

Tecnologias emergentes para produção de imagens sequenciais: Captura de Movimento

Emerging Technologies for Sequential Image Production: Motion Capture

Giovanna Florencio Antonio¹
giovanna.florencio@usp.br

Profª. Dra. Silvia Laurentiz²
laurentz@usp.br

Resumo

O seguinte artigo aborda uma pesquisa de iniciação científica sobre a técnica de animação por captura de movimento. Os resultados parciais do projeto foram apresentados no texto “Tecnologias emergentes para produção de imagens sequenciais: Rotoscopia e Captação de Movimento” (SIIMI 2025, prÁxIs) que tem como foco a base teórica e técnica construída a partir da revisão bibliográfica e da análise comparativa de obras e recursos. Este artigo tem como objetivo abordar a produção de um modelo experimental para captura de movimento, com base na fundamentação teórica elaborada, detalhando o processo de trabalho com softwares como *Clip Studio Paint*, *Live2D*, *VTube Studio* e *Animaze by Facerig*, analisando-o de uma perspectiva da acessibilidade de recursos, um aspecto fundamental na difusão da captura de movimento para novos públicos.

Palavras-chave: Animação, Arte, Captura de Movimento, Imagens Sequenciais

Abstract

The following article discusses an undergraduate research project on motion capture animation techniques. Partial results of the project were presented in the text “Emerging Technologies for Sequential Image Production: Rotoscoping and Motion Capture” (SIIMI 2025, prÁxIs), which focuses on the theoretical and technical basis built from a literature

¹ Giovanna Florencio é artista visual, educadora e graduanda dos cursos de Bacharelado e Licenciatura em Artes Visuais da ECA-USP. É parte do grupo de pesquisa Realidades - da realidade tangível à realidade ontológica, liderado pela Profª. Dra. Silvia Laurentiz, e desenvolveu uma pesquisa sobre captura de movimento e suas tecnologias. Atualmente, é bolsista no Programa de Iniciação à Docência da CAPES na área de artes.

² Silvia Laurentiz é Professora Associada do Departamento de Artes Plásticas e do Programa de Pós-Graduação em Artes Visuais da ECA-USP. É Doutora em Comunicação e Semiótica e Livre-docente em Artes. Coordena o Grupo de Pesquisa e Extensão Realidades..

review and comparative analysis of works and resources. This article aims to address the production of an experimental motion capture model, based on the theoretical framework developed, detailing the workflow with software such as Clip Studio Paint, Live2D, VTube Studio, and Animaze by Facerig, analyzing it from the perspective of resource accessibility, a fundamental aspect in the dissemination of motion capture to new audiences.

Keywords: Animation, Art, Sequential Images, Motion Capture

1. Introdução

A captura de movimento, segundo Menache (2000 como citado em Gomide, 2013, p.17) “é o processo de gravar um evento de movimento ao vivo, e traduzi-lo em termos matemáticos utilizáveis ao rastrear um número de pontos-chave no espaço através do tempo, e combiná-los para obter uma representação tridimensional única da performance”. Esse processo compreende uma série de técnicas que baseadas em diferentes princípios físicos a partir dos quais os sistemas utilizados podem ser classificados como ópticos, mecânicos e eletromagnéticos.

Ao longo da segunda metade do século XX e das primeiras décadas do século XXI, a captura de movimento gradualmente conquistou seu espaço em campos como a saúde, o esporte e, principalmente, o entretenimento, onde se estabeleceu como um recurso para a produção de animações em filmes e jogos. Porém, a partir 2015, é possível observar que a captura de movimento se tornou um dos aspectos centrais de novos fenômenos como o surgimento de *Youtubers* Virtuais, ou *VTubers*, indivíduos que utilizam “um avatar 2D ou 3D controlado por movimento para criar conteúdo em plataformas de distribuição de vídeo através de gravações ou transmissões ao vivo” (Jiang *et al.*, 2023, p.168).

