

1 Tên đề tài

Nghiên cứu ứng dụng nhận diện cử chỉ bàn tay trong thời gian thực và ứng dụng vào trò chơi AngryBirds.

2 Giới thiệu

2.1 Tổng quan

Bài báo cáo đề xuất sau đây, giới thiệu về việc một trò chơi tương tác với con người thông qua cử chỉ di chuyển động của bàn tay. Nghiên cứu này đề cập đến việc quan tâm ngày càng tăng về trải nghiệm chơi game sáng tạo và tương tác mới lạ. Bằng cách tận dụng khả năng của Mediapipe [2], một thư viện thị giác máy tính mạnh mẽ, người chơi sẽ tương tác với trò chơi dựa trên những cử chỉ chuyển động của bàn tay thay vì phụ thuộc vào các thiết bị đầu vào truyền thống như bàn phím và chuột.

2.2 Lập luận ý nghĩa, tầm quan trọng của vấn đề

Với việc phát triển gần đây của trí tuệ nhân tạo (AI), việc máy tính có thể nhận biết và hiểu được các hành động cử chỉ của con người một cách chính xác. Nhờ đó mà việc triển khai hệ thống nhận diện cử chỉ bàn tay trong thời gian thực trở nên chính xác và dễ tiếp cận hơn. Trong đó thị giác máy tính là việc máy tính có thể nhận biết và hiểu những sự vật xung quanh như con người. Học sâu có nghĩa là máy tính học thông qua việc mô phỏng cách não bộ người hoạt động. Tự động học và rút ra những đặc trưng từ dữ liệu, giúp máy tính hiểu và thực hiện các nhiệm vụ phức tạp.

Việc sử dụng chuột, bàn phím để tương tác với trò chơi hoàn toàn phổ biến hiện nay. Với mong muốn tìm ra giải pháp mới để thay thế, công nghệ nhận dạng cử chỉ bàn tay cho phép ta làm điều đó. Nó giúp người chơi điều khiển các hành động trong trò chơi thông qua những cử động tự nhiên của bàn tay. Việc cầm, ném, đánh, chạy hoặc nhảy... không còn chỉ là những lệnh từ bàn phím hay chuột mà trở thành những cử chỉ tự do, mang lại cho người chơi trải nghiệm tương tự như đang tham gia vào thế giới thực.

Chính vì những lý do ở trên, tôi đề xuất đề tài nghiên cứu ứng dụng nhận diện cử chỉ bàn tay trong thời gian thực và ứng dụng vào trò chơi video. Nhằm thay đổi trải nghiệm chơi game, mang đến cảm giác chơi game mới lạ so với các phương thức phổ biến hiện đang có và tăng tính hấp dẫn, tính chân thực.

2.3 Vấn đề kỹ thuật

Để máy tính có thể theo dõi các cử chỉ của bàn tay trong thời gian thực, mình sẽ sử dụng Mediapipe Hands để giải quyết vấn đề này. Nó là mã nguồn mở của Google có thể giúp nhận dạng và theo dõi (track) các chuyển động của bàn tay trong thời gian thực.

2.4 Mục tiêu

2.4.1 Mục tiêu chung của đề tài

Xây dựng một ứng dụng theo dõi cử chỉ của bàn tay, chuyển đổi những cử chỉ đó thành những lệnh cụ thể để mà điều khiển được các hành động trong game AngryBirds.

2.4.2 Mục tiêu kỹ thuật

Phát triển một mô hình nhận diện cử chỉ bàn tay chính xác và nhanh chóng sử dụng thư viện Mediapipe.

Sử dụng ngôn ngữ python, xây dựng hàm chuyển đổi những cử chỉ đó thành những lệnh cụ thể.

Tích hợp hệ thống nhận diện vào trò chơi Angry Birds sử dụng ngôn ngữ lập trình Python.

2.5 Tóm tắt phương pháp, nội dung nghiên cứu

Nội dung 1: tìm hiểu về công nghệ và cách áp dụng

- Tìm hiểu về mediapipe
- Làm quen với python



Hình 1: Mô hình mốc bàn tay

- Tìm cách kết hợp python với mediapipe để tạo ra một ứng dụng có thể nhận diện được bàn tay trong thời gian thực.

