação da aplicação multimídia, é a comunicação com outros dispositivos, para linguagens que permitem definir a apresentação da aplicação em múltiplos dispositivos. Nesse cenário, o formatador deverá ser capaz de gerenciar as conexões com os outros dispositivos para garantir a reprodução correta da aplicação.

3. Trabalhos relacionados

A incorporação de efeitos sensoriais em aplicações multimídia tradicionais vem sendo estudada em diferentes trabalhos [Waltl et al. 2013, Cha et al. 2009, Murray et al. 2014, Luque et al. 2014], com o objetivo de prover uma maior imersão do usuário na informação que está sendo apresentada. A sensação de imersão no conteúdo está diretamente relacionada com a capacidade do usuário de interagir com o ambiente de reprodução, e ser estimulado por ele.

Durante a autoria do conteúdo multimídia, o autor pode especificar diferentes efeitos sensoriais como efeitos de luz, vibração e dispersão de aromas para estimular outros sentidos dos seres humanos, além da visão e audição, já manipulados pelas aplicações tradicionais. Na maioria dos trabalhos atuais, a descrição das características dos efeitos sensoriais e da informação de sincronização dos efeitos é realizada com base no padrão MPEG-V [ISO/IEC 23005-3 2011]. O padrão fornece mecanismos como as descrições

de metadados de efeitos sensoriais (SEM - *Sensory Effect Metadata*) e uma linguagem de descrição de efeito (SEDL - *Sensory Effect Description Language*) que especifica certos atributos como duração e intensidade de um efeito.

[Waltl et al. 2013] apresenta um simulador que é capaz de apresentar conteúdo multimídia e simular efeitos sensoriais sincronizados com a aplicação. A arquitetura do simulador é dividida em três camadas, sendo uma delas a camada de entrada de dados, que 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 no metadado 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. O simulador dá suporte a efeitos de luz, vento, umidade, vibração, temperatura, jatos de água e aromas.

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. O framework 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 multimídia háptica apresentada em [Cha et al. 2009] é especificada utilizando MPEG-4 BIFS [Signes et al. 2000], que permite sintetizar e sincronizar os objetos de mídia audiovisual com os efeitos relacionados ao sentido do tato. Na fase de reprodução do conteúdo, o formatador 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. O formatador então processa a informação de interação do usuário e realiza alguma ação na apresentação do conteúdo.

A sincronização entre fluxo audiovisual e fluxos de aromas é estudada em [Murray et al. 2014], com o objetivo de analisar o impacto da utilização desse tipo de efeito e da perda de sincronização entre os fluxos na QoE do usuário do conteúdo. O trabalho não especifica o ambiente de formatação do conteúdo, mas aponta alguns desafios para implementação do mesmo. No trabalho, a disponibilização dos aromas no eixo temporal segue uma recomendação definida em [Nakamoto and Yoshikawa 2006], que determina que 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 é o tempo que o usuário demora para detectar a presença de odores, após serem emitidos pelos dispositivos. Além disso, o trabalho cita que este tempo de percepção está diretamente relacionado

com a distância do usuário ao dispositivo dispersor do aroma.

Quando o conteúdo multemídia é entregue através de uma rede de comunicação, o controle do jitter é muito importante, para evitar problemas como a perda de sincronização entre efeitos de aroma e o fluxo de mídia, e a sobreposição entre diferentes fluxos de aroma. Este controle pode ser realizado no servidor, quando o conteúdo é enviado em broadcast, ou pelo formatador, nos casos onde o conteúdo é entregue em modo pull.

A utilização de estímulos multisensoriais para complementar o conteúdo audiovisual principal de um sistema de televisão é apresentada em [Luque et al. 2014]. O trabalho 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 multisensorial. [Luque et al. 2014] estuda a viabilidade da inserção de efeitos sensoriais e interações multimodais no sistemas de TV atuais. Para o estudo, o trabalho implementa a transmissão de uma partida de futebol em que diferentes dispositivos atuadores são disparados de maneira síncrona para entregar uma combinação de estímulos sensoriais.

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 multisensorial 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 browser Web.

