

Estilização de nuvens de pontos em tempo real para interação entre espaços físico e virtual

Real-time point cloud stylization for interaction between physical and virtual spaces

Miguel Venâncio Bertol¹
Ana Clara Lemos de Camargo²
Laurita Ricardo de Salles³
Luciana de Oliveira Berretta⁴
Hugo Alexandre Dantas do Nascimento⁵

Resumo

Este trabalho aborda interações entre espaços físicos e espaços digitais imersivos por meio da construção dinâmica de nuvens de pontos em um ambiente de realidade virtual. Uma nuvem de pontos é uma representação computacional tridimensional de um objeto, de uma pessoa ou de outro elemento, gerada a partir da captura de dados do espaço físico ou produzida sinteticamente. O trabalho é parte de uma pesquisa maior que adota uma câmera RGB-D com sensor iToF para a captura de uma cena física na forma de nuvem de pontos tridimensional e a sua transmissão em tempo real para um ambiente de realidade virtual na web, onde ela é percebida pelos usuários visitantes. Até o momento, o esforço técnico priorizou a fidelidade dessa representação. O presente estudo propõe uma direção complementar: estilizar a nuvem de pontos usando transformações computacionais que ampliem as possibilidades de experiência no ambiente virtual. São investigadas transformações sobre a geometria, os atributos visuais e a dinâmica temporal da nuvem de pontos. São abordadas também possibilidades de correspondência entre estilizações visuais e manipulações sonoras. A base

¹ Miguel Venâncio Bertol, estudante de graduação do Bacharelado em Ciência da Computação do INF UFG, desde de 2024.

² Ana Clara Lemos de Camargo, estudante de graduação do Bacharelado de Engenharia de Computação na EMC UFG, desde de 2023.

³ Laurita Ricardo de Salles, artista pesquisadora na área de Arte e tecnologia e Artes Visuais. Professora Permanente do Programa de Pós-Graduação em Artes, Culturas e Tecnologias/Media Lab - UFG; também professora colaboradora voluntária na Universidade Federal do Rio Grande do Norte (UFRN) em pesquisa e extensão. Foi docente DE no Departamento de Artes da UFRN de 2010/2022. Pós-Doutora em Mídias Interativas pelo Media Lab / UFG.

⁴ Luciana de Oliveira Berretta, professora Associada do Instituto de Informática da Universidade Federal de Goiás e Professora Permanente do Programas de Pós-Graduação em Ciência da Computação da UFG, com formação em Ciência da Computação. Pesquisa sobre Realidade Virtual, Acessibilidade e Interação Humano Computador.

⁵ Hugo Alexandre Dantas do Nascimento, professor Titular do Instituto de Informática da Universidade Federal de Goiás (UFG) e Professor Permanente dos Programas de Pós-Graduação em Ciência da Computação e em Artes, Culturas e Tecnologias/Media Lab - UFG, com formação em Ciência da Computação. Possui forte atuação interdisciplinar e investiga temáticas que abordam Visualização de Informações, Otimização Combinatória, Interação Humano-Computador, Ciência de Dados, Engenharia de Software e Arte Tecnologia.

técnica do trabalho é um fluxo de captura, transmissão, transformação e renderização de dados no ambiente virtual. São discutidos cenários de aplicação para a proposta, os quais incluem arte tecnológica, comunicação entre espaços mediada por metáforas visuais e representação estilizada de gêmeos digitais. Ao reposicionar o foco da apresentação da nuvem de pontos de uma réplica real do objeto para um meio de expressão artística, o trabalho contribui para ampliar o repertório de interações possíveis entre o físico e o virtual.

Palavras-chave: nuvem de pontos, estilização, interações entre espaços, realidade virtual, performances tecnológicas.