Nội dung 2: Áp dụng vào việc điều khiển các hành động trong game AngryBirds

- Sử dụng python để xây dựng hàm chuyển đổi những cử chỉ tay đó thành những câu lệnh cụ thể.
- Áp dụng vào điều khiển trò chơi AngryBirds.

3 Nội dung nghiên cứu

3.1 Xây dựng ứng dụng nhận diện cử chỉ bàn tay trong thời gian thực thông qua webcam

3.1.1 Giới thiệu tổng quan về Medipipe

MediaPipe là một thư viện mã nguồn mở do Google phát triển, chủ yếu được thiết kế để cung cấp các công cụ và giải pháp cho xử lý dữ liệu hình ảnh và video. Thư viện này chủ yếu tập trung vào việc phát hiện và theo dõi các phần tử trong hình ảnh và video, như khuôn mặt, bàn tay, cử chỉ, v.v.

MediaPipe Hands là một phần của thư viện MediaPipe. Nó cung cấp một cách thuận tiện để xác định vị trí các điểm đặc trưng trên bàn tay (hình 1), cho phép theo dõi và phân tích cử chỉ tay trong thời gian thực.

3.1.2 Tìm hiểu về python

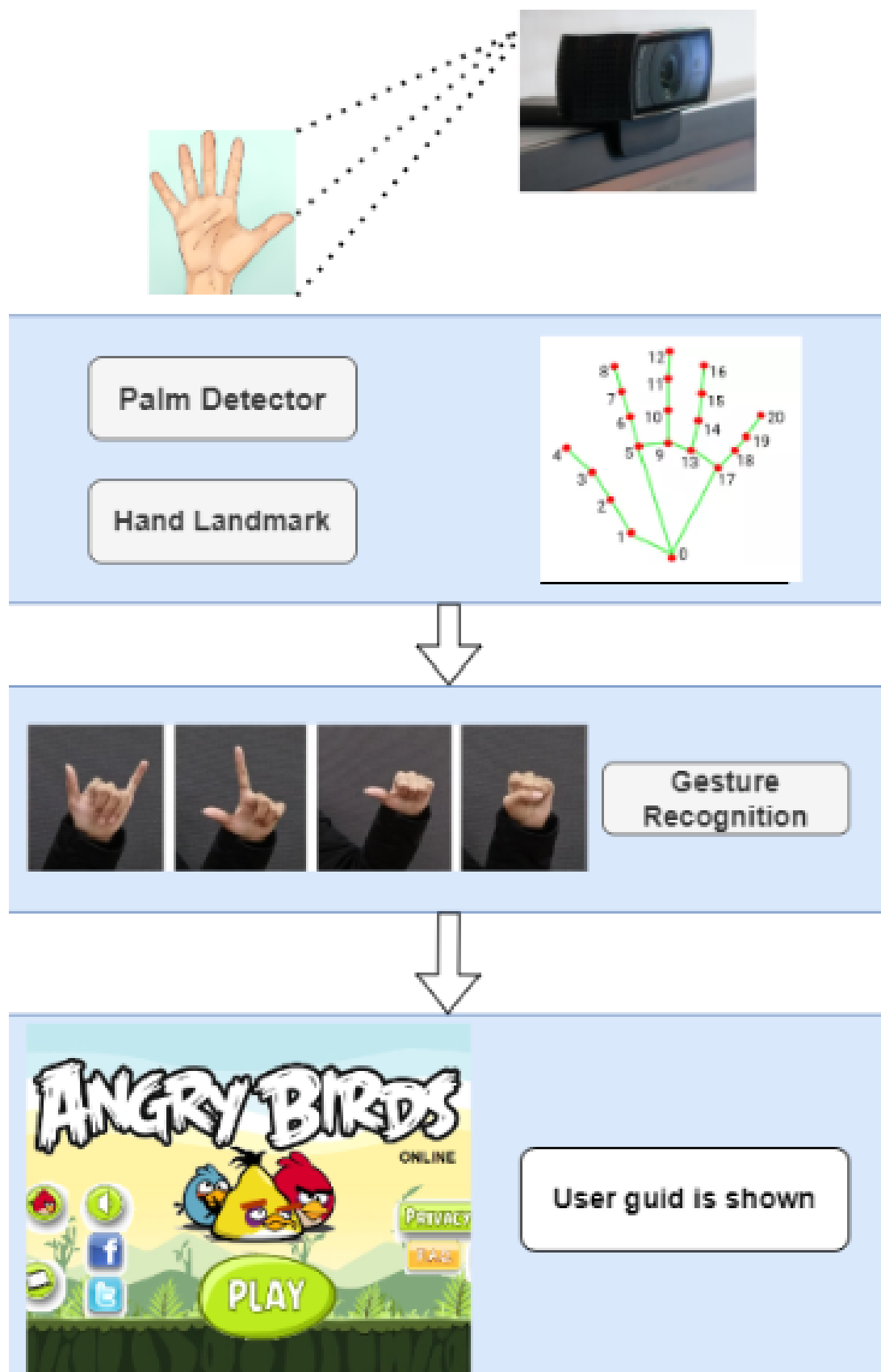
Python là một trong những ngôn ngữ phổ biến và được ưa chuộng trong lĩnh vực Trí tuệ nhân tạo (AI) do cú pháp đơn giản và dễ đọc, giúp người lập trình tập trung vào logic thay vì cú pháp phức tạp. Hệ sinh thái đa dạng của Python cung cấp các thư viện và frameworks mạnh mẽ như TensorFlow và PyTorch, giúp người phát triển dễ dàng xây dựng, đào tạo, và triển khai mô hình machine learning và deep learning.