O plugin proposto em [Waltl et al. 2014] sincroniza os objetos de mídia com os efeitos sensoriais através da análise do tempo de reprodução do conteúdo audiovisual. Se a marcação de tempo atual for igual ao instante de apresentação 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.

É importante destacar que os trabalhos atuais relacionados à aplicações multimídia com múltiplos efeitos sensoriais, estão mais focados em analisar a percepção humana da qualidade de experiência no consumo deste tipo de conteúdo, e não em especificar o formatador responsável por coordenar a reprodução do conteúdo.

Neste contexto, este trabalho tem o objetivo de levantar alguns requisitos para a implementação de um formatador multemídia que possibilite a apresentação do conteúdo respeitando as especificações de sincronismo espaço-temporal definidas pelo autor da aplicação, e que ofereça algumas funcionalidades adicionais, como adaptação de conteúdo e pré-busca de dados. Esses requisitos são listados na Seção 5.

4. Controle do sincronismo temporal pelo formatador

A reprodução de uma aplicação multimídia está estritamente relacionada com o controle do sincronismo temporal entre os objetos de mídia que a compõem. O processo de con-

trole pode ser guiado através de uma estrutura, denominada plano de apresentação, que contém os instantes de início e fim de apresentação de cada objeto. O plano também deve fornecer um mecanismo que preserve o histórico de transições ocorridas, possibilitando que quando a apresentação for reiniciada, as opções escolhidas e adaptações realizadas sejam recuperadas.

Em geral, o plano de apresentação é derivado de uma estrutura que modela o comportamento temporal da aplicação. Essa estrutura permite também a construção de outros planos de controle, como o plano de carregamento de exibidores, o plano de transmissão e o plano de pré-busca. Em [Costa et al. 2008] é proposta uma estrutura definida a partir de um grafo temporal hipermídia, denominada HTG (*Hypermedia Temporal Graph*), que modela o comportamento de aplicações onde a sincronização dos objetos é baseada em eventos.

Os eventos que podem ocorrer sobre os conteúdos das mídias são representados no HTG por meio de vértices, enquanto os relacionamentos são representados por arestas dirigidas. O grafo obtido a partir da análise da aplicação deve preservar todos os relacionamentos entre os eventos (determinísticos e não-determinísticos) definidos pelo autor da aplicação, e o estado atual de cada máquina de estados associada a um evento. Ao controlar os estados de um evento, é possível determinar, com precisão, qual o comportamento esperado de um evento, considerando o seu estado atual e a transição realizada.

No HTG, os vértices contêm triplas, compostas pela transição da máquina de estado do evento, que pode ser identificada pela ação que dispara a transição (*start, stop, abort, natural end, pause ou resume*), pelo tipo de evento e pelo identificador da interface que define o evento. Uma interface pode representar o conteúdo de uma mídia, uma âncora do conteúdo ou até mesmo uma propriedade do objeto. Quando identifica uma propriedade, seu valor também é acoplado ao identificador da interface.

A utilização do HTG está limitada apenas ao tipo de sincronização que é definida pela linguagem de especificação do documento. Como a estrutura utiliza o conceito de eventos para interligar os objetos de mídia de uma aplicação, ela pode ser empregada por qualquer linguagem cujo a especificação da sincronização é baseada em eventos.

Desse modo, o HTG poderia ser utilizado para modelar o comportamento temporal de aplicações multemídia especificadas em linguagens que dão suporte a esse tipo de sincronização. Porém, algumas extensões à estrutura são necessárias, como a inserção de um evento de tipo “pré-busca” nos eventos suportados, para dar suporte à pré-busca de atuadores com o propósito de garantir que um efeito seja entregue ao usuário no momento correto.

Como o HTG modela a aplicação a partir de triplas do tipo (transição de evento, evento, interface), para que o grafo hipermídia dê suporte à modelagem de ocorrência de eventos sobre efeitos sensoriais, eles poderiam ser especificados como um elemento de interface.

