

BỘ THÔNG TIN VÀ TRUYỀN THÔNG
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Nguyễn Ngọc Điệp

NGHIÊN CỨU PHƯƠNG PHÁP HỌC MÁY
CHO NHẬN DẠNG HOẠT ĐỘNG
SỬ DỤNG CẢM BIẾN MANG TRÊN NGƯỜI

Chuyên ngành: Hệ thống thông tin

Mã số: 62.48.01.04

TÓM TẮT LUẬN ÁN TIẾN SĨ KỸ THUẬT

Hà Nội - 2016

Công trình được hoàn thành tại:

Học viện Công nghệ Bưu chính Viễn thông

Người hướng dẫn khoa học:

1. PGS.TS. Từ Minh Phương
2. TS. Phạm Văn Cường

Phản biện 1:.....

Phản biện 2:.....

Phản biện 3:.....

Luận án được bảo vệ trước Hội đồng chấm luận án cấp Học viện
họp tại:

Học viện Công nghệ Bưu chính Viễn thông

Vào hồi giờ.....ngày.....tháng.....năm.....

Có thể tìm hiểu luận án tại: Thư viện Học viện Công nghệ Bưu
chính Viễn thông

MỞ ĐẦU

1. Tính cấp thiết của luận án

Nhờ vào sự tiến bộ của công nghệ và các nhu cầu ứng dụng trong những năm gần đây, nghiên cứu về nhận dạng hoạt động người đã có được sự phát triển mạnh mẽ. Có nhiều bài toán thực tế cần các giải pháp dựa vào nhận dạng hoạt động để tạo ra các cách thức tương tác đa dạng, chủ động cung cấp các dịch vụ trợ giúp người dùng hoàn thành công việc.

Hiện nay, có ba cách tiếp cận phổ biến để giải quyết bài toán nhận dạng hoạt động, bao gồm: nhận dạng hoạt động dựa trên thị giác máy tính, nhận dạng hoạt động dựa trên các cảm biến gắn trong môi trường xung quanh và nhận dạng hoạt động dựa trên cảm biến mang trên người. Hai cách tiếp cận đầu đều có hạn chế do các hoạt động của người dùng bị giới hạn trong một môi trường cố định và hệ thống cũng cần được triển khai, lắp đặt sẵn trong môi trường. Các hạn chế này là rào cản cho việc triển khai rộng rãi các ứng dụng nhận dạng hoạt động người trong thực tế. Cách tiếp cận thứ ba sử dụng cảm biến mang theo trên cơ thể người đã mở ra nhiều ứng dụng tiềm năng trong nhận dạng hoạt động do hoạt động của người dùng không bị giới hạn trong những không gian lắp sẵn các thiết bị. Cách tiếp cận này mang lại khả năng cung cấp những sự trợ giúp thông minh, giao tiếp ảo tại bất kỳ nơi đâu và bất kỳ khi nào, thông qua việc quan sát các hoạt động từ góc nhìn người dùng.

Bài toán nhận dạng hoạt động người nói chung và nhận dạng hoạt động dựa trên cảm biến mang theo người nói riêng

cũng có thể tiếp cận theo hai hướng, đó là nhận dạng dựa trên tri thức và nhận dạng dựa trên dữ liệu. Cách tiếp cận dựa trên tri thức có nhược điểm là cần nhiều chi phí về thời gian và kinh nghiệm của chuyên gia để xây dựng được tập luật suy diễn tốt, việc cập nhật tự động các luật là không khả thi do nguồn dữ liệu đầu vào thường không có cấu trúc và luôn biến động, đồng thời không có khả năng xử lý thông tin tạm thời và chưa rõ ràng. Cách tiếp cận dựa trên dữ liệu sử dụng các quy tắc nhận dạng được xây dựng dựa trên các tập dữ liệu hành vi người dùng lớn có sẵn và các kỹ thuật học máy, thống kê. Do dựa trên dữ liệu nên cách tiếp cận này đảm bảo được rằng hệ thống có thể cập nhật các quy tắc nhận dạng hoạt động một cách tự động mà không phụ thuộc vào tri thức chuyên gia cũng như có khả năng xử lý các thông tin tạm thời và chưa rõ ràng. Đồng thời, hệ thống cũng có khả năng thích nghi cao và tận dụng được nguồn dữ liệu có sẵn. Chính vì vậy, cách tiếp cận này được quan tâm nghiên cứu nhiều hơn so với phương pháp nhận dạng hoạt động dựa trên tri thức.

Đề tài “*Nghiên cứu phương pháp học máy cho nhận dạng hoạt động sử dụng cảm biến mang trên người*” được thực hiện trong khuôn khổ luận án tiến sĩ chuyên ngành hệ thống thông tin nhằm góp phần giải quyết một số vấn đề còn tồn tại trong phương pháp nhận dạng hoạt động sử dụng cảm biến mang trên người.

2. Mục tiêu, đối tượng và phạm vi nghiên cứu

Mục tiêu của luận án là nghiên cứu đề xuất một số phương pháp học máy nhằm tăng cường hiệu năng cho các hệ

thống nhận dạng hoạt động sử dụng cảm biến mang trên người, cụ thể là đề xuất các phương pháp trích xuất đặc trưng mới và hiệu quả cho nhận dạng hoạt động, với hai mục tiêu:

- Nghiên cứu và đề xuất phương pháp trích xuất đặc trưng có *tốc độ nhanh* cho các ứng dụng nhận dạng một số *hoạt động riêng lẻ*, với yêu cầu về *thời gian thực* và chạy được trên các thiết bị thông minh mang theo người.
- Nghiên cứu và đề xuất phương pháp trích xuất đặc trưng giúp nâng cao *độ chính xác* cho hệ thống cần nhận dạng *nhiều loại hoạt động*.

