

Tái tạo bàn tay 3D dựa trên mô hình thông qua học tự giám sát

Lưu Anh Dũng^{1,3} và Nguyễn Trọng Doanh^{2,3}

{¹ 19521392, ² 19521368}@gm.uit.edu.vn

³ Trường ĐH Công Nghệ Thông Tin - ĐHQG HCM

Mục tiêu

Chúng tôi đề xuất một mô hình mạng neural tái tạo bàn tay 3D thông qua học tự giám sát. Cụ thể, chúng tôi sẽ:

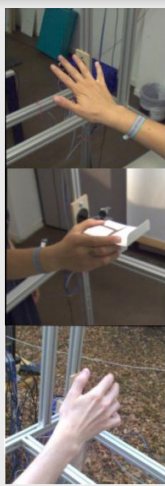
- Nghiên cứu và xây dựng mô hình mạng neural tái tạo bàn tay 3D từ một hình ảnh duy nhất mà không sử dụng bất kỳ dữ liệu đã gắn nhãn nào.
- Phát triển mô hình để đạt độ chính xác và tốc độ xử lý cao hơn các mô hình hiện có.
- Thử nghiệm mô hình trên một số bộ dữ liệu khó, qua đó cải thiện độ chính xác của mô hình.

Lý do chọn đề tài

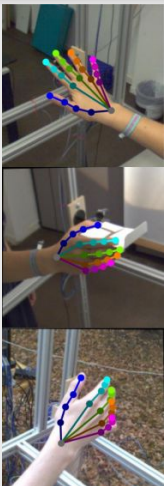
- Việc tái tạo 3D bàn tay người từ một hình ảnh là rất quan trọng đối với các tác vụ thị giác máy tính như nhận dạng hành động liên quan đến bàn tay, thực tế tăng cường (AR), dịch ngôn ngữ ký hiệu và tương tác giữa con người với máy tính.
- Đây là một vấn đề khó khăn vì tư thế, hình dạng phong phú của bàn tay và sự mơ hồ về chiều sâu trong tái tạo 3D.
- Các phương pháp hiện đại đều dựa vào gắn nhãn 3D rất tốn kém.

Overview

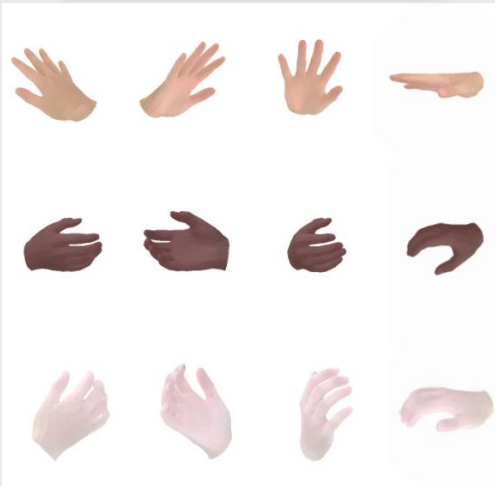
Hình ảnh input



Biểu diễn các khớp tay



Mô hình 3D bàn tay



Description

1. Phương pháp

Sử dụng bộ mã hoá tự động để nhận hình ảnh bàn tay làm đầu vào, và trả về hình dạng, kết cấu của tay và góc nhìn máy ảnh, sau đó tạo ra nhiều biểu diễn 2D trong không gian hình ảnh và thiết kế các hàm mất mát cho quá trình huấn luyện. Để hiện thực hoá phương pháp này:

- Một bộ mã hoá tự động sẽ được sử dụng để dự đoán các khớp và hình dạng 3D. Ngoài ra cũng cần nghiên cứu thêm giải pháp để mạng neural trả về bàn tay 3D với tư thế và kích thước hợp lý.
- Cần có một công cụ dự đoán điểm khoá 2D để trả về các điểm khoá 2D, và một sự mất mát mới về tính nhất quán 2D-3D để liên kết công cụ dự đoán điểm khoá 2D và mạng tái tạo 3D.

2. Nội dung

- Nghiên cứu và xây dựng mạng neural tái tạo bàn tay 3D thông qua học tự giám sát, sao cho kết quả thể hiện chính xác các khớp và kết cấu 3D của bàn tay từ một hình ảnh duy nhất mà không sử dụng bất kỳ dữ liệu đã gắn nhãn nào.
- Thử nghiệm mô hình trên các bộ dữ liệu lớn như FreiHAND, HO-3D,...
- So sánh độ chính xác và tốc độ xử lý với các mô hình state-of-the-art như OpenPose, MANO-CNN,...

3. Kết quả mong đợi

- Mô hình đạt độ chính xác hơn 80% khi thử nghiệm trên một số bộ dữ liệu lớn như FreiHAND, HO-3D,...
- Mô hình đạt độ chính xác cao hơn so với ít nhất 3 mô hình state-of-the-art.
- Mô hình được ứng dụng trong các ứng dụng thực tế ảo (VR), thực tế tăng cường (AR).

