BỘ GIÁO DỤC VÀ ĐÀO TẠO

**ĐẠI HỌC ĐÀ NẴNG**

**UNG NHO DÃI**

**NGHIÊN CỨU TRÍCH CHỌN ĐẶC TÍNH**

**TRONG NHẬN DẠNG HÀNH ĐỘNG NGƯỜI**

**TRONG KHÔNG GIAN 3D**

**LUẬN VĂN THẠC SĨ KỸ THUẬT**

**ĐÀ NẴNG - Năm 2015**

BỘ GIÁO DỤC VÀ ĐÀO TẠO

**ĐẠI HỌC ĐÀ NẴNG**

**UNG NHO DÃI**

**NGHIÊN CỨU TRÍCH CHỌN ĐẶC TÍNH**

**TRONG NHẬN DẠNG HÀNH ĐỘNG NGƯỜI**

**TRONG KHÔNG GIAN 3D**

**Chuyên ngành: KHOA HỌC MÁY TÍNH**

**Mã số: 60.48.01**

**LUẬN VĂN THẠC SĨ KỸ THUẬT**

**Người hướng dẫn khoa học: TS. PHẠM MINH TUẤN**

**ĐÀ NẴNG - Năm 2015**

**LỜI CẢM ƠN**

*Tôi chân thành cảm ơn*

**LỜI CAM ĐOAN**

*Tôi xin cam đoan :*

* 1. *Những nội dung trong luận văn này là do tôi thực hiện  
     dưới sự hướng dẫn trực tiếp của thầy giáo TS. Phạm Minh Tuấn.*
  2. *Mọi tham khảo dùng trong luận văn đều được trích dẫn rõ ràng và trung thực tên tác giả, tên công trình, thời gian, địa điểm công bố.*
  3. *Mọi sao chép không hợp lệ, vi phạm quy chế đào tạo, hay gian trá,  
     tôi xin chịu hoàn toàn trách nhiệm.*

*Tác giả*

**UNG NHO DÃI**

**MỤC LỤC**

**Lời cam đoan**

**Mục lục** i

**Danh mục các từ viết tắt** ii

**Danh mục các bảng biểu** v

**Danh mục các hình vẽ** vi

MỞ ĐẦU 3

NGHIÊN CỨU TỔNG QUAN 13

I. Cấu trúc Acclaim – asf/amc 13

I.1. Thí nghiệm chụp chuyển động 13

I.2. Cấu trúc tệp ASF 14

I.3. Cấu trúc tệp AMC 14

II. Dựng chuyển động 3D 15

II.1. Thư viện SlimDX 15

II.2. Dựng mô hình 3D từ tệp asf 15

II.3. Dựng chuyển động 3D từ tệp amc 15

III. Các phương pháp trích chọn đặc tính 15

III.1. Phương pháp phân tích thành phần chính – PCA 15

III.2. Phương pháp phân tích thành phần chính đa tuyến - MPCA 18

III.3. Phương pháp phân tích biệt thức tuyến tính – LDA 18

III.4. Các phương pháp lựa chọn đặc tính 18

IV. Máy vectơ hỗ trợ 18

IV.1. Cơ sở của thuật toán SVM 19

IV.2. Ứng dụng của mô hình SVM 26

IV.2.1. Tiêu đề mục con nhỏ hơn 27

GIẢI PHÁP ĐỀ XUẤT 29

I. Mô hình tổng quan 29

II. Trích chọn đặc tính 29

III. Giải thuật huấn luyện 29

III.1. Tiêu đề mục con 29

III.1.1. Tiêu đề mục con nhỏ hơn 29

TRIỂN KHAI VÀ ĐÁNH GIÁ KẾT QUẢ 31

I. Dữ liệu nhận dạng 31

II. Môi trường triển khai 31

III. Kết quả thực nghiệm 31

IV. Đánh giá 31

IV.1. Tiêu đề mục con 31

IV.1.1. Tiêu đề mục con nhỏ hơn 31

KẾT LUẬN 33

PHỤ LỤC ii

TÀI LIỆU THAM KHẢO v

DANH MỤC CÁC CHŨ VIẾT TẮT

|  |  |
| --- | --- |
| ACL | Agent Communication Language |
| AMS | Agent Management System |
| AP | Agent Platform |
| API | Application Programming Interface |
| CF | Communication Failitator |

DANH MỤC CÁC BẢNG BIỂU

[Bảng 1‑1. Các tính năng của tác tử di động. **Error! Bookmark not defined.**](#_Toc40771839)

DANH MỤC CÁC HÌNH VẼ

[**Hình 1** ii](#_Toc271487141)

# MỞ ĐẦU

1. **Lý do chọn đề tài**

**a. 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: loại thứ nhất sử dụng cảm biến (sensor-based) loại thứ hai sử dụng 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 lại 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 hình ảnh từ hệ thống các camera. Kỹ thuật chủ yếu được dùng để nhận dạng từ hình ảnh là thị giác máy tính (vision computer). 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.

