

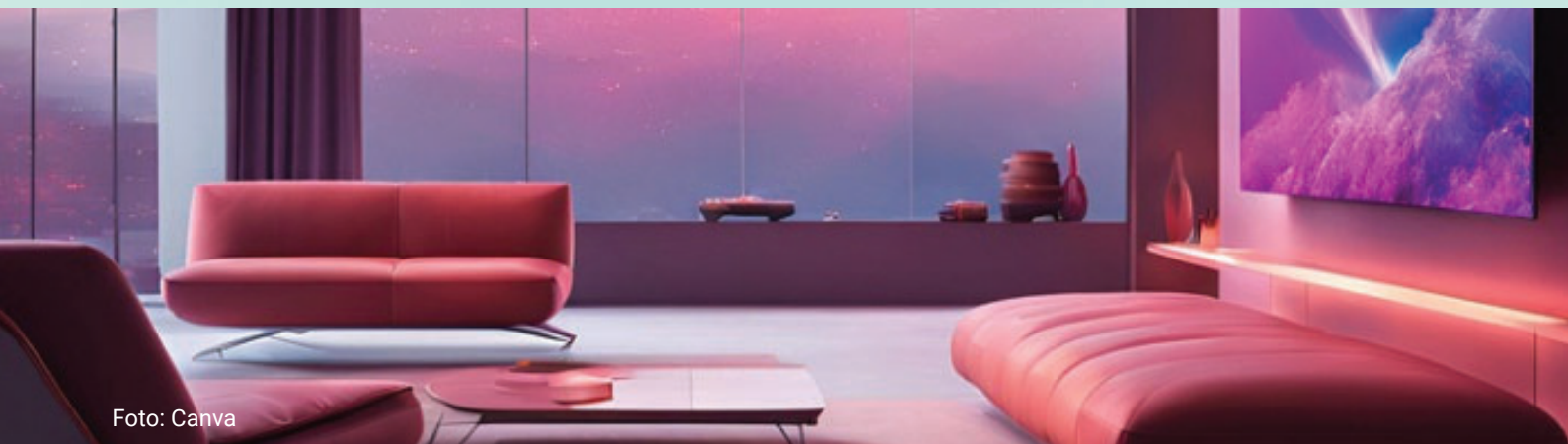


Foto: Canva

# TV Imersiva: Realidade Virtual e Interatividade Avançada para o Telespectador

Por Débora Christina Muchaluat-Saade e Joel André Ferreira dos Santos

*Professores do Laboratório MídiaCom da Universidade Federal Fluminense (UFF) e do Laboratório MultiSenS da Cefet/RJ explicam as novas funcionalidades que terá a TV 3.0 e que podem transformar os seus conteúdo em uma experiência realmente imersiva.*

A TV 3.0 traz uma nova geração de TV aberta para o telespectador brasileiro, com mais qualidade de vídeo e áudio, além de recursos inovadores, que podem aumentar a imersão e oferecer interatividade avançada ao público, transformando a TV de hoje em uma TV imersiva.

Na TV imersiva, o conteúdo audiovisual transmitido pela emissora via radiodifusão pode ser enriquecido com conteúdo adicional de realidade virtual transmitido pela Internet. O telespectador pode utilizar um HMD (*Head Mounted Display*) para consumir cenas imersivas com vídeos 360 graus sincronizados com outros objetos de mídia e também efeitos sensoriais. O receptor da TV 3.0 pode se conectar a dispositivos de Internet das Coisas (IoT - *Internet of Things*) disponíveis através da rede local nas residências, tais como lâmpadas inteligentes, difusores de aroma inteligentes, ventiladores inteligentes, etc. para renderizar efeitos de luz, aroma, vento, calor, frio, vibração, etc. transformando a sala de estar em um cinema 4D. A **Figura 1** ilustra este novo cenário da TV imersiva.

