Bom dia a todos, sou o João Gaspar/Guilherme Santos e o meu colega é o João Gaspar/Guilherme Santos de LECI. Viemos partilhar convosco as descobertas que fizemos após a leitura e análise do paper "Animation Fidelity in Self-Avatars: Impact on User Performance and Sense of Agency." escrito por Haoran Yun, Jose Luis Ponton, Carlos Andujar e Nuria Pelechano. Este trabalho foi apresentado na Conferência da IEEE (Instituto de Engenheiros Eletricistas e Eletrônicos) sobre Realidade Virtual e Interfaces de Utilizador 3D em Shanghai, em 2023, ocupando as páginas 286 a 296.

Optámos por este paper uma vez que o resumo se destacou de forma notável, capturando completamente a nossa atenção e despertando o nosso interesse nas questões abordadas. A escolha deste paper torna-se ainda mais relevante devido à sua abordagem centrada na fidelidade de animação em self-avatars e o consequente impacto no desempenho do utilizador e na sensação de agência.

Estamos particularmente entusiasmados com este tema, uma vez que reconhecemos a crescente importância da Realidade Virtual (RV) nos últimos anos. A recente introdução do Apple Vision Pro, e não menos importante, o Meta Quest 3, são dois exemplos significativos do rápido desenvolvimento nesta área, aumentando ainda mais o nosso fascínio pelo tema.

Deste modo, a escolha deste paper proporciona não apenas uma oportunidade para explorar aspetos específicos da animação em self-avatars, mas também nos permite contextualizar esses conceitos com as inovações tecnológicas mais recentes.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Introduction

Este paper discute a importância da fidelidade na animação de avatares em realidade virtual (RV), especificamente o avatar pessoal (self-avatar). Destacam-se os desafios causados pela limitação de rastreadores em dispositivos de consumo, resultando em representações simplificadas do corpo virtual, muitas vezes sem pernas. A qualidade do movimento do self-avatar é crucial para a sensação de presença na RV, especialmente em interações que requerem precisão de pose e posicionamento de extremidades.

O estudo compara diferentes métodos de animação, incluindo soluções baseadas em Cinemática Inversa (IK) com poucos dados de rastreamento, como UnityIK e FinalIK, e um sistema de captura de movimento baseado em um grande número de sensores inerciais (Xsens Awinda). Os resultados indicam que a fidelidade na animação afeta a Sensação de Incorporação (Sense of Embodiment) e o desempenho do utilizador. Métodos simples de IK podem diminuir a Sensação de Incorporação em comparação com soluções de alta qualidade, mas em interações ambientais, a latência e o erro mínimo na posição dos extremos podem ser mais críticos do que poses de alta qualidade.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

3. Animation Fidelity Study

Este estudo visa avaliar a importância da fidelidade na animação na performance dos utilizadores e na Sensação de Incorporação (SoE) ao realizar tarefas que exigem posicionamento cuidadoso e/ou poses precisas em ambientes virtuais. Três condições de fidelidade de animação foram estudadas: Unity Inverse Kinematics (UIK), FinalIK (FIK) e captura de movimento com Xsens (MoCap).

As condições diferem na qualidade da animação em termos de posicionamento de extremidades (mais preciso em UIK e FIK), ângulos de pose (mais preciso em MoCap devido ao maior número de sensores) e latência (maior para MoCap). Os participantes realizaram três tarefas em cada condição: "Step-over-spikes", "Pick-and-place" e "Copy-pose". Cada tarefa enfoca diferentes interações entre o corpo virtual e o ambiente virtual.

O estudo envolveu a comparação de três métodos de animação: UIK usando o solver IK integrado no Unity, FIK usando o solver VRIK da RootMotion, e MoCap usando o sistema Xsens Awinda com 17 sensores inerciais. Medidas de desempenho incluíram tempo de conclusão, colisões não intencionais e questionários de Sensação de Incorporação. Os resultados indicaram que a fidelidade da animação afeta a Sensação de Incorporação e o desempenho do utilizador, variando de acordo com a tarefa realizada.

Os participantes foram submetidos a três tarefas: "Step-over-spikes" para interação com o corpo inferior, "Pick-and-place" para interação com o corpo superior, e "Copy-pose" para avaliar poses estáticas. Métricas como volume e duração das colisões, tempo de conclusão e precisão na cópia de poses foram medidas.

O estudo também detalhou as condições experimentais, os métodos de animação utilizados em cada condição (UIK, FIK, MoCap), e os equipamentos usados para rastreamento (HMD, controladores, trackers e sensores inerciais). A latência dos sistemas de rastreamento foi avaliada, sendo menor para UIK e FIK em comparação com MoCap.

