

# **BÀI THU HOẠCH**

## **CHỦ ĐỀ: 3D RECONSTRUCTION**

### **A. Thông tin cá nhân**

MSSV: 19127273

Huỳnh Thị Mỹ Thanh

### **B. Bài thu hoạch**

#### **I. Problem statement**

- Depth estimation  
Input: rgb sequence  
Output: depth map (groundtruth)
- Reconstruction  
Có dc depth map => point cloud  
Input: point cloud  
Output: 3d object

#### **II. Challenge**

- Ước tính chuyển động của máy ảnh và đối tượng một cách chính xác.
- Xử lý sự che khuất của đối tượng bởi các đối tượng khác hoặc nền.
- Yêu cầu lượng dữ liệu lớn để đào tạo mô hình học sâu và việc thu thập dữ liệu đó rất khó khăn.
- Khi input đầu vào là point cloud ta không có thông tin vị trí camera.

#### **III. Method**

Work flow: 2D Images => Depth maps => 3D point cloud => mesh

- Depth estimation: nhiệm vụ đo khoảng cách của từng pixel so với máy ảnh và tạo depth map đại diện.  
Depth map: một phân diễn tả chiều sâu dựa trên hình ảnh (chuyển dần từ màu này đến màu khác từ dữ liệu chiều sâu gần nhất tới xa nhất)  
Kiểu màu JET  
Kiểu màu GRAYSCALE

**Depth sensor:** có 3 loại

- Stereo vision: sử dụng ít nhất 2 camera được đặt cách nhau một khoảng cách đã biết trước. Bằng cách so sánh các điểm ảnh được chụp từ hai camera khác nhau, phương pháp này tính toán khoảng cách hoặc độ sâu của các điểm ảnh trên đối tượng.
- Time of flight: Là phương pháp dựa trên nguyên tắc đo thời gian bay của ánh sáng trong không khí để tính toán khoảng cách từ nguồn ánh sáng đến đối tượng và trở lại. Phương pháp này sử dụng cảm biến TOF để đo thời gian bay của tia laser hoặc ánh sáng hồng ngoại. TOF được sử dụng trong các ứng dụng như đo khoảng cách, theo dõi chuyển động, và thực tế ảo.
- Structured light camera: Là phương pháp sử dụng một nguồn ánh sáng có cấu trúc nhất định, chẳng hạn như dãy tia laser hoặc mẫu chiếu, để chiếu lên đối tượng và sử dụng camera để quan sát mẫu chiếu hoặc dãy tia laser trên đối tượng. Bằng cách phân tích biến dạng của mẫu chiếu hoặc dãy tia laser, phương pháp này tính toán khoảng cách hoặc độ sâu của đối tượng.

**Monocular depth estimation:** predict depth dựa trên những dữ liệu đầu vào là dữ liệu 2 chiều.

- GLPDepth: độ sâu hoặc khoảng cách từ các điểm ảnh trong ảnh RGB đến các điểm ảnh tương ứng trong ảnh độ sâu hoặc ảnh khoảng cách. Để đạt được điều này, GLPDepth sử dụng mạng neural để học mối quan hệ giữa các điểm ảnh RGB và độ sâu hoặc khoảng cách tương ứng trong quá trình huấn luyện trên dữ liệu đào tạo. Sau khi được huấn luyện, mô hình GLPDepth có thể được sử dụng để dự đoán độ sâu hoặc khoảng cách của các điểm ảnh không được huấn luyện trong dữ liệu thực tế.
- Point cloud construction: nhiệm vụ từ depth map đưa ra point cloud.  
Gồm 2 bước chính:
  1. **Camera calibration:** là quá trình để xác định các thông số và đặc tính của một máy ảnh hoặc hệ thống máy ảnh

Phương pháp:

- direct linear transform là một phương pháp dựa trên đối tượng. Phương pháp này dựa trên việc tìm một ma trận chiếu sao cho tích của ma trận chiếu và các điểm ảnh tương ứng trên đối tượng là như nhau
- Zhang's method: là một phương pháp dựa trên đối tượng, dựa trên việc tìm một dãy biến đổi hình học của đối tượng trên mặt phẳng hình học và từ đó tính toán các thông số calibration của máy ảnh.