**b. 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 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). L. Fengjun và N. RamaKant 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). K. Dana, T. Wataru và N. Yoshihiko đã phát triển một mô hình học tăng cường mới dựa trên mô hình Markov ẩn [11].



**Hình 3. Kết quả nhận dùng bảy loại đặc tính do L. Fengjun và N. RamaKant [5] đề xuất**

**c. 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
* Tỉ 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.

1. **Mục tiêu và nhiệm vụ**

Mục tiêu 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 dữ liệu chuyển động 3D.

Các hoạt động cụ thể:

* Nghiên cứu hệ thống chụp chuyển động và phân tích dữ liệu chuyển động 3D
* 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.

1. **Đối tượng và phạm vi nghiên cứu**

**a. Đối tượng nghiên cứu**

* Cấu trúc dữ liệu chuyển động 3D (Acclaim)
* Các phương pháp trích chọn đặc tính (feature extraction) và lựa chọn đặc tính (feature selection)
* Mô hình SVM

**b. 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, LDA

1. **Phương pháp nghiên cứu**

**a. Nghiên cứu lý thuyết**

* Cấu trúc Acclaim (asf/amc)
* Các phương pháp trích chọn đặc tính: PCA, LDA
* Mô hình SVM

**b. Nghiên cứu thực nghiệm**

* Thực nghiệm trên mô hình nhận dạng với nhiều phương pháp khác nhau và thu thập, phân tích kết quả
* So sánh, đánh giá kết quả

1. **Ý nghĩa khoa học và thực tiễn của đề tài**

**a. Về mặt lý thuyêt**

* Lý thuyết về nhận dạng
* Lý thuyết về dữ liệu chuyển động 3D
* Các lý thuyết trích chọn đặc tính và lựa chọn đặc tính

**b. 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

1. **Bố cục của luận văn**

Ngoài phần mở đầu và kết luận, luận văn gồm có ba chương với các nội dung chính như sau:

**Chương 1: Nghiên cứu tổng quan**

Chương này trình bày tổng quan các vấn đề liên quan đến luận văn. Nội dung chủ yếu xoay quanh ba chủ đề chính. Thứ nhất là hoạt động nhận dạng hành động người. Thứ hai là mô hình chụp chuyển động. Cuối cùng là các phương pháp trích chọn và lựa chọn đặc tính.

**Chương 2: Giải pháp đề xuất**

Chương này tập trung vào trình bày quá trình phân tích và chuẩn hóa dữ liệu. Dựa trên các phương pháp trích chọn đặc tính và dữ liệu đã xử lý đề xuất, xây dựng mô hình nhận dạng hành động người sử dụng máy vectơ hỗ trợ.

**Chương 3: Thực nghiệm và đánh giá kết quả**

Dựa trên mô hình đề xuất ở chương hai, xây dựng chương trình thực nghiệm, sau đó phân tích và đánh giá kết quả.

CHƯƠNG 1

# NGHIÊN CỨU TỔNG QUAN

## Nhận dạng hành động người trong không gian 3D

Từ những năm 1980, nhận dạng hành động người luôn là lĩnh vực quan trọng trong các nghiên cứu về thị giác máy tính. Có rất nhiều hướng tiếp cận khác nhau đã được đề xuất và phần lớn trong số đó sử dụng hình ảnh từ các camera thông dụng. Các phương pháp thu thập dữ liệu chuyển động 3D đã ra đời và phát triển mạnh mẽ trong các thập niên gần đây. Phần này khái quát các kỹ thuật thu thập dữ liệu chuyển động 3D cũng như các phương pháp học máy phổ biến trong nhận dạng hành động người.

### Các phương pháp thu thập dữ liệu chuyển động 3D

Từ sau sự ra đời của các thiết bị cảm biến chiều sâu (depth sensor), hoạt đông nghiên cứu trên dữ liệu 3D đã có những chuyển biến tích cực. Nhìn chung có ba phương pháp chính trong việc thu thập dữ liệu chuyển động 3D. Đầu tiên là sử dụng hệ thống chụp chuyển động dựa vào các marker như là Mocap (Motion Capture). Kế đến là phương pháp sử dụng sóng âm (stereo) – chụp hình ảnh 2D từ nhiều hướng khác nhau để dựng thành mô hình 3D. Cuối cùng là phương pháp sử dụng cảm biến chiều sâu (range sensor). Mục này sẽ trình bày sơ lược từng phương pháp.