Như vậy, *các phương pháp học máy* sẽ được sử dụng để nghiên cứu và đề xuất các *phương pháp trích xuất đặc trưng mới* trong *các hệ thống nhận dạng hoạt động sử dụng cảm biến mang trên người*. Các phương pháp trích xuất đặc trưng này có thể phù hợp với các hệ thống trợ giúp cá nhân thông minh nhỏ gọn mang theo người, năng lực xử lý thấp và yêu cầu hoạt động theo thời gian thực, hoặc phù hợp với hệ thống cần nhận dạng nhiều hoạt động người phức tạp. Các hoạt động con người bao gồm các hoạt động trong cuộc sống hàng ngày, các hoạt động sản xuất, giải trí, thể thao, là các chuyển động của cơ thể có thể ghi nhận và phân biệt được bằng các cảm biến mang trên người.

3. Các đóng góp của luận án

Đóng góp thứ nhất của luận án là đề xuất một phương pháp trích xuất đặc trưng mới đơn giản, hiệu quả dựa trên tín hiệu cảm biến mang trên người cho các bài toán nhận dạng hoạt động riêng lẻ gọi là HALF. Các đặc trưng này có khả năng tính toán nhanh, phù hợp cho các hệ thống di động thông minh nhỏ

gọn với năng lực tính toán thấp, đặc biệt là các hệ thống đòi hỏi thời gian thực. Các đặc trưng này cho kết quả cao khi ứng dụng trong việc phân biệt các hoạt động *ngã* và các hoạt động khác, sử dụng cảm biến gia tốc.

Đóng góp thứ hai của luận án là đề xuất phương pháp trích xuất đặc trưng tự động có *độ chính xác cao cho nhiều loại hoạt động người* dựa trên học đặc trưng, gọi là MPF. Các đặc trưng đề xuất không những cải thiện được độ chính xác trong nhận dạng hoạt động người so với các đặc trưng đa mức kiểu cũ mà còn giúp loại trừ được các hạn chế về mặt tốc độ xử lý, đồng thời phù hợp để nhận dạng nhiều loại hoạt động người.

Đóng góp thứ ba của luận án là áp dụng phương pháp trích xuất đặc trưng đã đề xuất để xây dựng các ứng dụng nhận dạng hoạt động người có tính khả thi trong thực tế, bao gồm: *phát hiện ngã trong thời gian thực* và *xác thực người dùng sử dụng chữ ký 3D*. Hệ thống phát hiện ngã sử dụng các bộ cảm biến đeo trên người rẻ tiền, và dễ triển khai, có thể phát hiện ngã trong thời gian thực. Hệ thống xác thực trên điện thoại thông minh có thể xác thực người dùng thông qua hành động “ký tên” trong không gian (gọi là chữ ký 3D) trong thời gian thực.

3. Bố cục của luận án

Nội dung luận án được xây dựng thành bốn chương.

Chương 1. Giới thiệu tổng quan về nhận dạng hoạt động sử dụng cảm biến mang trên người, bao gồm giới thiệu chung về nhận dạng hoạt động người, các loại cảm biến, các hoạt động người và kiến trúc hệ thống nhận dạng hoạt động.

Cuối cùng là phân loại các phương pháp trích xuất đặc trưng cùng những ưu nhược điểm trong từng phương pháp. Từ những cơ sở nghiên cứu này sẽ xác định rõ hướng nghiên cứu của luận án.

Chương 2. Trình bày phương pháp trích xuất đặc trưng HALF. Nội dung trình bày trong chương này được tổng hợp dựa trên kết quả các công trình nghiên cứu số 1, số 2, số 4 và số 6.

Chương 3. Trình bày phương pháp trích xuất tự động MPF nhờ học đặc trưng. Nội dung trình bày trong chương này được tổng hợp từ kết quả công trình nghiên cứu số 7 của tác giả.

Chương 4. Trình bày hai ứng dụng nhận dạng hoạt động người. Nội dung trình bày được tổng hợp dựa trên kết quả các công trình nghiên cứu số 4, số 5 và số 6 của tác giả.

Cuối cùng là một số kết luận về luận án.

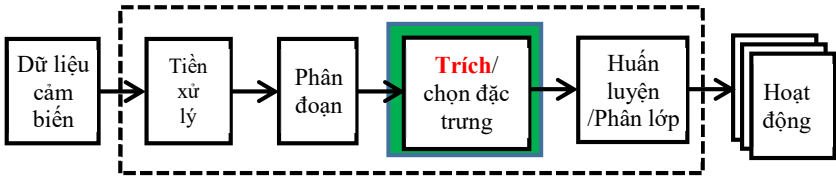
CHƯƠNG 1: TỔNG QUAN VỀ NHẬN DẠNG HOẠT ĐỘNG SỬ DỤNG CẢM BIẾN MANG TRÊN NGƯỜI

1.1. Kiến trúc hệ thống nhận dạng hoạt động người

Nhận dạng hoạt động người là quá trình giám sát và phân tích hành vi người dùng và trạng thái môi trường xung quanh nhằm suy diễn/nhận dạng các hoạt động đang xảy ra. Một trong những mục tiêu của nhận dạng hoạt động là cung cấp thông tin về hành vi của người dùng, từ đó cho phép hệ thống tính toán chủ động hỗ trợ người dùng trong công việc.

Một hệ thống nhận dạng hoạt động người nói chung nhận dữ liệu đầu vào là dữ liệu thô lấy ra từ các cảm biến và thông thường gồm các bước như sau: tiền xử lý, phân đoạn, trích chọn

đặc trưng (gồm trích xuất đặc trưng và lựa chọn đặc trưng), huấn luyện và phân lớp hoạt động (xem hình 1.1).



Hình 1.1. Kiến trúc hệ thống nhận dạng hoạt động người

1.2. Các phương pháp trích xuất đặc trưng

1.2.1. Trích xuất đặc trưng theo tri thức chuyên gia

Tín hiệu gia tốc là tín hiệu có mức độ dao động lớn, do đó rất khó có thể nhận dạng những mẫu cơ bản khi chỉ sử dụng các giá trị thô của tín hiệu chưa xử lý. Hầu hết các hệ thống HAR hiện thời sử dụng các đặc trưng theo miền thời gian hoặc tần số. Phương pháp trích chọn đặc trưng thường được sử dụng là tính toán các đại lượng thống kê trực tiếp trên dữ liệu thô đầu vào của cảm biến, độc lập trên từng khung dữ liệu được phân đoạn bởi cửa sổ trượt. Các đại lượng thống kê phổ biến nhất bao gồm: trung bình, độ lệch chuẩn, phương sai, năng lượng, trung bình đạo hàm, khoảng tứ phân vị, entropy, tương quan giữa các trục, skewness, độ nhọn (kurtosis). Trong miền tần số, các phương pháp như biến đổi Cô-sin rời rạc (Discrete Cosine Transform) và biến đổi Fourier (Fourier Transform) được áp dụng và mang lại những kết quả khả quan. Các đặc trưng thống kê được sử dụng nhiều nhất trong các nghiên cứu do tính toán đơn giản và hiệu năng cao trong rất nhiều bài toán nhận dạng hoạt động người.

