

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI

ĐỒ ÁN TỐT NGHIỆP

Khôi phục khuôn mặt 3D sử dụng deeplearning và
Photometric stereo

Trần Nhật Hoàng

hoang.tn204559@sis.hust.edu.vn

Ngành: Khoa học máy tính

Giảng viên hướng dẫn: TS. Nguyễn Thị Thu Hương

Chữ kí GVHD

Khoa: Khoa học máy tính

Trường: Công nghệ Thông tin và Truyền thông

HÀ NỘI, 01/2025

LỜI CẢM ƠN

Tôi xin được phép gửi lời cảm ơn chân thành và sâu sắc nhất đến gia đình, thầy cô và bạn bè đã giúp tôi hoàn thành đồ án tốt nghiệp này. Cảm ơn gia đình đã luôn tin tưởng và động viên tôi trong suốt thời gian học tập. Cảm ơn thầy cô là những người đã tận tâm hướng dẫn và dành thời gian cho tôi hoàn thành đồ án, cảm ơn những thầy cô đã hết mực giúp đỡ cho tôi. Cảm ơn bạn bè đã là những người cùng tôi chia sẻ niềm vui và khó khăn, luôn đồng hành cùng tôi dù trong những lúc bế tắc. Cảm ơn bản thân mình đã không từ bỏ, quyết tâm hoàn thành nhiệm vụ đến cuối cùng dù gặp nhiều khó khăn. Cảm ơn Đại học Bách Khoa Hà Nội vì tất cả. Sự chia sẻ, giúp đỡ của mọi người chính là động lực lớn lao giúp tôi hoàn thành đồ án với kết quả mong đợi.

LỜI CAM KẾT

Họ và tên sinh viên: Trần Nhật Hoàng

Điện thoại liên lạc: 0337808436

Email: hoang.tn204559@sis.hust.edu.vn

Lớp: Khoa học máy tính 02

Hệ đào tạo: Khoa học máy tính - cử nhân

Tôi – Trần Nhật Hoàng – cam kết Đồ án Tốt nghiệp (ĐATN) là công trình nghiên cứu của bản thân tôi dưới sự hướng dẫn của TS. Nguyễn Thị Thu Hương. Các kết quả nêu trong ĐATN là trung thực, là thành quả của riêng tôi, không sao chép theo bất kỳ công trình nào khác. Tất cả những tham khảo trong ĐATN – bao gồm hình ảnh, bảng biểu, số liệu, và các câu từ trích dẫn – đều được ghi rõ ràng và đầy đủ nguồn gốc trong danh mục tài liệu tham khảo. Tôi xin hoàn toàn chịu trách nhiệm với dù chỉ một sao chép vi phạm quy chế của nhà trường.

Hà Nội, ngày tháng năm

Tác giả ĐATN

Họ và tên sinh viên

TÓM TẮT NỘI DUNG ĐỒ ÁN

Đồ án này tập trung nghiên cứu và tìm hiểu các phương pháp khôi phục hình ảnh 3D, đặc biệt là đối với gương mặt con người. Khôi phục 3D từ lâu đã là một chủ đề nhận được sự quan tâm lớn trong nhiều lĩnh vực, từ nghiên cứu khoa học đến ứng dụng thực tiễn và tầm quan trọng của nó vẫn không ngừng gia tăng trong tương lai. Mặc dù hiện nay có nhiều phương pháp hiện đại đạt độ chính xác cao, nhưng chúng vẫn tồn tại những hạn chế nhất định, từ các kỹ thuật thủ công đến các giải pháp tiên tiến. Trong số đó, phương pháp SDM-UnPS[1] được đánh giá là một trong những cách tiếp cận hiệu quả để khôi phục hình ảnh 3D nhờ vào độ chính xác vượt trội và khả năng áp dụng rộng rãi trên nhiều môi trường và bề mặt khác nhau. Tuy nhiên, phương pháp này vẫn còn một số điểm yếu cần cải thiện, đặc biệt là yêu cầu cao về thời gian xử lý và tài nguyên tính toán, cũng như thách thức trong việc xử lý dữ liệu thời gian thực trong các môi trường tự nhiên. Nhận thấy những hạn chế này, tôi đề xuất cải tiến bằng cách giảm lượng thông tin đầu vào thông qua việc sử dụng duy nhất một hình ảnh chứa thông tin từ ba nguồn sáng màu đỏ (red), xanh lá (green), và xanh dương (blue). Hơn nữa, việc kết hợp nhiều mô hình khác nhau giúp khai thác tối đa các ưu điểm của từng phương pháp để giải quyết bài toán khôi phục gương mặt 3D một cách hiệu quả hơn. Dựa trên ý tưởng này, tôi đã thiết kế và triển khai một mô hình cải tiến, mang lại kết quả khả quan. Đồ án trình bày mô hình khôi phục gương mặt 3D mới, với số lượng tham số tính toán và chi phí tài nguyên được tối ưu đáng kể. Đồng thời, mô hình này còn giải quyết được vấn đề xử lý thời gian thực mà SDM-UnPS[1] gặp phải. Kết quả cuối cùng cho thấy mô hình được đề xuất có thể tái tạo gương mặt 3D chính xác với yêu cầu tính toán thấp hơn và khả năng hoạt động trong thời gian thực, vượt trội so với phiên bản gốc của SDM-UnPS[1].

