

**ĐẠI HỌC QUỐC GIA HÀ NỘI  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

**NGUYỄN VIỆT ANH**

**PHÁT HIỆN NGÃ SỬ DỤNG ĐẶC TRƯNG  
CHUYỂN ĐỘNG VÀ HÌNH DẠNG CƠ THỂ  
DỰA TRÊN CAMERA ĐƠN**

Ngành: Công nghệ thông tin

Chuyên ngành: Kỹ thuật phần mềm

Mã số: 60480103

**TÓM TẮT LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN**

**Hà Nội - 2016**

# MỤC LỤC

Danh mục hình vẽ.....	3
Danh mục bảng.....	4
MỞ ĐẦU .....	5
CHƯƠNG 1. TỔNG QUAN BÀI TOÁN PHÁT HIỆN NGÃ TỰ ĐỘNG .	8
1.1. Phát hiện ngã sử dụng thiết bị mang theo người .....	8
1.1.1. Gia tốc kế gắn trên cơ thể .....	8
1.1.2. Cảm biến tích hợp trên điện thoại thông minh .....	9
1.1.3. Xu hướng, ưu điểm và hạn chế.....	9
1.2. Phát hiện ngã dựa trên phân tích dữ liệu video .....	9
1.2.1. Phát hiện ngã sử dụng camera đơn .....	10
1.2.2. Phát hiện ngã sử dụng hệ multi camera .....	10
1.2.3. Phát hiện ngã sử dụng Camera độ sâu .....	10
CHƯƠNG 2. CƠ SỞ LÝ THUYẾT .....	11
2.1. Tổng quan về xử lý ảnh số .....	11
2.1.1. Ảnh kỹ thuật số.....	11
2.1.2. Xử lý ảnh số.....	12
2.1.3. Các phép toán chính trong xử lý ảnh.....	13
2.2. Tổng quan về thị giác máy tính.....	13
2.2.1. Khái niệm và sơ lược lịch sử .....	13
2.2.2. Ứng dụng .....	13
2.2.3. Hệ thống các kỹ thuật thị giác máy.....	13
2.2.4. Phân tích video (video analysis).....	14

2.2.5.	Bài toán phát hiện hành động (action detection) .....	14
CHƯƠNG 3.	PHƯƠNG THỨC ĐỀ XUẤT .....	16
3.1.	Tổng quan.....	16
3.2.	Phân tách vùng chuyển động.....	17
3.2.1.	Một số thuật toán trừ nền.....	17
3.2.2.	Áp dụng kỹ thuật trừ nền, phân tách vùng chuyển động .....	17
3.1.	Trích rút đặc trưng chuyển động .....	18
3.1.1.	Motion History Image.....	18
3.1.2.	Image Moments .....	18
3.1.1.	Trích rút đặc trưng chuyển động.....	18
3.2.	Trích rút đặc trưng hình dạng cơ thể.....	22
3.2.1.	Kỹ thuật fitting ellipse .....	22
3.2.2.	Áp dụng fitting ellipse đo lường đặc trưng hình dạng.....	22
3.3.	Phát hiện ngã.....	23
CHƯƠNG 4.	THÍ NGHIỆM VÀ ĐÁNH GIÁ .....	24
4.1.	Tập dữ liệu và phương pháp đánh giá hiệu quả thuật toán.....	24
4.1.1.	Tập dữ liệu thực nghiệm.....	24
4.1.2.	Phương pháp đánh giá độ hiệu quả của giải thuật .....	24
4.2.	Cài đặt thí nghiệm .....	24
4.3.	Kết quả và thảo luận.....	25
CHƯƠNG 5.	KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN.....	26
TÀI LIỆU THAM KHẢO .....		29

## Danh mục hình vẽ

Hình 1.2. Minh hoạ hệ thống phát hiện ngã tự động dựa trên phân tích video	9
Hình 2.1. Hệ thống phân tích ảnh số	12
Hình 3.1. Luồng hoạt động của hệ thống phát hiện ngã được đề xuất	16
Hình 3.9. Một ví dụ phân tách vùng chuyển động	17
Hình 3.11. So sánh phương thức xác định hướng chuyển động	19
Hình 3.12. Xác định $M_{rate}$ lỗi trong thời điểm gần kết thúc chuyển động	20
Hình 3.13. Ví dụ cho ước lượng độ lớn chuyển động	21
Hình 3.15. Minh hoạ sự thay đổi hình dạng cơ thể khi ngã	23
Hình 3.16. Quy ước về góc trong xác định hướng chuyển động và góc nghiêng cơ thể	23

