

Geração de expressões faciais realistas para animação de avatares hiper-realistas

Júlia Galvão Matos, e Thiago Luange Gomes e Michel Melo da Silva

ODS 10

Pesquisa

Introdução

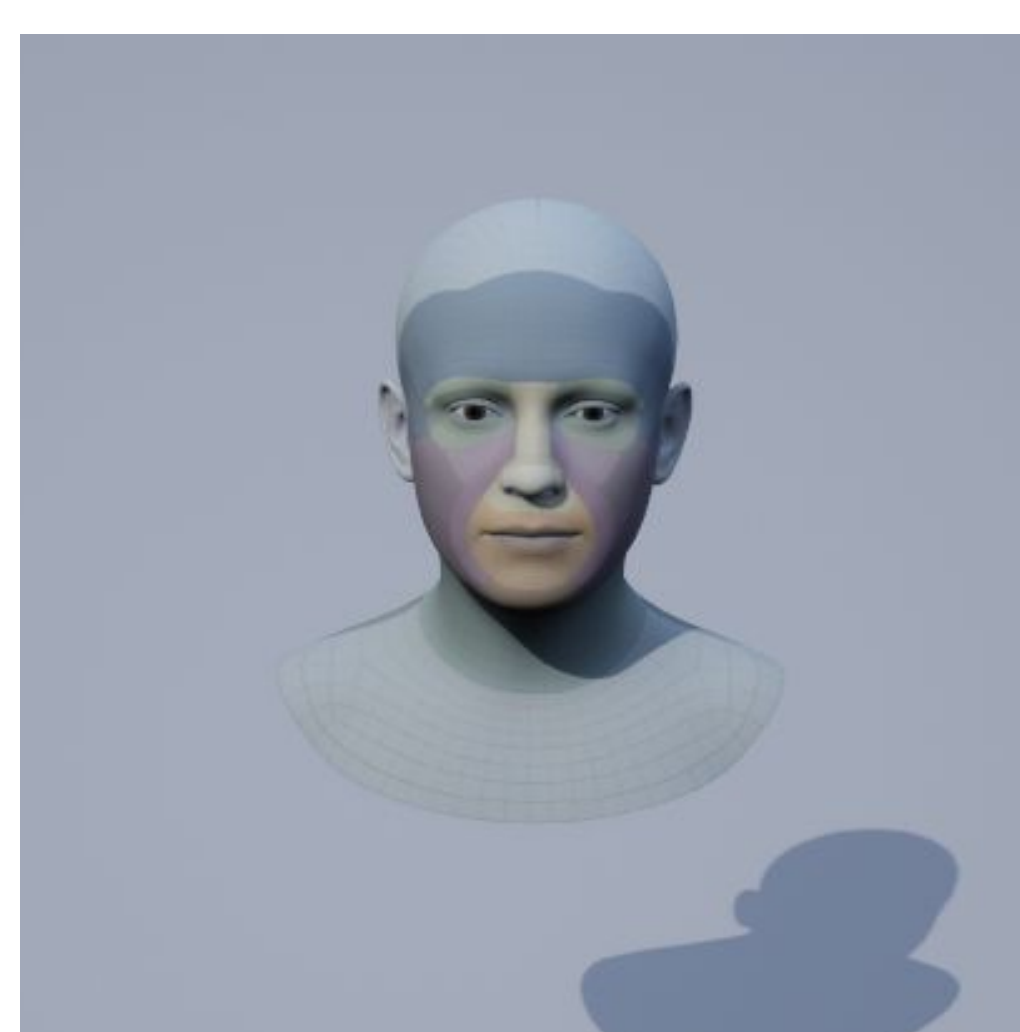
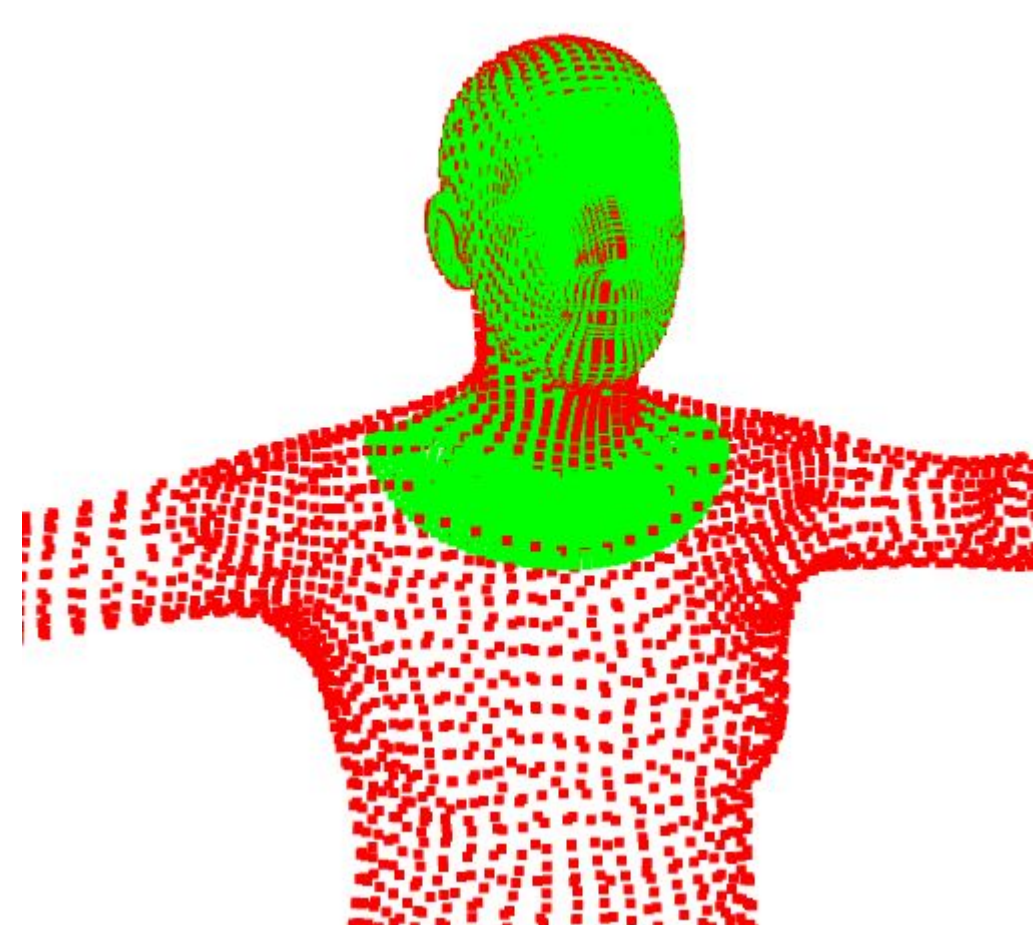
- Comunicação em Libras: expressões faciais são cruciais na comunicação das linguagens de sinais, adicionando profundidade e sentido à comunicação.
- Desafio dos Dados: escassez de datasets de expressões faciais de qualidade para tradução.
- Geração de Dados: promessa de superar a escassez, como o BEDLAM [2].
- Propósito: automatizar o processo de geração de dados sintéticos a fim de reproduzir expressões faciais realistas.

