

Ứng dụng học máy để nhận dạng tư thế bàn tay từ một chuỗi ảnh chiều sâu

Phạm Minh Lộc¹, Nguyễn Văn Hồ^{2,*}



Use your smartphone to scan this QR code and download this article

TÓM TẮT

Nhận dạng hình ảnh là một trong những vấn đề khó và được quan tâm nghiên cứu ứng dụng rất nhiều trong thời gian gần đây. Xu hướng cho phép máy tính nhận dạng tư thế bàn tay người theo thời gian thực đang dần trở nên phổ biến và được triển khai áp dụng trong nhiều lĩnh vực cũng như tích hợp trên nhiều sản phẩm. Nghiên cứu này đề xuất hướng tiếp cận giúp cho việc nhận dạng hình ảnh nhanh chóng và chính xác hơn. Với việc phát triển công nghệ, thiết bị camera ngày nay đã có nhiều thay đổi đáng kể, camera vừa có thể thu ảnh RGB vừa có thể thu ảnh theo chiều sâu. Điều này giúp việc nhận dạng trong điều kiện thiếu ánh sáng chính xác hơn, giúp cải thiện hiệu năng nhận dạng. Thuật toán "Random Forest" là một thuật toán được đánh giá hiệu quả, nhanh chóng và chính xác hơn các thuật toán nhận dạng trước đây được sử dụng để thực nghiệm mô hình. Chúng tôi đã tiến hành thu thập khoảng 10.000 mẫu dữ liệu ảnh bàn tay, sau đó huấn luyện mô hình và đánh giá mô hình với kết quả chính xác gần 80%. Bên cạnh đó, nghiên cứu trình bày một số kết quả nhận dạng cử chỉ của bàn tay người theo thời gian thực sử dụng thông tin thu được từ cảm biến Kinect, mô hình có khả năng huấn luyện dữ liệu khá lớn, tốc độ nhận dạng nhanh. Kết quả nghiên cứu có thể sử dụng trong các ứng dụng điều khiển phục vụ trong đời sống, trong sản xuất công nghiệp.

Từ khóa: ảnh chiều sâu, rừng ngẫu nhiên, học máy, nhận dạng hình ảnh

GIỚI THIỆU

Với sự đột phá ngày càng mạnh mẽ của trình độ khoa học ngày nay đã đạt được nhiều bước tiến quan trọng, khiến cho việc sử dụng máy tính trở nên tiện dụng và thân thuộc với con người hơn. Ngoài những thiết bị như chuột và bàn phím bây giờ cũng đã xuất hiện các thiết bị giao tiếp khác như camera và micro. Điều đó dẫn đến sự phát triển mạnh mẽ trong lĩnh vực nghiên cứu giao tiếp giữa con người với máy tính thông qua ngôn ngữ hình ảnh và âm thanh¹. Khi so sánh cách giao tiếp với máy tính dùng hình ảnh và âm thanh, thì phương pháp giao tiếp bằng hình ảnh sẽ thích hợp hơn phương pháp giao tiếp bằng âm thanh bởi vì xung quanh môi trường chúng ta đang sống có rất nhiều âm thanh ồn ào và tạp âm khiến cho việc giao tiếp bằng âm thanh là điều khó chính xác. Việc nghiên cứu này ngày càng trở nên cấp thiết bởi vì nó đem lại rất nhiều ứng dụng trong tất cả lĩnh vực khác nhau của xã hội như: hỗ trợ rất tốt trong giao tiếp cho người khuyết tật (câm, điếc), hỗ trợ điều khiển từ xa mà không cần thiết bị cảm biến phụ trợ (điều khiển robot, các loại game giải trí không cần thiết bị điều khiển). Những đề tài nghiên cứu trước đây vẫn còn hạn chế về công nghệ thiết bị camera (camera RGB), thuật toán xử lý ảnh chậm. Trong đề tài này sẽ đưa ra hướng

nghiên cứu mới, giúp cho việc nhận dạng nhanh chóng và chính xác hơn, cải thiện đáng kể hiệu năng. Với các camera vừa có thể thu ảnh RGB vừa có thể thu ảnh theo chiều sâu. Với kỹ thuật nhận dạng tư thế bàn tay từ chuỗi ảnh theo chiều sâu dựa trên thuật toán cây quyết định ngẫu nhiên được nghiên cứu và đề xuất, đây là phương pháp được nghiên cứu đi theo hai mục tiêu chính: tính toán hiệu quả và nhanh hơn. Một hình ảnh theo chiều sâu đầu vào sẽ được tách ra các thành phần bộ phận, đính nhãn với các bộ phận được định nghĩa trong không gian lưu trữ tạo thành bộ khung hình ảnh theo độ sâu của các bộ phận bàn tay. Dựa vào dữ liệu lưu trữ các kiểu khác nhau của mỗi thành phần được phân bố và suy ra chính xác các vị trí của mỗi bộ phận bàn tay. Kết quả nghiên cứu có thể được phát triển và ứng dụng trong các thiết bị nhận dạng tại các doanh nghiệp hoặc trong đời sống. Phần còn lại của bài báo gồm Phần Cơ sở lý thuyết và Các nghiên cứu liên quan trình bày về cơ sở lý thuyết và các nghiên cứu liên quan. Phần tiếp theo mô tả về Phương pháp nghiên cứu của bài báo. Kết quả nghiên cứu được đề cập ở phần kế tiếp. Và cuối cùng bài báo đưa ra Kết luận và đề xuất Các hướng phát triển trong tương lai.

¹Công ty CP Quốc tế Anh văn Hội Việt Mỹ (VUS), Tp. Hồ Chí Minh, Việt Nam

²Trường Đại học Kinh tế - Luật, Đại học Quốc gia Tp. Hồ Chí Minh, Tp. Hồ Chí Minh, Việt Nam

Liên hệ

Nguyễn Văn Hồ, Trường Đại học Kinh tế - Luật, Đại học Quốc gia Tp. Hồ Chí Minh, Tp. Hồ Chí Minh, Việt Nam

Email: honv@uel.edu.vn

Lịch sử

- Ngày nhận: 01-7-2021
- Ngày chấp nhận: 11-02-2022
- Ngày đăng: 11-4-2022

DOI: 10.32508/stdjelm.v6i2.870



Bản quyền

© ĐHQG Tp.HCM. Đây là bài báo công bố mở được phát hành theo các điều khoản của the Creative Commons Attribution 4.0 International license.



Trích dẫn bài báo này: Lộc P M, Hồ N V. Ứng dụng học máy để nhận dạng tư thế bàn tay từ một chuỗi ảnh chiều sâu. *Sci. Tech. Dev. J. - Eco. Law Manag.*; 6(2):2373-2385.

CƠ SỞ LÝ THUYẾT VÀ CÁC NGHIÊN CỨU LIÊN QUAN

Nhận dạng hình ảnh cử chỉ bàn tay

Tương tác giữa người và máy

Những thiết bị đầu vào và đầu ra đặc biệt đã được thiết kế trong những năm qua với mục đích làm cho phần giao tiếp giữa máy tính và con người được thực hiện dễ dàng. Hai thiết bị phổ biến nhất là bàn phím và con chuột. Ý tưởng để làm cho máy tính hiểu ngôn ngữ con người và phát triển giao diện người – máy thân thiện đang nhận được sự quan tâm của cộng đồng các nhà nghiên cứu^{2,3}. Làm cho máy tính hiểu được lời nói, nét mặt, cử chỉ của con người là một trong số những quan tâm đó. Trong tương tác giữa con người và máy tính, các hình dạng khác nhau của bàn tay có thể giả định để thao tác với các đối tượng hoặc có thể truyền tải rất nhiều thông tin. Do đó, bàn tay của con người có thể sử dụng làm “Thiết bị đầu vào” rất có giá trị. Trong thế giới thực, chúng ta có thể cầm, thả, di chuyển ... các đối tượng bằng các cử chỉ của bàn tay. Tương tự như vậy, khi tương tác với các thiết bị như máy tính, tivi, ô tô ... chỉ với cử chỉ của bàn tay là ta có thể điều khiển được hoạt động của nó. Để làm được điều này, bộ điều khiển của thiết bị phải được gắn một thiết bị cảm nhận như camera hay webcame, thiết bị sẽ thu nhận hình ảnh bàn tay, nhận dạng cử chỉ để phát lệnh điều khiển tương ứng.

Để nhận dạng cử chỉ, bước đầu tiên trong các hệ thống nhận dạng là phát hiện ra vị trí bàn tay bằng một thiết bị đặc biệt như camera để thu thập các quan sát chuyển động cử chỉ bàn tay. Sau đó tiến hành xử lý hình ảnh để lọc ra được ảnh cử chỉ bàn tay và đưa vào huấn luyện cho máy tính nhận dạng. Các cử động của bàn tay được phát hiện dựa trên tập dữ liệu về bàn tay được thu thập và huấn luyện từ trước. Hai phương pháp để thu nhận và huấn luyện dữ liệu hình ảnh bàn tay cho máy tính học:

1. Dùng găng tay chuyên biệt có gắn thiết bị cảm biến khớp tay, từ những thiết bị cảm biến này sẽ truyền tính hiệu để cho máy tính hiểu được vị trí và tư thế bàn tay.
2. Dùng găng tay sơn màu để phân biệt từ bộ phận bàn tay và kết hợp thu nhận hình ảnh bằng camera. Nhưng ở nghiên cứu này sẽ dùng thiết bị camera Kinect với khả năng thu thập thêm thông tin khoảng cách của cách điểm đến camera. Điều này giúp hình ảnh nhận được hiệu quả và chính xác hơn.