## **Danh mục bảng**

Bảng 4.1. Bảng mô tả các tập dữ liệu thực nghiệm	24
Bảng 4.2. Kết quả thực nghiệm	25

## MỞ ĐẦU

Theo số liệu của bộ Y tế [1], tỷ lệ người cao tuổi hiện chiếm 10,5%, dự đoán tăng lên 23% dân số cả nước năm 2040. Và tuy tuổi thọ trung bình tăng nhanh nhưng do chất lượng cuộc sống, chế độ dinh dưỡng và điều kiện chăm sóc y tế, số người cao tuổi có sức khỏe tốt chỉ chiếm khoảng 5% trong khi 95% còn lại không khỏe mạnh. Người cao tuổi thường mắc các chứng bệnh như tim mạch; phổi – phế quản; đái tháo đường; suy giảm trí tuệ... Đó là những chứng bệnh dễ dẫn đến đột quy.

Theo [2], đời sống gia đình của người cao tuổi đang thay đổi. Tỷ lệ người cao tuổi sống cùng con cái đang giảm nhanh, tỷ lệ hộ gia đình người cao tuổi sống cô đơn hoặc chỉ có vợ chồng người cao tuổi tăng lên đáng kể. Đó là hệ quả của việc di cư khi người trong độ tuổi lao động tập trung tại các thành phố lớn để tìm kiếm cơ hội việc làm hoặc thậm chí là di cư quốc tế.

Trong bối cảnh người cao tuổi sống cô đơn và không khỏe mạnh, người già gặp rất nhiều nguy hiểm khi đột quy hay ngã mà không được phát hiện, cấp cứu kịp thời. Luận văn này nghiên cứu về các phương thức phát hiện ngã tự động nhằm góp phần tìm ra giải pháp gia tăng sự an toàn cho người cao tuổi sống một mình.

Một thực trạng về điều kiện y tế khác là sự quá tải của bệnh viện khi thường xuyên xảy ra việc nhiều người bệnh nằm chung một giường. Phòng bệnh vốn chật chội lại càng chật chội bởi cứ mỗi một người ốm cần ít nhất một người nhà chăm sóc. Điều này gây ra mệt mỏi cho cả người bệnh và người chăm sóc, làm lãng phí sức lao động của xã hội khi người khỏe mạnh phải nghỉ làm, cũng như gây cản trở các y bác sĩ trong khi thăm khám. Nếu có một hệ thống giám sát bệnh nhân tự động sẽ giúp giảm bớt số người chăm sóc, dẫn đến giảm tải cho bệnh viện. Một phương thức hiệu quả giúp tự động giám sát, phát hiện ngã cũng sẽ góp phần giải quyết bài toán trên.

### **Mục đích nghiên cứu**

Mục tiêu nghiên cứu của luận văn là tìm hiểu, quan sát để tìm ra các đặc điểm của việc ngã, định nghĩa được sự kiện ngã. Từ đó đề xuất một phương thức phát hiện ngã dựa trên các quan sát quá trình ngã.

## **Đối tượng và phạm vi nghiên cứu**

Do đặc điểm là một quốc gia đang phát triển với mức thu nhập bình quân thấp, các phương thức phát hiện ngã tự động phải là các giải pháp chi phí thấp, dựa trên các tài nguyên phổ biến, luận văn này tập trung vào các phương thức phát hiện ngã dựa trên phân tích dữ liệu video thu được từ camera giám sát. Đối tượng nghiên cứu bao gồm lý thuyết về xử lý ảnh số, xử lý video số, thị giác máy tính, các đặc điểm của hành động ngã và cách thức phát hiện việc ngã.

## **Phương pháp nghiên cứu**

Phương pháp nghiên cứu khi thực hiện luận văn là tìm hiểu từ cơ sở lý thuyết chung về xử lý ảnh số, video số, thị giác máy tính, sau đó tìm hiểu về bài toán phát hiện ngã tự động từ các nghiên cứu đã được công bố và các kết quả đã đạt được. Từ đó cải tiến, đề xuất các kỹ thuật nhằm nâng cao hiệu quả phát hiện ngã.

## **Đóng góp mới của luận văn**

Luận văn này đã cải tiến một số kỹ thuật và đề xuất một phương thức phát hiện ngã tự động dựa trên phân tích dữ liệu video; cài đặt thành công thuật toán phát hiện ngã với kết quả rất khả quan với tốc độ tính toán đảm bảo hoạt động thời gian thực; công bố kết quả nghiên cứu với tiêu đề “Single camera based Fall detection using Motion and Human shape Features” tại hội thảo quốc tế The Seventh International Symposium on Information and Communication Technology (SoICT 2016).

