

REALIDADE ESTENDIDA (XR) E *IMMERSIVE LEARNING*

Prof. André Roberto Guerra

INTRODUÇÃO

O avanço da multimídia e da realidade virtual (RV), proporcionado pela maior potência dos computadores, permitiu a integração, em tempo real, de vídeo e ambientes virtuais interativos. Ao mesmo tempo, o aumento da largura de banda das redes de computadores também influencia positivamente na evolução da multimídia, permitindo a transferência de imagens e outros fluxos de informação com eficiência.

A realidade aumentada (RA) beneficiou-se dessa evolução, tornando viáveis, acessíveis e populares aplicações que antes só existiam nos ambientes acadêmico, de pesquisa ou industriais, baseados em plataformas sofisticadas. Ao mesmo tempo, pesquisas têm feito evoluir o potencial dessa tecnologia, trazendo-a para aplicações cotidianas, a ponto de chamar a atenção de grandes empresas como Microsoft, Apple, IBM, HP, Sony, Google, Facebook, dentre outras.

Na discussão até aqui sobre Realidade Estendida (XR) e *Immersive Learning*, caracterizamos os principais dispositivos de entrada e saída utilizados em RV, além de termos destacado os requisitos básicos necessários ao processo de desenvolvimento de sistemas e os conceitos de modelagem e programação relacionados aos ambientes virtuais. Dando continuidade ao conteúdo, teremos aqui como objetivos:

- demonstrar a arquitetura típica de um sistema de RA, assim como as tecnologias relacionadas;
- apresentar os fundamentos que norteiam a RA, com foco especial no conceito de rastreamento (*tracking*); e
- abordar as principais vantagens e desvantagens da utilização de RA verificadas atualmente.

E seguirá a seguinte estrutura:

- arquitetura típica de um sistema de RA;
- tecnologias de RA;
- fundamentos da RA: rastreamento (*tracking*);
- vantagens e desvantagens da RA; e
- conclusão e próximos passos.

Aproveitem! Bons estudos!

TEMA 1 – ARQUITETURA TÍPICA DE UM SISTEMA DE RA

Diferentemente da RV, que transporta o usuário para o ambiente virtual, a RA mantém o usuário em seu ambiente físico e transporta o ambiente virtual para o espaço do usuário, permitindo a interação com o mundo virtual de maneira mais natural e sem necessidade de treinamento ou adaptação. Novas interfaces multimodais estão sendo desenvolvidas para facilitar a manipulação de objetos virtuais no espaço do usuário, usando as mãos ou dispositivos mais simples de interação (Kirner; Tori, 2006).

O uso de rastreamento óptico de pessoas ou mãos e as técnicas de RA podem colocar elementos reais, como as mãos, para interagir com o ambiente virtual, eliminando os inconvenientes dos aparelhos tecnológicos. Além disso, é possível também enriquecer uma cena real, capturada por câmera de vídeo, por exemplo, com elementos virtuais interativos, de forma a permitir muitas aplicações inovadoras. Para exemplificar, pode-se citar a decoração, em tempo real, de um apartamento vazio (real) com um mobiliário virtual (Figura 1). Nesse caso, o usuário pode usar um capacete de visualização com uma câmera de vídeo acoplada, mostrando a visão real enriquecida com os elementos virtuais posicionados adequadamente pelo computador. O sistema é implementado de tal maneira que o cenário real e os objetos virtuais permanecem ajustados, mesmo com a movimentação do usuário no ambiente real (Kirner; Tori, 2006).

Figura 1 – Exemplo de utilização da RA para decoração de interiores



Créditos: Gorodenkoff / Adobe stock.

Assim, a RV e a RA permitem ao usuário retratar e interagir com situações imaginárias, como os cenários de ficção, envolvendo objetos reais e virtuais estáticos e em movimento. Permitem também reproduzir, com fidelidade, ambientes da vida real – como a casa virtual, a universidade virtual, o banco virtual, a cidade virtual etc., –, de modo que o usuário possa entrar nesses ambientes e interagir com seus recursos de forma natural, usando as mãos (com ou sem aparatos tecnológicos, como luvas) e, eventualmente, comandos de voz. Com isso, o usuário pode visitar salas de aula e laboratórios de universidades virtuais, interagindo com professores e colegas e realizando experimentos científicos; pode entrar no banco virtual e manusear o terminal de atendimento virtual, da mesma maneira que usa o equipamento real, inclusive conversar com o gerente, representado, no ambiente, por um humanoide virtual (avatar) (Kirner; Tori, 2006).

