

ALINHAMENTO DE MODELOS MÉDICOS 3D NO MICROSOFT HOLOLENS

Guilherme Vieira Silva Gonçalves¹, Paulo Sergio Silva Rodrigues²

^{1,2} Departamento de Ciência da Computação, Centro Universitário FEI, São Paulo, Brasil
unifgoncalves@fei.edu.br, psergio@fei.edu.br

Resumo: Com o advento da Realidade Aumentada, uma das áreas mais beneficiadas foi a área médica, que utiliza essa tecnologia, por exemplo, em centros cirúrgicos. No entanto, o alinhamento automático entre os objetos 3Ds e uma cena real é um desafio, devido ao custo computacional. Portanto, este projeto de pesquisa visa realizar o alinhamento de modelos 3Ds visualizados com o Microsoft Hololens com imagens reais sobre as quais os modelos estão sendo projetados, utilizando métodos de detecção de pontos fiduciais como o SIFT e Redes Neurais Convolucionais (CNN).

1. Introdução

Apesar do recente desenvolvimento da tecnologia na área da saúde, trata-se de uma área ainda com forte demanda em diversos campos. Entre os principais interesses estão as tecnologias como Realidade Aumentada.

Lançado no final de 2016, o Microsoft Hololens é atualmente um dos mais avançado hardware de captura da estrutura 3D de ambientes internos disponível ao usuário comum. Trata-se de um equipamento para Realidade Aumentada, através da projeção no ambiente real de hologramas de objetos previamente armazenados, não servindo, portanto, como modelador automático dessas estruturas. Essa iniciativa impulsionou a pesquisa e o mercado para novas e mais modernas aplicações, como por exemplo, na área médica e outras que são exemplificadas em Park, S et al. (2021)[5].

Os sensores de profundidade que estão acoplados no Hololens permitem a ele gerar uma malha tridimensional somente da perspectiva de visualização do ambiente. No entanto, as malhas produzidas pelo Hololens têm o objetivo apenas de projeções holográficas e não a reconstrução 3D dos ambientes internos. Para que isso seja possível, é necessário pós-processamento, sobretudo para criar a projeção texturizada de imagens 2D sobre as regiões das malhas. Essa tarefa é ainda extremamente árdua e imprecisa, sobretudo porque é necessário alinhar pontos de interesse da malha e das imagens de texturas de modo a criar a sensação de um ambiente contínuo.

Por outro lado, a detecção automática de pontos fiduciais em imagens 2D observou um grande avanço em décadas recentes, sobretudo com o desenvolvimento de algoritmos como o SIFT (Scale Invariant Feature Transform)[1], SURF (Speed Up Robust Feature)[2], e várias outras propostas variantes. Recentemente, as representações de imagens baseadas na rede neural convolucional (CNN) têm atraído interesse crescente na comunidade devido ao desempenho impressionante demonstrado[6].

No artigo de Herbert Bay et al.(2001) [3], foi realizado um levantamento bibliográfico abrangente de ambas as instâncias baseadas em SIFT e CNN para métodos de recuperação de imagens. A conclusão geral é que, embora os algoritmos baseados em SIFT sejam competitivos com os baseados em CNN para o Based Image Retrieval (CBIR)[4], a CNN pode superar o SIFT. No entanto, a principal desvantagem da CNN em relação ao SIFT é que as CNNs são dependentes de grandes bases de treinamento, porém ao utilizar o SIFT, pode facilitar a criação dessas bases de treinamento.

A junção de técnicas de mapeamento de texturas 2D com malhas 3D visando a reconstrução automática de ambientes de visualização imersiva é uma tarefa desafiadora por requerer muito trabalho manual dos modeladores.

No entanto, a construção de uma malha de maneira automática em tempo real permitida pelo Microsoft Hololens abre várias possibilidades de pesquisas em desenvolvimento de software para a área. O maior desafio no momento é a correspondência entre os pontos fiduciais da malha e os pontos fiduciais da imagem, de forma a criar texturas observadas na cena capturada que possam ser projetadas sobre a malha construída.

Assim, este projeto de Iniciação Científica tem por objetivo o alinhamento de objetos 3D, gerados a partir de uma malha espacial capturada com o Microsoft Hololens, em ambientes reais, através da detecção e correspondência de pontos fiduciais tanto na malha quanto nas imagens. A captura e correspondências serão realizadas através de recentes avanços em algoritmos de detecção de pontos fiduciais como aqueles baseados em CNN e SIFT.

2. Base de dados

Embora o problema de alinhamento de malhas (base para este projeto) seja um problema há décadas resolvido através de heurísticas, uma vez que se trata de um problema NP-Completo, na área médica ainda são poucos os trabalhos na literatura que abordam esse problema como casamento entre modelos 3D reconstruídos a partir de tomografias de órgãos humanos com superfícies reconstruídas a partir de imagens capturadas dentro de ambientes médicos.

Sendo assim, a quantidade de bases de dados que contenham imagens e informações visuais ou descritivas de centros cirúrgicos, disponíveis na literatura, para estudo e avanço dessa tecnologia, é quase nula. Assim, este projeto tem por objetivo construir a própria base de dados, através da coleta de informações em centros cirúrgicos de instituições parceiras do Centro Universitário FEI como o Hospital Alemão Oswaldo Cruz, Hospital Universitário da USP e Hospital da Santa Casa de Misericórdia de SP.