Cử chỉ bàn tay

Thật khó để có thể giải quyết với một định nghĩa cụ thể nào của cử chỉ do có nhiều ứng dụng để xuất và mỗi ứng dụng chỉ có thể xác định trên một miền cụ thể của cử chỉ. Cử chỉ phải được định nghĩa như những chuyển động của thân thể khi giao tiếp với các cá thể khác. Để giao tiếp thành công, người truyền và người nhận phải có cùng một tập hợp thông tin cho những cử chỉ đặc biệt. Trong nghiên cứu, cử chỉ được định nghĩa như một sự chuyển động của những ngón tay như là tín hiệu đặc biệt để liên lạc chính xác giữa người gửi và thiết bị nhận.

Nhận dạng cử chỉ bàn tay sử dụng ảnh theo chiều sâu

Hiện nay, ứng dụng tương tác dựa vào cử chỉ bàn tay được phát triển mạnh mẽ và áp dụng trong chơi game hoặc tương tác của con người và máy tính hoặc robot mà không cần thiết bị điều khiển, ứng dụng nhận dạng trong lĩnh vực an ninh, ứng dụng điều khiển trong các hội thảo từ xa và thậm chí là lĩnh vực chăm sóc sức khỏe hỗ trợ người tàn tật. Những công việc đó gần đây đã được đơn giản hóa rất nhiều bởi sự ra đời của máy ảnh đo chiều sâu theo thời gian thực⁴⁻⁶. Tuy nhiên, thậm chí với những công nghệ tốt nhất hiện nay vẫn còn những hạn chế nhất định. Đặc biệt cho đến khi có sự ra đời của Kinect, không phụ thuộc vào phần cứng máy tính của người sử dụng mà vẫn xử lý đầy đủ các hình dạng và kích cỡ của hình ảnh khi chuyển động tốc độ cao. Một số hệ thống đạt được tốc độ cao bằng cách theo dõi từ frame này sang frame khác của hình ảnh nhưng rất khó khăn để có thể tái khởi tạo hình ảnh một cách nhanh chóng, như vậy không phải là tối ưu nhất. Trong nghiên cứu này, chúng tôi tập trung vào nhận dạng tư thế bàn tay và các bộ phận của bàn tay bằng cách theo dõi và phát hiện chuyển động của bàn tay từ các hình ảnh RGB-D⁷⁻⁹. Hình ảnh RGB-G là hình ảnh có 3 thông số về màu sắc là RED, GREEN, BLUE và thông số về DEPTH là khoảng cách từ điểm ảnh tới thiết bị thu hình. Nghiên cứu phương pháp khởi tạo và phục hồi mỗi khung ảnh để bổ sung cho các thuật toán theo dõi¹⁰⁻¹² các chuyển động tư thế theo thời gian thực. Đối với dữ liệu huấn luyện, có thể tạo ra những hình ảnh theo chiều sâu và hình ảnh thực từ nhiều hình dạng và kích cỡ trong những tư thế bàn tay rất đa dạng tạo thành một cơ sở dữ liệu lớn lưu trữ các tư thế của bàn tay. Thuật toán cây quyết định ngẫu nhiên¹³ được sử dụng để huấn luyện giúp cải thiện tốc độ nhanh hơn khi sử dụng số lượng lớn hình ảnh huấn luyện. Tóm lại, việc phân biệt hình ảnh so sánh chiều sâu làm tăng năng suất và giữ ổn định các điểm ảnh mà

vấn duy trì sự tính toán hiệu quả cao hơn các phương thức truyền thống. Cuối cùng, không gian mô hình của mỗi pixel phân bố được tính toán bằng cách dựa trên sự thay đổi trung bình dẫn đến xuất ra chính xác hình ảnh.

Phương pháp thực hiện tối ưu của thuật toán hoạt động dựa trên cơ chế frame-by-frame dẫn đến sự khác biệt đáng kể hình dạng cơ thể và kích thước, phương pháp tiếp cận được thực hiện một cách tự nhiên và phân chia được trên khung ảnh. Chúng tôi đánh giá trên cả hai hình ảnh độ sâu và hình ảnh thực. Mặc dù không có ràng buộc về thời gian khai thác hoặc chuyển động học nhưng điểm ảnh xuất ra vẫn chính xác và ổn định.

Cấu tạo camera Kinect

Hệ thống cảm biến chiều sâu

Bao gồm bộ phát hồng ngoại (infrared light) và camera hồng ngoại¹⁴. Kích cỡ ảnh là 640x480 pixel, tốc độ chụp 30 khung hình mỗi giây. Có 2 phạm vi độ sâu: (1) Phạm vi mặc định có sẵn trong cả hai bộ cảm biến Kinect cho Windows và các cảm biến Kinect cho Xbox 360, (2) Phạm vi gần là chỉ có trong cảm biến Kinect cho Windows. Hình 1 mô tả cách thức hoạt động của Kinect. Để cảm biến hoạt động tốt nhất thì nên sử dụng khoảng cách từ 1,2m đến 3,5m. Khi đó, các tia hồng ngoại được chiếu qua bộ phát hồng ngoại đến đối tượng, sau đó camera hồng ngoại sẽ thu thập dữ liệu bị phản chiếu.

Camera RGB-D

Chức năng chính của camera¹⁵ là nhận biết 3 màu cơ bản là đỏ, xanh lá cây và xanh da trời (Red-Green-Blue). Quá trình chụp bao gồm việc chụp một ảnh màu (RGB) và thực hiện một phép đo độ sâu (D). Cảm biến hình ảnh kết hợp với cảm biến chiều sâu nằm ở gần nhau, cho phép sáp nhập bản đồ, cho ra hình ảnh 3D. Thông tin ảnh RGB-D được lưu trữ. Với kích cỡ 1280x960 pixel, tốc độ 12 khung hình/giây. Với kích cỡ 640x480 pixel, tốc độ chụp 30 khung hình/giây. Từ đây cho thấy khi thiết lập độ phân giải thấp hơn thì tốc độ chụp và truyền hình ảnh là cao hơn. Với dữ liệu sâu thu được, nó sẽ tạo ra một bản đồ về bộ xương người đứng ở trước cảm biến. Và với bộ xương đó, nó có thể xác định được cử chỉ, hành động của người sử dụng.

Động cơ, máy đo gia tốc, micro

Động cơ: nằm ở đế của cảm biến Kinect, có khả năng làm cho cảm biến hướng đầu lên trên cao và hạ xuống thấp. Với khả năng này sẽ giúp cho cảm biến có thể thu được hình ảnh cao hơn rất nhiều so với khi không có bộ động cơ.

Máy đo gia tốc: Kinect sử dụng nó để xác định người đứng trước, đứng sau trong khi đi.

Micro: bao gồm 4 micro nhỏ, hoạt động như nhau ở 16 bit với tốc độ lấy mẫu là 16kHz (số mẫu/giây).

Phân mềm hỗ trợ

Các thư viện có thể hỗ trợ cho camera Kinect như: OpenNI (được sử dụng phổ biến nhất), OpenKinect (Microsoft sản xuất), Microsoft Kinect dành cho hệ điều hành Windows.

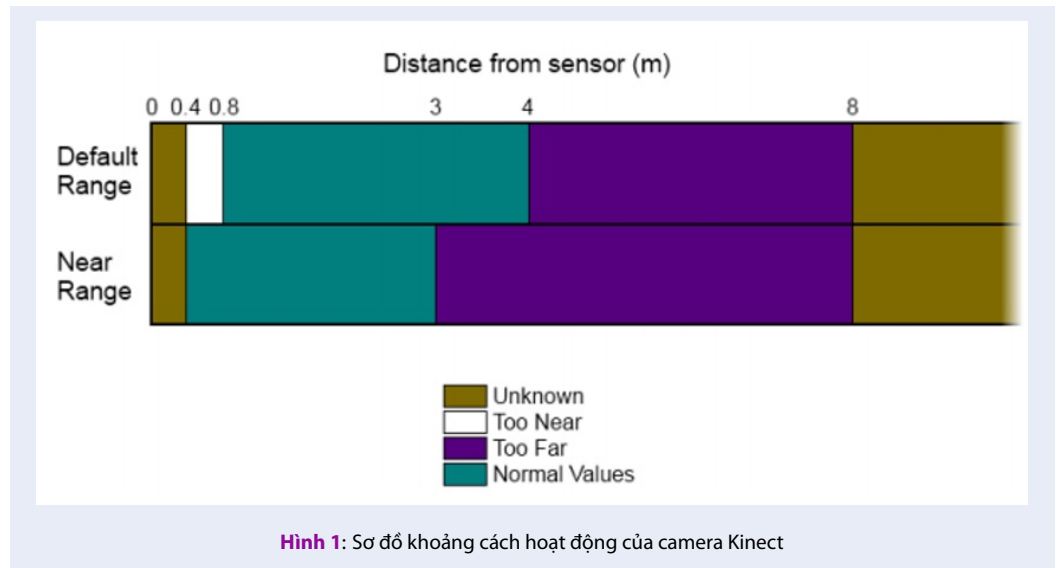
Tính toán độ sâu của Kinect

Cặp cảm biến IR camera và IR projector sẽ phối hợp với nhau để cho ra giá trị độ sâu ảnh bằng công nghệ Light Coding của PrimeSense.