Um sistema gráfico é composto tipicamente por módulos de entrada, processamento e saída de informações, que podem se dividir em várias tarefas, como mostra o Quadro 1.

Quadro 1 – Exemplo de tarefas necessárias para compor um sistema gráfico

MÓDULO DE ENTRADA
<ul style="list-style-type: none">• Captura de vídeo: responsável por capturar a cena real na qual serão inseridos os objetos virtuais;• Sensoriamento: quaisquer dispositivos que sejam usados para identificar objetos, observador e/ou seu posicionamento e ações.
MÓDULO DE PROCESSAMENTO
<ul style="list-style-type: none">• Monitoramento dos objetos: responsável por identificar, de forma única e precisa, uma indicação a um objeto virtual em uma configuração (posição e orientação) específica (o que se chama de registro – <i>registering</i>) e também por identificar como esse objeto virtual desloca-se no ambiente (o que se chama de rastreamento – <i>tracking</i>);• Gerenciamento da interação: responsável por identificar e determinar a resposta às ações de seleção ou manipulação dos objetos virtuais;• Processamento da aplicação: responsável por dar sentido às interações e promover mudanças na cena, conforme os objetivos da aplicação (jogo, ambiente industrial, aplicação de turismo etc.).

MÓDULO DE SAÍDA

- **Visualização:** responsável por renderizar visualmente o objeto virtual na condição especial requerida pela aplicação e, então, mostrá-lo ao usuário por um dispositivo de visualização apropriado;
- **Atuação:** responsável por renderizar parâmetros para dispositivos hápticos.

Fonte: Baseado em Hounsell, Tori e Kirner, 2021.

Os módulos de entrada e saída são fortemente dependentes do *hardware* utilizado. O módulo de processamento tem suas características dependentes das técnicas de *software* utilizadas. Um ciclo de processamento pode ser resumido em:

- captura de vídeo e execução do rastreamento de objetos;
- processamento do sistema de RA, incluindo leitura de dispositivos e simulação/animação;
- calibração, misturando o real com o virtual; e
- renderização sensorial, envolvendo os aspectos visuais, auditivos e hápticos (Hounsell; Tori; Kirner, 2021).

Como o sistema funciona em tempo real e deve apresentar tempo de latência igual ou menor que 100 ms, o processamento envolvido é muito maior que aquele considerado durante a discussão do processamento de sistemas de RV. No caso dos sistemas de RA, o processamento é uma das partes de um conjunto maior e mais complexo, que também envolve técnicas de multimídia (Hounsell; Tori; Kirner, 2021).

TEMA 2 – TECNOLOGIAS DE RA

Um dos componentes fundamentais da RA é uma câmera de vídeo para capturar a cena e, às vezes, identificar nela os locais de posicionamento dos elementos virtuais. Entretanto, além da entrada de vídeo, vários outros dispositivos podem ser usados com o objetivo de auxiliar na identificação dos elementos da cena e/ou posicionar os elementos em relação ao observador. Esses últimos também são fundamentais para auxiliar no processo de interação. O *hardware* de RA pode usar dispositivos típicos da RV, mas tem a preocupação de não obstruir as mãos, que devem atuar naturalmente no ambiente misturado. Técnicas de rastreamento visual, usando visão computacional e processamento

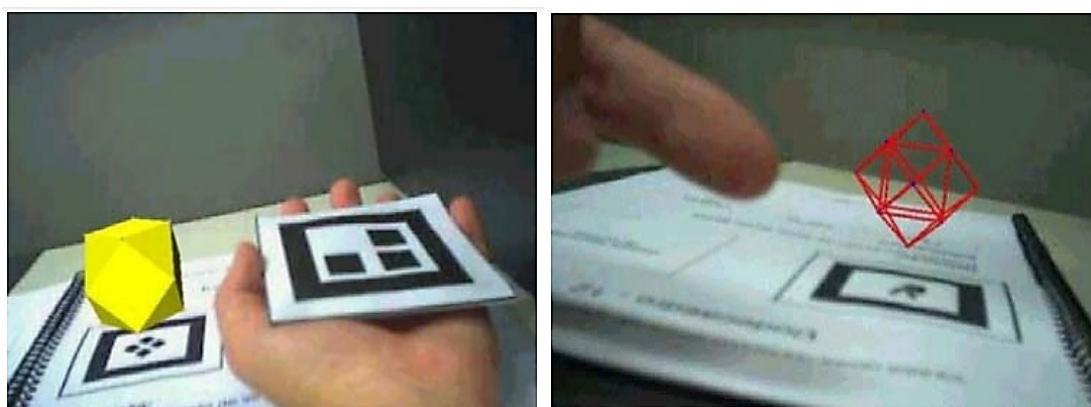
de imagens, são importantes, nesse caso. Assim, os seguintes dispositivos podem ser usados como *hardwares* de entrada para um sistema de RA (Hounsell; Tori; Kirner, 2021):