#### Phương pháp sử dụng sóng âm

#### Hệ thống chụp chuyển động (Mocap)

#### Phương pháp sử dụng range sensor

## Cấu trúc Acclaim – asf/amc

Acclaim là một công ty game (2006 – 2010) có rất nhiều nghiên cứu trong lĩnh vực theo dõi chuyển động. Họ phát triển và sử dụng cấu trúc riêng trong việc lưu trữ dữ liệu chuyển động 3D. Đó là một bộ gồm hai tệp văn bản. Sau này, họ công bố và cho phép sử dụng rộng rãi cấu trúc đó trên toàn thế giới. Có rất nhiều công ty, tổ chức sử dụng cấu trúc Acclaim làm dữ liệu đầu ra cho các hệ thống theo dõi chuyển động của họ.

Acclaim là một bộ gồm hai tệp văn bản có cấu trúc. Tệp thư nhất lưu trữ cấu trúc tổng quan của đối tượng, tiệp thứ hai chứa dư liệu chuyển động tương ứng của đối tượng theo thời gian. Nếu là chuyển động của con người thì tệp thứ nhất thể hiện một bộ xương được nối với nhau bởi các khớp, tệp thứ hai là vị trí của các khớp xương ứng với mỗi mốc thời gian. Tệp cấu trúc là ASF (Acclaim Skeleton File), tệp chuyển động là AMC (Acclaim Motion Capture).

### Thí nghiệm chụp chuyển động

Để theo dõi chuyển động của các đối tượng, các nhà nghiên cứu tại đại học Carnigie Mellon đã xây dựng hệ thống gồm 12 camera hồng ngoại MX-40 lắp đặt xung quanh một không gian hình chữ nhật có kích thước 3m x 8m. Các chuyển động diễn ra trong vùng này sẽ được ghi lại. Đối với các chuyển động đòi hỏi sự tỉ mỉ và chi tiết nhu chuyển động của tay, các camera có thể di chuyển vào gần hơn. Người ta sử dụng một bộ quần áo đặc biệt có gắn 41 marker bên trên, nếu muốn ghi lại chuyển động của ai đó thì người đó phải mặc bộ quần áo này và di chuyển trong vùng ghi nhận. Các camera sẽ nhận ra các marker bằng sóng hồng ngoại. Dữ liệu chuyển động thu được từ hệ thống camera được xử lý và cho ra kết quả cuối cùng là dữ liệu 3D. Có rất nhiều cấu trúc định dạng khác nhau để lưu trữ dữ liệu 3D như asf/amc, vsk/v, c3d, bvh, txt.

### Cấu trúc tệp ASF

Tệp cấu trúc ASF mô tả sự gắn kết giữa các xương trong cơ thể và bậc tự do (degrees of freedom) của các khớp. ASF chính là trạng thái ban đầu của dữ liệu chuyển động. Thành phần cụ thể trong ASF được mô tả như sau:

* Các chú thích được bắt đầu bởi dấu thăng (#).
* Các từ khóa bắt đầu bởi dấu hai chấm (:). Từ khóa có thể được dùng cho các giá trị toàn cục hoặc bắt đầu cho một mục dữ liệu.
* Từ khóa :version cho biết phiên bản hiện tại của tệp tin
* Từ khóa :name được dùng để đặt tên cho dữ liệu, tên này có thể khác với tên tệp.
* Mục :units định nghĩa đơn vị cho một số loại dữ liệu trong tệp. Nó củng có thể chứa giá trị mặc định cho các đại lượng.
* Mục :documentation lưu trữ các thông thêm về dữ liệu
* Mục :root định nghĩa một khớp xương đặc biệt của ASF, đây chính là nút gốc trong hệ thống cây với nút là các khớp và nhánh là các đoạn xương giữa các khớp. Từ khóa axis trong mục :root cho biết thứ tự quay của các trục tọa độ X, Y, Z của nút gốc. Từ khóa order cho biết các kênh chuyển động của nút gốc cũng như thứ tự xuất hiện của các đại lượng này trong tệp AMC. Hai từ khóa còn lại position và orientation lưu giữ tọa độ và phương hướng ban đầu của nút gốc, thường thì giá trị của các đại lượng này bằng không (0).
* Mục :bonedata là nơi chứa thông tin chi tiết của mỗi đoạn xương trong cây hệ thống. Các thông tin của mỗi đoạn xương đực đặt trong một cặp tù khóa begin và end. Vỡi mỗi cặp, chúng ta có: id, name, direction là hướng của xương, length là độ dài, axis là góc quay của trục tọa độ tương đối của mỗi đoạn xương, dof là độ tự do của đoạn xương, limits là giới hạn quay của các khớp.
* Mục cuối cùng :hierarchy định nghĩa sự liên kết giữa các khớp xương để tạo nên một bộ xương hoàn chỉnh.

### Cấu trúc tệp AMC

Tệp chuyển động AMC chứa các đại lượng có thể thay đổi trong hệ thống các khớp xương được định nghĩa ở tệp ASF. Là dữ liệu chuyển động theo thời gian nên trên tệp AMC được tạo thành bởi nhiều frame, mỗi frame thể hiện dữ liệu cho một thời điểm. Dữ liệu bên trong mỗi frame là giá trị theo thứ tự của đại lượng dof trên tệp ASF.