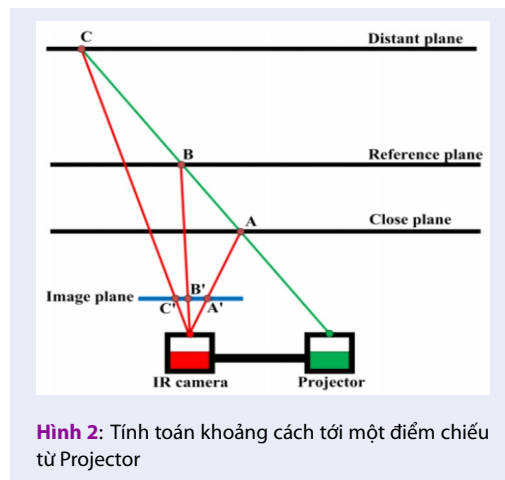
Khác với kỹ thuật Stereo Camera với việc dùng cặp camera giống nhau để xây dựng nên bản đồ độ sâu, kỹ thuật Time-Of-Flight (TOF) định nghĩa khoảng cách bằng ước lượng thời gian di chuyển của tia sáng đi và về trong không gian, kỹ thuật Light Coding dùng một nguồn sáng hồng ngoại chiếu liên tục kết hợp với một camera hồng ngoại để tính toán khoảng cách. Công việc tính toán này được thực hiện bên trong Kinect bằng chip PS1080 SoC của PrimeSense. Công nghệ mới này được cho là đáp ứng chính xác hơn, giá cả rẻ hơn cho việc sử dụng ở môi trường trong nhà.

Projector sẽ chiếu một chùm sáng hồng ngoại, tạo nên những đốm sáng ở không gian phía trước Kinect, tập hợp đốm sáng được phát ra này là cố định. Những đốm sáng này được tạo ra nhờ một nguồn sáng truyền qua lưới nhiễu xạ (diffraction gratings). Tập hợp các đốm sáng này được IR camera chụp lại, thông qua giải thuật đặc biệt được tích hợp trong PS1080 SoC cho ra bản đồ độ sâu. Bản chất của giải thuật này là các phép toán hình học dựa trên quan hệ giữa hai cảm biến IR camera và Projector. Hình 2 cho thấy rõ mẫu hình tập hợp các đốm sáng từ Projector và được chụp lại bởi IR camera.

Ta giả sử Projector phát đi một tia sáng dọc đường màu xanh lá, nó sẽ được chụp lại dưới dạng một đốm sáng bởi IR camera khi chạm vào bề mặt vật thể trong không gian. Ta xét ba mặt phẳng ở ba khoảng cách khác nhau: mặt phẳng gần Kinect (close plane), mặt phẳng ở xa Kinect (distant plane) và mặt phẳng tham chiếu (reference plane) ở giữa hai mặt phẳng trên. Trong đó, mặt phẳng tham chiếu ngấm được biết trước bên trong Kinect với đầy đủ thông tin về khoảng cách. Ngoài ra, ta cũng đề cập thêm mặt phẳng ảnh (Image Plane) của IR camera là mặt phẳng hình chiếu của các điểm trong không gian thu về bởi IR camera. Ta xét trong ba trường hợp khi tia sáng màu xanh lá chạm vào ba điểm trên ba mặt phẳng lần lượt là A, B,



Hình 1: Sơ đồ khoảng cách hoạt động của camera Kinect



Hình 2: Tính toán khoảng cách tới một điểm chiếu từ Projector

C, ba điểm này được chiếu lên mặt phẳng ảnh tương ứng là A', B', C'. Quan sát vị trí A', B' và C', ta có nhận xét: điểm A càng gần Kinect (hay close plane càng gần Kinect) thì A' càng xa B' về phía bên phải và ngược lại, điểm C càng xa Kinect (hay distant plane càng xa Kinect) thì C' càng xa B' về phía bên trái. Từ đó, khi ta biết trước hướng, điểm xuất phát của tia sáng từ Projector và vị trí B' là hình chiếu của điểm B trên mặt phẳng tham chiếu lên mặt phẳng ảnh, ta hoàn toàn có thể tính toán được độ sâu ảnh hay khoảng cách tới vật thể.

Kinect làm điều tương tự với tập hợp các đốm sáng còn lại phát đi từ projector, với mặt phẳng tham chiếu biết trước. Nó tìm điểm là tâm của đốm sáng mà IR camera chụp lại được và điểm tương đồng của đốm sáng đó trên mặt phẳng tham chiếu, để tìm khoảng chênh lệch giữa hai điểm này theo chiều ngang khi

chiếu về trên mặt phẳng ảnh và lưu ý là giá trị chênh lệch này được tính bằng đơn vị pixel. Tập hợp của tất cả các giá trị chênh lệch từ tập hợp đốm sáng, sẽ tạo nên bản đồ độ chênh lệch (disparity map), giá trị này càng lớn thì khoảng cách hay giá trị độ sâu ảnh (depth) càng lớn, từ đó mà ta xây dựng được bản đồ độ sâu (depth map) với giá trị tính bằng mét thực sự. Tuy nhiên, do tập hợp số lượng đốm sáng phát đi từ projector nhỏ hơn so với tổng số pixel trên mặt phẳng ảnh của IR camera nên một phần giá trị độ sâu ảnh còn lại sẽ được nội suy.

PHƯƠNG PHÁP NGHIÊN CỨU

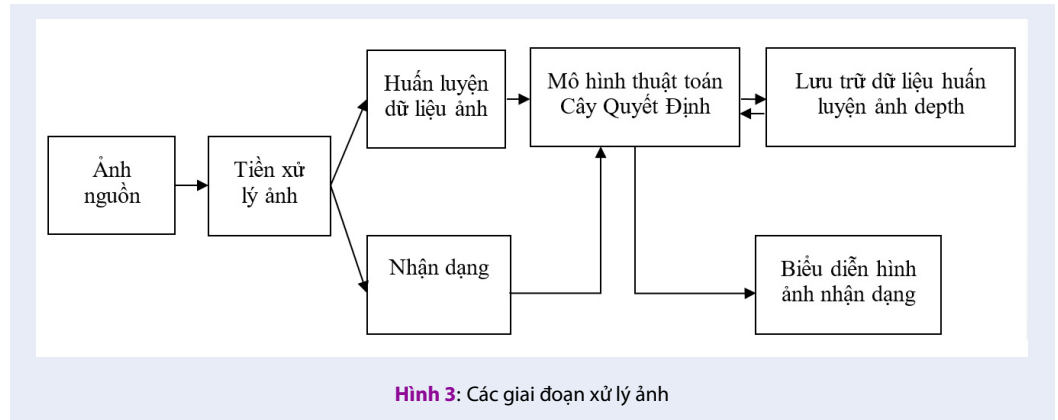
Tổng quan các giai đoạn của nhận dạng cử chỉ bàn tay

Quá trình xử lý ảnh được xem như là quá trình thao tác với ảnh đầu vào nhằm cho ra kết quả mong muốn¹⁶. Kết quả đầu ra của quá trình xử lý ảnh là một ảnh tốt hơn. Các giai đoạn xử lý ảnh như Hình 3, bao gồm: Thu nhận ảnh, Tiền xử lý, Phân đoạn, Trích lọc đặc trưng, Huấn luyện nhận dạng ảnh, Nhận dạng, Biểu diễn kết quả.

Huấn luyện ảnh theo chiều sâu

Thu thập dữ liệu huấn luyện bàn tay

Để có được bộ dữ liệu huấn luyện của bàn tay, ta cần có trên 10.000 mẫu tư thế bàn tay khác nhau và huấn luyện máy học. Mỗi bộ mẫu huấn luyện gồm: một file lưu ma trận ảnh theo độ sâu, một hình ảnh màu RGB phân biệt được các bộ phận chi tiết của bàn tay. Các bước thực hiện:



Hình 3: Các giai đoạn xử lý ảnh

1. Chuẩn bị găng tay vải thun, ôm sát bàn tay. Trên găng tay dùng sơn màu để đánh dấu những thành phần của bàn tay.
2. Viết chương trình kết nối camera Kinect có thể biểu diễn và lưu hình ảnh chiều sâu và ảnh RGB của mẫu bàn tay.
3. Chạy chương trình để lưu lại các tư thế khác nhau của bàn tay được mang găng tay.