- **GPS (Global Positioning System)**: pode-se registrar a posição de um elemento virtual em um espaço físico mais amplo por meio de suas coordenadas geográficas (latitude e longitude). Assim, o sistema de RA vai posicionar os elementos virtuais em uma posição geográfica específica, mas, para isso, precisa saber onde o observador está;
- **Sensores iniciais** (acelerômetros, magnetômetros e giroscópios): para auxiliar na identificação da forma como a cena é observada, pode-se utilizar sensores para controlar o ângulo de visão. Eles ainda podem auxiliar na identificação de ações do usuário para usá-las como forma de interação;
- **Sensores de profundidade**: acoplados ao sistema de captura de imagens (como em câmeras RGBD, do tipo *Kinect*, *Xtion*, *RealSense*), ou isolados (como o *Leap Motion*), os sensores de profundidade são úteis para identificar a configuração do cenário físico ou para identificar a presença e a configuração da mão do usuário. A calibração dessa identificação em relação ao sistema de visualização permite verificar formas de interação do usuário com os elementos virtuais;
- **Luvas de dados**: capturam a configuração da mão do usuário e podem ser usadas isoladamente, para servir como forma de interação baseada em gestos (configuração dos dedos); ou, quando acopladas a rastreadores, permitem o posicionamento da mão do usuário na cena e sua interação com os elementos virtuais;
- **Interfaces tangíveis**: todo e qualquer dispositivo físico significante para a aplicação (no caso, uma ferramenta, um instrumento etc.) com o qual o usuário possa interagir diretamente (segurar, tocar, empurrar), mas que, ao mesmo tempo, possa servir como sensor de entrada para o sistema (identificando qual foi a ação), é uma interface tangível que pode ser utilizada também em um sistema de RA. Nas interfaces tangíveis, a interação não fica limitada a sutis movimentos de dedos e dos olhos, envolvendo também o espaço físico e a corporalidade do usuário. Assim, mesas translúcidas, com sistemas de visão acoplados, ou objetos físicos com sensores acoplados podem ser integrados a um sistema de RA (Reis; Gonçalves, 2016).

Ao mesmo tempo em que a RA demanda recursos de *hardware*, também impõe desafios de *software*, à medida que são desenvolvidas aplicações mais complexas e potentes. A capacidade de processamento das unidades centrais (CPU) e das placas gráficas (GPU), para tratar as necessidades da RA, deve ser alta o suficiente para garantir a execução, em tempo interativo, das seguintes ações: tratamento de vídeo; processamento gráfico 3D; geração de imagens misturadas; incorporação de som; execução háptica; controle multimodal; varredura de dispositivos de entrada com ênfase no rastreamento etc. (Hounsell; Tori; Kirner, 2021).

O *software* de RA é usado, na fase de preparação do sistema, por meio de ferramentas de autoria de ambientes misturados; e, na fase de execução, como um suporte à interação em tempo real. O *ARToolKit* é um dos recursos mais populares da RA, sendo uma biblioteca de *software* baseada nas linguagens Java, C e C++, usada para o desenvolvimento de aplicações de RA. Esse ambiente de desenvolvimento baseia-se no uso de marcadores. O *ARToolKit* é de código aberto e possibilita alteração e ajustes para aplicações específicas (Figura 2) que, junto com o *ARTag*, são os *frameworks* de desenvolvimento mais presentes nas publicações científicas no Brasil até recentemente (Hounsell et al., 2014).

Figura 2 – RA usando o *ARToolKit*



Fonte: Baseado em Hounsell, Tori e Kirner, 2021, p. 53.