Trong trường hợp bài toán phân biệt ngã và các hoạt động khác thì các đặc trưng thống kê nêu trên lại không hiệu quả. Lý do là ngã và một số hoạt động tương tự ngã khác như đứng ngồi, ngồi nằm, nhảy,... có tín hiệu với tính chất tương tự nhau và dễ gây nhầm lẫn. Cách tiếp cận truyền thống dựa trên ngưỡng có khả năng phát hiện ngã nhanh, đơn giản và tương đối chính xác, tuy vậy lại thường bị tỉ lệ cảnh báo giả cao và trong nhiều tình huống lại không hiệu quả, dẫn tới khả năng khái quát hóa thấp. Để tăng tính khái quát hóa và hiệu quả phân biệt ngã thì các cách tiếp cận dựa trên học máy gần đây như đề xuất sử dụng các tập gồm nhiều đặc trưng thống kê phức tạp hơn. Tuy vậy, độ phức tạp của tập đặc trưng lớn dẫn tới tính khả thi của các phương pháp này bị hạn chế do khó có thể tính toán nhanh trên hệ thống có năng lực xử lý thấp và thời lượng pin ngắn như đa số các thiết bị trợ giúp cá nhân di động thông minh mang theo người, đặc biệt là các hệ thống đòi hỏi thời gian thực. Do đó, chúng khó có thể ứng dụng được trong thực tế hiện nay, khi người dùng cần sự thuận tiện cao với các thiết bị di động trợ giúp cá nhân thông minh, nhỏ gọn. Đây chính là một vấn đề tồn tại trong nghiên cứu nhận dạng hoạt động sử dụng cảm biến mang trên người, đó là *cần phải có các phương pháp biểu diễn đặc trưng mới, hiệu quả có thể phân biệt tốt ngã và các hoạt động khác có đặc tính dữ liệu tương tự, có khả năng tính toán nhanh để chạy được trên các hệ thống yêu cầu xử lý theo thời gian thực và bị hạn chế về tài nguyên*. Vấn đề tồn tại này sẽ được bàn luận và giải quyết trong chương 2.

1.4.2. Trích xuất đặc trưng tự động bằng học đặc trưng

Các đặc trưng dựa trên tri thức chuyên gia như thống kê thường được sử dụng trong các hệ thống nhận dạng hoạt động do tính đơn giản và hiệu quả. Tuy nhiên, chúng chỉ hoạt động hiệu quả khi ứng dụng để phân biệt một số hoạt động đơn giản, cụ thể mà không đáp ứng được khi cần phân biệt tập các hoạt động sống hàng ngày (ADL -Activities of Daily Living) nhiều và đa dạng hơn. Một cách tiếp cận cho vấn đề này là sử dụng một tổ hợp các đặc trưng thống kê phức tạp hơn. Tuy nhiên cách tiếp cận này làm tăng chi phí và gây khó khăn cho việc xây dựng hệ thống nhận dạng hoạt động người.

Một phương pháp khác hiệu quả là sử dụng học đặc trưng, nghĩa là tìm cách tối ưu hóa một hàm mục tiêu để “bắt được” sự phù hợp của các đặc trưng cho mọi dữ liệu. Cách tiếp cận này có thể tìm được một biểu diễn đặc trưng chung cho mọi dữ liệu, đồng thời rất khả quan trong việc cải thiện độ chính xác nhận dạng. Các phương pháp khác được thử nghiệm và có kết quả tốt trong nhận dạng hoạt động người như: PCA, ECDF dựa trên PCA, học đặc trưng sử dụng các kỹ thuật học sâu với CNN và học đặc trưng dựa vào phương pháp túi từ (BoW – Bag of Word). Học đặc trưng dựa vào BoW tạo ra các đặc trưng đa mức. Có thể hiểu rằng các đặc trưng này được xác định bằng tổ hợp của các đặc trưng cơ sở của chuyển động cơ thể người. Đặc trưng đa mức không quá đòi hỏi năng lực xử lý cao như các đặc trưng dựa trên học sâu. Đặc trưng đa mức được tính thông qua các bước: phân cụm dữ liệu sử dụng một phương pháp phân cụm để tính các thống kê sự xuất hiện của nhóm trên một cửa sổ trượt, sau đó các đặc trưng có ý nghĩa được tạo nên thông

qua các tính toán về quãng thời gian, tần số và sự kiện đồng xảy ra từ dữ liệu. Các mô hình BoW cho nhận dạng hoạt động thường được xây dựng theo ba bước sau: trích xuất đặc trưng, học từ vựng (vocabulary hay codebook) và cuối cùng là xây dựng phân phối tần suất (histogram) của các từ (chuyển động). Các đặc trưng được trích xuất thường là các đặc trưng cục bộ, được tính toán từ các phân đoạn nhỏ của cửa sổ thời gian. Từ vựng có thể được học bằng cách sử dụng thuật toán phân cụm không giám sát để nhóm các đặc trưng thành các cụm (cluster). Trung tâm của cụm tạo thành một từ (hay véc-tơ mã/code vector), và một tập các từ tạo ra từ tập dữ liệu tạo thành từ vựng. Tần suất xuất hiện của các từ có trong từ vựng tạo nên biểu đồ tần suất của các từ.