Abstract

This work addresses the interactions between physical spaces and immersive digital spaces through the dynamic construction of point clouds in a virtual reality environment. A point cloud is a three-dimensional computational representation of an object, a person, or any other element, generated from data capture of the physical space or synthetically produced. This study is part of a broader research project that utilizes an RGB-D camera with an iToF sensor to capture a physical scene as a 3D point cloud and transmit it in real time to a web-based virtual reality environment, where it is perceived by visiting users. Up to this point, the technical effort has prioritized the fidelity of this representation. The present study proposes a complementary direction: stylizing the point cloud using computational transformations that expand the possibilities of experience within the virtual environment. It investigates transformations applied to the geometry, visual attributes, and temporal dynamics of the point cloud. Furthermore, it explores potential correspondences between visual stylizations and sound manipulations. The technical foundation of this work consists of a pipeline for data capture, transmission, transformation, and rendering within the virtual environment. Application scenarios for the proposal are discussed, including technological art, communication between spaces mediated by visual metaphors, and stylized representation of digital twins. By shifting the focus of the point cloud presentation from a realistic replica of the object to a medium for artistic expression, this work contributes to expanding the repertoire of possible interactions between the physical and the virtual.

Keywords: *point cloud, stylization, interactions between spaces, virtual reality, technological performances.*

Introdução

Este trabalho situa-se em uma pesquisa sobre interações entre espaços físicos e espaços digitais imersivos, entendidas como performances tecnológicas, isto é, comportamentos, ações corporais e relações socioculturais mediadas pela tecnologia (Camargo et al., 2011; Chauí, 2000). A investigação é conduzida no Media Lab / UFG, vinculada ao Programa de Pós-Graduação em Ciência da Computação, e foi desenvolvida em dois projetos sequenciais de mesmo título, financiados pela Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG) e, na sequência, pelo Centro de Competências em Tecnologias Imersivas (AKCIT/EMBRAPII).

Um dos ambientes desenvolvidos nessa pesquisa é o cenário de sobreposição de espaços, que busca uma conexão bidirecional e simultânea entre uma sala física e sua contraparte virtual. Na direção físico para virtual, câmeras RGB-D capturam continuamente a cena (pessoas, objetos e o próprio espaço) na forma de nuvens de pontos coloridas, transmitidas em tempo real a um ambiente de realidade virtual na *web*, onde usuários visitantes percebem quem está na sala física e como essas pessoas se movem. Trata-se de uma instância de realidade mista, na qual elementos do mundo real e da virtualidade coexistem e se influenciam mutuamente (Rokhsaritalemi et al., 2020), aproximando-se de sistemas de telepresença volumétrica baseados em captura RGB-D (Orts-Escolano et al., 2016).

Até o momento, o esforço técnico desse cenário priorizou a fidelidade da representação: quanto mais a nuvem de pontos se aproximasse de uma réplica do espaço físico, melhor. O presente trabalho propõe uma direção complementar. Em vez de tratar a nuvem de pontos apenas como réplica, propomos tratá-la como material maleável, sujeito a transformações computacionais intencionais, abordagem que dialoga com a renderização não fotorrealística (NPR) e a estilização de nuvens de pontos (Wegen et al., 2024). A renderização não fotorrealística investiga há décadas como produzir imagens que se comunicam por meios distintos da reprodução fiel, e sua extensão à nuvem de pontos abre espaço para que a própria captura tridimensional, e não apenas a imagem final, seja objeto de manipulação estética. A contribuição do artigo é dupla: (i) um enquadramento conceitual que reposiciona a nuvem de pontos de réplica para meio de expressão; e (ii) uma camada de estilização extensível, integrada ao pipeline em tempo real, demonstrada por um primeiro efeito implementado, o *underwater*.

O cenário de sobreposição de espaços e o *pipeline* de nuvens de pontos

A proposta deste trabalho não cria um *pipeline* novo, mas se apoia em uma infraestrutura já implementada, voltada à direção físico para virtual do cenário de sobreposição de espaços. Compreender essa base é necessário porque a estilização

atua dentro dela, como uma operação inserida no fluxo existente. A Figura 1 apresenta o pipeline em quatro estágios.