As hipóteses formuladas incluíram que uma melhor fidelidade de animação resultaria em melhor desempenho em tarefas que exigem interação precisa com o ambiente, que a fidelidade de animação afetaria o desempenho na tarefa de copiar poses, e que afetaria a Sensação de Incorporação. Os resultados observados confirmaram a influência da fidelidade na animação nos aspectos mencionados.

O estudo representa uma contribuição ao comparar um sistema de captura de movimento baseado em IMU com soluções de IK para animação de self-avatars em RV, e explora a relação entre fidelidade na animação, desempenho do utilizador e Sensação de Incorporação.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Nesta secção, são apresentados os resultados do experimento, focando na performance do utilizador em tarefas de interação, tarefas relacionadas a poses e na Sensação de Incorporação (SoE).

4.1 Performance do Utilizador em Tarefas de Interação

Os resultados das tarefas que envolveram interação direta com o Ambiente Virtual (VE) foram analisados. Para as tarefas "Step-over-spikes" e "Pick-and-place", foram utilizadas medidas como volume e duração de colisões, frequência de colisões e tempo de conclusão. Testes de Shapiro-Wilk indicaram desvios significativos da normalidade para algumas métricas, levando à utilização de testes não paramétricos.

Step-over-spikes Task: Houve diferenças significativas entre as condições de fidelidade na animação para volume por colisão e frequência de colisões. MoCap apresentou valores significativamente mais altos do que FIK para todas as métricas, exceto duração por colisão, onde não houve diferença significativa. UIK teve uma frequência de colisões significativamente mais alta do que FIK.

Pick-and-place Task: Houve efeitos significativos da fidelidade na animação na duração por colisão e na frequência de colisões. UIK teve uma frequência de colisões significativamente mais alta do que FIK e MoCap, além de um tempo de conclusão mais longo do que FIK.

Esses resultados confirmam a hipótese [H1], indicando que a fidelidade na animação afeta o desempenho do utilizador em tarefas que exigem interação precisa com o ambiente.

4.2 Performance do Utilizador em Tarefas Relacionadas a Poses

Na tarefa "Copy-pose", que avalia a precisão na replicação de poses, foram analisadas métricas como Jaccard Distance (JD), Mean per Segment Angle Error (MPSAE) e Mean per Part Angle Error (MPPAE).

Houve diferenças significativas entre as condições de fidelidade na animação para JD e MPSAE, com UIK apresentando valores significativamente mais altos do que FIK e MoCap em ambas as métricas.

Na análise de MPPAE por parte do corpo, UIK teve erros significativamente mais altos para os braços, pernas e coluna, em comparação com FIK e MoCap.

Esses resultados validam a hipótese [H2], indicando que os erros de pose foram significativamente menores ao usar MoCap em comparação com soluções de IK.

4.3 Sensação de Incorporação (SoE)

Os resultados da SoE foram analisados considerando a pontuação geral e as subcomponentes de agência, propriedade e mudança.

Houve um efeito significativo da fidelidade na animação na pontuação geral da SoE. UIK teve uma pontuação pior em comparação com FIK e MoCap.

Na análise das subcomponentes, a agência foi significativamente menor em UIK em comparação com FIK e MoCap. A propriedade foi significativamente menor em UIK em comparação com FIK.

Esses resultados confirmam a hipótese [H3], indicando que a fidelidade na animação afeta a SoE.

Resumo dos Resultados

A fidelidade na animação afetou o desempenho do utilizador em tarefas de interação, com MoCap superando UIK e FIK em várias métricas.

Erros de pose foram significativamente menores com MoCap em comparação com soluções de IK.

A fidelidade na animação influenciou a SoE, com UIK apresentando pontuações inferiores em agência e propriedade.

Os resultados detalhados e valores estatísticos estão disponíveis na Tabela 2. Os gráficos de violino (Figuras 4, 5 e 6) proporcionam uma visualização adicional dos resultados.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

5. Discussão

5.1 Precisão da Pose do Corpo vs. Posicionamento do Efetuador Final

Conforme esperado, o fato de um fato de captura de movimento (MoCap) ser capaz de capturar a maioria dos movimentos humanos e representar poses com precisão é evidente na tarefa de "Copy-pose". MoCap superou significativamente as soluções de Inverse Kinematics (IK) para todas as métricas, alinhando-se melhor com as poses dadas. No entanto, a alta qualidade das poses não se traduziu em melhor desempenho em tarefas não diretamente relacionadas à pose, mas que exigem interações diretas com o Ambiente Virtual (VE). A deriva posicional de sistemas inerciais pode levar os efetuadores a se afastarem de suas posições reais, resultando em inconsistências.

5.2 Animação do Membro Superior para Interações Precisas com o Ambiente

Na tarefa "Pick-and-place", a solução de IK para o membro superior (UIK) teve um desempenho significativamente pior em termos de frequência de colisões em comparação com MoCap e FIK. No entanto, MoCap e FIK apresentaram desempenho semelhante. Resultados consistentes foram observados na tarefa "Copy-pose". A posição incorreta do cotovelo em UIK complicou a tarefa, levando a mais colisões e maior tempo de conclusão. Mesmo que MoCap forneça uma posição mais precisa do cotovelo, as posições imprecisas dos efetuadores finais podem resultar em mais colisões com obstáculos.