Các đặc trưng đa mức có được thông qua học đặc trưng dựa trên mô hình BoW dựa trên các phương pháp phân cụm không giám sát như k -means có các hạn chế về độ phức tạp tốc độ tính toán khi dữ liệu đầu vào có số chiều lớn, gây ra khó khăn cho việc áp dụng trong thực tế. Vấn đề này sẽ được phân tích và giải quyết trong chương 3, với đề xuất của tác giả về sử dụng các đặc trưng MPF để tăng tốc độ tính toán và độ chính xác.

CHƯƠNG 2: NHẬN DẠNG HOẠT ĐỘNG VỚI CÁC ĐẶC TRƯNG HALF

2.1. Thống kê về phân phối tần suất góc của các hoạt động

Xem xét các phân phối xác suất góc cho nhiều loại hoạt động với dữ liệu lấy từ 3 tập dữ liệu sử dụng trong phần thử nghiệm cho thấy: (1) Ba kiểu ngã là *ngã trước*, *ngã sau* và *ngã*

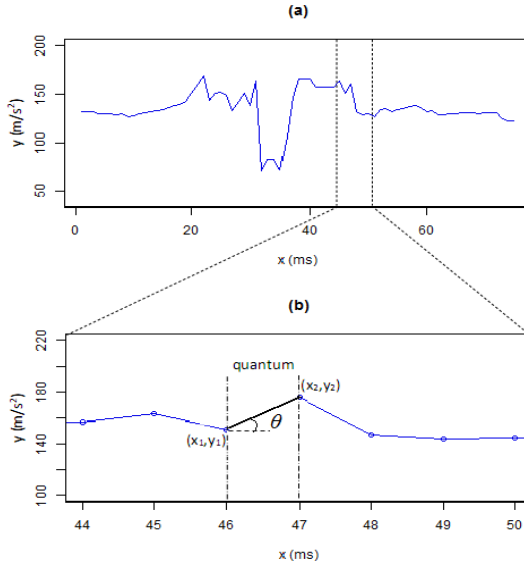
sang bên về cơ bản có phân phối xác suất tương tự nhau. (2) So sánh ngã với 6 hoạt động khác là *đứng*, *đi bộ*, *chạy bộ*, *nhảy*, *ngồi*, *đi cầu thang*, phân phối xác suất theo góc của các hoạt động này khác hẳn nhau, trừ *ngã* và *ngồi*. Điều này cho thấy việc sử dụng phân phối tần suất góc của các mảnh tín hiệu nói hai điểm có thể giúp phân biệt được nhiều hoạt động. *Ngồi* và *ngã* có phân phối xác suất gần giống nhau do có tính chất tín hiệu gia tốc tương tự nhau.

2.2. Các đặc trưng HALF

Đặc trưng HALF (Histograms of Angles of Line Fragments between Two Consecutive Points) được đề xuất với ý tưởng là với các loại chuyển động khác nhau, biểu đồ tín hiệu gia tốc sẽ có hình dạng thay đổi khác nhau. Đặc trưng về sự thay đổi gia tốc trong một cửa sổ thời gian của tín hiệu gia tốc có thể được nắm bắt thông qua thống kê, phân phối về tần suất góc của các đoạn nối hai điểm dữ liệu liên tiếp.

Thuật toán trích xuất đặc trưng như sau. Đầu tiên, mỗi khung tín hiệu được chia thành các mảnh nhỏ hơn (*quantum*) có độ dài l . Góc của mảnh được tạo bởi véc-tơ nối giữa điểm dữ liệu đầu và cuối của mảnh với trục ngang thời gian x như trên hình 2.6. Sau khi tính toán góc của mỗi mảnh, cần thực hiện tính một phân phối tần suất (*histogram*) của các góc nhờ việc chia toàn bộ khoảng giá trị góc sang dạng khoảng, được gọi là các *ngăn*, và đếm số các mảnh có góc thuộc từng *ngăn*. Khoảng giá trị của *ngăn* được học thông qua việc sử dụng các phương pháp phân cụm như k -means để nhóm các góc của mảnh trong các

khung trong tập huấn luyện thành M cụm. Phân phối tần suất đã tạo có thể được sử dụng làm đặc trưng cho khung đang xét.



Hình 2.6. Ví dụ về một mảnh (quantum) và hướng của nó

Để nắm được thông tin về vị trí tương đối của mảnh trong một cửa sổ, cần áp dụng thuật toán đã trình bày cho K đoạn dữ liệu nhỏ hơn liên tiếp nhau trên khung dữ liệu để tạo ra các véc-tơ đặc trưng cục bộ. Véc-tơ đặc trưng của cả khung được tạo ra bằng cách nối K véc-tơ đặc trưng cục bộ từ K đoạn trong khung.

Do tốc độ lấy mẫu cho các hoạt động sử dụng cảm biến gia tốc thường từ 50 Hz đến 100 Hz, nghiên cứu này chỉ sử dụng mảnh với độ dài bằng $l = 1$. Hai tham số còn lại có thể xác định được thông qua thực nghiệm hay tối ưu hóa.

Trong trường hợp gán nhãn dữ liệu đầu vào, độ phức tạp thuật toán bằng tổng độ phức tạp thuật toán của tất cả các

bước trừ bước phân cụm. Độ phức tạp cho tìm kiếm gần cụm chỉ mất $O(\log M)$ và thực tế số cụm M nhỏ hơn 100 nên độ phức tạp sẽ là $O(N)$.

2.3. Thử nghiệm

2.3.1. Phương pháp phát hiện người ngã

Thuật toán đề xuất thực hiện phát hiện ngã bằng cách phân loại một cửa sổ tín hiệu hoạt động thành “ngã” hoặc “không ngã”. Dữ liệu đầu vào là 3 luồng tín hiệu gia tốc theo ba trục x, y, z . Thuật toán phát hiện ngã gồm 3 bước: phân đoạn dữ liệu, trích xuất đặc trưng và phân lớp. Trong bước phân đoạn dữ liệu, cần thử nghiệm các khung cửa sổ trượt với các kích thước khác nhau để tìm ra giá trị tối ưu. Các đặc trưng HALF từ mỗi khung được trích xuất và sau đó được dùng làm dữ liệu đầu vào cho bộ phân lớp SVM dùng cho phát hiện ngã.