Por exemplo, em uma aplicação multimídia que especifica o início de um efeito de luz após 10s do início de apresentação de uma mídia vídeo, o HTG poderia modelar essa característica através da criação das duas triplas, definidas na Figura 1. A aplicação pode definir também a apresentação de um efeito de luz de cor específica, a partir da apresentação de uma mídia imagem. Neste caso, o HTG poderia modelar essa caracte-

terística como uma propriedade do efeito de luz, como ilustrado na Figura 2

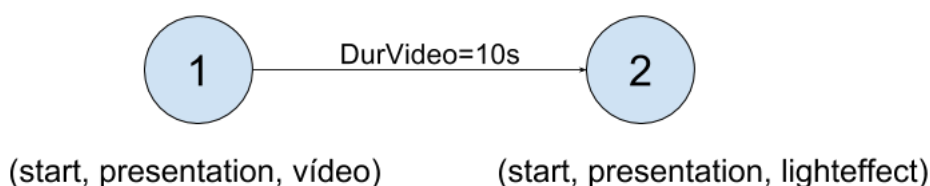


Figura 1. Exemplo de grafo para modelar sincronização entre vídeo e efeito de luz

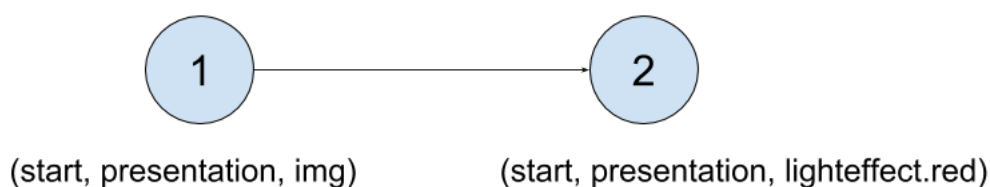


Figura 2. Exemplo de grafo para modelar sincronização entre imagem e efeito de luz vermelha

Desse modo, quando a aplicação é especificada com sincronização baseada em eventos, o formatador mulsemídia pode empregar o grafo hipermídia para modelar o comportamento temporal de aplicações multimídia com múltiplos efeitos sensoriais, e utilizá-lo como base para a construção do plano de apresentação do conteúdo, e de outros planos de controle, necessários para a reprodução correta da aplicação no dispositivo receptor.

Além da sincronização baseada em eventos, alguns modelos hipermídia possibilitam a especificação de sincronização baseada em restrições. Por exemplo, o autor da aplicação pode definir uma restrição entre dois objetos de mídias que especifica que ambos devem ser apresentados ao mesmo tempo. O estudo da modelagem do comportamento temporal de aplicações mulsemídia que dão suporte à sincronização baseada em restrições, é deixado como trabalho futuro.

5. Requisitos de um formatador mulsemídia

Um formatador mulsemídia é um componente do sistema multimídia responsável por coordenar a apresentação sincronizada de objetos de mídia e efeitos sensoriais. A reprodução de aplicações multimídia com múltiplos efeitos sensoriais demanda algumas extensões ao formatador multimídia tradicional, uma vez que este deve ser capaz de manipular efeitos e interações por parte de sensores, além da manipulação de conteúdo audiovisual.

Desse modo, a especificação de um formatador mulsemídia deve atender aos requisitos definidos para um formatador multimídia relacionados à manipulação de objetos de mídia tradicionais (áudio, vídeo, texto e imagem) e a novos requisitos relacionados à manipulação de efeitos sensoriais.

Neste contexto, esta seção apresenta alguns requisitos do formatador mulsemídia para a entrega de aplicações multimídia sincronizadas com efeitos sensoriais. Estes re-

quisitos foram levantados com base nos trabalhos relacionados ao desenvolvimento e reprodução de aplicações multimídia multisensoriais, e são listados a seguir.