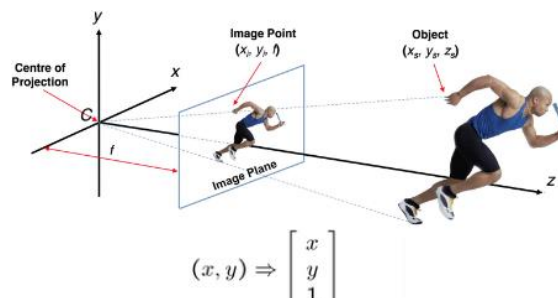
## 2. Reprojecting from image plane to real world 3D coordinate

Là quá trình chuyển đổi lại từ các điểm ảnh trên ảnh 2D vào tọa độ 3D trong không gian thực tế.

Phương pháp:

- Sử dụng ma trận chiếu: Trong phương pháp này, ma trận chiếu của máy ảnh được sử dụng để chuyển đổi các điểm ảnh trên ảnh 2D sang tọa độ 3D trong không gian thực tế. Ma trận chiếu bao gồm các thông số calibration của máy ảnh, chẳng hạn như ma trận camera (camera matrix) và ma trận chiếu của không gian thực tế (world-to-camera matrix), và nó có thể được tính toán từ quá trình camera calibration.

Homogeneous coordinates



- Mesh generation: đổ mesh lên.
  - PCA: xác định vector pháp tuyến hướng từ đám mây điểm gồm tọa độ không gian (x,y,z) để dựng pháp tuyến vuông góc với tiếp tuyến.  
Pháp tuyến vô hướng thành có hướng bằng thông tin vị trí camera:  
d là thông tin camera, thay đổi phương hướng của N nếu  $(D, N) \geq 90^\circ$
  - Biểu diễn tập điểm bằng hàm poisson:  
Cũng giống như phương trình (hàm) đường tròn:  $(x-0)^2 + (y-0)^2 = R^2$ . Khi xét 1 điểm bất kì (0,1) thay vào, nếu vế trái < vế phải thì điểm nằm bên trong đường tròn, VT=VP thì điểm nằm trên đường tròn và ngược lại. Đối với hàm đối tượng 3D indicator function cũng vậy và ta đi tìm 2 ẩn số: hàm  $\chi$  và ngưỡng isovalue  $\gamma$  để khi thế  $q(x,y,z)$  vào  $\chi$  nếu:

$$\begin{cases} \tilde{\chi}(q) > \gamma: \text{ngoài object} \\ \tilde{\chi}(q) < \gamma: \text{trong object} \\ \tilde{\chi}(q) = \gamma: \text{biên object} \end{cases}$$

**Ta tìm Indicator function  $\chi$**  sao cho đạo hàm của nó  $\nabla\chi$  xấp xỉ trường vector pháp tuyến (Oriented points  $N \rightarrow$ )

Sau đó plug điểm  $(x,y,z)$  vào indicator function, nếu  $\text{indicator} < \text{isovalue}$ : điểm inside,  $\text{indicator} > \text{isovalue}$ : điểm outside,  $\text{indicator} = \text{isovalue}$ : boundary

Cách tìm isovalue: sử dụng Octree

**Sau đó phủ lưới tam giác cho hàm indicator bằng marching cube:**

Sau khi đã biết điểm nào nằm ngoài hay trong object, phủ khối hộp vuông lớn bao trọn đối tượng, Khối hộp vuông lớn được tạo từ những khối hộp vuông nhỏ.

Duyệt theo từng khối vuông nhỏ.

Dựa theo trường hợp, phủ tam giác tương ứng với trường hợp đó.

#### **IV. Conclusion**

3D reconstruction là một lĩnh vực quan trọng trong công nghệ thông tin đồ họa và thị giác máy tính, có ứng dụng rộng rãi trong nhiều lĩnh vực như trong công nghiệp, y học, nghiên cứu khoa học, và giải trí.

Tuy nhiên, việc thực hiện 3D reconstruction vẫn đối mặt với nhiều thách thức, như độ chính xác của kết quả tái tạo, độ phức tạp của thuật toán, độ phức tạp của dữ liệu đầu vào, và tính thực tế trong ứng dụng thực tế. Do đó, việc lựa chọn phương pháp phù hợp, xử lý dữ liệu đầu vào chính xác và đánh giá kết quả là những yếu tố quan trọng trong quá trình 3D reconstruction.

Tuy nhiên, với sự phát triển nhanh chóng của công nghệ, 3D reconstruction tiếp tục đem lại nhiều tiềm năng và cơ hội cho các ứng dụng hiện tại và tương lai, giúp cải thiện và mở rộng các ứng dụng trong nhiều lĩnh vực khác nhau.