

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC ĐÀ NẴNG**

ĐỀ CƯƠNG
LUẬN VĂN THẠC SĨ KỸ THUẬT
NGÀNH KHOA HỌC MÁY TÍNH

TÊN ĐỀ TÀI

**NGHIÊN CỨU TRÍCH CHỌN ĐẶC TÍNH
TRONG NHẬN DẠNG HÀNH ĐỘNG NGƯỜI
TRONG KHÔNG GIAN 3D**

**Họ và tên học viên
Cán bộ hướng dẫn
Lớp Cao học**

**: Ung Nho Dãi
: TS. Phạm Minh Tuấn
: Khoá 27 (2013 - 2015)**

Đà Nẵng, 10/2014

I. TÍNH CẤP THIẾT CỦA ĐỀ TÀI

1. Bối cảnh chung

Từ những năm 80 của thế kỷ trước, nhận dạng hành động (activity recognition) [1] đã thu hút rất nhiều quan tâm, nghiên cứu của các nhà khoa học. Nó được sử dụng rộng rãi trong nhiều ứng dụng và trong các lĩnh vực khác như y học, xã hội học, giao tiếp người máy. Nhận dạng hành động được chia làm hai loại chính: nhận dạng dựa vào cảm biến (sensor-based) và nhận dạng dựa vào hình ảnh (vision-based) [2].

Hoạt động nhận dạng sử dụng cảm biến kết hợp sự đa dạng của mạng lưới thiết bị cảm biến với việc khai phá dữ liệu và học máy để mô hình hoạt động của con người. Cấu hình các thiết bị di động hiện nay đủ mạnh để thu thập dữ liệu từ nhiều loại cảm biến khác nhau và xử lý các dữ liệu đó để có thể đưa ra ước lượng về năng lượng cần thiết cho các hoạt động hàng ngày của con người. Các nhà nghiên cứu tin rằng, với sự phát triển mạnh mẽ của các loại thiết bị và các loại cảm biến, việc theo dõi và nhận dạng hoạt động của con người sẽ trở nên dễ dàng hơn [3].

Vấn đề quan trọng và thách thức nhất đối với nhận dạng hành động là nhận biết được hành động của con người thông qua các hình ảnh từ hệ thống các máy quay. Kỹ thuật chủ yếu được dùng để nhận dạng từ hình ảnh là thị giác máy tính. Có rất nhiều phương pháp đã được áp dụng trong nhận dạng hành động dựa vào hình ảnh như optical flow, bộ lọc Kalman, mô hình Markov ẩn, ... sử dụng các dữ liệu khác nhau từ camera, sóng âm (stereo) và hồng ngoại [1, 2, 4].



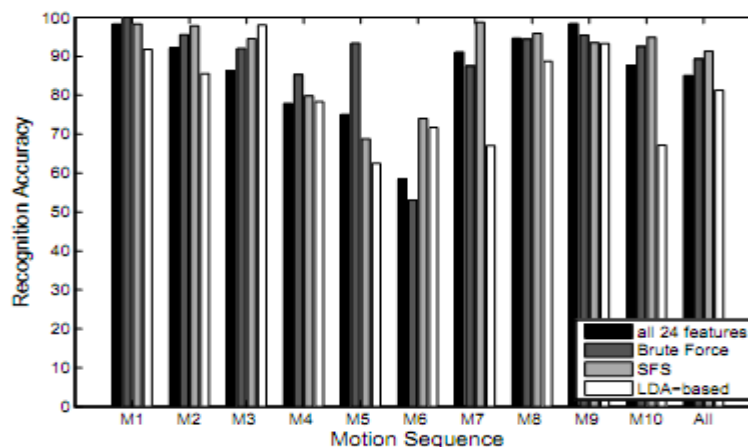
Hình 1. Microsoft Kinect Camera

Gần đây, một số nhà nghiên cứu đã sử dụng camera RGBD (Red, Green, Blue, Depth) như Microsoft Kinect [3] (xem hình 1) để nhận dạng hoạt động của con người. Dữ liệu thu được từ các thiết bị chuyên dụng này là dữ liệu chuyển động 3D của cơ thể người. Những dữ liệu này sẽ là dữ liệu huấn luyện hữu ích cho các mô hình nhận dạng hành động.