Sinh viên thực hiện

(Ký và ghi rõ họ tên)

ABSTRACT

This thesis discusses the research and exploration of 3D reconstruction methods for objects, particularly focusing on human faces. 3D reconstruction has been a topic of interest not only in the present but also in the past and will continue to be in the future. Although many methods for reconstruction are currently being widely applied with high accuracy, alongside their advantages, there remain limitations across techniques ranging from manual to modern approaches.

The SDM-UnPS method is recognized as one of the highly effective approaches for 3D image reconstruction. Despite its outstanding advantages in terms of accuracy and broad applicability (in various environments and surfaces), this method still has drawbacks that need to be addressed, including time consumption and resource demands, as well as the challenge of real-time processing in natural environments. Recognizing these limitations, improvements can be made by reducing the input data and performing real-world processing using a single image containing information from three colored light sources: red, green, and blue.

Furthermore, combining multiple models allows me to leverage the strengths of different approaches for the problem of 3D face reconstruction. Based on this observation, I proceeded to design a model incorporating these improvements and achieved promising results. This thesis presents a model design for 3D face reconstruction with significantly reduced computational parameters and resource costs. Additionally, it addresses the issue of real-time processing for the SDM-UnPS problem.

The final result of this thesis is a model capable of reconstructing 3D human faces with reduced computational requirements and real-time performance compared to the original SDM-UnPS model.

MỤC LỤC

CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI.....	1
1.1 Giới thiệu bài toán khôi phục ảnh 3D.....	1
1.2 Các giải pháp hiện tại và hạn chế	2
1.3 Mục tiêu và định hướng giải pháp	3
1.4 Đóng góp của đề án trong bài toán khôi phục 3D.....	4
1.5 Tóm tắt bố cục đề án.....	4
CHƯƠNG 2. NỀN TẢNG LÝ THUYẾT	6
2.1 Quá trình phát triển nghiên cứu về khôi phục ảnh 3D.....	6
2.2 Các bài toán khôi phục ảnh 3D truyền thống	6
2.2.1 Các phương pháp dựa trên các điều kiện ngoài.....	6
2.2.2 Nghiên cứu phản xạ Lambertian	8
2.2.3 Các phương pháp dựa trên ánh sáng	9
2.3 Bài toán Photometric Stereo	9
2.3.1 Bài toán Photometric Stereo truyền thống.....	9
2.3.2 Bài toán Photometric Stereo trong Deep learning.....	11
2.3.3 Khôi phục khuôn mặt 3D sử dụng photometric stereo.....	12
2.4 Khôi phục ảnh 3D dựa trên SDM - Universal Photometric Stereo.....	12
2.4.1 Phương pháp Universal Photometric Stereo	12
2.4.2 Phương pháp SDM - Universal Photometric Stereo tổng quan	13
2.4.3 Phương pháp SDM - Universal Photometric Stereo trên gương mặt người	13
2.5 Một vài mô hình xử lý ảnh trong computer vision	14
2.5.1 Mô hình CNN trong xử lý ảnh.....	14
2.5.2 Mô hình Transformer trong xử lý ảnh (ViT)	14