## Các phương pháp trích chọn đặc tính

### Phương pháp phân tích thành phần chính – PCA

Phương pháp phân tích thành phần chính (Principal Components Analysis - PCA) là một thuật toán thống kê sử dụng phép biến đổi trực giao để biến đổi một tập hợp dữ liệu từ một không gian nhiều chiều sang một không gian mới ít chiều hơn nhằm tối ưu hóa việc thể hiện sự biến thiên của dữ liệu (maximize the variability).

Các đặc tính tốt của PCA:

- Giúp giảm số chiều của dữ liệu

- Thay vì giữ lại các trục tọa độ của không gian cũ, PCA xây dựng một không gian mới ít chiều hơn, nhưng lại có khả năng biểu diễn dữ liệu tốt tương đương không gian cũ, nghĩa là đảm bảo độ biến thiên (variability) của dữ liệu trên mỗi chiều mới.

- Các trục tọa độ trong không gian mới là tổ hợp tuyến tính của không gian cũ, do đó về mặt ngữ nghĩa, PCA xây dựng các thuộc tính mới dựa trên các thuộc tính hiện có. Điểm hay là những thuộc tính này vẫn biểu diễn tốt dữ liệu ban đầu.

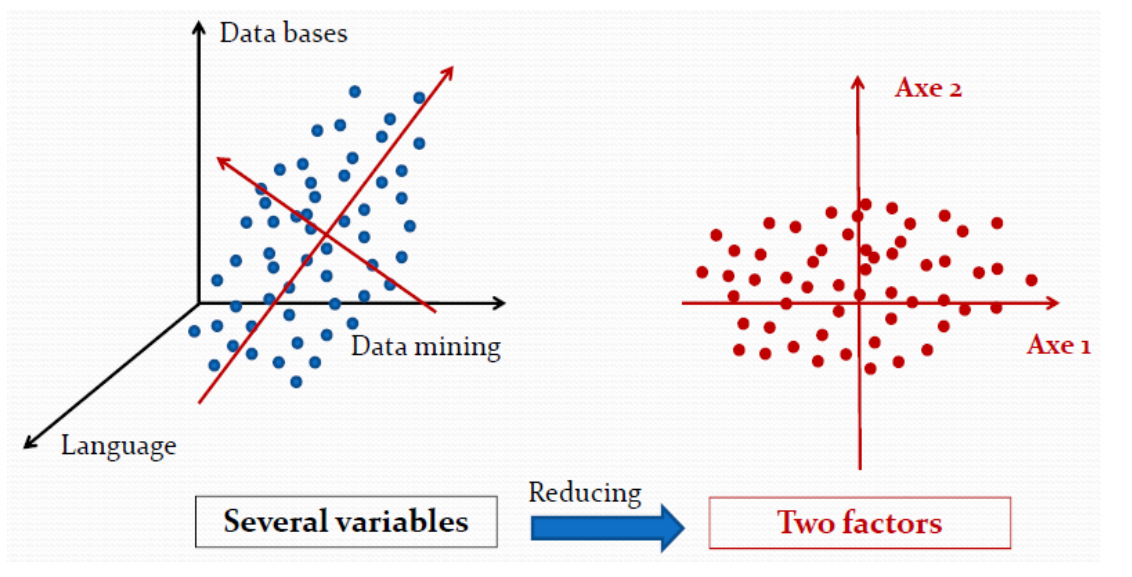
- Trong không gian mới, các liên kết tiềm ẩn của dữ liệu có thể được khám phá, mà nếu đặt trong không gian cũ thì khó phát hiện hơn, hoặc những liên kết như thế không thể hiện rõ.



Minh họa PCA. Phép chiếu lên các trục tọa độ khác nhau có thể cho cách nhìn rất khác nhau về cùng một dữ liệu

Một ví dụ kinh điển để minh họa PCA là hình ảnh về con lạc đà. Cùng là một con lạc đà nhưng nếu nhìn từ bên hông thì ta có được đầy đủ thông tin nhất, trong khi nhìn từ phía trước thì thật khó để nói nó là con lạc đà.

Một ví dụ thuyết phục hơn được minh họa như hình sau.



Giả sử tập dữ liệu ban đầu (tập điểm màu xanh) được quan sát trong không gian ba chiều (trục màu đen) như hình bên trái. Rõ ràng ba trục này không biểu diễn được tốt nhất mức độ biến thiên của dữ liệu. PCA do đó sẽ tìm hệ trục tọa độ mới (là hệ trục màu đỏ trong hình bên trái). Sau khi tìm được không gian mới, dữ liệu sẽ được chuyển qua không gian này để được biểu diễn như trong hình bên phải. Rõ ràng hình bên trái chỉ cần hai trục tọa độ nhưng biểu diễn tốt hơn độ biến thiên của dữ liệu so với hệ trục ba chiều ban đầu.