2. Các phương pháp trước đây

Những nghiên cứu gần đây trong lĩnh vực nhận dạng hoạt động người chủ yếu tập trung vào nghiên cứu và nhận dạng từ những video được quay bởi các camera thông dụng [2, 4]. Khó khăn lớn nhất đối với dữ liệu từ camera thông dụng là chỉ quay được ở một hướng, dẫn đến sự thiếu hụt dữ liệu, nếu kết hợp nhiều camera thì vẫn không đảm bảo thu được toàn bộ hoạt động, đồng thời giảm hiệu năng của quá trình nhận dạng. Mặc dù đã có rất nhiều nỗ lực trong những thập kỷ qua, lĩnh vực nhận dạng hoạt động người từ dữ liệu video vẫn còn nhiều khó khăn, thách thức.

Từ sau sự ra đời của các thiết bị cảm biến chiều sâu (depth sensor) [3], đã có một hướng tiếp cận mới trong nhận dạng hành động người, đó là sử dụng dữ liệu chuyển động 3D [5, 6, 7]. Trong 20 năm trở lại, một số phương pháp chính để thu thập dữ liệu 3D hoặc là sử dụng hệ thống chụp chuyển động dựa vào marker như là Mocap (Motion Capture) [8] hoặc là dùng sóng âm (stereo) - chụp hình ảnh 2D từ nhiều hướng khác nhau để dựng thành mô hình 3D [9].



Hình 2. So sánh kết quả giữa các phương pháp trích chọn đặc tính

Sau khi đã thu thập được dữ liệu 3D, có rất nhiều phương pháp đã được đề xuất để hoàn thành quá trình nhận dạng. Điểm chung của các phương pháp này là cố gắng làm giảm số lượng thuộc tính của dữ liệu nhận dạng trước khi xây dựng mô hình huấn luyện. D. Gehrig và T. Schultz đã nghiên cứu, thực nghiệm ba phương pháp trích chọn và lựa chọn đặc tính khác nhau (Brute Force, SFS, LDA) và đã giảm đáng kể hiệu năng của quá trình nhận dạng so với dữ liệu ban đầu [10] (xem hình 2).

Fengjun Lv và RamaKant Nevatia phân tích và đưa ra bảy loại đặc tính khác nhau dựa vào tư thế và sự kết hợp giữa các khớp xương, sau đó dựng mô hình huấn luyện và nhận dạng dùng Markov ẩn [5] (xem hình 3). Dana Kulic, Wataru Takano và Yoshihiko Nakamura đã phát triển một mô hình học tăng cường mới dựa trên mô hình Markov ẩn [11].

action	walk	run	j upward	j forward	stand	sit	bow	lie
Exp.1	94.1%	95.5%	92.2%	91.2%	91.8%	92.4%	89.8%	88.7%
Exp.2	89.0%	91.3%	87.3%	86.6%	87.9%	90.5%	86.0%	84.8%
action	stand2sit	sit2stand	stand2bow	bow2stand	stand2lie	lie2stand	sit2lie	lie2sit
Exp.1	89.7%	89.8%	89.0%	88.3%	92.4%	88.2%	91.2%	91.8%
Exp.2	84.7%	86.6%	84.8%	86.5%	88.7%	84.5%	86.6%	86.1%
action	wave hand	point	lower arm	lift arm	nod	shake head		
Exp.1	95.8%	94.2%	92.7%	92.3%	97.9%	96.7%		
Exp.2	91.3%	92.8%	89.2%	89.4%	95.1%	94.8%		

Hình 3. Kết quả nhận dạng bảy loại đặc tính do Fengjun Lv và RamaKant Neva đề xuất

3. Những vấn đề tồn tại

Việc xây dựng mô hình nhận dạng sử dụng dữ liệu chuyển động 3D vẫn còn nhiều điểm chưa tốt về hiệu năng cũng như chi phí.