- **R_1 : Suporte à sincronização de efeitos sensoriais, e sincronização entre efeitos e objetos de mídia**
O formatador deve ser capaz de apresentar os efeitos sensoriais sincronizados com os objetos de mídia ou sincronizados entre si [Cha et al. 2009, Murray et al. 2014]. Para isso, ele deve utilizar uma estrutura que orquestra a apresentação da aplicação, e possibilite posicioná-la em um instante qualquer no tempo. A sincronização entre objetos de mídia e efeitos sensoriais permite que uma mídia gere um estímulo no usuário, e que a informação obtida por um sensor dispare algum evento em um objeto de mídia.
- **R_2 : Suporte ao acionamento de atuadores**
Da mesma maneira que é necessário o carregamento de players para a apresentação de objetos de mídia audiovisual, o formatador deve ser capaz de acionar os dispositivos responsáveis pela reprodução dos efeitos sensoriais. [Waltl et al. 2013, Cha et al. 2009, Murray et al. 2014, Nakamoto and Yoshikawa 2006, Luque et al. 2014]
- **R_3 : Suporte à utilização de sensores**
Em aplicações multimídia com múltiplos efeitos sensoriais, a entrada de dados pode se dar através de dispositivos tradicionais, como seleção de um botão do controle remoto, ou através de sensores. Assim, o formatador deve ser capaz de utilizar os sensores para capturar interações do usuário ou até mesmo informações do ambiente. [Cha et al. 2009]
- **R_4 : Suporte a APIs para sensores**
O formatador deve fornecer uma interface de comunicação bem definida com os dispositivos atuadores e sensores. Isso possibilita que um formatador utilize diferentes dispositivos para apresentação de efeitos sensoriais, independente de seu fabricante. A API deve possibilitar que o formatador obtenha as informações dos sensores, e acione a execução dos efeitos. Essa característica não é observada nos trabalhos atuais [Waltl et al. 2013, Cha et al. 2009, Murray et al. 2014, Nakamoto and Yoshikawa 2006, Luque et al. 2014, Waltl et al. 2014], que em geral, propõem uma solução relacionada a dispositivos específicos.
- **R_5 : Suporte à pré-busca de atuadores**
Um efeito pode demorar um tempo para ser sentido pelo usuário, após ser disparado por um dispositivo. Para garantir que a sensação do efeito ocorra no instante definido pela sincronização com o objeto de mídia, o formatador deve ser capaz de realizar a pré-busca de atuadores, considerando o tempo em que o efeito demora para ser sentido pelo usuário [Rodrigues and Soares 2002, Kim et al. 2001]. Um exemplo de característica que deve ser considerada pelo formatador para pré-busca de atuadores, é a distância entre o dispositivo dispersor de aroma e o usuário. Um usuário localizado a uma distância maior do dispositivo, irá demorar mais tempo para receber o efeito, do que um usuário mais próximo.
- **R_6 : Suporte à geração automática de efeitos**
Geralmente, os efeitos sensoriais estão relacionados a um fluxo audiovisual principal. O formatador pode dar suporte à geração automática de efeitos através da análise do conteúdo audiovisual apresentado, como a extração de informações de cores para geração de efeitos de luz. [Waltl et al. 2013]

- ***R*₇: Suporte à adaptação de efeitos sensoriais**

Durante a fase de autoria da aplicação o autor pode especificar um conjunto de estímulos sincronizados com os objetos de mídia. Entretanto, certos usuários podem não estar interessados em receber um certo efeito, como por exemplo, efeito de umidade. Além disso, pode ocorrer de um atuador especificado na aplicação não estar disponível no ambiente do usuário. Assim, o formatador deve ser capaz de adaptar a aplicação em tempo de apresentação, tanto em relação a preferências do usuário, quanto ao contexto em que ele está inserido. [Abdelli and Badache 2011, Soares et al. 2000]

6. Conclusão e Trabalhos futuros

Aplicações multimídia com múltiplos efeitos sensoriais, é um novo conceito de aplicações, denominado *Mulsemídia*, que tem como objetivo entregar informação ao receptor utilizando diferentes tipos de efeito que podem estimular todos os cinco sentidos do ser humano.

Na literatura em geral, os estudos relacionados às aplicações com múltiplos efeitos sensoriais tem o objetivo de analisar o impacto da utilização de efeitos na apresentação do conteúdo. Poucos trabalhos propõem um formatador, e quando o fazem, definem formatadores muito específicos para um certo tipo de aplicação, com um conjunto específico de efeitos.