Tiền xử lý ảnh

Để việc nhận dạng ảnh chỉ tập trung vào đúng đối tượng quan tâm, ta cần loại bỏ những đối tượng không liên quan tới nhận dạng như phông nền xung quanh đối tượng quan tâm, các vật thể trong ảnh không liên quan đến quá trình nhận dạng. Thế nên công đoạn tiền xử lý là bước đầu tiên nhằm loại bỏ nhiễu, khắc phục những khiếm khuyết do bước thu nhận ảnh không tốt là việc làm quan trọng. Ta có nhiều phương pháp cho việc nâng cao chất lượng ảnh nói chung và tiền xử lý nói riêng. Ở đây, chúng tôi đề xuất lọc đối tượng quan tâm dựa theo phương pháp theo độ sâu của từng điểm pixel, nghĩa là mỗi điểm ảnh theo độ sâu của camera Kinect thu được sẽ cung cấp thông tin khoảng cách từ pixel đó đến camera, dựa trên thông số này sẽ lọc ra được đối tượng bàn tay với một chỉ số khoảng cách nhất định.

Phân đoạn hình ảnh

Phân đoạn ảnh là chủ đề nghiên cứu chính cho nhiều nghiên cứu về xử lý ảnh. Mục đích rõ ràng và nhiều ứng dụng vô tận: hầu hết các vấn đề phân tích hình ảnh và thị giác máy tính đòi hỏi phải có một giai đoạn phân đoạn để phát hiện các đối tượng hoặc phân chia hình ảnh thành các vùng có thể coi là đồng nhất theo một tiêu chuẩn nhất định, chẳng hạn như màu sắc, kết cấu v.v... Kết quả của việc phân đoạn ảnh là một tập các vùng chung bao trùm toàn bộ hình ảnh hay một

tập các đường nét được trích xuất từ hình ảnh. Mỗi một điểm ảnh trong tập điểm ảnh trong một vùng là tương tự nhau với sự luyến ý về một vài tính chất hoặc thuộc tính tính toán chẳng hạn như màu sắc, cường độ và kết cấu.

Đặc trưng của ảnh theo độ sâu

Đây là phương pháp so sánh chiều sâu đơn giản bằng cách lấy độ chênh lệch chiều sâu của vùng lân cận xung quanh điểm cần xét, cách này được lấy ý tưởng trong. Với phương pháp này chỉ với một tín hiệu phân biệt yếu nhưng kết hợp với thuật toán cây quyết định ngẫu nhiên cũng đủ phân biệt chính xác các vùng khác nhau của bàn tay. Xét một điểm ảnh theo độ sâu nhất định, phương pháp so sánh theo chiều sâu được tính theo công thức (1) như sau:

$$d = f(x + u_i, y + v_i) - f(x - u_i, y - v_i) \quad (1)$$

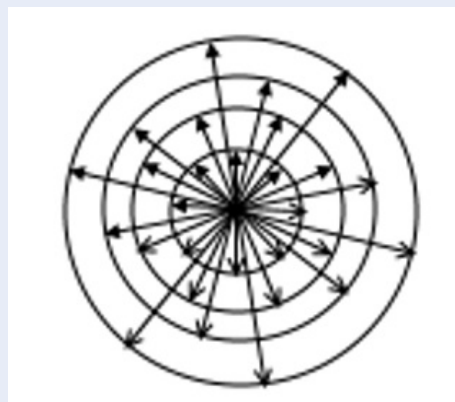
Trong đó, u_i và v_i là các cặp khoảng cách đến các điểm lân cận của điểm cần xét. Dựa trên các cặp u, v xác định được vị trí các điểm lân cận xung quanh điểm cần xét. Từ đó sẽ lấy được giá trị độ sâu của các điểm lân cận và so sánh với nhau để tính độ chênh lệch và hình thành những đặc trưng tại điểm đang xét (Hình 4).

Với cách làm này, ta sẽ lấy ngẫu nhiên một số lượng các điểm cần xét trên ảnh theo độ sâu (Hình 5). Từ đó, ta sẽ lấy được bộ dữ liệu đặc trưng cho từng điểm.

Xây dựng bộ dữ liệu huấn luyện ảnh theo chiều sâu

Với bộ dữ liệu mẫu bàn tay thu thập được, ta dùng thuật toán cây quyết định ngẫu nhiên để huấn luyện dữ liệu mẫu cho máy học. Các bước thực hiện:

1. Cài đặt thuật toán cây quyết định ngẫu nhiên cho chương trình huấn luyện.



Hình 4: Các vector đối xứng xung quanh 1 điểm



Hình 5: Lấy ngẫu nhiên đặc trưng của các điểm trên bàn tay

2. Chạy chương trình để huấn luyện toàn bộ dữ liệu thu thập được.

Cây quyết định ngẫu nhiên¹⁷⁻¹⁹ đã chứng minh là một thuật toán có khả năng nhanh chóng và hiệu quả trong việc phân loại nhiều lớp thành phần khi chạy đa tác vụ²⁰⁻²². Và có thể được thực hiện hiệu quả trên nền GPU (Graphics Processing Unit)²³. Như minh họa trong Hình 6.

Forest là một tập hợp dữ liệu của T cây quyết định, mỗi cây trong đó đều có nút chia và các nút lá. Mỗi nút bao gồm một tính năng f_θ và một τ ngưỡng. Để phân loại các điểm ảnh x trong hình ảnh, người ta bắt đầu từ gốc và sau đó thực hiện nhiều lần đánh giá để xác định đối tượng cần nhận dạng. Thành lập nhánh trái hoặc phải của cây quyết định ngẫu nhiên để so sánh với ngưỡng τ . Tại nút lá tìm được trong cây T, sẽ

phân phối lên hình ảnh bàn tay đưa vào huấn luyện một nhãn c được lưu trữ trước đó theo công thức (2). Các nhãn phân phối được tính trung bình với nhau trong tất cả các cây trong tập hợp dữ liệu để cung cấp cho các phân loại cuối cùng cho hình ảnh.

$$P(c|I,x) = \frac{1}{T} \sum_{t=1}^T P_t(c|I,x) \quad (2)$$

Huấn luyện: Mỗi cây được tập huấn về một tập khác nhau được lấy ngẫu nhiên hình ảnh tổng hợp. Một tập hợp ngẫu nhiên lấy khoảng 3000 mẫu pixels từ mỗi hình ảnh được lựa chọn để đảm bảo phân bố đều trên các bộ phận bàn tay. Mỗi cây được huấn luyện sử dụng theo thuật toán²⁰

Ngẫu nhiên để xuất một tập hợp dữ liệu thành phân $\varphi = (\theta, \tau)$ (thông số tính năng θ và ngưỡng τ).

Phân vùng các bộ ví dụ $Q = \{(I, x)\}$ trong nhánh trái và nhánh phải các tập hợp con của mỗi φ được tính theo công thức (3) và (4):

$$Q_l(\varphi) = \{(I,x) | f_\theta(I,x) < \tau\} \quad (3)$$

$$Q_r(\varphi) = Q \setminus Q_l(\varphi) \quad (4)$$

Tính toán φ sao cho là lớn nhất trong thông tin, công thức (5) và (6):

$$\varphi^* = \underset{\varphi}{\operatorname{argmax}} G(\varphi) \quad (5)$$

$$G(\varphi) = H(Q) - \sum_{s \in \{l,r\}} \frac{|Q_s(\varphi)|}{|Q|} H(Q_s(\varphi)) \quad (6)$$

Trong đó, $H(Q)$ được tính toán trên được tính toán trên biểu đồ bình thường của một bộ phận của bàn tay được phủ nhãn (x) cho tất cả $(I, x) \in Q$. Nếu đạt được lớn nhất $G(\varphi^*)$ là đặc trưng và chiều sâu trong cây là dưới mức tối đa, sau đó tái phân bố cho các tập con trái $Q_l(\varphi^*)$ và tập con phải $Q_r(\varphi^*)$.