3.1.3 Quy trình làm việc

Ở hình 2 dưới đây mô tả về cách mà ứng dụng hoạt động.

3.2 Xây dựng hàm chuyển đổi những cử chỉ tay thành câu lệnh điều khiển được sử dụng trong trò chơi

Xây dựng hàm chuyển đổi sẽ được xây dựng bằng ngôn ngữ lập trình python. Hàm này có nhiệm vụ chuyển đổi những cử chỉ tay thành các câu lệnh nhằm mục đích điều khiển các hành động trong game Angry Birds:



Hình 2: Quy trình làm việc

- Sau khi có được các mốc tọa độ của bàn tay mà Mediapipe cung cấp thông qua việc theo dõi bàn tay từ webcam. Ta đem so sánh vị trí của các tọa độ thu được để xác định rõ cử chỉ đó có nghĩa là gì. Chẳng hạn như từ hình mốc bàn tay ở trên (Hình 1), hệ thống của chúng ta sẽ cho chúng ta biết trạng thái là "Open" vì tất cả các ngón tay đều đang được duỗi thẳng.
- Ta gán những cử chỉ đó cho các lệnh điều khiển cần thiết trong trò chơi. Chẳng hạn như khi ta nắm tay lại thì hành động đó có thể được hiểu là tấn công, hay là đang cầm vật gì đó...

3.3 Áp dụng vào trò chơi AngryBirds

Trong trò chơi AngryBirds, ta phải sử dụng chuột để điều khiển các hành động như kéo, thả, bấm chuột trái để kích hoạt nội tại,... Giờ đây ta có thể sử dụng cử chỉ bàn tay cụ thể để thay thế cho việc sử dụng chuột.

4 Tài liệu tham khảo

[1] Ninh Văn Thọ, "Ứng dụng và tương lai của thị giác máy tính", tạp chí khoa học và công nghệ, 11-08-2022.

[2] Mediapipe, "<https://developers.google.com/mediapipe/solutions/vision/gesture-recognizer>"

[3] Md Rafid Islam and Ratun Rahman and Akib Ahmed and Rafsan Jany, arXiv, 2022, "NFS: A Hand Gesture Recognition Based Game Using MediaPipe and PyGame"

Tổ chức / cá nhân đề xuất

(Họ, tên và chữ ký - đóng dấu đối với tổ chức)

Nguyễn Minh Phương