A base para o funcionamento da RA é a capacidade de o sistema computacional identificar onde os elementos virtuais vão aparecer na cena. Para os elementos 1D e 2D, essa funcionalidade impacta pouco no funcionamento da aplicação, pois leva a um posicionamento em relação a tela, que pode ser fixo. Entretanto, quando o elemento é 3D, independentemente de sua qualidade

gráfica, essa integração torna-se fundamental, pois o objeto virtual tem que ser coerente com o resto do mundo 3D real (Hounsell; Tori; Kirner, 2021).

Para alcançar tal funcionalidade, no caso dos objetos 3D, o sistema de RA tem que executar as tarefas de monitoramento, que se dividem em:

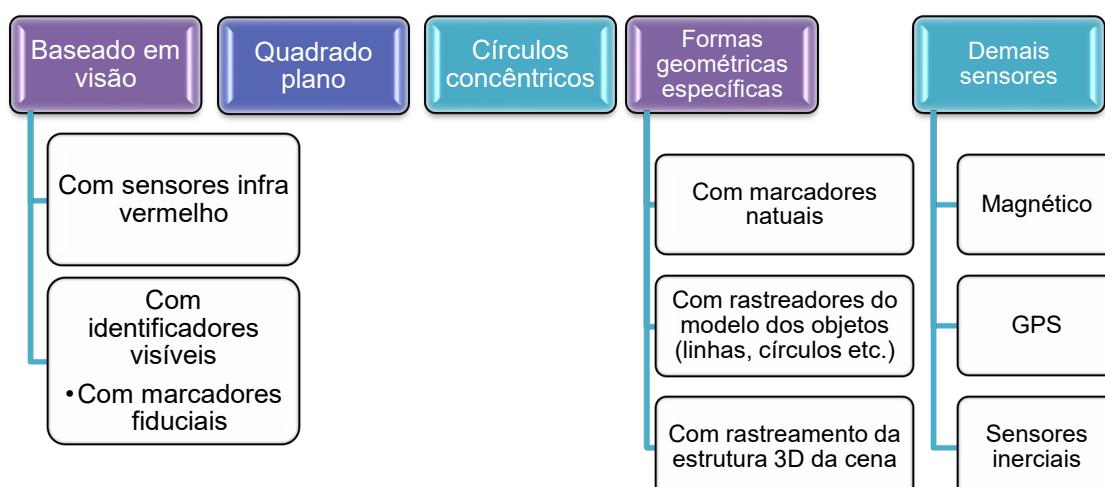
- **Registro:** relaciona-se com a capacidade que o sistema de RA deve ter para identificar qual é o elemento virtual que deve aparecer e em qual posição e orientação relativa ao restante da cena deve estar (tanto em relação a objetos reais como a virtuais e ao observador);
- **Rastreamento:** a capacidade que o sistema de RA deve ter de identificar como um elemento virtual presente na cena move-se e para onde (Hounsell; Tori; Kirner, 2021).

Erros no processo de registro podem fazer com que um objeto fique aparecendo e desaparecendo da cena (também chamado de *popping*), ou em posição incompatível com o resto da cena. Por outro lado, erros no processo de rastreamento podem fazer com que o objeto 3D acumule erros, realizando movimentação estranha ou acabando em um local inadequado (Hounsell; Tori; Kirner, 2021).

TEMA 3 – FUNDAMENTAS DA RA: RASTREAMENTO (TRACKING)

O rastreamento é tão fundamental que os tipos de RA podem ser classificados quanto a esse aspecto: RA baseada em visão; e RA baseada em sensores (Figura 3) (Billinghurst *et al.*, 2015).

Figura 3 – Tecnologias mais recentes de rastreamento

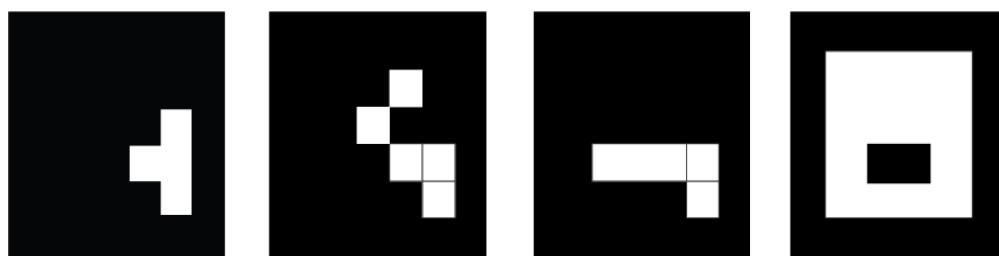


Fonte: Baseado em Hounsell, Tori e Kirner, 2021.