3. Metodologia

A primeira etapa consiste construção e no processamento da base de dados que será utilizada para o treinamento de uma CNN. Esta base, mais bem descrita na Seção 2, conterà vídeos e imagens de ambientes cirúrgicos que serão capturadas em hospitais parceiros da FEI, e com imagens tomográficas dos pacientes que foram operados nos centros cirúrgicos capturados, que serão utilizadas futuramente para reconstrução volumétrica.

A segunda etapa consiste em capturar a sequência de malhas tridimensionais e a sequência de frames em formato de vídeos que corresponde às sequências de malha geradas pelo Microsoft Hololens. Após estas capturas, serão aplicados filtros baixa como pré-processamento inicial, porém, nos frames poderá ser efetuado uma pré segmentação das cenas com objetivo de induzir as etapas seguintes nas regiões de interesse. Essa ideia ainda será testada, uma vez que não há na literatura trabalho similar realizado com o Hololens.

A terceira etapa recebe a sequência de malhas suavizadas e aplica um algoritmo para a detecção de Pontos Fiduciais da Malha. A ideia desse algoritmo é também descrita na Figura 1 e consiste no processamento por frame (Figura 1, item a) para detecção de pontos fiduciais com o algoritmo SIFT (Figura 1, item b). Esse algoritmo realiza um pré-processamento dos pontos fiduciais iniciais que são alimentados em uma CNN (Figura 1, item c). Essa CNN tem a função de filtrar e classificar os pontos chaves finais (Figura 1, item d) que são posteriormente utilizados para a reconstrução 3D do ambiente médico, assim, acelerando assim o processo de percepção dos pontos estratégicos da estrutura da cena. Essa estratégia visa corresponder em cada malha pontos similares e, a partir daí, unir a sequência de malhas processadas como uma única malha do ambiente 3D. E a sequência de frames suavizado, serão processados pelo algoritmo SIFT gerando como saída os pontos fiduciais de cada frame, chamado aqui de Frame Panorâmico. Assim os pontos correlatos podem ser superpostos gerando uma única imagem panorâmica 3D.

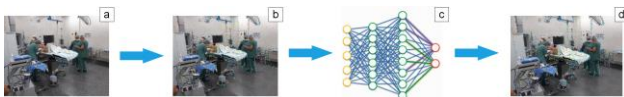


Figura 1: Estrutura geral do algoritmo proposto. (Fonte: Autor).

A última etapa consiste na correspondência entre a malha 3D e o Frame Panorâmico, que será feita manualmente onde cada ponto da malha 3D é associada a um único ponto do frame panorâmico, visto que a correspondência automática não é efetiva. Assim, projetaremos os frames individuais do frame panorâmico em formato de textura sobre a malha 3D, gerando assim uma Reconstrução 3D semiautomática do sinal de entrada. Uma exemplificação do alinhamento está na Figura 2.



Figura 2: Desenho da sobreposição de um modelo humano 3D com o paciente em uma sala cirúrgica. (Fonte: Brother UK)

4. Conclusões Parciais e Etapas Futuras

O vigente projeto de pesquisa encontra-se em desenvolvimento e está em fase de implementação dos algoritmos mencionados na seção 3 e na formação e processamento da base de dados mencionada na seção 2.

As etapas futuras consistem em continuar implementando os passos da metodologia exposta na seção 3, analisar os resultados e levar para dentro de centros cirúrgicos de instituições parceiras de FEI para realizar testes em campo, projetando modelos 3D reconstruídos a partir de tomografias de órgãos humanos com superfícies reconstruídas a partir de imagens capturadas dentro dos centros cirúrgicos, auxiliando o cirurgião. Podendo também o alinhamento ser feito com a articulação do paciente, para auxílio à fisioterapia ou prevenção e previsão de queda.

5. Referências

- [1] D. G. Lowe. Object recognition from local scale-invariant features. In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, volume 2, pages 1150–1157 vol.2, 1999.
- [2] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). Computer Vision and Image Understanding, 110:346–359, 2008.
- [3] Denis Feurer and Fabrice Vinatier. The time-sift method : detecting 3-d changes from archival photogrammetric analysis with almost exclusively image information. CoRR, abs/1807.09700, 2018
- [4] Henning Müller, Wolfgang Müller, David McG. Squire, Stéphane Marchand-Maillet, Thierry Pun, Performance evaluation in content-based image retrieval: overview and proposals, Pattern Recognit. Lett. 22 (5) (2001) 593–601.
- [5] Park, S., Bokijonov, S., & Choi, Y. (2021). Review of Microsoft HoloLens Applications over the Past Five Years. Applied Sciences.
- [6] Zeng, Wenjun. (2020). Toward human-centric deep video understanding. APSIPA Transactions on Signal and Information Processing.

Agradecimentos

Ao Centro Universitário FEI pelo empréstimo dos equipamentos como o Hololens. À FAPESP pelo incentivo financeiro. Ao HU-USP/INTEC pela parceria no desenvolvimento do projeto.

¹ Aluno de IC FAPESP. Projeto com vigência de 04/2022 a 03/2023.