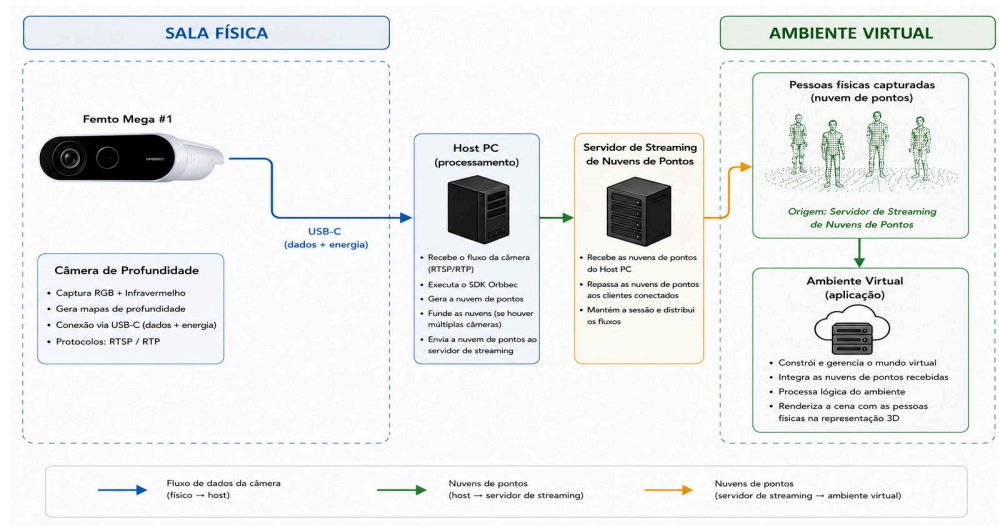


Figura 1, diagrama do pipeline de nuvens de pontos

O *pipeline* tem quatro estágios encadeados. No primeiro, uma câmera Orbbec Femto Mega, do tipo RGB-D com sensor iToF (indirect Time-of-Flight), captura a cena, e o Orbbec SDK v2 gera a nuvem de pontos a partir de dados de profundidade e cor. No segundo, um host PC executa uma aplicação de linha de comando (CLI) em C++ responsável pela geração, pelo *downsample* e pela codificação dos pontos, que são então enviados por TCP. No terceiro, um servidor de *streaming* em Node.js com Socket.IO valida o *token* de autenticação, interpreta o protocolo e redistribui os quadros por salas (*rooms*), uma para cada nuvem. No quarto, o *frontend* em React, A-Frame e Three.js descobre as nuvens de pontos próximas por busca espacial, cria entidades correspondentes no mundo virtual e as renderiza.

O contrato de dados que une essas camadas é simples e fixo: cada ponto trafega com a sua posição tridimensional e a sua cor, sem compressão no caminho ativo. No *frontend*, todos os pontos de uma nuvem compartilham um único *buffer* enviado à GPU, o que reduz cópias e mantém baixo o custo de renderização. A estilização proposta neste trabalho atua nessa última camada, no *frontend*, sem alterar a CLI de captura nem o servidor de *streaming*.

Toda essa infraestrutura foi projetada com um objetivo implícito: reproduzir o espaço físico com a maior fidelidade possível. As escolhas de *downsample*, de transmissão e de renderização convergem para que a nuvem exibida no ambiente virtual seja o mais próxima possível de uma réplica da cena capturada. É justamente essa premissa que a próxima seção propõe revisar.

Da fidelidade à expressão: a estilização como camada de transformação

Propomos deslocar o foco da fidelidade para a expressão. A meta de réplica fiel é legítima, mas corresponde a apenas uma das relações possíveis entre o dado capturado e sua representação. A nuvem de pontos deixa de ser tratada como réplica para ser tratada como material maleável, sobre o qual se pode operar com intenção estética. A fidelidade passa a ser um ponto de partida, e não o destino obrigatório da representação.