## **Kết cấu luận văn**

Ngoài phần mở đầu và phần tham khảo, luận văn này được tổ chức thành 5 chương với các nội dung chính như sau:

- **Chương 1: Tổng quan bài toán phát hiện ngã tự động**
  - Giới thiệu chung về bài toán
  - Các nghiên cứu đã công bố liên quan đến bài toán
- **Chương 2: Cơ sở lý thuyết**
  - Tổng quan về xử lý ảnh số

- Tổng quan về thị giác máy tính
- Tổng quan về phân tích video
- Tổng quan bài toán phát hiện hành động trong dữ liệu video
- **Chương 3: Phương thức đề xuất**
  - Tổng quan về phương thức đề xuất
  - Trình bày phương thức tách vùng chuyển động trong video
  - Trình bày về trích rút đặc trưng chuyển động
  - Trình bày về trích rút đặc trưng hình dạng cơ thể
  - Trình bày về quan sát các đặc trưng, đưa ra kết luận về việc ngã
- **Chương 4: Thí nghiệm và đánh giá**
  - Mô tả tập dữ liệu dùng để thí nghiệm
  - Trình bày phương pháp đánh giá độ hiệu quả của phương thức
  - Trình bày về cài đặt cấu hình thí nghiệm
  - Trình bày về kết quả thí nghiệm, giải thích về kết quả thí nghiệm
- **Chương 5: Kết luận và hướng phát triển**



## CHƯƠNG 1.

# TỔNG QUAN BÀI TOÁN PHÁT HIỆN NGÃ TỰ ĐỘNG

Một hệ thống phát hiện ngã tự động có thể được định nghĩa như một hệ thống trợ giúp với nhiệm vụ chính là báo động khi có sự kiện ngã xảy ra. Hệ thống này phải đảm bảo hoạt động thời gian thực để giảm thiểu thời gian người ngã nằm trên sàn từ sau thời điểm ngã đến khi được người chăm sóc phát hiện. Khoảng thời gian này là yếu tố chủ chốt quyết định mức độ nghiêm trọng sau ngã. Rất nhiều người già không thể tự di chuyển hoặc gọi trợ giúp sau khi ngã và đối mặt với các mối nguy hiểm cho sức khỏe. Trong các nghiên cứu được công bố gần đây, có thể phân loại các hướng nghiên cứu về bài toán phát hiện ngã thành các nhóm chính: Phát hiện ngã dựa trên thiết bị cảm biến mang theo người; dựa trên cảm biến tích hợp trên điện thoại di động thông minh; dựa trên camera độ sâu (depth camera); và dựa trên camera thông thường. Phần tiếp theo của chương này sẽ tóm lược khái quát các hướng nghiên cứu chính kể trên.

### 1.1. Phát hiện ngã sử dụng thiết bị mang theo người

Thiết bị mang theo người có thể được định nghĩa là các thiết bị cảm biến điện tử nhỏ có thể cầm theo, hoặc đính trên quần áo. Phần lớn các thiết bị phát hiện ngã mang theo người sử dụng cảm biến đo gia tốc. Trong đó có thể kết hợp cảm biến khác như con quay hồi chuyển để thu thập thông tin về vị trí của người mang. Việc sử dụng các cảm biến kể trên có thể giúp đánh giá dáng đi, sự cân bằng, mức độ chuyển động và vị trí cơ thể của người mang, giúp dự đoán về việc ngã. Xu hướng sử dụng thiết bị đeo được tăng lên trong những năm gần đây do sự phổ biến của các cảm biến giá rẻ được tích hợp sẵn trong điện thoại thông minh.

#### 1.1.1. Gia tốc kế gắn trên cơ thể

Thông tin về sự gia tăng tốc độ chuyển động trong quá trình ngã được thu thập dựa trên sử dụng các gia tốc kế ba trục độc lập được gắn trên các vị trí khác nhau của cơ thể. Sau đó, các kỹ thuật thường được áp dụng để xác định ngã bao gồm: i) sử dụng ngưỡng, trong đó việc ngã được ghi nhận nếu độ gia tăng vận tốc đạt ngưỡng xác định trước; ii) sử dụng học máy (machine learning) để phân loại giữa ngã và không phải ngã.