Objetivos

- Desenvolver metodologia automatizada para animação facial em avatares fotorrealistas.
- Automatizar a aplicação e interpolação de expressões faciais no Unreal Engine.
- Propor um pipeline de transferência de expressões do SMPL-X [1] para MetaHuman, de um modelo paramétrico para um avatar realista.
- Finalidade: gerar dados sintéticos de expressões faciais para aprimorar sistemas de tradução de línguas de sinais.

Material e Métodos ou Metodologia

- A primeira fase engloba scripts para aplicar expressões da biblioteca padrão do Unreal.
- Extração de dados de vértices, esqueletos e pesos de deformação do SMPL-X [1] e MetaHuman.
- Alinhamento de nuvens de pontos e mapeamento dos K-vizinhos mais próximos.
- Deformação efetiva com média ponderada.



Apoio Financeiro



Resultados e/ou Ações Desenvolvidas

- Pipeline de Automação Unreal: sucesso na aplicação e interpolação suave de expressões estáticas no MetaHuman.



- Pipeline de Transferência: sucesso na extração de dados, alinhamento e mapeamento de vértices dos modelos humanos.
- Desafio/Falha: distorções e artefatos visuais no Metahuman deformado “vale da estranheza”.



Conclusões

- Transferência parcial de expressões estáticas SMPL-X [1] para Metahuman.
- Limitação: não foi possível inferir emoções de vídeos e animar os avatares de maneira dinâmica.
- Desafios futuros:
 - Refinar algoritmos de transferência levando fatores de topologia e não-linearidade em conta.
 - Integrar inferência de expressão facial de vídeo.
- Impacto: contribuir para acessibilidade e comunicação para usuários de linguagens de sinais.

Bibliografia

[1] PAVLAKOS, G. et al. Expressive Body Capture: 3D Hands, Face, and Body from a Single Image. 2019.

[2] BLACK, M. J. et al. Bedlam: A synthetic dataset of bodies exhibiting detailed lifelike animated motion. In: IEEE Conf. on Comp. Vis. and Patt. Rec. [S.l.: s.n.], 2023. p. 8726–8737.