O rastreamento baseado em visão vem tornando-se muito popular em RA, devido a não exigir *hardware* adicional (além da câmera que, possivelmente, já esteja sendo usada) e à capacidade de processamento nos dispositivos móveis ser suficiente para analisar imagens. Os primeiros sistemas de rastreamento baseados em visão utilizavam emissores ou refletores de luz (tipo “olho de gato”), anexados aos objetos sendo rastreados. Isso tornava o rastreamento bem mais fácil, pela facilidade de controlar (e identificar) o emissor/refletor. Trabalhar com refletores mostrara-se mais fácil, pois eles resolvem os problemas de energia e sincronização em relação às fontes de luz. Se a fonte de luz era infravermelha, ela não era visível e, portanto, não influenciava o resto da cena. Apesar da precisão e da robustez dessa solução, ela requer uma fonte de luz específica (Hounsell; Tori; Kirner, 2021).

Trabalhar com luz visível para rastreamento baseado em visão facilita o processo, pois não requer nem fonte de luz nem sensor específicos, tornando a preparação do ambiente bem mais fácil, a qual é então feita procurando marcações (naturais ou artificiais) existentes na cena. Quando as marcações são artificiais e, portanto, adicionadas à cena, são chamadas de *fiduciais*. Marcadores fiduciais (Figura 4) são criados de forma a serem facilmente identificados pelo *software* de visão. Eles podem ser desde formas geométricas específicas em cores, anéis circulares concêntricos e coloridos, até uma simples marca quadrangular (Fiala, 2005).

Figura 4 – Exemplo de marcadores fiduciais utilizados em aplicações de RA



Fonte: Zorral e Silva, 2021, p. 132.

A RA pode, ainda, utilizar outros sensores, que não os óticos, para executar o rastreamento. No rastreamento magnético, tem-se um dispositivo que transmite um campo magnético alternado, o qual é captado por um ou mais receptores. Disso, então, são calculadas a posição e a orientação do receptor em relação ao transmissor. O rastreamento magnético é rápido e não requer visada direta entre os dispositivos, que são leves e pequenos. Entretanto, o volume de trabalho (área

na qual a precisão mantém-se aceitável) é limitado, e o sistema é sensível a interferências eletromagnéticas (Hounsell; Tori; Kirner, 2021).

O rastreamento por GPS é apropriado para espaços abertos e amplos. Um projeto adequado de uma aplicação de RA pode permitir que a precisão de 3 metros, atualmente disponível, seja suficiente. A vantagem desse rastreamento é que ele tende a melhorar em um futuro próximo. Por fim, os sensores inerciais têm as vantagens de não terem volume de trabalho limitado, não requererem linha de visada entre transmissor e receptor e não sofrerem interferências eletromagnéticas, acústicas ou óticas. Assim, podem ser amostrados a altas frequências, com quase nenhum atraso. No entanto, tais sensores são muito sensíveis ao acúmulo e à propagação de erros, tanto de posição como de orientação, ao longo do tempo, requerendo calibração periódica. Por essa razão, a composição híbrida de modos de rastreamento pode levar um sistema a ser mais versátil, robusto e preciso (Hounsell; Tori; Kirner, 2021).

TEMA 4 – VANTAGENS E DESVANTAGENS DA RA

Algumas das vantagens e aplicabilidades da RA confundem-se com as da RV, mas é possível encontrar algumas que são próprias da RA. Dentre elas, destacam-se (Wang; Ong; Nee, 2016):

- não é necessário fazer toda a modelagem do mundo virtual, o que normalmente demanda esforço manual, aumentando a dificuldade de integração com os sistemas de CAD, e também esforço computacional, para a renderização;
- o usuário pode agir no real (usar ferramentas, atuar sobre dispositivos, manipular objetos, mover-se em torno do objeto etc.) de forma natural, com suas propriedades responsivas ou hápticas (de peso/inércia, textura, rigidez etc.), o que dá maior sensação de realismo e de imersão no mundo enriquecido, trazendo o benefício tanto do real como do virtual;
- pode-se explorar novos elementos (virtuais) e sua interação com o ambiente (real) sem a necessidade de construir ou desenvolver os elementos, economizando tempo e recursos;
- proporciona um ambiente seguro, flexível, controlado e intuitivo, para que se experimentem interações físicas.