Một điểm rất hay nữa của PCA là các trục tọa độ trong không gian mới luôn đảm bảo trực giao đôi một với nhau, mặt dù trong không gian ban đầu, các trục có thể không trực giao.

Thuật toán PCA:

Về cơ bản, thuật toán PCA gồm có ba bước: tiền xử lí, xây dựng không gian mới, chuyển dữ liệu từ không gian ban đầu sang không gian mới. Cho ma trận . Các bước của PCA lần lượt như sau:

Bước 1: Tiền xử lí

Dữ liệu ban đầu có thể có giá trị thay đổi bất thường. Ví dụ một thuộc tính (feature) có gián trị thay đổi trong khoảng (0, 1) nhưng trên thuộc tính khác lại biến thiên trong đoạn (-100, 100). Rõ ràng cần phải có một bước tiền xử lí để chuẩn hóa các giá trị trên các cột của ma trận X. Có hai cách tiền xử lí thường được dùng cho PCA là Centered PCA và Normed PCA.

Centered PCA mang tất cả các thuộc tính (các cột của X) về cùng một gốc tọa độ:

,

(1a)

Trong đó n là số dòng của X, là mean của cột thứ j của X, được tính như sau:

Normed PCA mang tất các các feature về cùng một gốc tọa độ, đồng thời chuẩn hóa về cùng một quãng có độ lệch chuẩn (standard deviation) bằng 1:

,

(1b)

Trong đó là độ lệch chuẩn của cột thứ j trong X.

Thông thường, Normed PCA hay được dùng, sau bước tiền xử lí, ma trận sẽ là đầu vào cho bước tiếp theo.

Bước 2: Xây dựng không gian mới

Tính ma trận hiệp phương sai (covariance) của các thuộc tính trong :

(2)

Do là tích của ma trận với một chuyển vị của nó nên là ma trận positive semidefinite kích thước . Hơn nữa có p trị riêng .

Tiếp theo, PCA tìm trị riêng và vector riêng tương ứng của V, sắp xếp theo thứ tự giảm dần của trị riêng. Giả sử p trị riêng của V là (3), và vector riêng tương ứng là , …, (4). Khi đó các trục của không gian mới chính là các vector riêng ở trên, đương nhiên các vector riêng hoàn toàn độc lập tuyến tính (nghĩa là trực giao đôi một).

Bước 3: Chuyển dữ liệu từ không gian ban đầu sang không gian mới

Thông thường không gian mới không được xây dựng bằng tất cả các p vector riêng trong (4), mà thông thường chỉ từ k vector riêng đầu tiên.

Như vậy gọi  **|**

Khi đó tọa độ các điểm trong hệ tọa độ mới là

(5)

### Phương pháp phân tích biệt thức tuyến tính – LDA

### Các phương pháp lựa chọn đặc tính

## Máy vectơ hỗ trợ

Máy vectơ hỗ trợ (SVM – Support Vector Machine) là một phương pháp phân lớp dựa trên lý thuyết học thống kê và khoa học máy tính, gồm một tập hợp các phương pháp học máy liên quan đến nhau để phân lớp dữ liệu (classification) và phân tích hồi quy (regression analysis). SVM dạng chuẩn nhận dữ liệu vào và phân loại chúng vào hai lớp khác nhau. Do đó SVM là một thuật toán phân loại nhị phân. Với một bộ dữ liệu huấn luyện thuộc hai thể loại cho trước, thuật toán SVM xây dựng một mô hình SVM để phân loại các ví dụ khác vào hai thể loại đó. Một mô hình SVM là một cách biểu diễn các điểm trong không gian và lựa chọn ranh giới giữa hai thể loại sao cho khoảng cách từ các điểm tới ranh giới là xa nhất có thể. Các điểm mới cũng được biểu diễn trong cùng một không gian và được thuật toán dự đoán thuộc một trong hai thể loại tùy vào vị trí của nó nằm ở phía nào của ranh giới.

Siêu phẳng

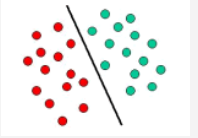
Mẫu cần nhận dạng

Siêu phẳng phân hoạch tập mẫu

Ý tưởng chính của SVM là tìm một siêu phẳng phân cách tối ưu để tách hai lớp, phân hoạch tập mẫu dựa vào các phân lớp của chúng sao cho khoảng cách giữa hai lớp đó đạt cực đại. Khoảng cách này được xác định bởi các vectơ hỗ trợ (SV – support vector). Các SV này được lọc ra từ tập mẫu huấn luyện bằng cách giải một bài toán tối ưu lồi, điều đó chính là tìm ra miền phân bố của từng lớp, để từ đó xác định được phân lớp của một mẫu cần nhận dạng.

