

Criando Aplicações Multissensoriais Interativas para a Futura TV 3.0

Pesquisadores da Universidade Federal Fluminense elaboram proposta que vira tese de Doutorado onde se propõe a criação de um editor para produção de aplicações multimídia multissensoriais interativas. Este editor gera código NCL 4.0, seguindo a proposta para TV 3.0, compatível com o middleware Ginga padrão para TV Digital Brasileira. NCL 4.0 segue para a terceira fase do Projeto TV 3.0 do Fórum SBTVD

**Por Douglas Paulo de Mattos e
Débora C. Muchaluat Saade**

Introdução

As aplicações multimídia estão disponíveis em diferentes dispositivos, como smartphones, computadores, tablets e TVs digitais. Não só um grande volume destas aplicações é consumido, mas também autores de diferentes perfis (profissionais e não profissionais) participam da produção de aplicações multimídia. O conteúdo multimídia tradicional possui somente mídia audiovisual, que envolve apenas dois sentidos humanos, a visão e a audição. Porém, a maior parte da comunicação humana não é verbal e o mundo ao nosso redor pode ser compreendido utilizando os cinco sentidos: visão, audição, tato, paladar e olfato.

Todos esses estímulos sensoriais têm sido explorados pela comunidade científica de multimídia através do conceito denominado MulSeMedia (*Multiple Sensorial Media*) [Ghinea et al., 2014]. Este conceito oferece aos usuários novas sensações, explorando outros sentidos além da visão e da audição em aplicações multimídia interativas. Além disso, o uso de vários efeitos sensoriais também fornece aos usuários um novo conteúdo imersivo, aumentando a qualidade de experiência (QoE) das aplicações multimídia.

Nesse contexto, as aplicações multimídia podem ser integradas a múltiplos efeitos sensoriais como luz, vento e cheiro, por exemplo. A título de exemplo de áreas em que pode-se introduzir esses tipos de aplicações, pode-se citar os cinemas 4D, onde as cadeiras se movimentam de forma sincronizada com o con-



teúdo audiovisual dos filmes. De fato, tecnologias multimídia estão se deslocando de nichos altamente especializados para setores mais populares com dispositivos se tornando cada vez mais acessíveis [Saleme et al., 2019]. Outro domínio que evidencia essa transição é o dos jogos digitais, em que as aplicações multimídia interativas podem fornecer um ambiente imersivo aumentando a realidade do jogo e a QoE dos jogadores. Na indústria de entretenimento, os efeitos sensoriais também são aplicados a simuladores (por exemplo, voo e direção) para tornar a experiência de simulação mais realista.

Em sistemas de TV digital, efeitos sensoriais também podem ser oferecidos. Na chamada de propostas para o futuro padrão de TV digital no Brasil, denominada TV 3.0 pelo Fórum do Sistema Brasileiro de TV Digital Terrestre (SBTVD), um dos requisitos (AP14) da camada de codificação de aplicações do novo sistema é o suporte à TV imersiva, incluindo efeitos sensoriais (AP14.1). **Vislumbrando oferecer essa funcionalidade ao novo sistema, a Universidade Federal Fluminense (UFF) submeteu ao Fórum SBTVD uma proposta intitulada NCL 4.0, que estende o padrão atual DTVPlay [ABNT, 2021] com este objetivo. NCL 4.0 foi avaliado e tem potencial para ser integrada ao futuro padrão a ser especificado pelo Fórum SBTVD já que passou a terceira Fase do Projeto TV 3.0.**

Editor STEVE 2.0

Considerando a produção de aplicações multimídia multissensoriais interativas, um ambiente gráfico para autoria desse tipo de aplicação é bastante desejável, pois facilita a criação e incentiva a adoção dessa tecnologia em larga escala. O Laboratório de Pesquisa MídiaCom¹ da UFF desenvolveu um ambiente gráfico denominado STEVE 2.0 [de Mattos et al., 2020] para aprimorar a fase de produção de tais aplicações. Esta ferramenta é capaz de exportar as aplicações multissensoriais para documentos seguindo a proposta de NCL 4.0 [Josue, 2021]. Esta proposta estende a linguagem declarativa NCL (*Nested Context Language*), padrão atual do middleware Ginga, para suportar efeitos sensoriais além de outras funcionalidades. NCL 4.0 também inclui suporte à interface avançada do usuário, oferecendo interação multimodal via voz, gestos, fixação do olhar etc., além de suporte à identificação e interação [Barreto, 2021].