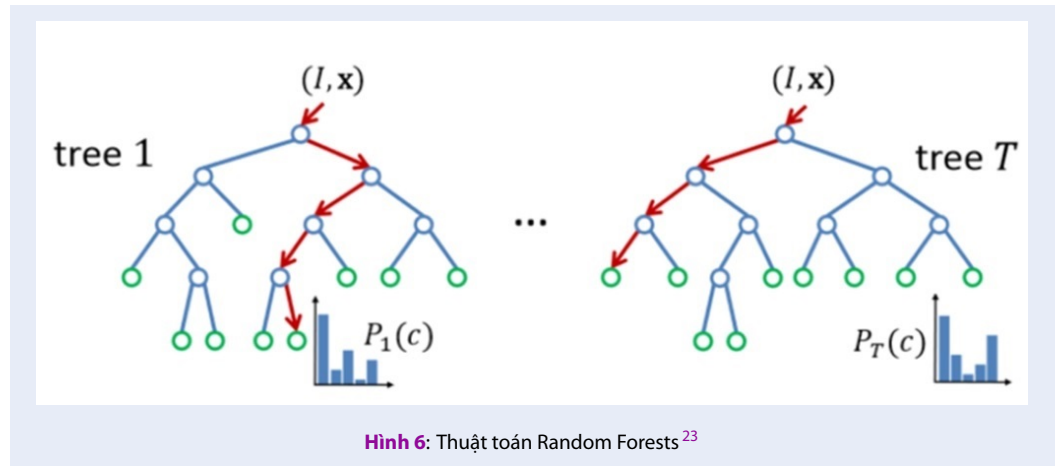
Nhận dạng và biểu diễn tư thế bàn tay

Nhận dạng bàn tay từ ảnh theo chiều sâu

Đây là bước kiểm tra nhận dạng của thuật toán cây quyết định ngẫu nhiên sau khi được huấn luyện với bộ dữ liệu bàn tay. Các bước thực hiện:

1. Viết chương trình kết nối camera Kinect, hiển thị ảnh theo độ sâu của bàn tay.
2. Kiểm tra thuật toán với bàn tay không mang găng tay.

Chương trình cài đặt thuật toán cây quyết định ngẫu nhiên sẽ nhận được tư thế bàn tay từ các ảnh chiều sâu và hiển thị hình ảnh RGB phân biệt các bộ phận của bàn tay được lấy ra từ bộ huấn luyện.



Hình 6: Thuật toán Random Forests²³

Thuật toán xử lý và hiển thị hình ảnh bàn tay

Sau khi nhận dạng bàn tay, hình ảnh hiển thị là hình bàn tay có dán nhãn các bộ phận của bàn tay. Với những thuật toán theo dõi đối tượng, ta có thể theo dõi các chuyển động khác nhau của bàn tay với các nhãn phân biệt từng bộ phận của bàn tay. Từ đó, ta có thể phát triển được ứng dụng để nhận dạng cử chỉ bàn tay.

KẾT QUẢ NGHIÊN CỨU THỰC NGHIỆM VÀ THẢO LUẬN

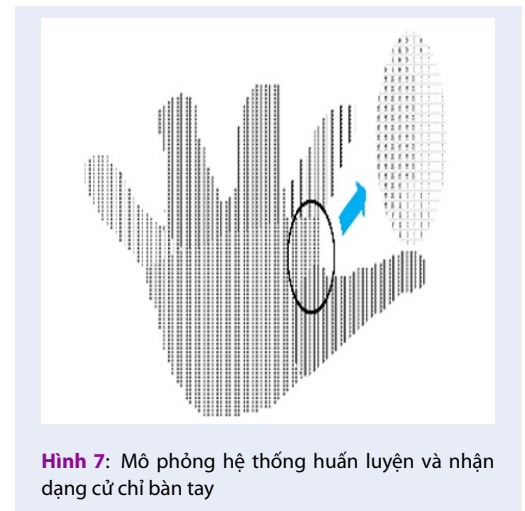
Mô tả dữ liệu huấn luyện

Bộ dữ liệu dùng để huấn luyện và thử nghiệm là bộ dữ liệu chúng tôi thu thập được khoảng 10000 mẫu bàn tay. Mỗi mẫu gồm 1 ảnh màu bàn tay và 1 ảnh lưu dữ liệu độ sâu với độ phân giải 320×240 , hình được chụp trong bất kể điều kiện ánh sáng và ở nhiều vị trí khác nhau. Hình ảnh bàn tay chụp được đảm bảo phải lưu được giá trị ảnh chiều sâu của bàn tay.

Với mỗi mẫu dữ liệu thì hình ảnh bàn tay đã được đánh nhãn cho từng bộ phận của bàn tay và nó được xử lý để loại những đối tượng không cần thiết, chỉ giữ lại đối tượng bàn tay. Cũng tương tự như vậy file ảnh lưu độ sâu chỉ lưu giá trị độ sâu của đối tượng bàn tay và loại bỏ những giá trị của đối tượng khác. Hình 7 mô phỏng hệ thống huấn luyện và nhận dạng cử chỉ bàn tay.

Mô tả hệ thống

Đây là một hệ thống mà quá trình xử lý bao gồm quá trình huấn luyện với ảnh đầu vào được xử lý và chọn lọc ra những đặc trưng nhận dạng (Hình 8). Kết quả đầu ra của quá trình xử lý là quá trình nhận dạng hình ảnh để đưa ra hình ảnh được nhận dạng chính xác từng bộ phận của đối tượng.

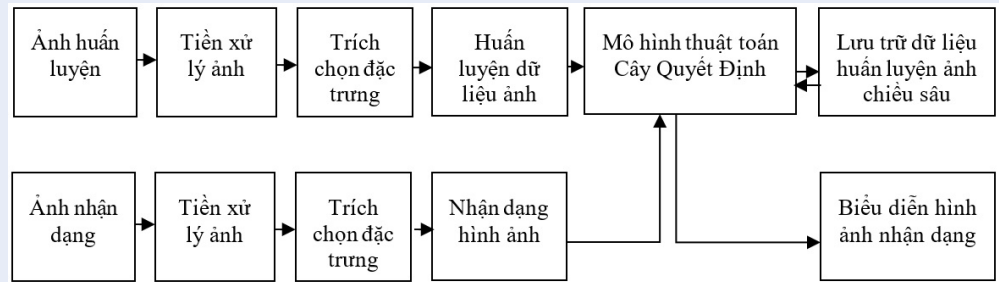


Hình 7: Mô phỏng hệ thống huấn luyện và nhận dạng cử chỉ bàn tay

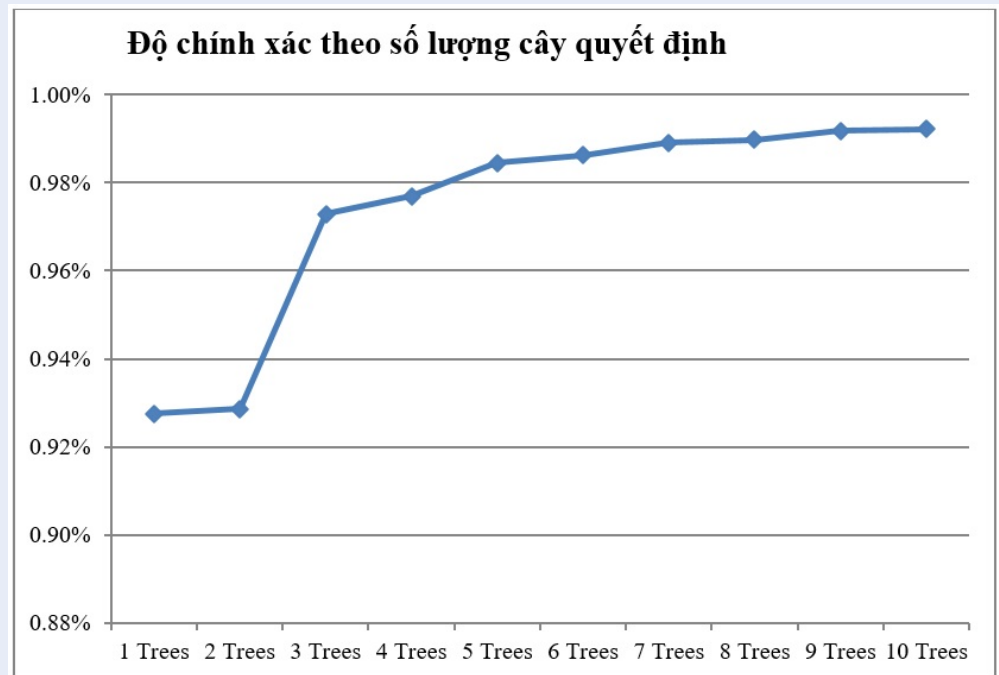
Độ chính xác của thuật toán Random Forest Theo số lượng cây huấn luyện

Để phân loại các điểm ảnh x trong hình ảnh, người ta bắt đầu từ gốc và sau đó thực hiện nhiều lần đánh giá để xác định đối tượng cần nhận dạng. Ở đây, chúng tôi đã đánh giá độ chính xác của thuật toán khi cấu huấn luyện từ 1 cho đến nhiều cây quyết định. Từ đó ta có thể thấy được hiệu quả và nhanh chóng của thuật toán cây quyết định ngẫu nhiên.

Theo Hình 9, độ chính xác của thuật toán cây định ngẫu nhiên còn bị ảnh hưởng rất lớn với số lượng cây huấn luyện khác nhau. Ở đây, theo thử nghiệm số lượng cây huấn càng nhiều thì chính xác trong việc phân loại mô hình càng cao. Nhưng với số lượng cây huấn luyện càng nhiều thì độ phức tạp trong nhận dạng cũng như thời gian xử lý càng lớn.