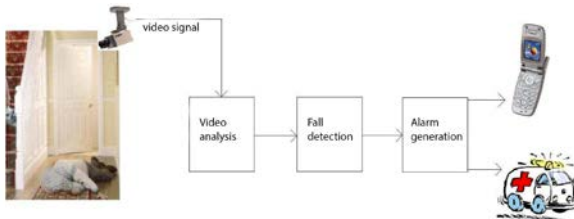
### 1.1.2. Cảm biến tích hợp trên điện thoại thông minh

Ngày nay, điện thoại di động thông minh dần trở lên rất phổ biến và thường được tích hợp sẵn một cách phong phú các loại cảm biến như gia tốc kế, la bàn số, GPS, con quay hồi chuyển, micro và camera. Một số nghiên cứu đã khai thác ưu thế kể trên để đưa ra các phương thức phát hiện ngã dựa trên sử dụng điện thoại thông minh. Trong đó, một số thuật toán đơn giản sử dụng kỹ thuật phân ngưỡng như [23, 27, 31, 44, 46]. Một số nghiên cứu khác sử dụng các kỹ thuật học máy như [39, 45].

### 1.1.3. Xu hướng, ưu điểm và hạn chế

## 1.2. Phát hiện ngã dựa trên phân tích dữ liệu video

Ngày nay, các thiết bị camera ngày càng trở lên phổ biến với giá thành thấp, vì vậy hệ thống phát hiện ngã dựa trên camera có chi phí không cao và dễ dàng triển khai. Hướng tiếp cận này dựa trên phân tích dữ liệu video thu được qua một hoặc nhiều camera giám sát. Các camera này được lắp đặt xung quanh môi trường sinh hoạt thường ngày của người già. Có thể thấy, lợi thế lớn nhất là người sử dụng không cần phải mang theo thiết bị. Tuy vậy các phương pháp sử dụng camera giám sát bị giới hạn bởi môi trường trong nhà và không có cách nào hoạt động khi người già rời khỏi phòng, nơi có triển khai các camera. Các phương thức phát hiện ngã dựa trên camera có thể được coi là nhóm các phương thức sử dụng thị giác máy, phân biệt với các phương thức còn lại. Các phương thức sử dụng thị giác máy lại có thể chia thành ba nhóm nhỏ: nhóm sử dụng camera RGB đơn; nhóm dựa trên phân tích dữ liệu 3-D sử dụng hệ nhiều camera RGB; nhóm dựa trên phân tích dữ liệu 3-D sử dụng camera độ sâu (depth camera).



Hình 1.2. Minh họa hệ thống phát hiện ngã tự động dựa trên phân tích video

### **1.2.1. Phát hiện ngã sử dụng camera đơn**

Phát hiện ngã sử dụng camera RGB đơn được nghiên cứu rộng rãi do việc cài đặt hệ thống rất dễ dàng với chi phí thấp. Các đặc trưng phổ biến được khai thác là đặc trưng hình dạng cơ thể, đặc trưng chuyển động, và việc thiếu vắng chuyển động sau ngã.

Nhìn chung, vì những ưu điểm đã nêu, số lượng nghiên cứu phát hiện ngã dựa trên phân tích dữ liệu chuỗi video thu được từ một camera đơn là rất lớn, áp dụng nhiều kỹ thuật đa dạng. Các đặc trưng được sử dụng thường tập trung vào thông tin hình dạng cơ thể và thông tin chuyển động.

### **1.2.2. Phát hiện ngã sử dụng hệ multi camera**

Một nhóm các phương pháp phát hiện ngã dựa trên thị giác là sử dụng thông tin 3-D thu được từ một hệ các camera được kết hợp cùng với nhau. Nhiều nghiên cứu thực hiện việc cân chỉnh các camera như [16-19] giúp việc tái tạo lại mô hình 3-D của đối tượng một cách chính xác. Tuy nhiên tiến trình cân chỉnh camera thường phức tạp và tiêu tốn thời gian.

Việc sử dụng hệ multi camera mang đến lợi thế là cho phép dựng lại mô hình 3-D của đối tượng, trích rút được các thông tin 3 chiều, vốn là các thông tin hết sức phù hợp cho việc phát hiện ngã.

### **1.2.3. Phát hiện ngã sử dụng Camera độ sâu**

Camera độ sâu là loại camera có khả năng ghi nhận thông tin khoảng cách từ đối tượng đến bề mặt cảm biến, tương tự như mắt người.

Các hệ thống phát hiện ngã dựa trên camera độ sâu có cùng lợi thế về khai thác thông tin ba chiều như khi sử dụng hệ multi camera, nhưng khác với hệ multi camera, sử dụng camera độ sâu không cần cấu hình phức tạp, không tốn chi phí tính toán cho tiến trình cân chỉnh.

