

BÁO CÁO ĐỒ ÁN CUỐI KỲ

Môn học: CS519 - PHƯƠNG PHÁP LUẬN NCKH

Lớp: CS519.M11 - CS519.M11.KHCL

GV: PGS.TS. Lê Đình Duy

Trường ĐH Công Nghệ Thông Tin, ĐHQG-HCM



3D-MOMENTS FROM NEAR-DUPLICATE PHOTOS

Trần Thành
Nguyễn Duy Phúc
y

20521924
18521255
yy

Tóm tắt

- Lớp: CS519.N11
- Link Github của nhóm: <https://github.com/google-research/3d-moments>
- Link YouTube video:
- Thành viên nhóm bao gồm:



Trần Thành



Nguyễn Duy Phúc

yy

Giới thiệu

- Việc chụp ảnh với số lượng lớn và bất kì đâu cực kì tiện lợi với chiếc smartphone.
- Có những khoảnh khắc đáng nhớ khiến chúng ta muốn lưu giữ kỹ càng hơn so với một bức ảnh 2D có thể.
- Đơn giản là muốn tạo một khoảnh khắc để chia sẻ cùng bạn bè trên mạng xã hội.

Mục tiêu

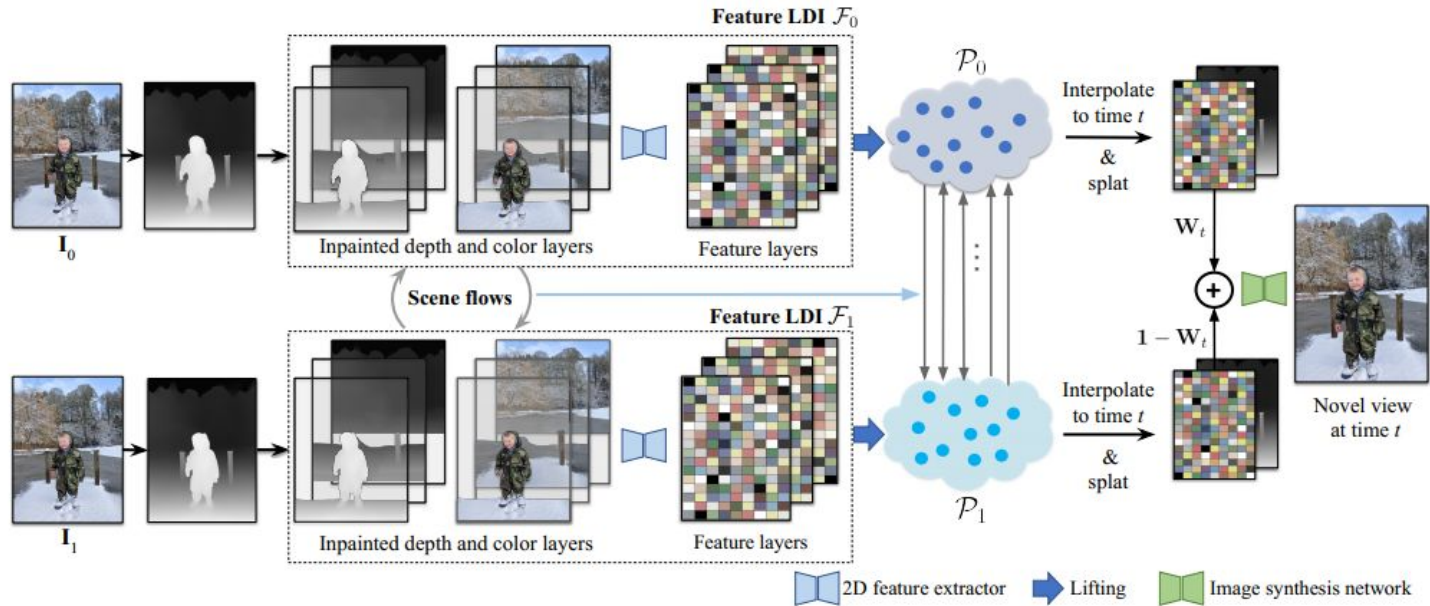
- Xây dựng một phương thức để tái hiện lại những khoảnh khắc 3D đáng nhớ thông qua những bức ảnh gần trùng lặp của hành động
- Xây dựng một mô hình mới dựa trên các LDI đặc trưng được tăng cường bằng các scene flow để tạo 3D-moments

Nội dung và Phương pháp

- Nội dung:
 - Thông tin đầu vào bao gồm 2 bức ảnh chụp gần trùng lặp của một hành động, ra khi đưa qua mô hình, thu được đoạn video ngắn 3D (space-time video)
 - Xây dựng mô hình trích xuất LDI đặc trưng từ mỗi tấm hình input, sau đó kết hợp 2 LDI đặc trưng để tạo ra thành phẩm cuối cùng

Nội dung và Phương pháp

- Phương pháp:



Nội dung và Phương pháp

- Xây dựng bộ dữ liệu:
 - Dùng hai bộ dữ liệu Vimeo-90K và Mannequin-Challenge
- Cách đánh giá mô hình: Cực tiểu hóa:
 - Hàm perceptual loss
 - Hàm L loss giữa ảnh dự đoán(predicted) và ảnh thực (ground-truth) trong mạng lưới học giám sát

Kết quả dự kiến

- Xây dựng được một mô hình tạo khoảng khắc 3D để so sánh với những mô hình làm tuần tự trước đây, đánh giá, cải tiến... Từ đó được áp dụng rộng rãi vào các công cụ chỉnh sửa ảnh.
- Xây dựng được thuật toán trích xuất LDI đặc trưng từ những scene flow để phục vụ cho việc tạo ra 3D moments

Tài liệu tham khảo

- [1] Aayush Bansal, Minh Vo, Yaser Sheikh, Deva Ramanan, and Srinivasa Narasimhan. 4d visualization of dynamic events from unconstrained multi-view videos. In CVPR, pages 5366–5375, 2020.
- [2] Wenbo Bao, Wei-Sheng Lai, Chao Ma, Xiaoyun Zhang, Zhiyong Gao, and Ming-Hsuan Yang. Depth-aware video frame interpolation. In CVPR, June 2019.
- [3] Mojtaba Bermana, Karol Myszkowski, Hans-Peter Seidel, and Tobias Ritschel. X-fields: Implicit neural view-, light-and time-image interpolation. ACM TOG, 39(6), 2020.
- [4] Michael Broxton, John Flynn, Ryan Overbeck, Daniel Erickson, Peter Hedman, Matthew Duvall, Jason Dourgarian, Jay Busch, Matt Whalen, and Paul Debevec. Immersive light field video with a layered mesh representation. ACM TOG, 39(4), July 2020.
- [5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In CVPR, pages 770–778, 2016.