Para viabilizar esse deslocamento sem reescrever o *pipeline*, introduzimos a estilização como uma camada de transformação posicionada entre a transmissão dos dados e a renderização final. Cada quadro recebido pode ser transformado antes de ser exibido, atuando sobre três eixos complementares: a geometria, que diz respeito à forma e à estrutura espacial dos pontos; os atributos visuais, que abrangem cor, opacidade, tamanho e intensidade; e a dinâmica temporal, responsável pela coerência e pela suavização do efeito entre quadros sucessivos. No eixo da geometria cabem operações como deslocamento, ondulação e fragmentação espacial. No eixo dos atributos visuais cabem a alteração cromática, a opacidade variável e o tamanho dinâmico dos pontos. No eixo da dinâmica temporal cabem a persistência de quadros anteriores, o eco temporal e a suavização de transições. Esses três eixos compõem uma taxonomia da estilização: oferecem um vocabulário para descrever qualquer efeito, e um mesmo efeito pode combinar mais de um deles (Wegen et al., 2024). A Figura 3 ilustra essa camada inserida no fluxo, e a Tabela 1 detalha os eixos com definições e exemplos de operações.

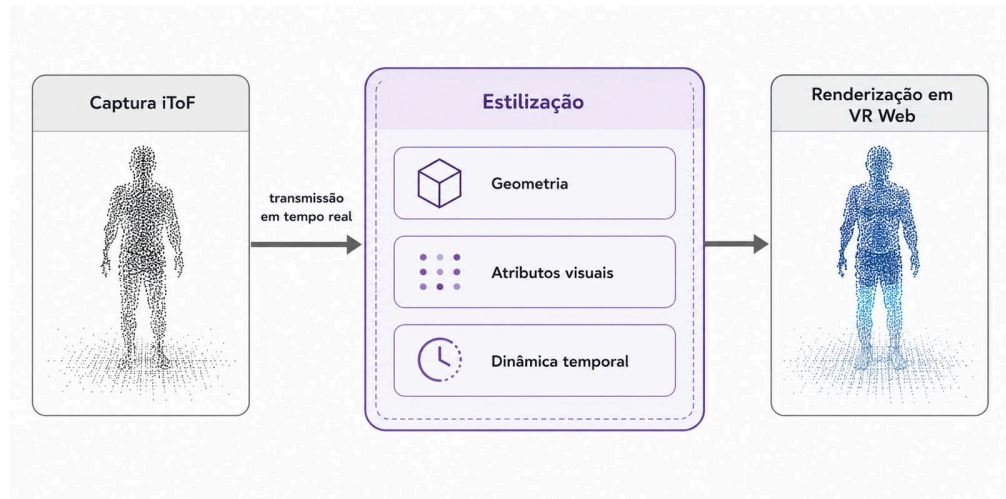


Figura 2, estilização como camada de transformação

Eixo	O que transforma	Exemplos de operação	<i>Underwater</i> atua?
Geometria	Forma e estrutura espacial dos pontos	Deslocamentos, ruído posicional, expansão e contração	SIM
Atributos Visuais	Cor, opacidade, tamanho, intensidade	Recolorização, mapas de profundidade/cor, alpha	NÃO
Dinâmica Temporal	Coerência e suavização entre quadros	Envelopes, easing, fragmentação temporal	SIM

Tabela 1, eixos de transformação da estilização, com definições, exemplos de operação e indicação de quais eixos o efeito *underwater* aciona.

Estudo de caso: o efeito *underwater*

Como primeira validação da camada de estilização, implementamos o efeito *underwater*. Quando uma pessoa aparece na nuvem de pontos, os pontos associados ao seu corpo passam a oscilar de forma coerente, como um corpo suspenso em água,

combinando uma ondulação tridimensional e uma leve expansão a partir do centro do corpo. A Figura 4 compara a cena sem estilização e com o efeito aplicado.



Figura 3, antes e depois do efeito *underwater*⁶

Na taxonomia da seção anterior, o *underwater* atua sobre dois eixos. Ele opera na geometria, ao deslocar espacialmente os pontos, e na dinâmica temporal, que garante a coerência da oscilação entre quadros. Os atributos visuais permanecem intactos: cor, opacidade e tamanho dos pontos não são alterados. A Figura 5 destaca essa posição do efeito dentro da taxonomia.