## CHƯƠNG 2. CƠ SỞ LÝ THUYẾT

Hướng tiếp cận sử dụng các phương pháp phân tích dữ liệu chuỗi video thu được qua camera để phát hiện ngã tự động nằm trong lớp bài toán phát hiện hành động (action detection) của lĩnh vực thị giác máy (computer vision), thuộc ngành khoa học máy tính (computer science). Lĩnh vực thị giác máy cố gắng mô phỏng lại những gì bộ não con người làm được với dữ liệu hình ảnh gửi về từ võng mạc, nghĩa là hiểu được ngữ cảnh dựa trên dữ liệu hình ảnh. Nó chủ yếu liên quan đến việc phân đoạn (segmentation), nhận diện (recognition), tái xây dựng mô hình 3D của đối tượng (reconstruction) và việc kết hợp các công việc đó cho mục đích hiểu ngữ cảnh.

Thị giác máy ứng dụng các kỹ thuật của xử lý ảnh số (digital image processing) với các mô hình học máy (machine learning) cũng như một số phương thức toán học để thực hiện mục tiêu nói trên. Có thể nói, Thị giác máy cùng với xử lý ảnh và trí tuệ nhân tạo, mà cụ thể là học máy, có rất nhiều phần giao thoa với nhau. Phần cơ sở lý thuyết sẽ trình bày một cách khái quát về xử lý ảnh số và thị giác máy, đồng thời giới thiệu một số kỹ thuật, giải thuật cơ bản của các lĩnh vực này mà có liên quan trực tiếp hoặc gián tiếp đến bài toán của luận văn này.

### 2.1. Tổng quan về xử lý ảnh số

Ngày nay, các lĩnh vực như y tế, thiên văn học, vật lý, hóa học, viễn thám, chế tạo, v.v.. và rất nhiều lĩnh vực khác nữa ngày càng lưu trữ, hiển thị, cung cấp ảnh số với số lượng vô cùng lớn. Thách thức đặt ra cho giới khoa học là làm sao trích rút ra được các thông tin có giá trị từ ảnh số nguyên gốc một cách nhanh chóng. Đó là mục đích chính của lĩnh vực xử lý ảnh số: chuyển đổi ảnh số thành thông tin.

#### 2.1.1. Ảnh kỹ thuật số

Ảnh kỹ thuật số là dữ liệu được các thiết bị ghi hình kỹ thuật số như máy ảnh số, camera số ghi lại từ phép chiếu hình ảnh ba chiều của vật thể từ thế giới thực lên mặt phẳng hai chiều. Ánh sáng từ nguồn sáng phản xạ trên bề mặt vật thể, đi qua thấu kính đến bề mặt cảm biến điện tử. Cảm biến này tiếp nhận ánh sáng

và chuyển đổi thành tín hiệu điện tử dạng tương tự. Sau đó bộ phận chuyển đổi tương tự - kỹ thuật số thực hiện việc lấy mẫu (sampling) để chuyển tín hiệu tương tự sang tín hiệu số và lưu xuống thiết bị lưu trữ. Ảnh số bao gồm một lưới các điểm ảnh (pixel), được lưu trữ dưới dạng mảng hai chiều. Trong đó, mỗi điểm ảnh là một thành phần ảnh nhỏ nhất biểu diễn giá trị cường độ sáng tại vị trí của nó. Giá trị của mỗi điểm ảnh là rời rạc. Mảng hai chiều lưu trữ dữ liệu ảnh gồm một số lượng hữu hạn số hàng và số cột.

### 2.1.2. Xử lý ảnh số

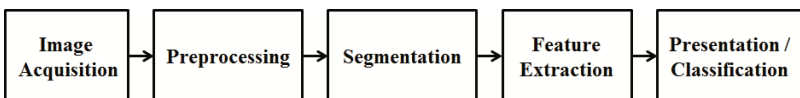
Xử lý ảnh số là quá trình áp dụng các phương thức, thuật toán để tác động vào và biến đổi ảnh ban đầu thành ảnh mới có chất lượng tốt hơn theo một tiêu chí xác định trước, hoặc trích rút các thông tin có ích từ dữ liệu ảnh.

Về mặt toán học, ảnh số có thể được coi là một hàm rời rạc hai biến  $f(x,y)$  với  $x, y$  là tọa độ của điểm ảnh. Giá trị hàm số  $f(x,y)$  chính là giá trị cường độ điểm ảnh tại vị trí  $x,y$ . Miền giá trị của  $f$  là:  $0 \leq f \leq f_{max}$ . Với  $f_{max}$  là giá trị lớn nhất của điểm ảnh. Với mã hóa 8bit,  $f_{max} = 255$ . Quá trình xử lý ảnh là quá trình thực hiện các phép biến đổi trên  $f(x,y)$ . Vì vậy có thể nói xử lý ảnh số là một dạng của xử lý tín hiệu số.