Do các dữ liệu huấn luyện cho mô hình phát hiện ngã mất cân bằng và dễ gây ra hiện tượng quá vừa dữ liệu, nghiên cứu này sử dụng phương pháp SDC của tác giả Akbani, giúp SVM có thể hoạt động hiệu quả hơn.

Các độ đo được sử dụng gồm độ chính xác (accuracy), độ nhạy (sensitivity), độ đặc hiệu (specificity).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2.1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2.2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (2.3)$$

2.3.2. Tập dữ liệu thử nghiệm

Phương pháp đề xuất được đánh giá thử nghiệm trên 3 tập dữ liệu ngã công khai, được sử dụng nhiều trong các nghiên

cứu về ngã, bao gồm: tập dữ liệu DLR, MobiFall2, tFall. Các tập dữ liệu trên chứa nhiều dữ liệu cảm biến về các sự kiện ngã và hoạt động thường ngày khác. Sau quá trình tiền xử lý, thu được 2771 khung ADL và 36 khung dữ liệu ngã từ tập dữ liệu DLR, 1832 khung ADL và 288 khung dữ liệu ngã từ MobiFall2, và 7816 khung ADL và 503 khung dữ liệu ngã từ tFall.

2.3.3. Kết quả thử nghiệm và đánh giá

Phương pháp đã đề xuất (pp3) được so sánh với ba phương pháp khác được đề xuất bởi Phạm và cộng sự (pp1), Jantaraprim và cộng sự (pp2) và Özdemir (pp4), trên cả 3 tập dữ liệu. Các kết quả được mô tả trong bảng 2.3 và được trình bày với cấu trúc độ chính xác trung bình \pm độ lệch chuẩn.

Bảng 2.3. Các kết quả phát hiện ngã trên 3 tập dữ liệu

Phương pháp	DLR	MobiFall2	tFall
(pp1)	98,11% \pm 0,20%	96,51% \pm 0,41%	95,62% \pm 0,53%
(pp2)	98,83% \pm 0,50%	96,22% \pm 0,34%	96,18% \pm 0,46%
(pp3)	99,93% \pm 0,08%	99,20% \pm 0,45%	98,70% \pm 0,32%
(pp4)	99,27% \pm 0,19%	98,76% \pm 0,23%	97,92% \pm 0,41%

Kết quả chỉ ra rằng phương pháp đã đề xuất có kết quả cao nhất trên cả 3 tập dữ liệu. Mức tăng của độ chính xác với phương pháp đề xuất so với (pp1) và (pp2) là đáng kể trên cả 3 tập dữ liệu, còn so với (pp4) là xấp xỉ. Đồng thời, theo số liệu tốc độ như bảng 2.4 và phân tích về độ phức tạp thuật toán, có

thể thấy được rằng phương pháp trích xuất đặc trưng HALF có khả năng tính toán nhanh do độ phức tạp thấp, phù hợp với các thiết bị trợ giúp cá nhân thông minh, có yêu cầu xử lý thời gian thực.

Bảng 2.4. Tốc độ phát hiện ngã trên tập dữ liệu tFall

Phương pháp	Tập dữ liệu tFall (1000 mẫu)
(pp1)	2,86 giây
(pp2)	1,15 giây
(pp3)	3,01 giây
(pp4)	15,02 giây

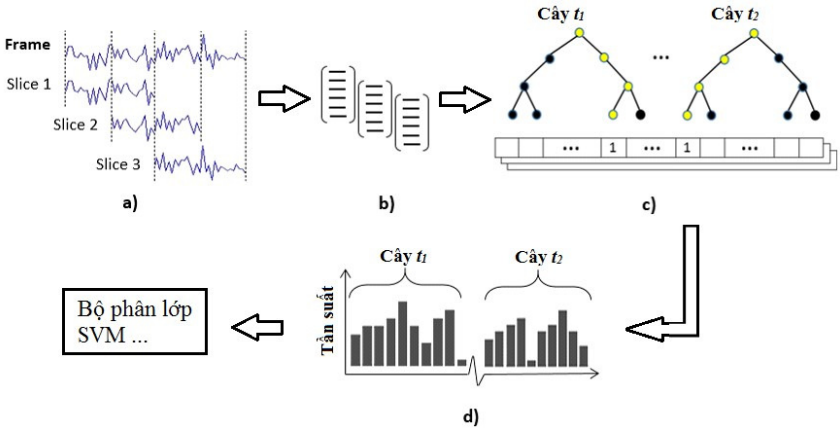
CHƯƠNG 3: HỌC CÁC ĐẶC TRƯNG ĐA MỨC MPF

3.1. Giới thiệu

Chương này đề xuất một phương pháp thay thế để xây dựng tập từ vựng chuyên động (motion vocabulary) và thực hiện gán các biểu diễn đặc trưng bằng cách sử dụng một kết hợp lượng nhỏ các cây quyết định hay được gọi là motion primitive forests (MPF). MPF hoạt động giống như các cây phân nhóm thông qua việc nhóm các mô tả cục bộ tương tự trong các nút lá. Quá trình này được định hướng bởi các nhãn của hoạt động và nhanh hơn nhiều so với k -means. Lợi thế của việc sử dụng cụm cây quyết định trong việc học từ vựng (vocabulary/codebook) đã được minh chứng trong các nghiên cứu xử lý ảnh, nhưng đây là lần đầu tiên được áp dụng cho nhận dạng hoạt động dựa trên cảm biến. Đề xuất thứ hai là xây dựng các đặc trưng cục bộ mới đơn giản, phù hợp khi sử dụng với MPF. Các đặc trưng này là các giá trị bao gồm những giá trị thô

chưa xử lý của các điểm dữ liệu hoặc các giá trị tổng, hiệu của 2 điểm dữ liệu. Những giá trị này có thể tính toán nhanh, đồng thời có thể giúp cho nhận dạng hoạt động chính xác như các đặc trưng cục bộ phức tạp khác.