Hình 8: Mô phỏng hệ thống huấn luyện và nhận dạng chữ chỉ bàn tay



Hình 9: Độ chính xác của thuật toán theo số lượng cây huấn luyện

Theo độ sâu của cây huấn luyện

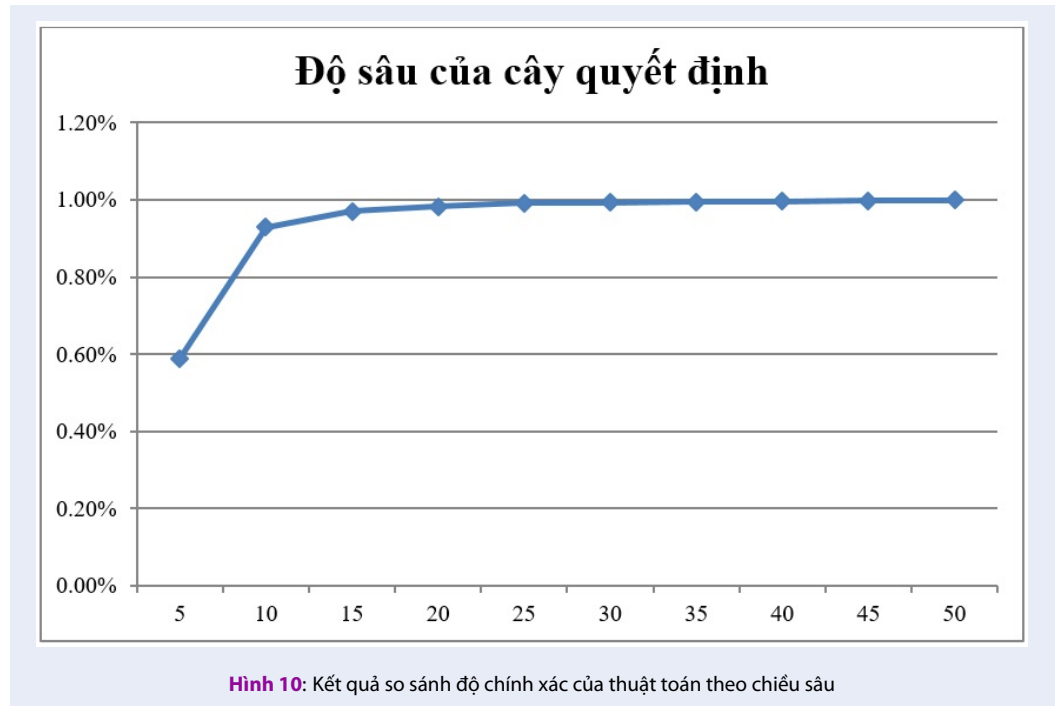
Độ sâu của cây quyết định cũng ảnh hưởng đến độ chính xác khi chúng tôi kiểm tra bằng cách sử dụng 10000 hình ảnh theo chiều sâu để nhận dạng. Trong tất cả các thông số của thuật toán cây quyết định ngẫu nhiên, độ sâu đường như cũng có ảnh hưởng quan trọng vì nó trực tiếp tác động đến khả năng phân loại mô hình. Ở đây, chúng tôi chỉ thiết lập 10 cây huấn luyện và quan sát khi thay đổi khoảng chiều sâu từ 5 cho đến 50. Quan sát, chúng tôi thấy độ chính xác lớn hơn và kết quả tốt hơn có thể đạt được khi đào tạo cây huấn luyện với độ sâu lớn. Nhưng tại một thời gian chạy với độ sâu càng lớn thì thời gian và chi phí tính toán nhận dạng càng lớn. Ta có tham khảo lưu đồ như Hình 10.

Độ chính xác của ảnh chiều sâu

Ảnh theo chiều sâu là ảnh lưu các giá trị mô tả khoảng cách từ camera tới các đối tượng được thu ảnh. Bộ ảnh thu thập được chia ra thành nhiều nhóm con theo mỗi loại chữ và đưa vào tập huấn luyện. Qua bước huấn luyện, chúng tôi tạo ra 01 tệp để ghi lại các đặc trưng đã trích chọn từ dữ liệu chiều sâu của các tư thế bàn tay cơ bản. Từ đó, chúng tôi sử dụng nó làm dữ liệu nhận dạng chữ chỉ bàn tay.

Khả năng phân biệt các bộ phận của bàn tay

Với dữ liệu được đánh nhãn phân biệt các bộ phận khác nhau của bàn tay và đưa vào huấn luyện với thuật toán cây quyết định ngẫu nhiên (Hình 11), ta có thể



nhận dạng được các thành phần của bàn tay khi đưa dữ liệu theo chiều sâu của bàn tay vào nhận dạng. Nhưng mức độ phân biệt chi tiết rõ ràng từng thành phần của bàn tay dính nhân còn phụ thuộc từng tư thế của bàn tay khác nhau.

Trong quá trình thực nghiệm nhận dạng trên nhiều mẫu ảnh, chúng tôi nhận thấy rằng bất kể điều kiện ánh sáng khác nhau thì kết quả nhận dạng vẫn rất tốt đạt độ chính xác đạt 60 - 100%. Yếu tố ở đây làm ảnh hưởng đến độ chính xác của quá trình nhận dạng phụ thuộc nhiều bởi các tư thế bàn tay khác nhau. Với những tư thế bàn tay có bề mặt quay vào hướng camera càng lớn thì camera đo sâu sẽ lấy được giá trị theo độ sâu càng nhiều và dẫn đến khả năng nhận dạng tư thế bàn tay tốt hơn. Tuy nhiên, với những ảnh theo chiều sâu mà bề mặt vật thể có độ bằng phẳng cao thì hệ thống nhận dạng rất kém vì nhận dạng vật thể dựa trên giá trị theo chiều sâu tức khoảng cách từ 1 pixel đến camera kinect. Hình 12 biểu diễn đồ thị biểu diễn độ chính xác của từng mẫu bàn tay.

Khả năng phân biệt hình dạng cử chỉ bàn tay

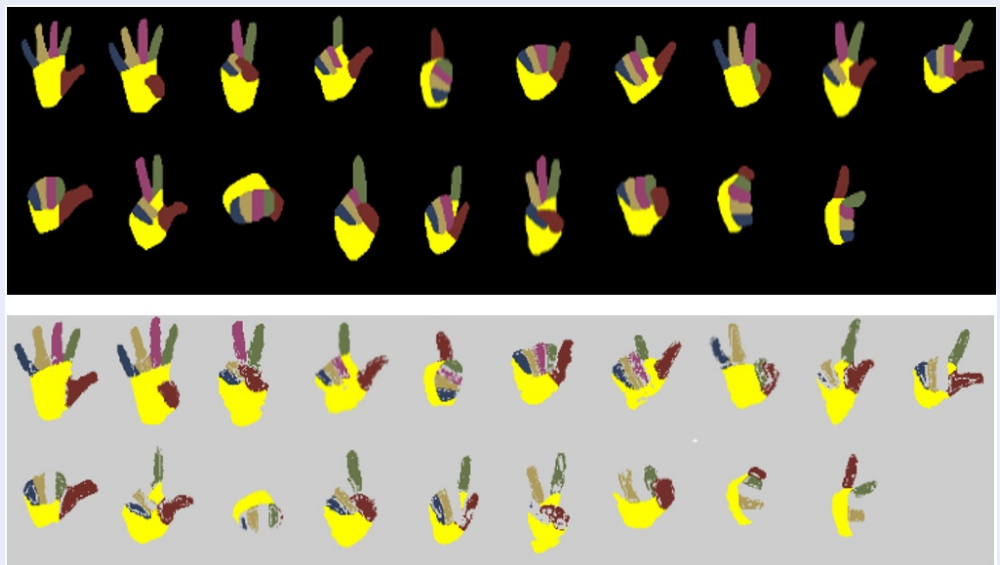
Với ảnh theo chiều sâu, khả năng nhận dạng hình dáng của vật thể trong mọi điều kiện ánh sáng khác nhau là rất tốt. Hệ thống chỉ bị nhiễu khi camera không đo được khoảng cách từ camera đến vật thể, đó là những vật thể bị che khuất hoặc nằm ngoài vùng đo của camera. Thế nên, trong mọi điều kiện ánh sáng khác nhau và vật thể không bị che khuất thì đều có

thể nhận dạng được rõ ràng hình dạng của bàn tay (Hình 13).