Essa relação pode ser atribuída em parte a oferta de recursos acessíveis como *softwares* que realizam a captura de movimentos por meio de *webcams* e câmeras de dispositivos móveis, os quais permitem que um público maior tenha contato com essa tecnologia produzindo suas próprias experimentações sem a necessidade de um grande investimento em infraestrutura.

A pesquisa de iniciação científica apresentada no presente artigo buscou investigar tecnologias emergentes de captura de movimento, dentre os quais estão recursos que contribuem para o ampliação do seu acesso. Para isso, foi realizado o levantamento da bibliografia atualmente disponível sobre seus aspectos históricos e técnicos,

assim como o mapeamento de recursos como dispositivos e *softwares* e a realização de um estudo de caso. A fundamentação teórica da pesquisa foi já apresentada no texto “Tecnologias emergentes para produção de imagens sequenciais: Rotoscopia e Captação de Movimento” (SIIMI 2025, prÁxIs), desse modo, o texto atual abordará o processo de criação de um modelo experimental para animação com captura de movimento, produzido a partir dos recursos estudados.

2. Recursos

Após o levantamento da bibliografia e a investigação dos recursos disponíveis para a animação com captura de movimento, constatou-se que a técnica mais viável para ser aplicada a um modelo experimental, em relação a acessibilidade de recursos, é a captura óptica sem marcadores. O baixo custo, os requisitos técnicos e a curva de aprendizado foram os principais aspectos considerados no processo de escolha, uma vez que o processo de animação poderia ser realizado em tempo real por meio dos *softwares* gratuitos *VTube Studio* e *Animaze by Facerig*, utilizando *webcams* e câmeras de celular e oferecendo ao usuário uma interface simples e intuitiva.

Para a elaboração do modelo, foram utilizados os *softwares* *Clip Studio Paint* e *Live2D Cubism*, para a ilustração e o *rigging*³ do modelo respectivamente. Nesse caso os programas utilizados são pagos e a escolha se deu com base na familiaridade prévia da autora com o fluxo de trabalho em ambos, porém o mesmo processo também é possível por meio de suas versões gratuitas e recursos similares de código aberto.

3. Metodologia

3.1. Design

Durante o processo de confecção do modelo, experiências anteriores, como a elaboração de trabalhos da graduação e a participação no 11º Encontro Internacional dos Grupos de Pesquisa: Convergências entre Arte, Ciência, Tecnologias &

³ *Rigging* é processo da animação através do qual diferentes partes de um modelo são dispostas em uma hierarquia, criando uma estrutura a partir da qual o modelo será posicionado e movimentado.

Realidades Mistas, para os quais foram produzidos modelos (figuras 1 e 2) similares ao da pesquisa apresentada, foram um grande diferencial, permitindo o desenvolvimento de um fluxo de trabalho coeso onde cada etapa foi elaborada de forma consciente.



Figuras 1 e 2: Modelos anteriores produzidos por Giovanna⁴

Esses modelos, assim como símbolos e temáticas presentes em trabalhos anteriores de uma das autoras foram as inspirações principais para o design do modelo atual (figuras 3 e 4). Buscou-se incorporar uma maior riqueza de detalhes e elementos animados a um desenho com linhas bem definidas e uma iluminação simplificada, evitando o uso de gradientes, para que a tridimensionalidade do modelo fosse majoritariamente transmitida por seus movimentos, criando a sensação de que o desenho estaria ganhando vida própria. É possível também observar a influência direta do trabalho de *Vtubers* como BigNoseBug e RaeSpark que utilizam elementos similares em seus designs.

⁴ Imagens de autoria própria (2026).