- Dữ liệu chuyển động 3D là dữ liệu phức tạp, có số lượng thuộc tính lớn
- Chi phí tính toán lớn
- Hiệu năng không cao
- Tỷ lệ nhận dạng đúng thấp, đặc biệt với những hoạt động phức tạp

Do vậy giải pháp đề xuất là nghiên cứu, thực nghiệm và sánh các phương pháp trích chọn và lựa chọn đặc tính khác nhau, từ đó đưa ra phương pháp hiệu quả nhất cho bài toán nhận dạng hoạt động người, góp phần nâng cao hiệu năng cũng như độ chính xác của mô hình.

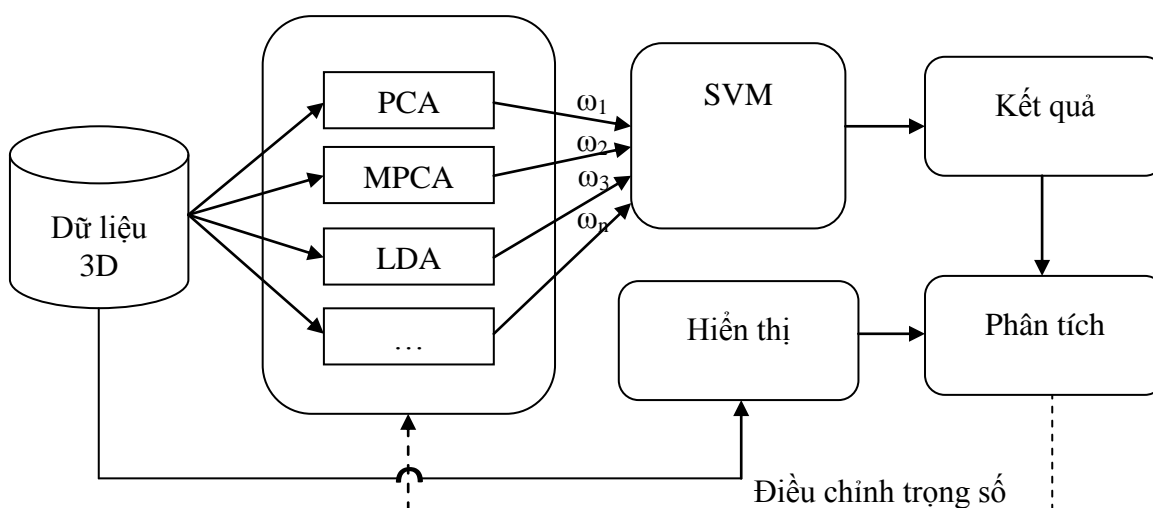
II. MỤC ĐÍCH NGHIÊN CỨU

Mục đích của đề tài là nghiên cứu đề xuất phương pháp trích chọn đặc tính tối ưu nhất cho mô hình nhận dạng hành động người dựa trên cấu trúc Acclaim.

Các hoạt động cụ thể:

- Nghiên cứu cấu trúc Acclaim, xây dựng mô hình chuyển động 3D từ cấu trúc Acclaim
- Sử dụng các phương pháp trích chọn đặc tính (feature extraction) và lựa chọn đặc tính (feature selection) làm giảm độ phức tạp và độ lớn của dữ liệu
- Sử dụng mô hình SVM (Support Vector Machine) xây dựng mô hình nhận dạng hành động người dựa trên dữ liệu trích chọn được; trên cơ sở đó thực nghiệm và so sánh kết quả giữa các phương pháp nhằm đề xuất phương án tốt nhất

Mô hình đề xuất



Hình 2. Mô hình đề xuất

III. ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU

1. Đối tượng nghiên cứu

- Cấu trúc Acclaim
- Các thư viện đồ họa 3D (OpenGL, DirectX)
- Các phương pháp trích chọn đặc tính
- Mô hình SVM

2. Phạm vi nghiên cứu

- Dữ liệu đầu vào của đề tài này là dữ liệu trên cấu trúc Acclaim (asf/amc) được xây dựng và công bố bởi Đại học Carnegie Mellon (CMU) [8]
- Mô hình nhận dạng sử dụng trong đề tài này là mô hình SVM
- Phương pháp trích chọn đặc tính sử dụng trong đề tài này là: PCA, MPCA, LDA