Xử lý ảnh thông thường bao gồm các bước:

- Quá trình thu nhận ảnh
- Phân tích và biến đổi ảnh gồm tiền xử lý, phân đoạn và trích rút đặc trưng ảnh.
- Biểu diễn kết quả như là ảnh kết quả, hoặc các báo cáo thu được từ việc phân tích ảnh

Một hệ thống xử lý ảnh gồm 5 thành phần: Thu nhận ảnh; tiền xử lý; phân đoạn ảnh; trích rút đặc trưng (mức thấp) ảnh; và mô tả, phân loại ảnh. Sơ đồ hệ thống được minh họa như Hình 2.1 dưới đây.



Hình 2.1. Hệ thống phân tích ảnh số

### 2.1.3. Các phép toán chính trong xử lý ảnh

## 2.2. Tổng quan về thị giác máy tính

### 2.2.1. Khái niệm và sơ lược lịch sử

Từ những năm 70, những nghiên cứu đầu tiên về hệ thống thị giác máy đã xuất hiện. Thời điểm đó, giới khoa học coi thị giác máy là một thành phần nhận thức trực quan của kế hoạch đầy tham vọng nhằm cho máy móc bắt trước con người, thực hiện những hành vi thông minh, tự động đưa ra các quyết định chính xác và hợp lý. Bài toán được đặt ra là, khi kết nối camera với máy tính, làm thế nào để máy tính có thể mô tả những gì nó nhìn thấy? Ngày nay, chúng ta biết rằng vấn đề đó là không hề đơn giản.

Thị giác máy có liên quan chặt chẽ và có nhiều phần giao thoa với các lĩnh vực như xử lý ảnh số, trí tuệ nhân tạo và toán học. Vậy thị giác máy là lĩnh vực liên ngành nhằm giúp cho máy tính có thể hiểu các thông tin mức cao (thông tin ngữ nghĩa) trong ảnh và video số. Từ góc độ khoa học ứng dụng, nó được coi là nghiên cứu để xây dựng lên các tác vụ tự động thực hiện các công việc như hệ thống thị giác con người có thể làm.

Thị giác máy hiện nay bao gồm nhiều nhánh như: tái tạo khung cảnh; phát hiện sự kiện; theo vết; nhận diện đối tượng; ước lượng chuyển động; khôi phục hình ảnh, v.v.. Đối tượng nghiên cứu của thị giác máy rất rộng. Một đặc điểm chung của các bài toán thuộc lĩnh vực này là có rất nhiều cách giải quyết theo nhiều cách tiếp cận khác nhau. Mỗi giải pháp đạt được một kết quả nhất định cho một trường hợp cụ thể. Rất khó để so sánh giải pháp nào là tốt hơn.

### 2.2.2. Ứng dụng

Ngày càng có nhiều ứng dụng thực tế áp dụng các kỹ thuật của thị giác máy, bao gồm: Nhận dạng ký tự (ORC); Kiểm tra tự động bằng máy (Machine inspection); Xây dựng mô hình 3D; Ứng dụng trong y tế; Giám sát; Nhận dạng vân tay và sinh trắc học; Phát hiện khuôn mặt; v.v..

### 2.2.3. Hệ thống các kỹ thuật thị giác máy

Hình ảnh thu được qua bước thu nhận, khử nhiễu nhờ các kỹ thuật thuộc xử lý ảnh sẽ là đầu vào cho các bước tiếp theo của hệ thống thị giác máy. Các kỹ thuật

thị giác máy có thể chia thành các mức: Thị giác mức thấp (low-level); thị giác mức trung (middle-level); và thị giác mức cao (high-level).

#### **2.2.4. Phân tích video (video analysis)**

Phân tích video có thể được định nghĩa là việc phân tích nội dung video để hiểu được ngữ cảnh mà nó mô tả. Đây là một thành phần đặc biệt của một số lĩnh vực công nghệ như camera giám sát (surveillance), robot (robotics), và đa phương tiện (multimedia). Các phương thức trong phân tích video được thúc đẩy bởi các nhu cầu phát triển các thuật toán máy tính có thể bắt trước được khả năng của hệ thống thị giác con người. Người nghiên cứu trong lĩnh vực này cần phải có rất nhiều kiến thức nền tảng khác nhau như xử lý ảnh/tín hiệu số; khoa học máy tính; lý thuyết hệ thống; xác suất thống kê; và toán ứng dụng.