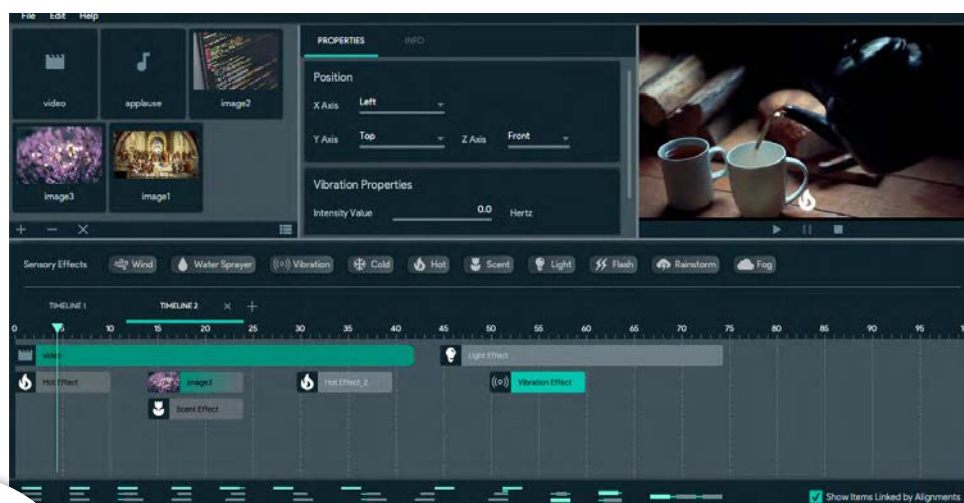
STEVE 2.0 usa o paradigma de sincronização temporal baseado em eventos para definir o comportamento temporal das aplicações multimídia multissensoriais.

A ferramenta fornece graficamente relações temporais causais para que os usuários possam sincronizar tanto mídias tradicionais como efeitos sensoriais no tempo. Além disso, STEVE exibe janelas de alertas sobre inconsistências de sincronização temporal ao utilizar estas relações causais. Os autores também podem criar relações de interatividade para ativar, por exemplo, efeitos sensoriais devido às interações dos telespectadores com as aplica-

¹ <http://www.midiacom.uff.br>

ções. O editor também fornece edição da visão espacial das aplicações para permitir que os autores editem características de renderização, por exemplo, intensidade, tipo de aroma, ou cor da luz, e localização de efeitos sensoriais no ambiente físico em que se encontra o receptor de TV digital. Além disso, pode-se editar as propriedades de apresentação de conteúdo multimídia tradicional, como por exemplo, transparência de uma imagem ou vídeo, volume do áudio, tamanho e posição das mídias visuais. STEVE 2.0 ainda permite aos autores verificar o comportamento temporal e espacial fornecendo uma visão gráfica temporal e espacial de forma sincronizada. Com uma interface gráfica baseada em visão temporal, STEVE torna-se parecido com outras ferramentas de edição de vídeo comerciais, o que facilita aos autores sem habilidades de programação definir suas próprias aplicações multissensoriais.

A **Figura 1** mostra a interface do STEVE 2.0. Para ilustrar suas funcionalidades, um exemplo de aplicação multimídia multissensorial interativa, demonstrando como STEVE pode auxiliar os usuários a criar aplicações multissensoriais graficamente e exportá-las em NCL 4.0, visando à proposta de TV imersiva no contexto do projeto TV 3.0. Esse exemplo descreve um ambiente imersivo, como por exemplo, um jogo educativo com vários efeitos sensoriais para atividades lúdicas com crianças. Nesta aplicação multissensorial, o conteúdo audiovisual é apresentado na TV e há atuadores físicos disponíveis no ambiente.



Assista aqui vídeo
com demonstração
da aplicação



Figura 1: Interface Gráfica do STEVE 2.0

No início da aplicação, é exibida uma sala de estar representada por um item de mídia de imagem (*background_image*) conforme ilustrado na **Figura 3**. Após 5s do início da aplicação, o vídeo e o áudio que ilustram a lareira em chamas são iniciados. Ao mesmo tempo, um efeito de luz também passa a representar a luminosidade das chamas.