### Cơ sở của thuật toán SVM

Xét một ví dụ của bài toán phân lớp như hình vẽ, ở đó ta phải tìm một đường thẳng sao cho bên trái nó toàn là các điểm đỏ, bên phải nó toàn là các điểm xanh. Bài toán mà dùng đường thẳng để phân chia này được gọi là phân lớp tuyến tính (linear classification).



Hàm tuyến tính phân biệt hai lớp như sau:

(1)

Trong đó:

là vector trọng số hay vector chuẩn của siêu phẳng phân cách, T là kí hiệu chuyển vị.

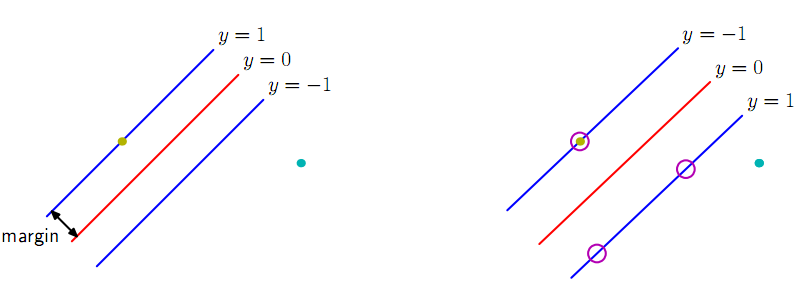
là độ lệch

là véc tơ đặc trưng, làm hàm ánh xạ từ không gian đầu vào sang không gian đặc trưng.

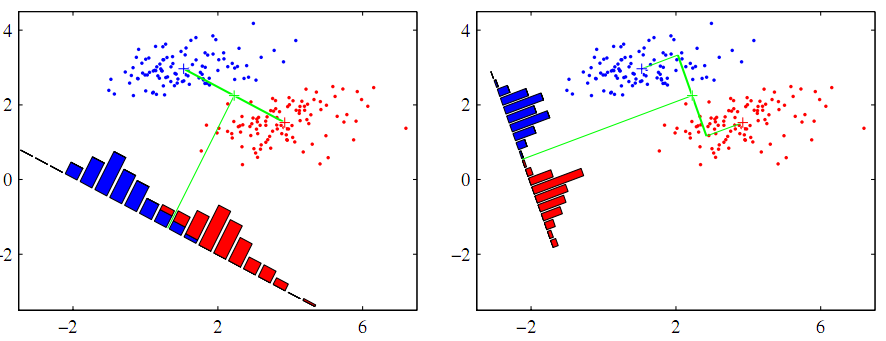
Tập dữ liệu đầu vào gồm N mẫu input vector {x1, x2,...,xN}, với các giá trị nhãn tương ứng là {t1,…,tN} trong đó .

Giả sử tập dữ liệu của ta có thể phân tách tuyến tính hoàn toàn (các mẫu đều được phân đúng lớp) trong không gian đặc trưng (feature space), do đó sẽ tồn tại giá trị tham số w và b theo (1) thỏa cho những điểm có nhãn và cho những điểm có , vì thế mà cho mọi điểm dữ liệu huấn luyện.

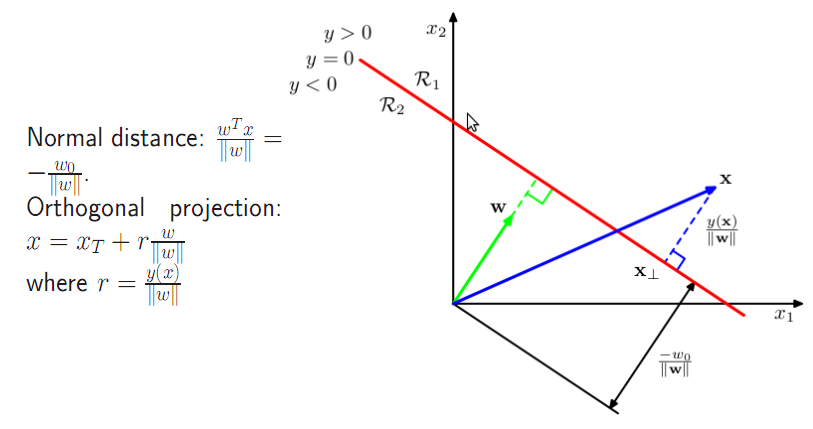
SVM tiếp cận giải quyết vấn đề này thông qua khái niệm gọi là lề, đường biên (margin). Lề được chọn là khoảng cách nhỏ nhất từ đường phân cách đến mọi điểm dữ liệu hay là khoảng cách từ đường phân cách đến những điểm gần nhất.



Trong SVM, đường phân lớp tốt nhất chính là đường có khoảng cách margin lớn nhất (tức là sẽ tồn tại rất nhiều đường phân cách xoay theo các phương khác nhau, và ta chọn ra đường phân cách mà có khoảng cách margin là lớn nhất).