IV. PHƯƠNG PHÁP NGHIÊN CỨU

1. Nghiên cứu lý thuyết

- Cấu trúc Acclaim (asf/amc)
- Các thư viện đồ họa 3D: OpenGL, DirectX
- Các phương pháp trích chọn đặc tính: PCA, MPCA, LDA
- Mô hình SVM

2. Nghiên cứu thực nghiệm

- Thực nghiệm trên mô hình nhận dạng với nhiều dữ liệu khác nhau và thu thập, phân tích kết quả
- So sánh, đánh giá các phương pháp trích chọn đặc tính khác nhau

V. Ý NGHĨA KHOA HỌC VÀ THỰC TIỄN CỦA ĐỀ TÀI

1. Về mặt lý thuyết

- Lý thuyết về nhận dạng
- Lý thuyết về cách sử dụng cấu trúc Acclaim
- Các lý thuyết trích chọn đặc tính trong dữ liệu 3D

2. Về mặt thực tiễn

- Góp phần hoàn thiện và nâng cao hiệu quả, độ chính xác của mô hình nhận dạng hành động người
- Xây dựng được ứng dụng nhận dạng hành động người với độ chính xác cao

VI. DỰ KIẾN BỐ CỤC CỦA LUẬN VĂN

MỤC LỤC

MỞ ĐẦU

- Tính cấp thiết của đề tài
- Mục đích nghiên cứu
- Đối tượng và phạm vi nghiên cứu
- Phương pháp nghiên cứu
- Ý nghĩa khoa học và thực tiễn của đề tài
- Bố cục đề tài

NỘI DUNG

Chương I. Nghiên cứu tổng quan

1.1 Cấu trúc Acclaim – asf/amc

- 1.1.1 Thí nghiệm chụp chuyển động
- 1.1.2 Cấu trúc tệp ASF
- 1.1.3 Cấu trúc tệp AMC

1.2 Dựng chuyển động 3D

- 1.2.1 Thư viện OpenGL, DirectX
- 1.2.2 Dựng mô hình 3D từ tệp ASF
- 1.2.3 Dựng chuyển động 3d từ tệp AMC

1.3 Các phương pháp trích chọn đặc tính

- 1.3.1 Phương pháp phân tích thành phần chính – PCA
- 1.3.2 Phương pháp phân tích thành phần chính đa tuyến - MPCA
- 1.3.3 Phương pháp phân tích biệt thức tuyến tính - LDA
- 1.3.4 Phương pháp lựa chọn đặc tính (feature selection)