Imagine uma cena de filme que mostra um campo de lavanda na região da Provence na França e um aroma de lavanda é difundido no ambiente. Ou mesmo uma cena de novela, onde os personagens estão tomando café da manhã e um aroma de café é sentido pelo telespectador. Em outro cenário, em

um filme de terror, por exemplo, a luz do ambiente é desligada e pisca algumas vezes quando a luz da cena também tem esse mesmo efeito. Esses efeitos sensoriais podem ser criados e utilizados através das novas APIs (*Application Programming Interfaces*) da Camada de Codificação de Aplicações da TV 3.0 tanto para aplicações NCL [1] quanto para aplicações HTML5 [2], as duas linguagens declarativas disponíveis para especificar aplicações e serviços interativos. Os efeitos sensoriais serão renderizados se os dispositivos inteligentes estiverem disponíveis na casa do telespectador, caso contrário, a programação audiovisual tradicional continua normalmente.

Outra funcionalidade avançada para o

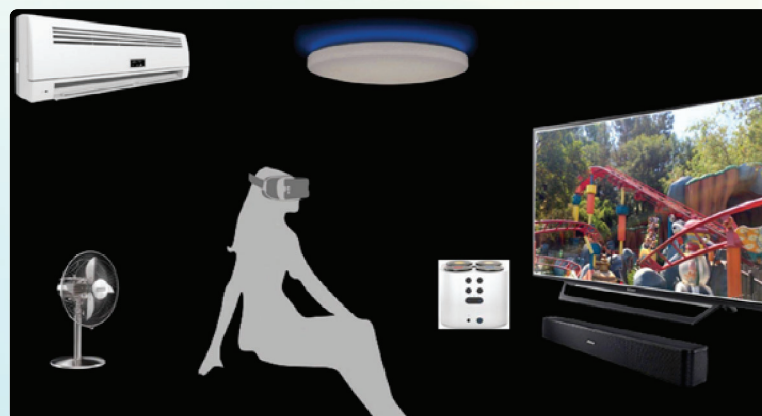


Figura 1 - TV imersiva

telespectador é a possibilidade de interação multimodal, além da interação pelo controle remoto. Na TV 3.0, comandos de voz, gestos, expressões faciais, toque, e até mesmo fixação do olhar na tela podem ser utilizados pelo telespectador para interatividade avançada [3]. Para comandos de voz, gestos ou reconhecimento de expressões faciais, o receptor de TV 3.0 deve oferecer microfone e câmera, que podem ser embarcados no receptor ou disponibilizados por algum dispositivo externo conectado ao receptor. Para interação através de fixação do olhar, um dispositivo rastreador ocular deve estar conectado ao receptor. Apesar de alguns receptores de TV atuais oferecerem comandos de voz ou gestos, essa facilidade é proprietária de cada dispositivo e não está disponível para os serviços interativos oferecidos pelas emissoras de TV aberta. Oferecendo esses outros modos de interação, a interatividade e participação nos programas veiculados será muito mais natural e facilitada, como ilustrado na **Figura 2**. Com a TV 3.0, essas facilidades podem ser exploradas e utilizadas em qualquer programa de TV através de novas APIs NCL e Web Services da Camada de Codificação de Aplicações.

A oferta de conteúdo de realidade virtual na TV 3.0 se dá por meio de cenas 360, que são cenas montadas a partir de um vídeo 360° e enriquecidas com imagens/vídeos 2D, áudio e objetos 3D. Estas cenas são descritas com a linguagem NCL 360, mantendo uma notação similar para a criação da aplicação sendo executada na TV e a cena 360. Um exemplo de cena 360 é apresentado na **Figura 3**.

Cenas 360 podem ser usadas para oferecer um conteúdo alternativo ao apresentado na TV, como um passeio de asa delta durante um programa de turismo ou similar a este, como uma versão 360 de um concerto assistido na TV, com elementos visuais que complementam a experiência. O conteúdo de

uma cena pode ser controlado pela TV e vice-versa. Assim, por exemplo, a interação realizada na TV pode influenciar o que é visto na cena 360, bem como a direção do olhar na cena 360 pode influenciar o que é visto na TV.