Ta có công thức tính khoảng cách từ điểm dữ liệu đến mặt phân cách như sau:



Do ta đang xét trong trường hợp các điểm dữ liệu đều được phân lớp đúng nên cho mọi n. Vì thế khoảng cách từ điểm xn đến mặt phân cách được viết lại như sau:

(2)

Lề là khoảng cách vuông góc đến điểm dữ liệu gần nhất xn từ tập dữ liệu, và chúng ta muốn tìm giá trị tối ưu của w và b bằng cách cực đại khoảng cách này. Vấn đề cần giải quyết sẽ được viết lại dưới dạng công thức sau:

(3)

Chúng ta có thể đem nhân tử ra ngoài bởi vì w không phụ thuộc n. Giải quyết vấn đề này một cách trực tiếp sẽ rất phức tạp, do đó ta sẽ chuyển nó về một vấn đề tương đương dễ giải quyết hơn. Ta sẽ scale và cho mọi điểm dữ liệu, từ đây khoảng cách lề trở thành 1, việc biến đổi này không làm thay đổi bản chất vấn đề.

(4)

Từ bây giờ, các điểm dữ liệu sẽ thỏa ràng buộc:

(5)

Vấn đề tối ưu yêu cầu ta cực đại được chuyển thành cực tiểu , ta viết lại công thức:

(6)

Việc nhân hệ số ½ sẽ giúp thuận lợi cho lấy đạo hàm về sau.

**Lý thuyết Nhân tử Lagrange:**

Vấn đề cực đại hàm f(x) thỏa điều kiện sẽ được viết lại dưới dạng tối ưu của hàm Lagrange như sau:

Trong đó x và λ phải thỏa điều kiện Karush-Kuhn-Tucker (KKT) như sau:

Nếu là cực tiểu hàm f(x) thì hàm Lagrange sẽ là

Để giải quyết bài toán trên, ta viết lại theo hàm Lagrange như sau:

(7)

Trong đó là nhân tử Lagrange.

Lưu ý dấu (–) trong hàm Lagrange, bởi vì ta cực tiểu theo biến w và b, và là cực đại theo biến a.

Lấy đạo hàm L(w,b,a) theo w và b ta có:

(8)

(9)

Loại bỏ w và b ra khỏi L(w,b,a) bằng cách thế (8), (9) vào. Điều này sẽ dẫn ta đến vấn đề tối ưu:

(10)

Thỏa các ràng buộc:

(11)

(12)

Ở đây hàm nhân (kernel function) được định nghĩa là .

Để phân lớp cho 1 điểm dữ liệu mới dùng mô hình đã huấn luyện, ta tính dấu của y(x) theo công thức (1), nhưng thế w trong (8) vào:

(13)

Thỏa các điều kiện sau:

(14)

(15)

(16)

Vì thế với mọi điểm dữ liệu, hoặc là hoặc là . Những điểm dữ liệu mà có sẽ không xuất hiện trong (13) và do đó mà không đóng góp trong việc dự đoán điểm dữ liệu mới.

Những điểm dữ liệu còn lại được gọi là support vector, chúng thỏa , đó là những điểm nằm trên lề của siêu phẳng trong không gian đặc trưng.

Support vector chính là cái mà ta quan tâm trong quá trình huấn luyện của SVM. Việc phân lớp cho một điểm dữ liệu mới sẽ chỉ phụ thuộc vào các support vector.

Giả sử rằng ta đã giải quyết được vấn đề (10) và tìm được giá trị nhân tử a, bây giờ ta cần xác định tham số b dựa vào các support vector xn có . Thế (13) vào:

(17)

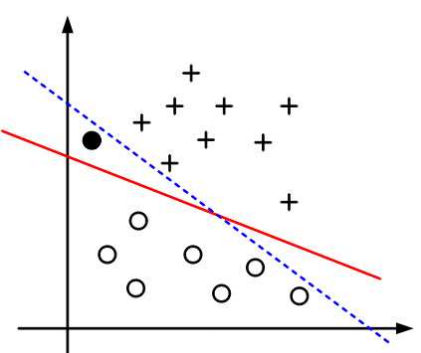
Trong đó S là tập các support vector. Mặc dù ta chỉ cần thế một điểm support vector xn vào là có thể tìm ra b, nhưng để đảm bảo tính ổn định của b ta sẽ tính b theo cách lấy giá trị trung bình dựa trên các support vector.

Đầu tiên ta nhân tn vào (17) (lưu ý , và giá trị b sẽ là:

(18)

Trong đó Ns là tổng số support vector.

Ban đầu để dễ trình bày thuật toán ta đã giả sử là các điểm dữ liệu có thể phân tách hoàn toàn trong không gian đặc trưng . Nhưng việc phân tách hoàn toàn này có thể dẫn đến khả năng tổng quát hóa kém, vì thực tế một số mẫu trong quá trình thu thập dữ liệu có thể bị gán nhãn sai, nếu ta cố tình phân tách hoàn toàn sẽ làm cho mô hình dự đoán quá khớp.