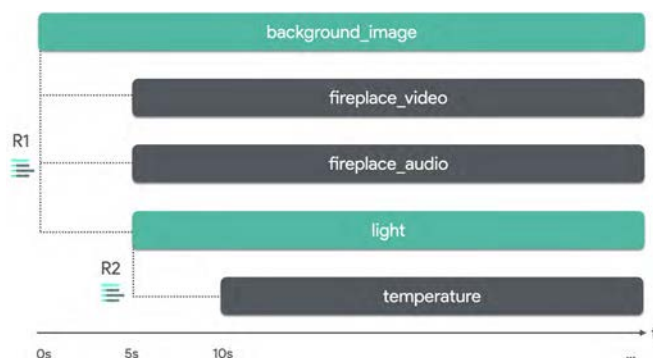
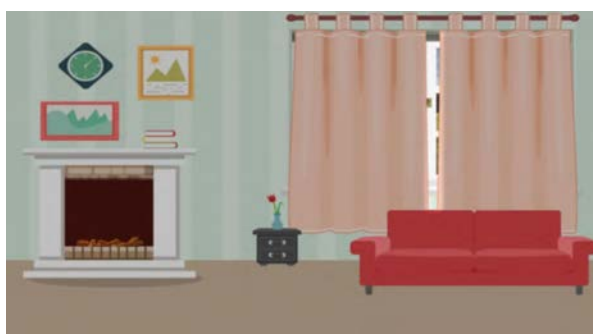


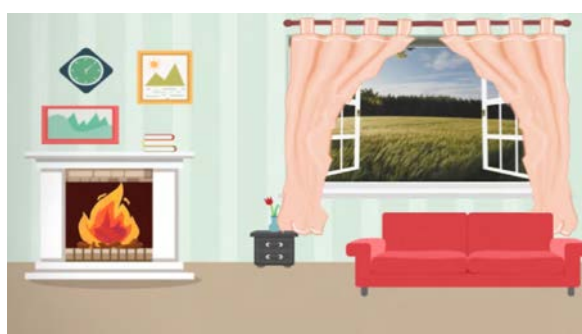
Figura 2: Visão Temporal 1

A **Figura 4** ilustra o segundo cenário da aplicação exemplo, em que mais dois itens de mídia e três efeitos sensoriais são iniciados. Entretanto, esta segunda linha do tempo somente se inicia quando o usuário da aplicação interage pressionando a tecla verde do controle remoto e apenas se uma expressão lógica também é satisfeita. Os itens de *window video* e *window audio* representam a abertura da janela e a vista externa, resultando na cena

ilustrada na **Figura 3b**. Os efeitos sensoriais acionados são luz (*light*, vento *wind* e frio *temperature* para representar a luminosidade de fora e o vento frio da janela, respectivamente. A avaliação da expressão lógica é definida para avaliar se a janela está aberta (*windowsOpen = 1*) ou fechada (*windowsOpen = 0*). Se estiver fechada, a relação de interatividade define o valor *windowsOpen* para 1 e inicia os itens de mídia e efeitos sensoriais.



(a) Sala de estar (*background image*) no início da aplicação



(b) Sala de estar após a interação do usuário

Figura 3: Cena da sala de estar antes e após a interação do usuário

Voltando à interface gráfica do STEVE 2.0 apresentada na **Figura 1**, um repositório de conteúdo multimídia tradicional é oferecido no canto superior esquerdo, uma região no centro permite que usuários editem propriedades de apresentação de mídias e características de renderização de efeitos sensoriais, uma região

para reproduzir o conteúdo audiovisual está disponível no canto superior direito e uma região dedicada à visão temporal da aplicação está em destaque na parte central-inferior. Além disso, STEVE 2.0 fornece uma lista de tipos de efeitos sensoriais acima da visão temporal, conforme mostrado na **Figura 1**.

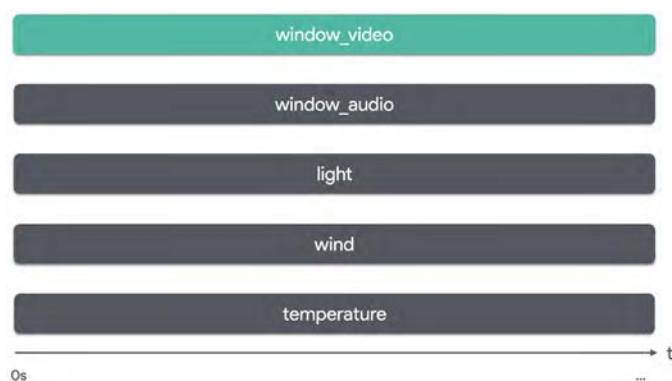
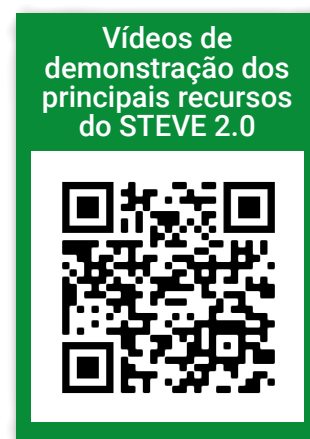