Toda essa integração é possibilitada pelas APIs **Multi-device** e **Remote Device** [4,5] da Camada de Codificação de Aplicações da TV 3.0. Com essas APIs, um dispositivo na rede local pode se registrar com a TV e, a partir de então, receber ações e notificar eventos durante a execução de um conteúdo informado pela TV. No caso de cenas 360, o PresentationEngine Guaraná [6] é uma aplicação que roda em dispositivos HMD e é capaz de se conectar à TV, trocar metadados conforme as APIs mencionadas anteriormente e renderizar cenas 360.

A nova geração da TV aberta brasileira também é inovadora em relação aos padrões para codificação de vídeo adotados. O codificador de vídeo obrigatório é o VVC (**Versatile Video Coding**) (H.266) [7] e opcionalmente pode-se utilizar codificadores de vídeo volumétrico para conteúdo imersivo como o V3C: **Video-based Point Cloud Compression** (V-PCC) – ISO/IEC 23090-5 [8], e MPEG **ImmersiveVideo** (MIV) – ISO/IEC 23090-12 [9]. Os codificadores de vídeo volumétrico oferecem a noção de profundidade dos objetos na cena, permitindo que o telespectador navegue espacialmente pela cena para explorar melhor o conteúdo em diferentes pontos de vista.

Oferecendo as novas funcionalidades para transformar a TV em uma TV imersiva, a TV 3.0 se coloca como um sistema inovador, utilizando tecnologia de ponta, disponível no mercado e integrada ao novo sistema através de atividades de pesquisa e desenvolvimento realizadas pelas universidades públicas brasileiras em parceria com o setor de radiodifusão

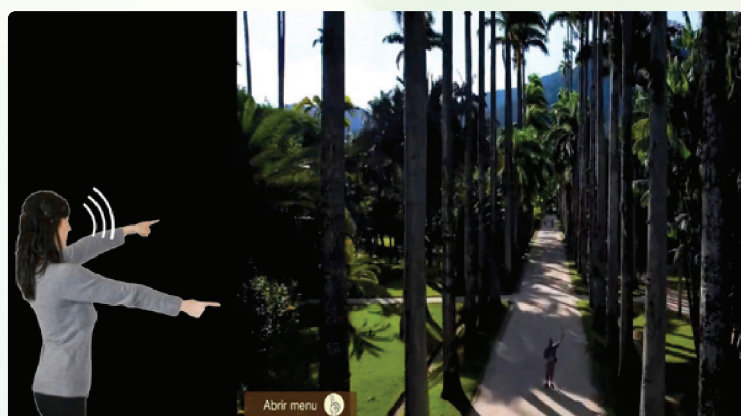


Figura 2 - Interação avançada através de voz e gestos



Figura 3 - Cena 360 de um concerto de música clássica

## Referências

Ivanov, Moreno, Muchalut-Saade. Automatic Preparation of Sensory Effects. In: MMSys '24: ACM Multimedia Systems Conference 2024, 2024, Bari Italy. Proceedings of the ACM Multimedia Systems Conference 2024. New York: ACM, 2024.

W3C, HTML5 Living Standard, disponível em <http://www.w3.org/TR/html5/>

Barreto, Abreu, Josué, Montevicchi, Valentim, Muchalut-Saade. Providing multimodal and multi-user interactions for digital tv applications. Multimedia Tools and Applications, v. 82, p. 4821-4846, 2023.

Oliveira, Macedo, Josué, Muchalut-Saade, dos Santos. TV 3.0: A Ginga-NCL and Common Core Webservices Extension for Multidevice Support. Workshop O Futuro da TV Digital Interativa - Simpósio Brasileiro de Sistemas Multimídia e Web, 2023.

dos Santos, Vieira, Josué, Oliveira, Muchalut-Saade. Multidevice Support in the Next Generation of the Brazilian Terrestrial TV System.

International Conference on Interactive Media Experiences - Latin America Workshop, 2024.

Souza, Silva, Delgado, Rodrigues, Mendes, Amorim, Guedes, dos Santos. Interactive 360-degree Videos in Ginga-NCL Using Head-Mounted-Displays as Second Screen Devices. Simpósio Brasileiro de Sistemas Multimídia e Web, 2020.