1.4 Mô hình SVM

- 1.4.1 Khái niệm và cơ sở của thuật toán SVM
- 1.4.2 Ứng dụng của mô hình SVM

Chương 2. Giải pháp đề xuất

2.1 Mô hình tổng quan

2.2 Trích chọn đặc tính

2.3 Giải thuật huấn luyện

Chương 3. Triển khai và đánh giá kết quả

3.1 Dữ liệu nhận dạng

3.2 Môi trường triển khai

3.3 Kết quả thực nghiệm

3.4 Đánh giá

KẾT LUẬN VÀ KIẾN NGHỊ

DANH MỤC TÀI LIỆU THAM KHẢO

PHỤ LỤC

VII. TÀI LIỆU THAM KHẢO

- [1] J.K Aggarwal, Lu Xia (2014), “Human Activity Recognition from 3D Data-A Review”, *Pattern Recognition Letters, Elsevier B.V, USA*.
- [2] Aggarwal J.K, Ryoo M.S (2011), “Human Activity Analysis: A Review”, *ACM Comput. Surv*, page 16.
- [3] Kohei Arai, Rosa Andrie Asmara (2013), “3D Skeleton model derived from Kinect Depth Sensor Camera and its application to walking style quality evaluations”, *IJARAL – International Journal of Advanced in Artificial Intelligence*.
- [4] Turaga P, Chellappa R, Subrahmanian V.S, Udrea O (2008), “Machine Recognition of Human Activities: A survey”, *Circuits Syst. Video Technol. IEEE Trans 18*, pages 1473-1488.
- [5] Fengjun Lv, Ramakant Nevatia (2006), “Recognition and Segmentation of 3D Human Action Using HMM and Multi-class AdaBoost”, *Lecture Notes in Computer Science Volume 3954, 2006, pp 359-372*.
- [6] Rizwan Chaudhry, Ferda Ofli, Gregorij Kurillo, Ruzena Bajcsy, René Vidal (2013), “Bio-inspired Dynamic 3D Discriminative Skeletal Features for Human Action Recognition”, *CVPR-2013*.
- [7] Raviteja Vemulapalli, Felipe Arrate, Rama Chellappa (2014), “Human Action Recognition by Representing 3D Skeletons as Points in a Lie Group”, *CVPR-2014*.
- [8] <http://mocap.cs.cmu.edu/>
- [9] Vasileios Argyriou, Maria Petrou, Svetlana Barsky (2010), “Photometric Stereo with an Arbitrary Number of Illuminants”, *Computer Vision and Image Understanding 114*, pages 887-900.
- [10] D. Gehrig, T. Schultz (2008), “Selecting Relevant Features for Human Motion Recognition”, *ICPR 2008, IEEE*.
- [11] Dana Kulic, Wataru Takano, Yoshihiko Nakamura (2008), “Incremental Learning, Clustering and Hierarchy Formation of Whole Body Motion Patterns using Adaptive Hidden Markov Chains”, *The International Journal of Robotics Research vol. 27 no. 7 761-784*.
- [12] Gita Sukthankar, Katia Sycara (2005), “A Cost Minimization Approach to Human Behavior Recognition”.

- [13] Ahmad Jalal, Shaharyar Kamal, Daijin Kim (2014), “A Depth Video Sensor-Based Life-Logging Human Activity Recognition System for Elderly Care in Smart Indoor Environments”, *Sensors-2014*.
- [14] Lasitha Piyathilaka, Sarah Kodagoda (2013), “Human Activity Recognition for Domestic Robots”.
- [15] Mi Zhang, Alexander A. Sawchuk (2012), “Motion Primitive-Based Human Activity Recognition Using a Bag-of-Features Approach”.
- [16] Md. Zia Uddin, Nguyen Duc Thang, Jeong Tai Kim, Tae-Seong Kim (2011), “Human Activity Recognition Using Body Joint-Angle Features and Hidden Markov Model”, *ETRI Journal, Volume 33*.

VII. KẾ HOẠCH DỰ KIẾN TRIỂN KHAI ĐỀ TÀI

<i>STT</i>	<i>Thời gian</i>	<i>Nội dung thực hiện</i>	<i>Kết quả dự kiến</i>
1	Tháng 10/2014	Lập đề cương	Đề cương luận văn
2	Từ 01/11/2014 đến 15/11/2014	Nghiên cứu cấu trúc Acclaim	Chương 1, mục 1.1
3	Từ 16/11/2014 đến 30/11/2014	Dựng chuyển động 3D từ dữ liệu Acclaim	Chương 1, mục 1.2
4	Từ 01/12/2014 đến 15/12/2014	Nghiên cứu mô hình SVM	Chương 1, mục 1.3
5	Từ 16/12/2014 đến 06/01/2015	Nghiên cứu các phương pháp trích chọn đặc tính	Chương 1, mục 1.4
6	Từ 07/01/2015 đến 28/01/2015	Xây dựng mô hình nhận dạng hành động	Chương 2, mục 2.1
7	Từ 29/01/2015 đến 11/02/2015	Chạy dữ liệu huấn luyện cho mô hình	Chương 2, mục 2.2, 2.3
8	Từ 02/03/2015 đến 02/04/2015	Thực nghiệm, trích chọn đặc tính dựa theo mô hình nhận dạng	Chương 3
9	Từ 03/04/2015 đến 03/05/2015	Kết luận, hoàn chỉnh luận văn	Luận văn hoàn chỉnh

Ý KIẾN CỦA CÁN BỘ HƯỚNG DẪN

.....

.....

.....

.....

.....

.....

Đà Nẵng, ngày tháng năm 2014
Cán bộ hướng dẫn

TS. Phạm Minh Tuấn

Ý KIẾN CỦA HỘI ĐỒNG CHẤM ĐỀ CƯƠNG

.....

.....

.....

.....

.....

.....

.....

.....

Đà Nẵng, ngày tháng năm 2014
Cán bộ duyệt đề cương