3.2. Phương pháp học đặc trưng MPF



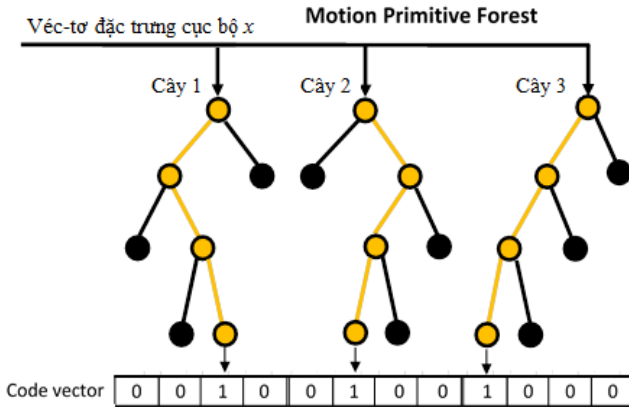
Hình 3.1. (a) 3 lát tín hiệu (slice) được phân đoạn từ một khung hoạt động (frame) với độ chồng lấn là 50%, (b) các đặc trưng được trích xuất từ mỗi lát tín hiệu, (c) MPF với vai trò là từ vựng chuyển động, (d) phân phối tần suất của chuyển động gốc (motion primitive histogram)

Đầu vào là chuỗi dữ liệu cảm biến liên tục 3 chiều. Bằng cách sử dụng một cửa sổ trượt, chuỗi dữ liệu cảm biến liên tục được phân đoạn thành các khung có kích thước bằng nhau, trong đó chiều dài của chúng có thời lượng dài hơn khoảng thời gian cho một hoạt động bất kỳ. Mỗi khung sau đó sẽ được phân chia thành các lát tín hiệu (hay các slice) bằng nhau (có thể chồng lấn) để mỗi lát tín hiệu nhỏ hơn nhiều so với

khung (frame) của nó (hình 3.1(a)). Từ mỗi lát tín hiệu, trích xuất các đặc trưng để tạo thành một véc-tơ đặc trưng cục bộ. Trong giai đoạn huấn luyện, nếu một khung chứa một hoạt động thì nhãn của hoạt động đó sẽ được gán cho toàn bộ khung cũng như các lát tín hiệu của nó. Trong pha dự đoán, bài toán trở thành dự đoán nhãn lớp hoạt động cho mỗi khung.

Các véc-tơ đặc trưng cục bộ từ tất cả các lớp hoạt động (và lớp nền) được huấn luyện, sau đó được gộp lại với nhau và lượng tử hóa để tạo thành từ vựng chuyển động gốc (motion primitive vocabulary) hay là bộ mã (codebook). Đây là quá trình gán một chỉ số cho mỗi véc-tơ đặc trưng cục bộ để các véc-tơ đặc trưng tương tự nhau sẽ có xác suất cao có cùng chung chỉ số. Phương pháp được đề xuất ở đây sử dụng MPF (motion primitive forests), là rừng ngẫu nhiên được sử dụng để phân cụm và ánh xạ các véc-tơ đặc trưng cục bộ thành chuyển động gốc (hình 3.1(c)).

Rừng ngẫu nhiên MPF sẽ xử lý mỗi nút lá từ mỗi cây như là một chuyển động gốc riêng biệt. Nói cách khác, các nút lá xác định một phân vùng, và mỗi nút lá tương ứng với một cụm các véc-tơ đặc trưng cục bộ tương tự nhau. Đối với mỗi véc-tơ đặc trưng cục bộ đầu vào, MPF trả về một tập các chỉ số của các nút lá, mỗi chỉ số nằm trong một cây. Các chỉ số nút lá này được dùng để tạo thành véc-tơ mã (code vector) như minh họa trong hình 3.3, trong đó “1” là véc-tơ đầu vào đã được ánh xạ tương ứng với chuyển động gốc.



Hình 3.3. Véc-tơ mã (code vector) tạo ra từ một rừng ngẫu nhiên với 3 cây cho một véc-tơ đặc trưng cục bộ. Đường dẫn đến các nút lá kết quả được hiển thị bằng màu vàng.

Các rừng ngẫu nhiên MPF, có tốc độ phân cụm nhanh hơn k -means và phương pháp gán theo hàng xóm gần nhất. Đồng thời có khả năng sinh ra các chuyển động gốc có độ phân biệt cao hơn kể cả khi có số lượng lớn các lát tín hiệu không nằm trong các khung có chứa hoạt động (background slice). Kết quả này có được chủ yếu là do các nhãn hoạt động được dùng trong việc định hướng xây dựng các cây quyết định. Trong bước tiếp theo, các chuyển động gốc được học bởi MPF được kết hợp với mô hình túi từ thông qua tổng hợp các véc-tơ mã của tất cả các lát tín hiệu thuộc về một khung để tạo thành véc-tơ phân phối tần suất các chuyển động gốc cho mỗi khung này (Hình 3.1(d)). Véc-tơ phân phối tần suất sau khi đã chuẩn hóa được sử dụng làm đặc trưng của khung trong bộ phân lớp cuối cùng.

Đặc trưng cục bộ đề xuất yêu cầu rất ít hoặc không cần tính toán bởi vì chúng là giá trị thô chưa xử lý của các điểm dữ

liệu hoặc là các tổng hay hiệu của giá trị của 2 điểm dữ liệu. Các đặc trưng đơn giản này, khi được kết hợp với MPF, sẽ cung cấp độ chính xác có thể so sánh được với các đặc trưng phức tạp hơn, trong khi chi phí tính toán lại thấp hơn nhiều.

Độ phức tạp về thời gian chạy cho việc gán nhãn một đặc trưng cục bộ của một rừng với T cây là $O(T \log N)$. Trong trường hợp ở đây số cây nhỏ hơn 10 (xem thực nghiệm) nên có thể bỏ qua, dẫn tới độ phức tạp yêu cầu là $O(\log N)$, nhanh hơn nhiều so với k -means với độ phức tạp thuật toán là $O(kp)$.

3.3. Tập dữ liệu thử nghiệm

Các tập dữ liệu thử nghiệm bao gồm Activity Prediction (AP), Opportunity (OP) và Skoda (SK), chứa dữ liệu gia tốc cho các hoạt động hàng ngày và trong sản xuất, được sử dụng rộng rãi trong nhiều nghiên cứu về nhận dạng hoạt động người.