Figura 3 e 4: Design final do modelo e detalhes do rosto.⁵

3.2. Preparação da arte

Durante o processo de ilustração é necessário ter em mente as etapas seguintes da elaboração do modelo, uma vez as partes que o compõem devem ser separadas de modo que possibilitem os movimentos pensados para a animação. Essa separação pode ser feita de dois modos diferentes: durante a própria ilustração, ou seja, construindo o personagem, desde o início, por meio de camadas que formam o desenho completo ao serem sobrepostas, ou começando a partir de uma ilustração pronta, cortando as partes da imagem, dividindo as em camadas e as editando em função dos movimentos que serão realizados. Em ambos os casos, o resultado será um arquivo composto por diferentes camadas que correspondem cada uma a uma parte do modelo (figura 5), sendo exportado na maioria das vezes como um arquivo de documento do *photoshop* (.psd), uma vez que ele permite a preservação e edição dessas camadas e é comumente aceito em programas como o *Live2d*.

⁵ Imagens de autoria Giovanna (2026).



orpioria

Figura 5: Camadas que compõem o modelo.⁶

3.3. Rigging

Durante a etapa de *rigging*, é construída a estrutura que permitirá que as diferentes partes do modelo se movimentem, estabelecendo relações de hierarquia entre elas. Nessas relações as partes são comumente chamadas de “pais” e “filhos”. Quando um elemento pai é alterado são provocadas alterações também nos elementos filhos, ao mesmo tempo, as alterações em elementos filhos, não afetam os elementos pais. Um modo de visualizar esse conceito é pensando na relação entre a cabeça e as partes do rosto. Por exemplo, é possível movimentar partes do rosto sem mover a cabeça, ao piscar ou abrir a boca, porém, sempre que a cabeça se move, o rosto necessariamente se move junto, fazendo com que a cabeça se torne o elemento pai do rosto. É importante destacar também que um elemento pode ser pai e filho ao mesmo tempo. Ao importar o modelo em um *software* de *rigging*, como *Live 2D*, é necessário criar uma malha para cada parte que o compõe. A malha, de modo simplificado, é uma

⁶ Imagem de autoria de Giovanna (2026).

série de pontos que permitirão a deformação da imagem, quanto mais pontos mais detalhada e fluida será a deformação, porém uma malha com muitos pontos não necessariamente é melhor, uma vez que o modo como esses pontos são posicionados afeta diretamente o seu comportamento.

As movimentações e alterações da malha são produzidas por meio de deformadores, ferramentas que, no *software* utilizado, são divididas em dois tipos: deformadores de distorção, que permitem que um elemento seja comprimido e expandido a partir dos pontos da malha; e deformadores de rotação, que rotacionam a malha como um todo a partir de um eixo.

Por fim, é através do uso de parâmetros (figuras 6 e 7) que é definido o escopo dos movimentos que serão realizados pelo modelo. Nessa etapa, valores numéricos são associados a diferentes posições e deformações, estabelecendo *keyforms*, ou formas chave, que representam pontos principais do movimento. O *software* então realiza a interpolação dessas formas, calculando as posições intermediárias entre elas.

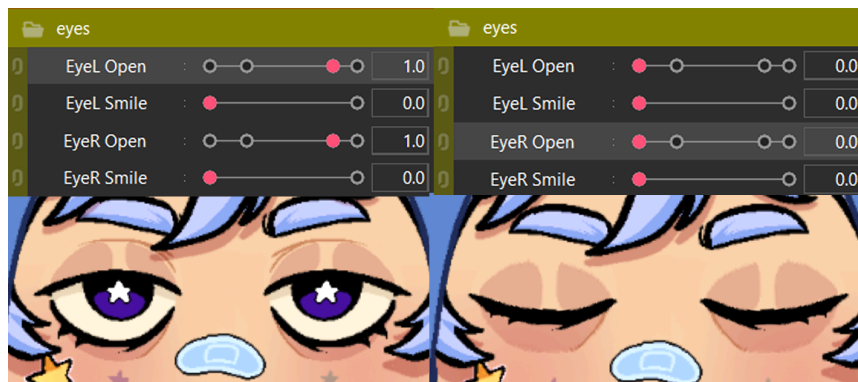


Figura 6: Demonstração do funcionamento dos parâmetros no software Live2D.⁷

Para o modelo da pesquisa apresentada, foram utilizados cerca de 60 parâmetros, responsáveis por definir a movimentação da cabeça, pescoço, elementos do rosto, mechas de cabelo, torso, braços e pernas. A maior parte dos parâmetros é referente à metade superior do modelo, onde estão concentrados movimentos mais complexos como expressões faciais e movimentações secundárias como o balançar do cabelo ao acompanhar a cabeça. Enquanto a metade inferior apresenta um escopo menor de movimentos apenas acompanhando a posição e a rotação do torso.