Outra questão que ainda é pouco explorada na literatura, é a utilização de sensores como dispositivos de entrada para aplicação. Ao permitir a utilização de sensores como meio de interação do usuário com a aplicação, é possível ampliar a gama de serviços oferecidos por aplicações multisensoriais, empregando-as em atividades de simulação ou até mesmo como auxílio no tratamentos de certas doenças.

Este trabalho levantou alguns requisitos para formatadores que dão suporte à apresentação de aplicações multimídia com múltiplos efeitos sensoriais. Como o formatador deve manipular tanto objetos de mídia tradicionais quanto efeitos sensoriais, o conjunto de requisitos do formatador *mulsemídia* deve englobar tanto os requisitos ligados à formatação multimídia quanto os requisitos levantados por este trabalho.

A partir dos requisitos levantados, como trabalho futuro, será implementado um formatador que permite a reprodução de aplicações multimídia com múltiplos efeitos sensoriais. Estes requisitos serão validados e novos requisitos poderão ser levantados a partir da prototipação e análise do formatador implementado. Outro trabalho futuro, é a especificação de uma estrutura para modelar o comportamento temporal da aplicação, que dê suporte tanto a sincronização baseada em eventos, quanto sincronização baseada em restrições.

Referências

- Abdelli, A. and Badache, N. (2011). Context-aware adaptation of multimedia documents for consistent presentations. *Multimedia systems*, 17(6):465–486.
- Buchanan, M. C. and Zellweger, P. T. (1993). Automatic temporal layout mechanisms. In *Proceedings of the first ACM international conference on Multimedia*, pages 341–350. ACM.

- Candan, K. S., Prabhakaran, B., and Subrahmanian, V. (1998). Retrieval schedules based on resource availability and flexible presentation specifications. *Multimedia Systems*, 6(4):232–250.
- Cha, J., Ho, Y.-S., Kim, Y., Ryu, J., and Oakley, I. (2009). A framework for haptic broadcasting. *IEEE MultiMedia*, 16(3):16–27.
- Chung, S. M. and Pereira, A. L. (2005). Timed petri net representation of smil. *IEEE MultiMedia*, 12(1):64–72.
- Consortium, W. W.-W. W. et al. (2008). Synchronized multimedia integration language–smil 3.0 specification, w3c recommendation. DOI: <http://www.w3.org/TR/SMIL3>.
- Costa, R. M. d. R., Moreno, M. F., and Soares, L. F. G. (2008). Intermedia synchronization management in dtv systems. In *Proceedings of the eighth ACM symposium on Document engineering*, pages 289–297. ACM.
- ISO/IEC 23005-3 (2011). Information technology – Media context and control – Part 3: Sensory information.
- Kim, M. Y. and Song, J. (1995). Multimedia documents with elastic time. In *Proceedings of the third ACM international conference on Multimedia*, pages 143–154. ACM.
- Kim, S.-W., Kim, S.-J., Jeong, T. I., and Yoo, S. (2001). The optimal retrieval start times of media objects for the multimedia presentation. *Information and Software Technology*, 43(4):219–229.
- 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.
- 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, 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.
- Rodrigues, R. F. and Soares, L. F. G. (2002). A framework for prefetching mechanisms in hypermedia presentations. In *Multimedia Software Engineering, 2002. Proceedings. Fourth International Symposium on*, pages 278–285. IEEE.
- Signes, J., Fisher, Y., and Eleftheriadis, A. (2000). Mpeg-4’s binary format for scene description. *Signal Processing: Image Communication*, 15(4):321–345.
- Soares, L. F. G., Rodrigues, R. F., and Muchaluat Saade, D. C. (2000). Modeling, authoring and formatting hypermedia documents in the hyperprop system. *Multimedia Systems*, 8(2):118–134.
- TIEN, T. and Roisin, C. (2002). A multimedia model based on structured media and sub-elements for complex multimedia authoring and presentation. *International Journal of Software Engineering and Knowledge Engineering*, 12(05):473–500.

- Van Rossum, G., Jansen, J., Mullender, K. S., and Bulterman, D. C. (1993). Cmifed: a presentation environment for portable hypermedia documents. In *Proceedings of the first ACM international conference on Multimedia*, pages 183–188. ACM.
- 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.
- 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.