CHƯƠNG 3. PHƯƠNG PHÁP ĐỀ XUẤT.....	16
3.1 Tổng quan giải pháp.....	16
3.2 Cải tiến về độ tính toán của bài toán dựa vào thay đổi độ phức tạp của dữ liệu	20
3.3 Cải tiến model bài toán gốc bằng sự kết hợp mô hình CNN và Transformer	21
3.4 Xử lý đánh giá bài toán ở môi trường và thời gian thực tế	23
3.5 Tổng hợp giải pháp và phát triển	23
CHƯƠNG 4. ĐÁNH GIÁ THỰC NGHIỆM.....	26
4.1 Bộ dữ liệu sử dụng trong quá trình huấn luyện và đánh giá	26
4.1.1 Bộ dữ liệu PreTrain	26
4.1.2 Bộ dữ liệu FineTune	27
4.2 Tính toán hiệu năng và tài nguyên sử dụng.....	27
4.3 Triển khai mô hình khôi phục 3D.....	28
4.3.1 Tiền xử lý dữ liệu	28
4.3.2 Tham số được sử dụng trong quá trình huấn luyện và kết quả của quá trình huấn luyện	28
4.4 Đánh giá kết quả so sánh với mô hình gốc.....	30
4.5 Đánh giá kết quả thực nghiệm.....	35
CHƯƠNG 5. KẾT LUẬN	42
5.1 Kết luận	42
5.2 Hướng phát triển trong tương lai	42
TÀI LIỆU THAM KHẢO.....	43

DANH MỤC HÌNH VẼ

Hình 2.1	Phương pháp khôi phục 3D ToF	7
Hình 2.2	Phương pháp khôi phục 3D sử dụng Lidar	7
Hình 2.3	Phương pháp khôi phục 3D sử dụng nhiều góc nhìn quanh vật thể	8
Hình 2.4	Phản xạ bề mặt Lambert	8
Hình 2.5	Phương pháp tái tạo bản đồ sâu dựa vào đổ bóng	9
Hình 2.6	Phương pháp tái tạo dựa vào phản xạ ánh sáng	9
Hình 2.7	Mô hình CNN base	14
Hình 2.8	Mô hình ViT base	15
Hình 3.1	Model SDM-UnPS[1]	16
Hình 3.2	Minh họa cho ý tưởng đề xuất	21
Hình 3.3	Thay đổi phương pháp triết suất thông tin đặc trưng[1]	22
Hình 3.4	Đầu vào của bài toán với một ảnh chứa thông tin 3 đèn màu	23
Hình 3.5	Down-sample	24
Hình 3.6	Bản đồ vector pháp tuyến bề mặt dự đoán	25
Hình 4.1	Phương pháp tạo ảnh	26
Hình 4.2	Cấu trúc bộ dữ liệu	27
Hình 4.3	Proposed method test: trong tập dữ liệu pretrain	31
Hình 4.4	SDM method test: trong tập dữ liệu finetune	32
Hình 4.5	Proposed method test: mới đầu khi finetune	33
Hình 4.6	Proposed method test: sau khi finetune	34
Hình 4.7	Data input - ảnh chụp dưới 3 đèn màu R-G-B	35
Hình 4.8	Mask input	36
Hình 4.9	SDM method không sử dụng mask	36
Hình 4.10	SDM method sử dụng mask	37
Hình 4.11	Proposed method không sử dụng mask	37
Hình 4.12	Proposed method sử dụng mask	38

DANH MỤC BẢNG BIỂU

Bảng 4.1	So sánh CNN và Transformer trong Photometric Stereo	40
Bảng 4.2	So sánh CNN và Transformer trong 3D Face Reconstruction .	41

DANH MỤC THUẬT NGỮ VÀ TỪ VIẾT TẮT

Thuật ngữ	Ý nghĩa
2D	Tow-Dimensional image
3D	Three-Dimensional image
CNN	Convolutional Neural Network
SDM-UnPS	Scalable, Detailed and Mask-Free Universal Photometric Stereo
UPS	Universal Photometric Stereo Network using Global Lighting Contexts
ViT	Vision Transformer