⁷ Imagem de autoria de Giovanna (2026).

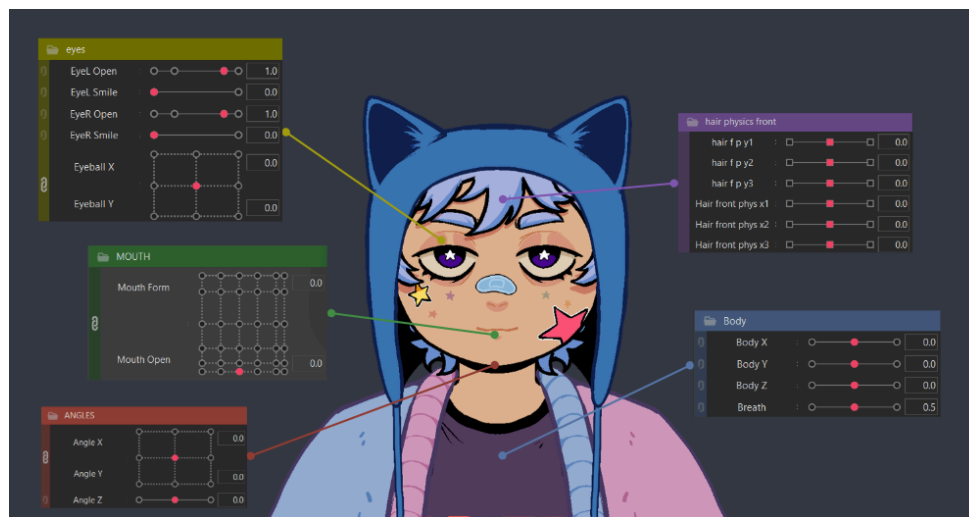


Figura 7: Parâmetros responsáveis pela movimentação da cabeça e das partes do rosto.⁸

3.4. Testes de animação

Com o rigging concluído, os arquivos do modelo e suas configurações podem então ser exportados e implementados em um *software* para dar início ao processo de animação em tempo real. Os programas escolhidos para a realização dos teste de animação foram *VTube Studio* e *Animaze by Facerig*, um dos principais fatores que levaram a essa escolha foi o fato de que ambos possuem suporte para o mesmo formato de arquivo, permitindo a portabilidade do mesmo modelo sem a necessidade de refazer o *rigging* e possibilitando que a comparação entre os resultados e a experiência do usuário sejam afetados apenas pelo *software* utilizado e não por outras etapas do processo.

VTube Studio é um programa disponível para computador e dispositivos móveis criado pelo artista e desenvolvedor independente Vincent Diener, em 2021 (Denchisoft,[s.d]), que permite que usuários configurem e controlem avatares bidimensionais criados no *software Live2d Cubism*. De modo similar, *Animaze by Facerig*, lançado no mesmo ano, oferece essa possibilidade a seus usuários, porém o *software* desenvolvido pela norte-americana *Holotech Studios Inc.* possui como suas

⁸ Imagem de autoria de Giovanna (2026).

principais diferenças uma variedade de modelos prontos customizáveis e o suporte para modelos 3D.

Ambos os programas oferecem ao usuário a possibilidade de controlar avatares por meio da captura óptica de movimentos sem marcadores, utilizando apenas uma webcam ou a câmera de um celular, além de possuírem uma série de integrações com outros sistemas de captura, como o “Mocopi”, da empresa japonesa *Sony*, que utiliza sensores mecânicos inerciais para captar movimentos do corpo todo .