Để chống lại sự quá khớp, chúng ta chấp nhận cho một vài điểm bị phân lớp sai.

Để làm điều này, ta dùng các biến slack variables cho mọi điểm dữ liệu.

cho những điểm nằm trên lề hoặc phía trong của lề

cho những điểm còn lại.

Do đó những điểm nằm trên đường phân cách sẽ có

Còn những điểm phân lớp sai sẽ có



Công thức (5) sẽ viết lại như sau:

(20)

Mục tiêu của ta bây giờ là cực đại khoảng cách lề, nhưng đồng thời cũng đảm bảo tính mềm mỏng cho những điểm bị phân lớp sai. Ta viết lại vấn đề cần cực tiểu:

(21)

Trong đó C > 0 đóng vai trò quyết định đặt tầm quan trọng vào biến hay là lề.

Bây giờ chúng ta cần cực tiểu (21) thỏa ràng buộc (20) và . Theo Lagrange ta viết lại:

(22)

Trong đó và là các nhân tử Lagrange.

Các điều kiện KKT cần thỏa là:

(23)

(24)

(25)

(26)

(27)

(28)

Với n = 1,…,N

Lấy đạo hàm (22) theo w, b và {}:

(29)

(30)

(31)

Thế (29), (30), (31) vào (22) ta được:

(32)

Từ (23), (26) và (31) ta có:

Vấn đề cần tối ưu giống hệt với trường hợp phân tách hoàn toàn, chỉ có điều kiện ràng buộc khác biệt như sau:

(33)

(34)

Thế (29) vào (1), ta sẽ thấy để dự đoán cho một điểm dữ liệu mới tương tự như (13).

Như trước đó, tập các điểm có không có đóng góp gì cho việc dự đoán điểm dữ liệu mới.

Những điểm còn lại tạo thành các support vector. Những điểm có và theo (25) thỏa:

(35)

Nếu theo (31) có , từ (28) suy ra và đó là những điểm nằm trên lề.

Những điểm có có thể là những điểm phân lớp đúng nằm giữa lề và đường phân cách nếu hoặc có thể là phân lớp sai nếu

Để xác định tham số b trong (1) ta sẽ dùng những support vector mà có vì thế :

(36)

Lần nữa, để đảm bảo tính ổn định của b ta tính theo trung bình:

(37)

Trong đó M là tập các điểm có

Để giải quyết (10) và (32) ta dùng thuật toán Sequential Minimal Optimization (SMO) do Platt đưa ra vào 1999.

### Ứng dụng của mô hình SVM

xxx

Tiêu đề mục con nhỏ hơn

* xxx
* xxx
* Bullet 1

Bullet 2

* Bullet 3

**Ví dụ 1**

* xxx

**Hình 1**

* xxx

Bảng 1

* Xxx

CHƯƠNG 2

# GIẢI PHÁP ĐỀ XUẤT

xxx

## Mô hình tổng quan

## Trích chọn đặc tính

## Giải thuật huấn luyện

* Xxx

### Tiêu đề mục con

xxx

Tiêu đề mục con nhỏ hơn

* xxx
* xxx

Chú ý sau mỗi chương cần thêm dấu ngắt chương (Section) :

**Insert-Break... Section Break Pages** = Next Page

CHƯƠNG 3

# TRIỂN KHAI VÀ ĐÁNH GIÁ KẾT QUẢ

xxx

## Dữ liệu nhận dạng

## Môi trường triển khai

## Kết quả thực nghiệm

## Đánh giá

* Xxx

### Tiêu đề mục con

xxx

Tiêu đề mục con nhỏ hơn

* xxx
* xxx

# KẾT LUẬN

xxx

xxx

# PHỤ LỤC

Tấ cả các phụ lục nằm ở đây và được đánh số thứ tự cùng tên phụ lục :

xxx

PHỤ LỤC A

xxx

# 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 andits 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), “Recongnition and Segmentation of 3D Human Action Using HMM and Multi-class AdaBoost”, *Lecture Notes in Computer Science Vlolumn 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*.

**TÓM TẮT LUẬN VĂN THẠC SĨ**

Phần này viết tên đề tài (chữ nhỏ 13pt) và :

tóm tắt nội dung LUẬN VĂN, viết ngắn gọn và rõ ràng  
(15 đến 20 dòng) cho biết :

Đề tài đã được đặt ra như thế nào (bối cảnh), mục đích (giải thích rõ hơn tên đề tài) và nhiệm vụ phải thực hiện (các mục tiêu cụ thể và kết quả cần có)

HV đã giải quyết vấn đề gì (đã nghiên cứu lý thuyết, thực tiễn như thế nào, đã đề xuất được những giải pháp (biện pháp) hay sáng kiến gì ?).

HV đã giải quyết đến đâu (nêu một số kết quả tiêu biểu).