3.4. Thử nghiệm và kết quả

Độ đo được dùng là độ chính xác tổng thể, được tính bằng tỷ lệ số khung được phân loại đúng trên tổng số khung.

a) So sánh với phương pháp học đặc trưng sử dụng k -means

Thử nghiệm này để so sánh MPF với phương pháp học đặc trưng dựa trên k -means, trong việc xây dựng từ vựng và tìm kiếm hàng xóm gần nhất để gán chuyển động gốc. Kết quả so sánh cho thấy mức cải thiện về độ chính xác tăng lên khá nhiều khi dùng phương pháp MPF. Đồng thời, kết quả cũng cho thấy rằng trong khi MPF có thời gian chạy tăng không đáng kể khi kích thước từ vựng tăng lên, thì thời gian chạy của k -means lại tăng tuyến tính. Với kích thước là 200, thời gian chạy của k -

means đã lớn hơn nhiều so với MPF (580 giây so với 54 giây, trên tập dữ liệu SK).

d) So sánh với các phương pháp khác

Thử nghiệm này so sánh độ chính xác nhận dạng của MPF khi sử dụng các đặc trưng cục bộ đơn giản với 2 phương pháp được cho là tốt nhất hiện nay. Phương pháp thứ nhất là học đặc trưng sử dụng PCA và ECDF. Phương pháp thứ hai sử dụng CNN với partial weight sharing cho cả việc học đặc trưng và phân lớp.

Bảng 3.2. So sánh độ chính xác trong phân lớp

	AP	OP	SK
MPF + 1-NN	97,93%	84,17%	92,58%
MPF + SVM	98,48%	85,68%	95,08%
PCA + ECDF	95,75%	79,39%	90,48%
CNN	96,21%	77,34%	86,73%

Bảng 3.2 cho thấy cả hai biến thể của MPF (với 1-NN và SVM), luôn đạt được độ chính xác cao hơn PCA+ECDF và CNN trong cả ba tập dữ liệu. SVM cho kết quả tốt hơn so với 1-NN, điều này là hợp lý vì số chiều của dữ liệu đầu vào (số chiều các véc-tơ phân phối tần suất) là khá cao. Những cải tiến về độ chính xác của MPF so với PCA+ECDF và CNN là đặc biệt đáng chú ý trên các tập dữ liệu OP và SK (khác biệt giữa MPF và CNN là gần 10%). Kết quả này là do các nhân hoạt động được dùng trong việc định hướng xây dựng các cây quyết định trong MPF. Nhờ vậy, MPF có khả năng sinh ra các chuyển động gốc có độ phân biệt cao hơn kể cả khi có số lượng lớn các

lát tín hiệu không nằm trong các khung có chứa hoạt động (background slice).

CHƯƠNG 4: XÂY DỰNG MỘT SỐ ỨNG DỤNG NHẬN DẠNG HOẠT ĐỘNG

4.1. Phát hiện người ngã trong thời gian thực

4.1.1. Giới thiệu

Phần này trình bày một hệ thống tự động phát hiện ngã sử dụng bộ cảm biến gia tốc, dựa trên các đặc trưng đề xuất trong chương 2. Cảm biến gia tốc này được tích hợp sẵn trong Wii Remote có giá thành tương đối rẻ và thời lượng pin sử dụng rất khả quan. Hệ thống cho tỉ lệ phát hiện ngã cao và chi phí thấp hơn so với các hệ thống phát hiện ngã khác.

4.1.2. Phần cứng và môi trường thu thập dữ liệu

Trong hệ thống này, các đối tượng đeo Wii Remote trên hông do cảm biến đeo bên hông có thể cung cấp các tính năng tốt cho việc phát hiện nhiều loại hoạt động, ví dụ như ngã, chạy, đi bộ, leo cầu thang, v.v.

4.1.3. Thuật toán phát hiện ngã

Hệ thống sử dụng đặc trưng đề xuất trong chương 2 và phương pháp phát hiện ngã đã đề xuất trong phần 2.3.1. Luồng tín hiệu cảm biến đầu vào sau khi tiền xử lý sẽ được phân đoạn thành các khung dữ liệu, thực hiện trích xuất đặc trưng và đưa ra dự đoán. Hệ thống được huấn luyện sử dụng tập dữ liệu ngã mô tả ở phần 4.1.4, với các tham số là số ngăn M và số đoạn trong một khung K được tối ưu hóa trên tập dữ liệu đầu vào, sử dụng tập tối ưu dành riêng (hold-out validation set). Việc lựa

chọn độ dài cửa sổ trượt cho bước phân đoạn cũng được tiến hành để tìm ra độ dài phù hợp. Độ dài phù hợp nhất được chọn là 1,8 giây.

4.1.4. Tập dữ liệu thử nghiệm

Tập dữ liệu huấn luyện cho ứng dụng là dữ liệu được thu thập từ 12 đối tượng. Các đối tượng được yêu cầu thực hiện 12 hoạt động hàng ngày bao gồm đi bộ, nhảy, đi lên cầu thang, đi xuống cầu thang, chạy, vươn người, lau nhà, đánh máy, ngồi xuống lúc đang đứng, đứng lên lúc đang ngồi, đánh răng, hút bụi và 12 tư thế ngã khác nhau. Số hoạt động ngã là 144 và các hoạt động còn lại thuộc về 12 hoạt động hàng ngày khác (không phải ngã).

4.1.5. Kết quả thử nghiệm

Thử nghiệm cho kết quả khá cao, với độ chính xác là $96,25\% \pm 0,38\%$, độ nhạy là $97,67\% \pm 0,23\%$ và độ đặc hiệu là $84,34\% \pm 2,32\%$.

4.1.6. Phần mềm

Phần mềm trên máy tính được xây dựng để thu nhận tín hiệu cảm biến gia tốc trên Wii Remote qua Wii dongle, truyền qua Bluetooth và qua đó phát hiện ngã trong thời gian thực.