Ao importar o modelo dentro dos *softwares* citados, as configurações podem ser feitas de modo automático, com a possibilidade de serem alteradas pelo usuário a qualquer momento. Os dois possuem interfaces simples contando com uma barra lateral para selecionar o modelo utilizado, editar os elementos presentes na cena em que ele será inserido e configurações de captura, animação, modelo e recursos adicionais, enquanto a maior parte da tela é ocupada pela cena que será exibida na gravação ou transmissão. É notável também que apesar de possuir uma interface bem simples, as configurações podem ser um pouco mais difíceis no *VTube Studio* para usuários que não compreendem o processo de rigging e a estrutura do modelo.

Após a realização dos testes por meio de uma *webcam* com a iluminação de uma luminária de mesa, foi possível observar que a responsividade da captura dos dois é boa nessas condições. Não houve problemas significativos, porém o *VTube Studio* precisou de ajustes em alguns parâmetros para responder corretamente. As diferenças na qualidade de animação, ao executar as mesmas sequências de movimentos e expressões faciais, foram muito sutis, porém é possível observar que o *Animaze* aplica um efeito desfoque de movimento que parece diminuir levemente a resolução do modelo. Um vídeo de demonstração pode ser acessado em: <https://youtu.be/mdS4UrDw6sM?si=XWfqwglVuqpiGIwj> (giovanna florencio, 2026).

Porém uma inconveniência durante os testes foi o limite de 40 minutos por sessão da versão gratuita do *Animaze*, que obriga o usuário a reiniciar o programa durante o uso, lembrando-o que para usar o programa livremente e com recursos completos é necessário uma assinatura. *VTube Studio* também conta com alguns complementos pagos e exige uma licença para alguns casos de uso comercial, porém eles são opcionais para a grande maioria dos usuários.

4. Considerações finais

Portanto, é possível afirmar que ambos são recursos viáveis para a experimentação com a captura de movimento óptica de forma gratuita, ou com um baixo investimento, possibilitando a aproximação de um público muito mais amplo com uma tecnologia ainda pouco discutida dentro e fora do ambiente acadêmico. Ao exigirem uma infraestrutura muito menor do que outros sistemas, recursos como esses permitem que pesquisadores, artistas e educadores ampliem esse diálogo.

O desenvolvimento do modelo de um projeto experimental, por sua vez, possibilitou não apenas a aplicação prática dos conceitos estudados, mas também uma análise aprofundada das etapas envolvidas na criação de personagens para animação com captura de movimento. Buscou-se também evidenciar a importância de compreender o modo como se integram os diferentes conhecimentos técnicos e artísticos envolvidos em cada etapa do modelo, para garantir o seu funcionamento adequado.

5. Referências

DenchiSoft. (s.d.). *Impressum*. <https://denchisoft.com/impressum/>

Gomide, J. V. B. (2013). *Captura digital de movimento no cinema de animação* [Dissertação de mestrado, Universidade Federal de Minas Gerais]. <http://hdl.handle.net/1843/JSSS-94EPAW>

giovanna florencio. (2026, 23 de janeiro) *Demonstração de modelo para captura de movimento* [Vídeo]. YouTube. <https://youtu.be/mdS4UrDw6sM?si=XWfqwglVuqpjGIwj>. Acesso em: 29 de maio de 2026

Jiang, H., Pan, D., & Lu, J. (2023). Better Technology, but Less Realism: The Perplexing Development and Application of VTuber Technology. In *2023 International Conference on Culture-Oriented Science and Technology (CoST)* (pp. 168–173). IEEE. <https://doi.org/10.1109/CoST60524.2023.00042>

Sony. (s.d.). *About mocopi*. <https://www.sony.co.jp/en/Products/mocopi-dev/en/documents/Home/Aboutmocopi.html>