##### ***1) Các tác vụ cơ bản trong phân tích video***

Các tác vụ trong phân tích video có thể phân loại theo: mức thấp; mức trung; mức cao, mặc dù ranh giới giữa các nhóm là tương đối mơ hồ. Ví dụ, một số nhà nghiên cứu phân chia tách cạnh, phân đoạn vào nhóm mức thấp; tái tạo 3D là mức trung; trong khi tác vụ nhận diện (recognition) được xếp vào nhóm mức cao.

##### ***2) Ứng dụng của phân tích video***

Các ứng dụng phổ biến của phân tích video bao gồm: Camera giám sát (surveillance); Truyền thông và internet; Thông tin di động; Thực tế ảo; Robot hoạt động dựa trên thị giác; Ứng dụng trong sinh học; v.v..

#### **2.2.5. Bài toán phát hiện hành động (action detection)**

Phát hiện hành động trong chuỗi video là công việc không chỉ nhận diện loại hành động mà còn phải xác định nó xảy ra ở đâu (vị trí không gian trên ảnh) và khi nào (vị trí thời gian trên chuỗi video).

Bài toán phát hiện hành động đặt ra rất nhiều thách thức. Thứ nhất, các hành động của con người hết sức đa dạng như đi lại thông thường, chạy, nhảy, làm việc nhà, múa, v.v.. mà trong số đó có rất nhiều hành động khó mà phân định rõ ràng với nhau. Các hành động, không những là tự thân mà còn có các hành động mang tính tương tác như ôm, bắt tay, hay các hành vi bạo lực tấn công người

khác. Ngoài việc tương tác giữa người và người, còn có các hành động có sự tương tác giữa người với đối tượng bất kỳ, như lau nhà, bê đồ, v.v.. Thứ hai, trong chuỗi video rất dễ xảy ra việc che khuất một phần cơ thể, gây khó khăn trong việc xác định tư thế, hình dạng. Mặt khác, trang phục cũng có sự khác biệt, đa dạng. Với trang phục là quần âu, hình dáng và sự chuyển động sẽ rất khác so với mặc váy rộng. Thứ ba, rất khó để phát triển một giải thuật phù hợp cho đa số các trường hợp phạm vi bài toán phát hiện hành động do sự khác nhau về môi trường như điều kiện chiếu sáng, sự đổ bóng, sự phản xạ ánh sáng và kết cấu của hậu cảnh. Thứ tư, ngoài những thách thức thuộc vấn đề thị giác máy, còn có các thách thức mang tính ngữ cảnh và tính triết học khi con người có thể thực hiện một hành động bằng cách không hành động gì cả.

Phát hiện hành động từ dữ liệu video – mà trọng tâm là hành động của con người, là công việc quan trọng trong rất nhiều các ứng dụng như video giám sát, phân tích nội dung video thông minh. Nhằm làm giảm bớt khó khăn của bài toán này, các hướng tiếp cận đa phần dựa trên giả định rằng vị trí không gian xảy ra hành động là biết trước, không có hoặc có rất ít sự thay đổi về tỉ lệ và góc nhìn cùng với hậu cảnh tĩnh và rõ ràng. Do đó, thông tin hình dạng cơ thể người có thể được trích rút một cách đáng tin cậy.

Chìa khoá cho bài toán phát hiện hành động là việc biểu diễn một hành động theo cách nào. Các đề xuất sử dụng đặc điểm tư thế và sự chuyển động tại các khớp của cơ thể cho việc biểu diễn hành động là rất hữu dụng, tuy nhiên, có rất nhiều khó khăn khi thực hiện.

Có hai dạng bài toán phát hiện hành động, bao gồm ngoại tuyến (offline) và trực tuyến (online), cho các miền ứng dụng khác nhau. Hai dạng bài toán này đặt ra các yêu cầu và các hướng tiếp cận khác nhau.



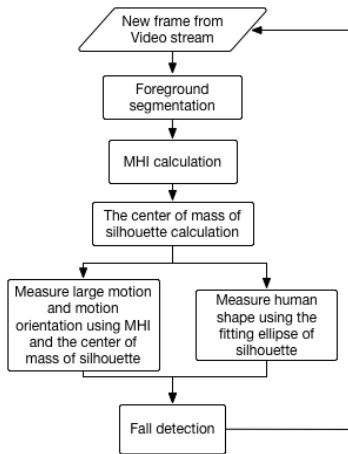
## CHƯƠNG 3.