4.2. Xác thực người dùng với chữ ký 3D – SIGVER3D

4.2.1. Vấn đề xác thực cho các hệ thống mang trên người

Tác giả đề xuất một phương thức xác thực mới cho người dùng thiết bị động với chữ ký 3D, gọi là SigVer3D. Phương thức này có cả ưu điểm về tính an toàn của chữ ký và thuận tiện cho người dùng khi muốn “đăng nhập” vào thiết bị di động khi không có các thiết bị nhập liệu truyền thống như

màn hình cảm ứng, hay bàn phím hoặc không thuận tiện. Phương pháp đề xuất sử dụng các dữ liệu cảm biến gia tốc như sau. Người dùng cảm thiết bị di động và vẽ chữ ký của họ trong không gian để đăng nhập. Khi người dùng thực hiện ký, các dữ liệu gia tốc được sinh ra và được hệ thống sử dụng để kiểm tra người dùng. Một lý do để chọn chữ ký 3D làm mật khẩu là nó dễ nhớ và dễ thực hiện, đồng thời khó giả mạo bởi người khác do sử dụng các đặc trưng sinh trắc học và vật lý bổ sung để làm tăng tính an toàn cho việc xác thực chữ ký.

4.2.2. Phân cứng và môi trường thu thập dữ liệu

Hệ thống này được cài đặt trên một điện thoại di động thông minh Samsung S3 có trang bị một cảm biến gia tốc gắn trong dùng để cảm nhận gia tốc theo ba trục.

4.2.3. Thuật toán xác thực người dùng

Bài toán xác thực người dùng ở đây thực chất là bài toán phân lớp nhị phân: phát hiện chữ ký giả mạo và chữ ký chính chủ. Luồng dữ liệu lấy được từ cảm biến gia tốc trên điện thoại được tiền xử lý để loại bỏ nhiễu, thực hiện trích xuất đặc trưng và đưa ra dự đoán. Các đặc trưng được sử dụng là các đặc trưng đã đề xuất trong chương 2 với các tham số là số ngăn M và số đoạn trong một khung K được tối ưu hóa trên tập dữ liệu đầu vào, sử dụng tập tối ưu dành riêng (hold-out validation set). Các đặc trưng trích xuất sẽ được sử dụng làm dữ liệu đầu vào cho bộ phân lớp SVM.

Đánh giá một hệ thống kiểm tra đòi hỏi sự phân tích của hai loại lỗi: tỷ lệ từ chối sai (FRR), tỷ lệ chấp nhận sai (FAR).

$$FRR = \frac{FN}{FN + TP} \quad (4.1)$$

$$FAR = \frac{FP}{FP + TN} \quad (4.2)$$

Trong đó, FP là dương tính giả (chữ ký giả mạo được chấp nhận), TP là dương tính thực (chữ ký gốc được phát hiện), FN là âm tính giả (chữ ký gốc bị từ chối không đúng), và TN là âm tính thực (chữ ký giả mạo bị phát hiện). Các số đo khác là diện tích đường cân bằng lỗi (AUC) và tỉ lệ lỗi tương đương (EER) cũng được sử dụng

4.2.4. Tập dữ liệu thử nghiệm

Có 30 đối tượng và tổng cộng 1800 mẫu chữ ký 3-D trong đó 300 là gốc và 1500 mẫu là chữ ký 3-D giả mạo.

4.2.5. Kết quả thử nghiệm

Kết quả thử nghiệm cho thấy hệ thống xác thực có độ chính xác tương đối cao và khá tin cậy. Giá trị AUC trung bình đạt được với giá trị là 98,3%, và giá trị EER = FRR = FAR = 1,4%. Kết quả này đã cho thấy rằng chữ ký động với việc sử dụng các đặc trưng sinh trắc học và vật lý bổ sung đã làm tăng tính an toàn cho việc xác thực. Hai người có đặc tính sinh lý khác nhau, kể cả bất chước động tác ký giống nhau cũng khó có thể thành công.

4.2.6. Phần mềm

Phần mềm xác thực chữ ký 3D được xây dựng trên điện thoại di động có hệ điều hành Android phiên bản 4.2, có trang bị cảm biến gia tốc.

DANH MỤC CÁC CÔNG TRÌNH CÔNG BỐ

TẠP CHÍ KHOA HỌC

- [1] Pham Cuong, Diep Nguyen Ngoc, Phuong Tu Minh (2013), A wearable sensor based approach to real-time fall detection and fine-grained activity recognition, *Journal of Mobile Multimedia*, Rinton Press. **9**(1-2), p. 15–26.
- [2] Đinh Khắc Quân, Phạm Văn Cường, Nguyễn Ngọc Điệp, Từ Minh Phương (2013), Tự động phát hiện người ngã trong thời gian thực sử dụng bộ cảm biến gia tốc, *Tạp chí Khoa học và công nghệ*, Viện Hàn lâm KHCN Việt Nam. **51**(1A), 108-116.
- [3] Nguyễn Ngọc Điệp, Hà Quang Tấn (2015), Nhận dạng cử chỉ người trong thời gian thực bằng thiết bị di động, *Tạp chí Khoa học và công nghệ*, Viện Hàn lâm KHCN Việt Nam. **53**(2C), 1-13.

HỘI NGHỊ KHOA HỌC

- [4] Diep Nguyen Ngoc, Pham Cuong, Phuong Tu Minh (2014), A classifier based approach to real-time fall detection using low-cost wearable sensors, *Proceedings of the 5th SoICT*, ACM. p. 14–20.
- [5] Diep Nguyen Ngoc, Pham Cuong, Phuong Tu Minh (2015), SigVer3D: Accelerometer Based Verification of 3-D Signatures on Mobile Devices, *Proceedings of the 6th International Conference KSE 2014*, Springer. p. 353–65.
- [6] Diep Nguyen Ngoc, Pham Cuong, Phuong Tu Minh (2016), An Orientation Histogram based Approach for Fall Detection using Wearable Sensors, *Proceedings of the 14th PRICAI*, Springer. p. 354-366.
- [7] Diep Nguyen Ngoc, Pham Cuong, Phuong Tu Minh (2016), Motion Primitive Forests for Human Activity Recognition using Wearable Sensors, *Proceedings of the 14th PRICAI*, Springer. p. 340-353.