ITU-T, H.266 :Versatile video coding, disponível em <https://www.itu.int/rec/T-REC-H.266>

ISO/IEC 23090-5:2023 Information technology – Coded representation of immersive media Part 5: Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC), disponível em <https://www.iso.org/standard/83535.html>

ISO/IEC 23090-12:2023 Information technology – Coded representation of immersive media Part 12: MPEG immersive video, disponível em <https://www.iso.org/standard/79113.html>



**Profa.**

**Debora Christina  
Muchalut Saade**

É graduada em Engenharia de Computação (1992), mestrado em Informática (1996) e doutorado em Informática pela Pontifícia Universidade Católica do Rio de Janeiro (2003). É professora titular da Universidade Federal Fluminense (UFF). Integrou o corpo docente do Departamento de Engenharia de Telecomunicações até maio de 2009 e desde então faz parte do corpo docente do Instituto de Computação. É bolsista de produtividade DT nível 1D do CNPq e Cientista do Nosso Estado (CNE) pela FAPERJ. Foi Jovem Cientista do Nosso Estado (JCNE) de 2009-2011 pela FAPERJ. Foi vice-coordenadora do Programa de Pós-graduação em Computação da UFF, avaliado com conceito 6 pela CAPES, de 2017 a 2021, e também foi coordenadora do Curso Superior de Tecnologia em Sistemas de Computação da UFF de 2014 a 2019. Foi coordenadora da Comissão Especial de Computação Aplicada à Saúde (CE-CAS) da Sociedade Brasileira de Computação (SBC) de 2017 a 2019. Fundou o Laboratório MídiaCom na UFF em 2003 com foco em pesquisas em redes de computadores e multimídia ([www.midiacom.uff.br](http://www.midiacom.uff.br)) e desde então é uma das coordenadoras do laboratório. É membro do Módulo Técnico do Fórum do Sistema Brasileiro de TV Digital. Suas áreas de interesse são multimídia, mulemídia, redes de computadores, redes sem fio, redes elétricas inteligentes, Internet das Coisas, televisão digital interativa e saúde digital. Participou do desenvolvimento da linguagem NCL - Nested Context Language - adotada como padrão ABNT NBR 15606-2 no middleware GINGA do Sistema Brasileiro de TV Digital e como recomendação internacional do ITU-T H.761 para serviços IPTV.

**Contato:** [debora@midia.com.uff.br](mailto:debora@midia.com.uff.br)



**Prof.**

**Joel André Ferreira  
dos Santos**

É graduado em Engenharia de Telecomunicações pela Universidade Federal Fluminense (UFF) (2009), com sanduíche (2007-08) na Universität Ulm, Mestrado (2012) e Doutorado em Ciência da Computação (2016), com bolsa Nota DEZ pela FAPERJ e sanduíche (2014-15) no Inria-Grenoble. Docente do Cefet/RJ atuante nos três níveis de ensino. É membro do Departamento de Ciência da Computação, da Coordenação de Informática e professor permanente do Programa de Pós-graduação em Ciência da Computação (PPCIC). Como docente do PPCIC orientou dissertações ganhadoras do prêmio de melhor dissertação do CTD - WebMedia 2019 e menção honrosa no Prêmio Artur Ziviani de Teses e Dissertações do SBCAS 2021. Atuou como vice-coordenador (2016-21) e atual coordenador (2021-) do PPCIC. Lidera desde 2020 o Grupo de Pesquisa MultiSenS - Laboratório de Espaços e Aplicações Multimídia Multissensoriais Imersivas. Foi coordenador geral do XXV Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia 2019), coordenador do Comitê de Programa da XXVII edição do evento realizada em 2021. Foi vice-coordenador (2020-21) e coordenador (2022-23) da Comissão Especial de Sistemas Multimídia e Web (CE-WebMedia) da SBC (Sociedade Brasileira de Computação). Foi editor convidado da Journal on Interactive Systems - JIS (2024). Atua na área de Multimídia desde 2006 em diversos projetos relacionados a TV digital, autoria e validação de aplicações multimídia e aplicações multimídia multissensorial. É proponente da tecnologia Guaraná (APIs NCL 360 VR e Multidevice) adotadas na TV 3.0.

**Contato:** [joel.afs@gmail.com](mailto:joel.afs@gmail.com)