Figura 4: Visão Temporal 2



Os usuários podem selecionar um desses efeitos sensoriais e arrastá-los para a linha do tempo (ou visão temporal) para sincronizá-los temporalmente com outros nós. Para fazer essa sincronização temporal, os usuários podem usar as relações temporais causais que STEVE fornece graficamente, apresentadas abaixo da linha do tempo, conforme mostrado na **Figura 1**. Pode-se interpretar cada uma destas relações como um alinhamento temporal entre dois ou mais itens da aplicação, seja efeito sensorial ou mídia tradicional. Além disso, a região de reprodução do conteúdo audiovisual também exibe ícones que representam efeitos sensoriais sincronizados temporalmente com o conteúdo multimídia tradicional. Por exemplo, na **Figura 1**, um efeito de calor é ilustrado por um ícone de fogo, mostrado sincronizado com início do primeiro vídeo da aplicação.

A **Figura 5** exibe a visão temporal do STEVE 2.0 que corresponde à visão temporal da **Figura 2**. Para definir esse comportamento temporal no STEVE 2.0, os autores devem primeiro arrastar os itens de mídia e os efeitos sensoriais do repositório de mídia e da barra de efeito sensorial, respectivamente, para a linha do tempo. Depois disso, eles devem sincronizá-los usando os botões de relação temporal.

Além das relações temporais síncronas, STEVE 2.0 fornece a definição de relações assíncronas para especificar interações do

usuário. Para criar a interação do usuário do segundo cenário da aplicação, os autores precisam especificar a tecla de interatividade e quais itens de mídia ou efeitos sensoriais irão parar e iniciar após a interação do usuário. Na aplicação exemplo apresentada aqui, os autores precisam definir a tecla **Green** como a tecla de interatividade e selecionar os nós **window_video** e **window_audio**, e os efeitos **light**, **wind** e **hot** para iniciarem após a interação dos usuários com a tecla **Green** do controle remoto, como demonstrado na **Figura 6**. Após definir essa relação de interatividade, STEVE cria automaticamente uma nova linha do tempo, **Linha do tempo 2** como mostrado na **Figura 5b**, para conter os novos itens que iniciam devido à interação do usuário. Portanto, para cada relação de interatividade definida, STEVE 2.0 cria uma nova linha do tempo baseada em eventos.

STEVE 2.0 também é integrado a uma solução de aprendizado de máquina para aprimorar a definição de efeitos sensoriais de conteúdos visuais. Este componente, denominado STEVEML (**STEVE Machine Learning**) [Abreu et al., 2020], possibilita a extração de efeitos sensoriais automaticamente de conteúdos visuais como vídeo e imagem. Com essa integração, os autores podem selecionar itens de vídeo ou imagem na visão temporal de STEVE e solicitar a extração dos efeitos sensoriais para a mídia selecionada de acordo com os tipos de efeito desejados pelo autor.



(a) Linha do Tempo 1



(b) Linha do Tempo 2

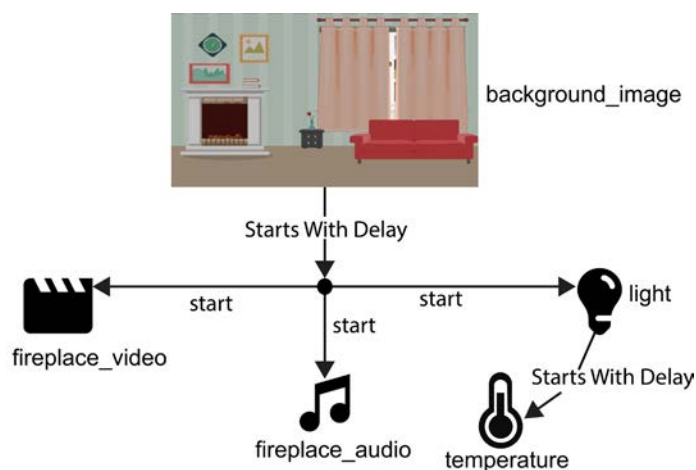
Figure 5: Linhas do Tempo Baseadas em Eventos do STEVE para a Aplicação Exemplo

Após a extração, os autores também podem fazer ajustes manuais na sincronização temporal dos efeitos extraídos, permitindo um refinamento posterior. Ao finalizar a aplicação, o

autor pode exportar o projeto em NCL 4.0, através do menu principal do STEVE 2.0, gerando automaticamente o código da aplicação em XML.