5.3 Diferenças de Desempenho entre Braços e Pernas

A análise dos erros de pose na tarefa "Copy-pose" indica que as poses dos braços foram menos precisas do que as poses das pernas. Isso pode ser explicado pela maior amplitude de movimento possível na parte superior do corpo. A pontuação média relatada pelos participantes também reflete uma percepção de menor precisão nos movimentos dos braços em comparação com as pernas. Portanto, sugere-se focar na animação de alta qualidade para os membros superiores, enquanto uma animação de qualidade inferior pode ser suficiente para as pernas.

5.4 Alta Sensação de Agência com um Pequeno Conjunto de Dispositivos de Rastreamento

Embora os dados do questionário não tenham mostrado diferenças estatisticamente significativas entre FIK e MoCap, MoCap obteve melhores resultados em métricas objetivas na tarefa "Copy-pose". Isso sugere que a Sensação de Agência (SoA) não está apenas relacionada à pose, mas também à interação com o VE. A capacidade de sentir o mesmo nível de controle sobre avatares animados por um fato de captura de movimento de alta qualidade ou por um pequeno conjunto de dispositivos de rastreamento (HMD, dois controladores e três trackers) e uma solução de IK de alta qualidade é destacada. Este achado está em linha com pesquisas anteriores que indicam que um número reduzido de trackers pode proporcionar uma ilusão de plausibilidade semelhante a sistemas mais complexos.

5.4 Considerações Finais

A precisão da pose do corpo (MoCap) supera as soluções de IK em tarefas de replicação de pose.

A alta qualidade das poses não garante melhor desempenho em tarefas que exigem interações diretas com o ambiente virtual.

A animação de alta qualidade para membros superiores é crucial para tarefas precisas de interação, enquanto para membros inferiores, uma animação de qualidade inferior pode ser suficiente.

Uma alta Sensação de Agência pode ser alcançada com um conjunto reduzido de dispositivos de rastreamento, desde que a animação seja de alta qualidade.

Os resultados e considerações apresentadas são fundamentadas nos dados fornecidos no texto. Para detalhes específicos, consulte as seções relevantes e as figuras mencionadas.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

6. Conclusões e Trabalho Futuro

Realizamos um estudo do usuário para examinar o impacto da fidelidade de animação do avatar no desempenho do usuário e na Sensação de Agência (SoA). Nossos resultados sugerem que o sistema de captura de movimento baseado em IMU teve um desempenho superior às soluções de Cinemática Inversa (IK) em aplicações que exigem precisão na pose. No entanto, soluções de IK superaram os sistemas de captura de movimento baseados em IMU quando há interações diretas com o Ambiente Virtual (VE). Nestes casos, a colocação precisa do efetuador final e baixa latência podem ser mais críticas do que uma correspondência exata de pose devido à propriocepção. O estudo também sugere que uma solução de IK de alta qualidade com entrada esparsa (6 trackers) pode atingir níveis semelhantes de SoA em comparação com uma captura de movimento baseada em IMU com entrada densa (17 trackers). Acreditamos que esses resultados oferecem insights sobre como a fidelidade da animação afeta o desempenho do usuário e a percepção, indicando direções futuras para aprimorar a fidelidade de animação do auto-avatar em Realidade Virtual (VR). Nosso trabalho também destaca as limitações da tecnologia atual para alcançar uma animação correta do auto-avatar (como latência, imprecisão de efetuadores finais e pose corporal), motivando assim pesquisas futuras para superar esses problemas.

Uma limitação do nosso experimento é que o avatar robótico não coincidiu precisamente com a forma do participante. A falta de correspondência precisa nas extremidades do avatar pode ter impactado na reprodução de autorreferências sugeridas por alguns alvos de "copy-pose". Estudos adicionais são necessários para investigar o papel da fidelidade de animação e autorreferência quando o avatar coincide mais precisamente com a forma do usuário.

Para futuras pesquisas, gostaríamos de explorar se os participantes poderiam ter um desempenho melhor usando um sistema de captura de movimento óptico, fornecendo tanto pose quanto posição global precisas. Essa nova condição permitirá separar o problema de deriva posicional da precisão da pose corporal, possibilitando um estudo mais aprofundado dos resultados perceptuais. Acreditamos que estudos futuros que integrem rastreamento de mãos, como o RotoWrist [30], ou métodos orientados por dados para animação do auto-avatar, seriam valiosos para proporcionar uma compreensão mais aprofundada de como a fidelidade da animação impacta a Sensação de Presença (SoE) e o desempenho do usuário em VR.