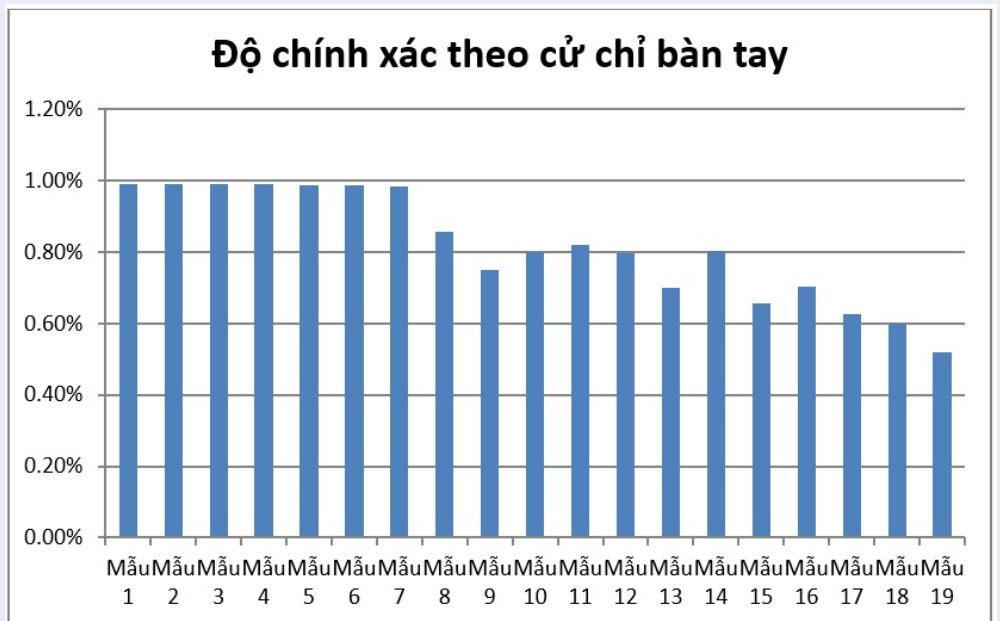
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Bài viết nghiên cứu những thuật toán trong nhận dạng xử lý ảnh theo chiều sâu để áp dụng vào việc nhận dạng tư thế cử chỉ và phân biệt các thành phần của bàn tay. Từ đó, ta có thể áp dụng để đưa các kỹ thuật nhận dạng đếm số lượng ngón tay qua cử chỉ bàn tay hoặc có thể điều khiển cánh tay robot ứng dụng trong sản xuất và trong đời sống. Cụ thể, những đóng góp chính của nghiên cứu là:

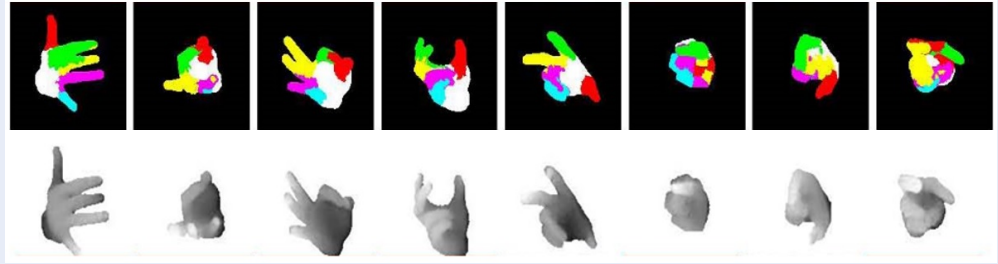
1. Trong quá trình thực nghiệm nhận dạng trên nhiều mẫu ảnh, chúng tôi nhận thấy với những ảnh chụp trong mọi điều kiện ánh sáng với camera theo chiều sâu thì kết quả nhận dạng sử dụng thuật toán cây quyết định ngẫu nhiên đạt độ chính xác đạt gần 80% và tốc độ nhận dạng cũng được cải thiện đáng kể.
2. Về mặt ưu điểm của phương pháp này là nhận dạng cử chỉ bàn tay từ ảnh theo chiều sâu nên trong bất kỳ điều kiện ánh sáng nào cũng có thể nhận dạng được. Ngoài ra, trong phương pháp này sử dụng thuật toán cây quyết định ngẫu nhiên nên có khả năng huấn luyện dữ liệu lớn, tốc độ nhận dạng nhanh.
3. Về khuyết điểm của phương pháp này là việc đánh nhãn để phân biệt các bộ phận trên trên



Hình 11: Nhận dạng các thành phần của bàn tay



Hình 12: Đồ thị biểu diễn độ chính xác của từng mẫu bàn tay



Hình 13: Nhận dạng cử chỉ bàn tay theo ảnh chiều sâu

bàn tay khi nhận dạng có hiện tượng nhiễu. Đặc biệt là với những tư thế bàn tay có diện tích camera thu được nhỏ thì mức độ nhiễu cao.

Với kết quả đạt được từ phương pháp nhận dạng ảnh chiều sâu sử dụng thuật toán cây quyết định ngẫu nhiên, chúng tôi đã hiểu rõ cơ sở lý thuyết, ưu điểm và nhược điểm về kỹ thuật xử lý và nhận dạng ảnh theo chiều sâu và xây dựng được bộ dữ liệu ảnh theo chiều sâu của các tư thế bàn tay. Từ đó, làm tiền đề để chúng tôi có thể tiếp tục cải thiện khuyết điểm của phương pháp này và có thể áp dụng nó trong các ứng dụng điều khiển phục vụ trong đời sống, trong sản xuất công nghiệp. Đây là những hướng có thể dự định phát triển trong tương lai:

1. Xây dựng các mô phỏng tương tác giữa người và môi trường ảo, ứng dụng tương tác giữa bàn tay con người với các vật thể ba chiều, giả lập các hiện tượng vật lý
2. Xây dựng chơi game dùng cơ chế giao tiếp thông qua cử chỉ, ứng dụng để có thể kết nối điều khiển robot, ứng dụng dùng điều khiển các cánh tay robot trong các khu công nghiệp.
3. Xây dựng ứng dụng phục vụ cho y học: điều khiển robot từ xa thông qua việc bắt chước cử chỉ bàn tay của phẫu thuật viên

DANH MỤC TỪ VIẾT TẮT

RGB: Red Green Blue
TOF: Time-Of-Flight
GPU: Graphics Processing Unit

XUNG ĐỘT LỢI ÍCH

Nhóm tác giả xin cam đoan rằng không có bất kỳ xung đột lợi ích nào trong công bố bài báo.

ĐÓNG GÓP CỦA CÁC TÁC GIẢ

Tác giả Phạm Minh Lộc tham gia vào việc đưa ra ý tưởng, thu thập dữ liệu, thực nghiệm mô hình và đánh giá kết quả. Tác giả Nguyễn Văn Hồ tham gia vào việc đưa ra ý tưởng viết bài, tổng hợp và viết bản thảo.

LỜI CẢM ƠN

Nhóm tác giả xin gửi lời cảm ơn chân thành đến Tiến sĩ Nguyễn Đức Thắng đã có nhiều ý kiến hỗ trợ và hướng dẫn nhóm hoàn thành bài báo này.

TÀI LIỆU THAM KHẢO

1. Lao F, Brown-Brandt T, Stinn JP, Liu K, Teng G, Xin H. Automatic recognition of lactating sow behaviors through depth image processing. *Computers and Electronics in Agriculture*. 2016 Jul 1;125:56-62; Available from: <https://doi.org/10.1016/j.compag.2016.04.026>.
2. Ghamisi P, Yokoya N, Li J, Liao W, Liu S, Plaza J, Rasti B, Plaza A. Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art. *IEEE Geoscience and Remote Sensing Magazine*. 2017 Dec;5(4):37-78; Available from: <https://doi.org/10.1109/MGRS.2017.2762087>.
3. Monga V, Li Y, Eldar YC. Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Processing Magazine*. 2021 Feb 25;38(2):18-44; Available from: <https://doi.org/10.1109/MSP.2020.3016905>.
4. Grest D, Woetzel J, Koch R. Nonlinear body pose estimation from depth images. In *Joint Pattern Recognition Symposium 2005 Aug 31 (pp. 285-292)*. Springer, Berlin, Heidelberg; Available from: https://doi.org/10.1007/11550518_36.
5. Shotton J, Sharp T, Kipman A, Fitzgibbon A, Finocchio M, Blake A, Cook M, Moore R. Real-time human pose recognition in parts from single depth images. *Communications of the ACM*. 2013 Jan 1;56(1):116-24; Available from: <https://doi.org/10.1145/2398356.2398381>.
6. Wu M, Chen L. Image recognition based on deep learning. In *2015 Chinese Automation Congress (CAC) 2015 Nov 27 (pp. 542-546)*. IEEE;
7. Shaikh MB, Chai D. RGB-D Data-based Action Recognition: A Review. *Sensors*. 2021 Jan;21(12):4246; Available from: <https://doi.org/10.3390/s21124246>.
8. Jalal A, Sarwar MZ, Kim K. RGB-D Images for Objects Recognition using 3D Point Clouds and RANSAC Plane Fitting. In *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST) 2021 Jan 12 (pp. 518-523)*. IEEE; Available from: <https://doi.org/10.1109/IBCAST51254.2021.9393166>.
9. Liu Y, Ma R, Li H, Wang C, Tao Y. RGB-D Human Action Recognition of Deep Feature Enhancement and Fusion Using Two-Stream ConvNet. *Journal of Sensors*. 2021 Jan 7;2021; Available from: <https://doi.org/10.1155/2021/8864870>.
10. Ioffe S, Forsyth DA. Probabilistic methods for finding people. *International Journal of Computer Vision*. 2001 Jun;43(1):45-68; Available from: <https://doi.org/10.1023/A:1011179004708>.