⁶ Demonstração em vídeo do efeito *underwater* sobre dados reais capturados em tempo real: <https://youtu.be/COfUYcjgNR0>

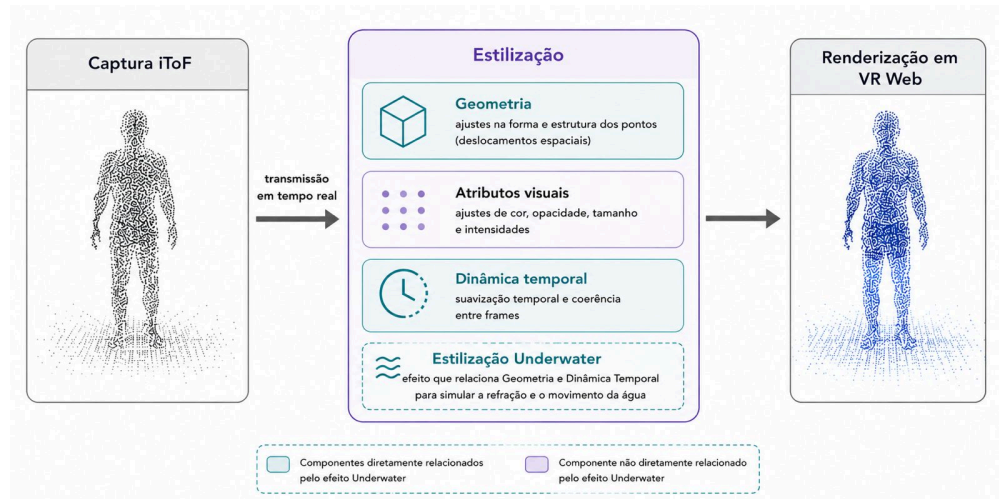


Figura 4, *underwater* como instância na taxonomia

O algoritmo recebe como entrada as posições XYZ de cada ponto e produz posições deslocadas. Antes de qualquer deformação, o efeito aprende um modelo de cena estável: a profundidade do fundo é estimada por índice de ponto ao longo do tempo, por meio de uma média móvel exponencialmente ponderada, com uma janela de *bootstrap* de cerca de 120 quadros, o equivalente a aproximadamente quatro segundos a 30 fps. Com o fundo conhecido, cada quadro detecta as regiões a transformar comparando a profundidade atual com a do fundo, o que produz uma máscara dos pontos pertencentes à pessoa. A Figura 6 sintetiza a construção do efeito.

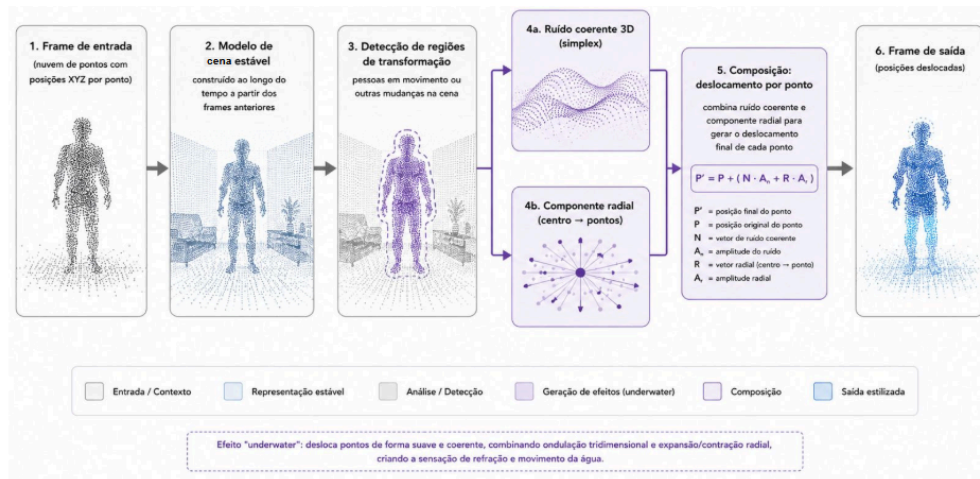


Figura 5, construção do efeito *underwater* em 6 passos

O deslocamento de cada ponto da pessoa combina duas contribuições: um ruído coerente tridimensional, gerado por ruído simplex (Perlin, 1985; Gustavson, 2005), que produz a ondulação, e um componente radial que empurra os pontos para fora do centro do corpo. A composição segue a expressão:

$$P' = P + (N \cdot A_n + R \cdot A_r)$$

em que N é o vetor de ruído, R é o vetor radial e A_n e A_r são as amplitudes respectivas. Um envelope ADSR (Roads, 1996), tomado da síntese sonora, modula a intensidade global do efeito na entrada e na saída da pessoa, evitando transições bruscas quando alguém surge na cena ou a abandona. Quando o envelope retorna a zero, os pontos voltam à posição real recebida. Os parâmetros perceptuais que governam a aparência do efeito, a amplitude da ondulação, o empurrão radial e os tempos do envelope, estão sumarizados na Tabela 2. As taxas de aprendizado do fundo, os limiares de detecção e os tempos de recuperação controlam a robustez da detecção, e não a aparência; seus valores de calibração acompanham o código.

Parâmetro perceptual	Valor	O que controla na aparência
Amplitude da ondulação	~5 cm	Quão longe cada ponto oscila em torno da sua posição real
Empurrão radial	~10 cm	Quanto o corpo se expande a partir do próprio centro
Ataque / decaimento	0,2 s / 0,3 s	Tempo até o efeito atingir intensidade plena quando a pessoa entra
Sustentação	70%	Intensidade mantida enquanto a pessoa permanece na cena
Relaxamento	1,5 s	Tempo de retorno à

		forma real quando a pessoa sai
--	--	-----------------------------------

Tabela 2, principais parâmetros perceptuais do efeito underwater e o que cada um controla na aparência.

O *underwater* é uma instância demonstrativa, e não o objetivo final do trabalho. Seu valor está em mostrar que a camada de estilização funciona de ponta a ponta sobre dados reais capturados em tempo real. A contribuição que importa é a camada que torna esse efeito, e outros, possíveis.

Implementação

A estilização roda em produção, em tempo real, dentro do navegador, mantendo a frequência da câmera de 30 fps. O estilo aplicado a cada nuvem de pontos é selecionado por meio de metadados do *backend*, de modo que cada câmera ou entidade presente no espaço virtual pode receber um efeito diferente, configurado pela API sem necessidade de alterar o código do cliente. Como a seleção viaja junto dos metadados da nuvem, um mesmo ambiente pode exibir simultaneamente cenas estilizadas de formas distintas.

O estilizador é aplicado no cliente, em CPU, no caminho quente de cada quadro, logo após a cópia dos dados recebidos e antes do envio à GPU. Cada entidade do mundo virtual recebe a sua própria instância do estilizador, de forma que duas nuvens com o mesmo efeito não compartilham estado, como buffers de fundo, máscaras ou envelope. A arquitetura é extensível: cada efeito é uma unidade isolada, registrada em um catálogo de estilos e dotada de estado próprio. A cada quadro, o efeito recebe os pontos da nuvem e marcações de tempo e pode alterar suas posições e cores; reinicia o estado acumulado quando a identidade da transmissão muda; e expõe parâmetros opcionais de ajuste. Um efeito pode, assim, ser criado sem reescrever a CLI de captura nem o servidor de *streaming*, o que torna essa a camada mais barata para experimentação artística.

Essa escolha de implementação traz uma limitação que convém registrar, retomada nas Considerações finais e próximos passos. O efeito ainda se apoia no material de pontos padrão da biblioteca de renderização, que não permite definir tamanho ou opacidade por ponto via shader, o que restringe a manipulação dos atributos visuais à granularidade do material inteiro. Além disso, o estilizador executa no *main thread* do navegador, competindo com o laço de renderização pelo mesmo tempo de processamento.