Por outro lado, as principais desvantagens da RA estão associadas à forma pela qual se promove a integração entre os dispositivos, o processamento e a tarefa em questão. Ou seja, não existem soluções prontas de como abordar uma determinada área. Muita pesquisa ainda precisa ser feita para analisar as formas mais intuitivas e naturais dessa integração, e isso tem-se refletido em sistemas com certas limitações. Como o foco ainda tem-se voltado para a interface com o sistema de RA, a oportunidade de explorar sistemas mais inteligentes ou, ainda, de explorar os próprios objetos reais da cena como recurso da aplicação tem sido esquecido (Wang; Ong; Nee, 2016).

As técnicas e soluções de rastreamento ainda estão em sua infância e, portanto, limitam as soluções possíveis. Isso traz, aos desenvolvedores, a responsabilidade de saber explorar os recursos de rastreamento atualmente disponíveis, mesmo que, para algumas aplicações, tais requisitos não sejam tão exigentes, como é o caso das aplicações de “fusão real-virtual fraca” (Ling, 2017).

TEMA 5 – CONCLUSÃO E PRÓXIMOS PASSOS

Em termos de aplicações, espera-se que, agora que a RA está madura e disponível para uma plateia mais ampla de desenvolvedores, os especialistas de domínios (conhecedores da área-problema) passem a ser os produtores de soluções com o uso da RA e não apenas os acadêmicos e pesquisadores, como ocorria até então. Assim, acredita-se que o foco desses profissionais passará a ser nas funcionalidades do sistema, e não apenas nos recursos de RA. Nesse sentido, a tendência é que a inteligência artificial, a percepção semântica e a internet das coisas tornem-se recursos integrados à RA, com a finalidade de promover experiências cada vez mais significativas e relevantes aos usuários.

Como a RA é fortemente dependente das funcionalidades do rastreamento (*tracking*), provavelmente as pesquisas nessa área continuarão a aparecer em busca de eficiência, robustez e flexibilidade. Uma das tarefas mais desafiadoras nas pesquisas em RA é o monitoramento do ponto de vista do observador. O objetivo é que o sistema de RA seja capaz de entender a composição da cena para que o objeto virtual, sendo inserido, seja coerente e possa interagir com ela.

Para finalizar, iremos abordar mais profundamente os conceitos de *Immersive Learning* e suas aplicações aos usuários na atualidade. Por fim, trataremos do funcionamento da RV imersiva e de como seus recursos estão

presentes no cotidiano dos milhões de usuários já presentes e ativos no mundo virtual.

REFERÊNCIAS

- BILLINGHURST, M. *et al.* A survey of augmented reality. **Foundations and Trends® in Human-Computer Interaction**, v. 8, n. 2-3, p. 73-272, 2015.
- FIALA, M. ARTag, a fiducial marker system using digital techniques. In: COMPUTER VISION AND PATTERN RECOGNITION, 2005. **Anais** [...]. IEEE, 2005. p. 590-596.
- HOUNSELL, M. S. *et al.* A brief history of Virtual Reality in Brazil: a survey over the publications in the “Symposium on Virtual and Augmented Reality”. **SBC Journal on Interactive Systems**, v. 5, n. 3, p. 78-92, 2014.
- HOUNSELL, M. S.; TORI, R.; KIRNER, C. Realidade Aumentada. In: _____. **Introdução a realidade virtual e aumentada**. 3. ed. Porto Alegre: SBC, 2021. p. 36-74.
- KIRNER, C.; TORI, R. Fundamentos de Realidade Aumentada. In: TORI, R.; KIRNER, C.; SISCOUTTO, R. (eds.). **Fundamentos e tecnologia de realidade virtual e aumentada**. Porto Alegre: SBC, 2006. p. 22-38.
- LING, H. Augmented Reality in Reality. **IEEE MultiMedia**, v. 24, n. 3, p. 10-15, 2017.
- REIS, A. V.; GONÇALVES, B. S. Interfaces Tangíveis: Conceituação e Avaliação. **Estudos em Design**, v. 24, n. 2, p. 92-111, 2016.
- WANG, X.; ONG, S. K.; NEE, A. Y. C. A comprehensive survey of augmented reality assembly research. **Advances in Manufacturing**, v. 4, n. 1, p. 1-22, 2016.
- ZORZAL, E. R.; SILVA, R. L. S. Software. In: TORI, R.; HOUNSELL, M. S. **Introdução a realidade virtual e aumentada**. 3. ed. Porto Alegre: SBC, 2021. p. 127-139.