11. Bourdev L, Malik J. Poselets: Body part detectors trained using 3d human pose annotations. In 2009 IEEE 12th International Conference on Computer Vision 2009 Sep 27 (pp. 1365-1372). IEEE; Available from: <https://doi.org/10.1109/ICCV.2009.5459303>.
12. Park C, Huang JZ, Ji JX, Ding Y. Segmentation, inference and classification of partially overlapping nanoparticles. IEEE transactions on pattern analysis and machine intelligence. 2012 Aug 1;35(3):1-; PMID: 22848127. Available from: <https://doi.org/10.1109/TPAMI.2012.163>.
13. Kotsiantis SB. Decision trees: a recent overview. Artificial Intelligence Review. 2013 Apr;39(4):261-83; Available from: <https://doi.org/10.1007/s10462-011-9272-4>.
14. Smisek J, Jancosek M, Pajdla T. 3D with Kinect. In Consumer depth cameras for computer vision 2013 (pp. 3-25). Springer, London; Available from: https://doi.org/10.1007/978-1-4471-4640-7_1.
15. Liu J, Liu Y, Zhang G, Zhu P, Chen YQ. Detecting and tracking people in real time with RGB-D camera. Pattern Recognition Letters. 2015 Feb 1;53:16-23; Available from: <https://doi.org/10.1016/j.patrec.2014.09.013>.
16. Liang H, Yuan J, Thalmann D. 3D fingertip and palm tracking in depth image sequences. In Proceedings of the 20th ACM international conference on Multimedia 2012 Oct 29 (pp. 785-788); Available from: <https://doi.org/10.1145/2393347.2396312>.
17. Shepherd BA. An Appraisal of a Decision Tree approach to Image Classification. In IJCAI 1983 Aug 8 (pp. 473-475);
18. Breiman L. Random forests. Machine learning. 2001 Oct;45(1):5-32; Available from: <https://doi.org/10.1023/A:1010933404324>.
19. Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, Rigol-Sanchez JP. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing. 2012 Jan 1;67:93-104; Available from: <https://doi.org/10.1016/j.isprsjrs.2011.11.002>.
20. Lepetit V, Laguerre P, Fua P. Randomized trees for real-time key-point recognition. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) 2005 Jun 20 (Vol. 2, pp. 775-781). IEEE;
21. Moosmann F, Triggs B, Jurie F. Fast discriminative visual codebooks using randomized clustering forests. In Twentieth Annual Conference on Neural Information Processing Systems (NIPS'06) 2006 Dec 4 (pp. 985-992). MIT Press;
22. Shotton J, Johnson M, Cipolla R. Semantic texture forests for image categorization and segmentation. In 2008 IEEE conference on computer vision and pattern recognition 2008 Jun 23 (pp. 1-8). IEEE; Available from: <https://doi.org/10.1109/CVPR.2008.4587503>.
23. Sharp T. Implementing decision trees and forests on a GPU. In European conference on computer vision 2008 Oct 12 (pp. 595-608). Springer, Berlin, Heidelberg; Available from: https://doi.org/10.1007/978-3-540-88693-8_44.

Applying machine learning for hand posture recognition from a series of depth images

Pham Minh Loc¹, Nguyen Van Ho^{2,*}



Use your smartphone to scan this QR code and download this article

ABSTRACT

Image recognition one of the most difficult issue, is becoming an applied research in recent times. The possibility of allowing computers to recognize human's hand posture in real-time is gradually becoming popular and involved in many fields. It is also integrated into so many products. Thus, this study proposes an approach for faster and more accurate image recognition. With the development of technology, this camera device has had many significant changes; the camera can capture both RGB images and images in depth. This makes identification in low-light conditions more accurate, which improves recognition performance. The Random Forest algorithm is evaluated to be more efficient, faster, and more accurate than previous recognition algorithms used for model testing. We collected about 10,000 hand image data samples, then trained and evaluated the model with nearly 80% accuracy results. In addition, the study presents some results of gesture recognition of human hands in real-time using information obtained from Kinect sensor; the model has the ability to train quite large data and the recognition speed is faster. Research results can be used in control applications serving in life and industrial production.

Key words: depth image, machine learning, random forest, image recognition

¹Vietnam USA Society English Centers,
Ho Chi Minh City, Việt Nam

²Faculty of Information Systems,
University of Economics and Law,
VNU-HCM, Việt Nam

Correspondence

Nguyen Van Ho, Faculty of Information
Systems, University of Economics and
Law, VNU-HCM

Email: honv@uel.edu.vn

History

- Received: 01-7-2021
- Accepted: 11-02-2022
- Published: 11-4-2022

DOI : 10.32508/stdjelm.v6i2.870



Copyright

© VNUHCM Press. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.



Cite this article : Loc P M, Ho N V. **Applying machine learning for hand posture recognition from a series of depth images.** *Sci. Tech. Dev. J. - Eco. Law Manag.*; 6(2):2373-2385.

Tạp chí Phát triển Khoa học và Công nghệ Đại học Quốc gia Tp. Hồ Chí Minh



Tạp chí Phát triển Khoa học và Công nghệ

ISSN: 1859-0128

Hình thức xuất bản: In và trực tuyến

Hình thức truy cập: Truy cập mở (Open Access)

Ngôn ngữ bài báo: Tiếng Anh

Tỉ lệ chấp nhận đăng 2021: 72%

Phí xuất bản: Miễn phí

Thời gian phản biện: 43 ngày

Lập chỉ mục (Indexed): Google Scholar, Scilit



SCAN ME



**Tạp chí Phát triển Khoa học và Công nghệ -
Khoa học Tự nhiên**

ISSN: 2588-106X

Hình thức xuất bản: In & trực tuyến

Hình thức truy cập: Truy cập mở

Ngôn ngữ bài báo: Tiếng Việt

Tỉ lệ chấp nhận đăng 2021: 75%

Phí xuất bản: Miễn phí

Thời gian phản biện: 30-45 ngày

Lập chỉ mục (Indexed): Google Scholar, Scilit



SCAN ME



**Tạp chí Phát triển Khoa học và Công nghệ -
Kỹ thuật và Công nghệ**

ISSN: 2615-9872

Hình thức xuất bản: In & trực tuyến

Hình thức truy cập: Truy cập mở

Ngôn ngữ bài báo: Tiếng Việt

Tỉ lệ chấp nhận đăng 2021: 61%

Phí xuất bản: Miễn phí

Thời gian phản biện: 50 ngày

Lập chỉ mục (Indexed): Google Scholar, Scilit



SCAN ME



**Tạp chí Phát triển Khoa học và Công nghệ -
Kinh tế-Luật và Quản lý**

ISSN: 2588-1051

Hình thức xuất bản: In & trực tuyến

Hình thức truy cập: Truy cập mở

Ngôn ngữ bài báo: Tiếng Việt

Tỉ lệ chấp nhận đăng 2021: 65%

Phí xuất bản: Miễn phí

Thời gian phản biện: 45 ngày

Lập chỉ mục (Indexed): Google Scholar, Scilit



SCAN ME



**Tạp chí Phát triển Khoa học và Công nghệ -
Khoa học Xã hội và Nhân văn**

ISSN: 2588-1043

Hình thức xuất bản: In & trực tuyến

Hình thức truy cập: Truy cập mở

Ngôn ngữ bài báo: Tiếng Việt

Tỉ lệ chấp nhận đăng 2021: 62%

Phí xuất bản: Miễn thu phí đối với tác giả là CBVC của ĐHKHXHNV, ĐHQG-HCM; Tác giả khác: 500.000 VNĐ/bài

Thời gian phản biện: 75 ngày

Lập chỉ mục (Indexed): Google Scholar, Scilit



SCAN ME



**Tạp chí Phát triển Khoa học và Công nghệ -
Khoa học Trái đất và Môi trường**

ISSN: 2588-1078

Hình thức xuất bản: In & trực tuyến

Hình thức truy cập: Truy cập mở

Ngôn ngữ bài báo: Tiếng Việt và tiếng Anh

Tỉ lệ chấp nhận đăng 2021: 87%

Phí xuất bản: liên hệ tòa soạn

Thời gian phản biện: 45 ngày

Lập chỉ mục (Indexed): Google Scholar, Scilit



SCAN ME



**Tạp chí Phát triển Khoa học và Công nghệ -
Khoa học Sức khỏe**

ISSN: 2734-9446

Hình thức xuất bản: In & trực tuyến

Hình thức truy cập: Truy cập mở

Ngôn ngữ bài báo: Tiếng Việt

Tỉ lệ chấp nhận đăng 2021: 70%

Phí xuất bản: Miễn phí

Thời gian phản biện: 30 ngày

Lập chỉ mục (Indexed): Google Scholar, Scilit



SCAN ME

Tạp chí Phát triển Khoa học và Công nghệ, Đại học Quốc gia Tp.HCM

25 năm xuất bản học thuật (1997-2022)

Tòa soạn: Nhà điều hành Đại học Quốc gia Tp.HCM, P. Linh Trung, TP. Thủ Đức, TP. HCM

Email: stj@vnuhcm.edu.vn; tcptkcn@vnuhcm.edu.vn; Website: <http://www.scienceandtechnology.com.vn>