Cenários de aplicação

A camada de estilização habilita um conjunto de cenários de aplicação:

- Arte tecnológica: o corpo capturado torna-se material expressivo, e a estilização passa a ser, ela própria, o gesto artístico apresentado ao público no ambiente virtual.
- Comunicação entre espaços por metáforas visuais: o estilo aplicado carrega significado próprio. Uma sala pode parecer respirar, congelar ou dissolver conforme o estado que se queira comunicar de um espaço a outro, usando a aparência da nuvem como linguagem.
- Gêmeos digitais estilizados: a visualização de espaços ganha uma camada estética que enriquece a interação, em lugar de perseguir somente a réplica exata da cena (Grieves, 2014).

Atravessando esses cenários está a possibilidade de uma correspondência sonora, que associa manipulações de áudio às estilizações visuais e explora relações sinestésicas entre o que se vê e o que se ouve (Whitelaw, 2008). O vocabulário do envelope ADSR, já presente no controle temporal do efeito *underwater*, sugere uma ponte natural entre a dimensão visual e a sonora.

Considerações finais e próximos passos

O trabalho reposiciona a nuvem de pontos de réplica para meio de expressão, integrando a um pipeline de tempo real uma camada de estilização extensível e

demonstrável. Esse resultado vem acompanhado de limitações que reconhecemos abertamente. A estilização roda em CPU, no main thread, e ainda não conta com shader customizado, o que, somado ao material de pontos padrão, impede o controle de tamanho e opacidade por ponto. Até o momento, há apenas um efeito implementado além do comportamento neutro, e o áudio físico, embora previsto no protocolo de transporte, ainda não foi integrado nas pontas de captura e de consumo.

Os próximos passos são:

- novas estilizações, incluindo transformações de cor, fragmentação temporal e efeitos aplicados a partes específicas do corpo;
- maior controle sobre as regiões de transformação, permitindo mais intenção artística sobre o que é afetado e por quanto tempo;
- correspondência sonora, com manipulação de áudio sincronizada à estilização visual;
- estudos de recepção sobre como o público responde a diferentes estilizações.

Ao tratar a nuvem de pontos como material maleável, em lugar de mera cópia do espaço físico, o trabalho amplia o repertório de interações possíveis entre o físico e o virtual e oferece uma base técnica e conceitual para que a estilização se torne parte da linguagem do cenário de sobreposição de espaços.

Referências

Camargo, R., Capel, H., & Reinato, E. (2011). *Performances culturais*. Hucitec.

Chauí, M. (2000). *Convite à filosofia*. Ática.

Grieves, M. (2014). *Digital twin: Manufacturing excellence through virtual factory replication* [White paper]. Florida Institute of Technology.

Gustavson, S. (2005). *Simplex noise demystified*. Linköping University.

Orts-Escolano, S., Rhemann, C., Fanello, S., Chang, W., Kowdle, A., Degtyarev, Y., Kim, D., Davidson, P. L., Khamis, S., Dou, M., Tankovich, V., Loop, C., Cai, Q., Chou, P. A., Mennicken, S., Valentin, J., Pradeep, V., Wang, G., Kang, S. B., &

Izadi, S. (2016). Holoportation: Virtual 3D teleportation in real-time. In *Proceedings of the 29th Annual ACM Symposium on User Interface Software and Technology* (pp. 741–754). ACM. <https://doi.org/10.1145/2984511.2984517>

Perlin, K. (1985). An image synthesizer. *ACM SIGGRAPH Computer Graphics*, 19(3), 287–296. <https://doi.org/10.1145/325165.325247>

Roads, C. (1996). *The computer music tutorial*. MIT Press.

Rokhsaritalemi, S., Sadeghi-Niaraki, A., & Choi, S.-M. (2020). A review on mixed reality: Current trends, challenges and prospects. *Applied Sciences*, 10(2), 636.

Wegen, O., Scheibel, W., Trapp, M., Richter, R., & Döllner, J. (2024). A survey on non-photorealistic rendering approaches for point cloud visualization. *IEEE Transactions on Visualization and Computer Graphics*, 31(9), 4511–4533. <https://doi.org/10.1109/TVCG.2024.3402610>

Whitelaw, M. (2008). Synesthesia and cross-modality in contemporary audiovisuals. *The Senses and Society*, 3(3), 259–276. <https://doi.org/10.2752/174589308X331314>