Vídeos de
demonstração
do STEVEML



Workflow do Steve / Fonte: Autores

Considerações Finais

Experimentos remotos foram realizados com o STEVE 2.0 e seus resultados se mostraram promissores [Paulo de Mattos, 2021]. A partir de análises dos resultados, a ferramenta mostrou-se capaz de aprimorar a fase de autoria de aplicações multimídia multissensoriais interativas, permitindo que até mesmo usuários sem conhecimento de programação criem suas próprias aplicações. Dessa forma, STEVE 2.0 possui um grande potencial de ser uma ferramenta crucial na fase de autoria de aplicações multissensoriais interativas para a futura TV 3.0.

Dessa forma, STEVE 2.0 possui um grande potencial de ser uma ferramenta crucial na fase de autoria de aplicações multissensoriais interativas para a futura TV 3.0.

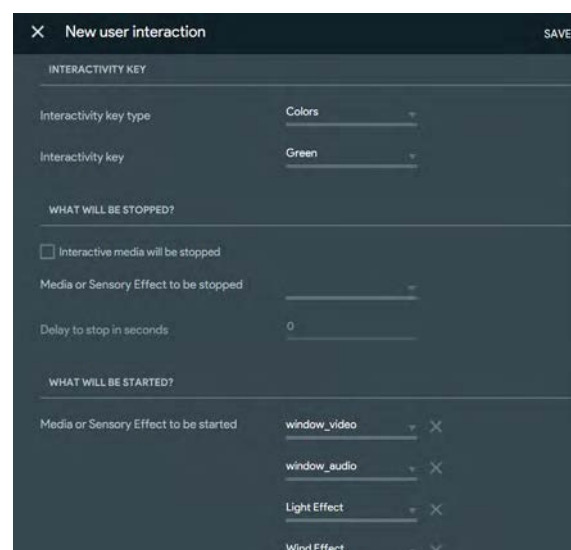


Figura 6: Definindo Relações de Interatividade no STEVE 2.0

Referências

[ABNT, 2021] ABNT (2021). Parte 2: Ginga-ncl para receptores fixos e móveis - linguagem de aplicação xml para codificação de aplicações. NBR 15606- 2:2021.

[Abreu et al., 2020] Abreu, R., Mattos, D., Santos, J. A. d., Ghinea, G., and Muchaluat-Saade, D. C. (2020). *Towards content-driven intelligent authoring of mulsemedia applications*. IEEE MultiMedia.

[Barreto, 2021] Barreto, F. (2021). *Uma Proposta de Extensão do Middleware Ginga-NCL para Interação Multimodal e Suporte Multiusuário em Ambientes Hipermedia*. PhD thesis, Universidade Federal Fluminense.

[de Mattos et al., 2020] de Mattos, D. P., Muchaluat-Saade, D. C., and Ghinea, G. (2020). *An approach for authoring mulsemedia documents based on events*. In *2020 International Conference on Computing, Networking and Communications (ICNC)*, pages 273{277. IEEE.

[Ghinea et al., 2014] Ghinea, Timmerer, Lin, and Gulliver (2014). *Mulsemedia: State of the art, perspectives, and challenges*. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 11(1s):17.

[Josu_e, 2021] Josu_e, M. (2021). *Preparação de Objetos de Mídia e Efeitos Sensoriais para Formação de Documentos Mulsemídia*. Ph.d. thesis, Universidade Federal Fluminense.

[Paulo de Mattos, 2021] Paulo de Mattos, D. (2021). *An Approach for Authoring Mulsemedia Applications Based on Events*. Tese de doutorado, Universidade Federal Fluminense.

[Saleme et al., 2019] Saleme, E. B., Covaci, A., Mes_n, G., Santos, C. A., and Ghinea, G. (2019). *Mulsemedia diy: a survey of devices and a tutorial for building your own mulsemedia environment*. *ACM Computing Surveys (CSUR)*, 52(3):1-29.



Débora Christina Muchaluat Saade é engenheira de Computação, mestre e doutorada em Informática pela PUCRJ. É professora titular da Universidade Federal Fluminense (UFF). Fundou e coordenada o Laboratório MídiaCom na UFF em 2003 com foco em pesquisas em redes de computadores e multimídia. É membro do Módulo Técnico do Fórum SBTVD.

Contato: debora@midiacom.uff.br



Douglas Paulo de Mattos é doutor em computação na área de sistemas multimídia pela Universidade Federal Fluminense (UFF). Foi pesquisador na Brunel University London (2019). Mestre (2016) em Computação pela UFF na área de Sistemas Multimídia, criando novas tecnologias para facilitar a autoria de aplicações hipermedia para TV Digital e Web.

Contato: dougpmattos@gmail.com