### PHƯƠNG THỨC ĐỀ XUẤT

#### 3.1. Tổng quan

Luận văn đề xuất một phương thức phát hiện ngã dựa vào việc phân tích thông tin về chuyển động và sự thay đổi hình dáng cơ thể. Phương thức được đề xuất áp dụng kỹ thuật Motion History Image (MHI) cho trích rút các đặc trưng chuyển động. Phân tích hình dáng cơ thể bằng kỹ thuật Ellipse Fitting, đo lường tỉ lệ giữa chiều cao và bề ngang, góc nghiêng so với mặt phẳng nằm ngang của cơ thể. Luận văn cũng đề xuất sử dụng kỹ thuật Moment ảnh để xác định trọng tâm người, sử dụng tốc độ di chuyển của trọng tâm hỗ trợ cho việc trích rút đặc trưng chuyển động. Mô tả chi tiết các kỹ thuật sẽ được trình bày trong các phần tiếp theo.

Luồng hoạt động của phương thức đề xuất được mô tả như Hình 3.1 dưới đây:



Hình 3.1. Luồng hoạt động của hệ thống phát hiện ngã được đề xuất

Mục đích và chi tiết kỹ thuật của các khối sẽ được trình bày trong các phần tiếp theo của chương này.

## 3.2. Phân tách vùng chuyển động

Tách vùng chuyển động là công việc xác định tiền cảnh – vùng ảnh tương ứng với cơ thể, đạt được thông qua các kỹ thuật trừ nền (background subtraction). Trừ nền là phương thức so sánh khung hình hiện tại với mô hình nền, qua đó xác định được đối tượng chuyển động. Kỹ thuật này ngày càng được áp dụng rộng rãi cho việc phát hiện đối tượng chuyển động, nhất là trong ứng dụng giám sát bằng camera. Lý do chính để sử dụng kỹ thuật này là vì nó đơn giản, chính xác và chi phí tính toán thấp. Tuy nhiên các phương thức trừ nền đối mặt với một số thách thức trong trường hợp như môi trường thay đổi (độ chiếu sáng, nguồn sáng) hay sự di chuyển của bóng của vật thể, sự đồng nhất về màu sắc của đối tượng chuyển động.

### 3.2.1. Một số thuật toán trừ nền

### 3.2.2. Áp dụng kỹ thuật trừ nền, phân tách vùng chuyển động

Kết quả thử nghiệm nhiều kỹ thuật trừ nền khác nhau cho thấy GMM là giải thuật phù hợp với bài toán phát hiện ngã, đảm bảo được tốc độ xử lý thời gian thực. Vì vậy luận văn này lựa chọn áp dụng giải thuật GMM.

Vùng đánh dấu chuyển động tương ứng với cơ thể người bị phân chia thành nhiều đốm không kết nối với nhau do ảnh hưởng bởi nhiễu. Vì thế tác giả đề xuất sử dụng bộ lọc Gaussian làm trơn ảnh đầu vào, giúp giảm bớt ảnh hưởng của nhiễu ngẫu nhiên, sử dụng các phép biến đổi hình thái học để loại bỏ đốm nhỏ, liên kết các đốm lớn sau khi thực hiện trừ nền.



*Hình 3.9. Một ví dụ phân tách vùng chuyển động. Từ trái qua phải, ảnh đầu tiên là ảnh đầu vào từ chuỗi video; ảnh thứ 2 là kết quả của giải thuật trừ nền GMM; ảnh thứ 3 thể hiện kết quả loại bỏ các đốm nhỏ bằng thực hiện phép co; ảnh cuối cùng sau khi thực hiện phép giãn nở, các đốm lớn đã được kết nối và thể hiện tốt hình dạng cơ thể*

### 3.1. Trích rút đặc trưng chuyển động

Luận văn sử dụng các đặc trưng chuyển động bao gồm độ lớn và hướng của chuyển động cho phát hiện ngã. Các đặc trưng được trích rút dựa trên kỹ thuật Motion History Image (MHI), kết hợp với chuyển động của trọng tâm cơ thể, xác định qua kỹ thuật Image Moments. Chi tiết về các kỹ thuật được áp dụng như sau:

#### 3.1.1. Motion History Image

MHI được giới thiệu lần đầu tiên tại [4], là một kỹ thuật biểu diễn chuyển động trong chuỗi ảnh đơn giản nhưng hết sức mạnh mẽ. MHI biểu diễn vị trí và quỹ đạo của chuyển động trong chuỗi video. Giá trị của mỗi điểm ảnh trong MHI là một hàm số của lịch sử chuyển động tại vị trí điểm ảnh đó. Giá trị điểm ảnh càng lớn, chuyển động diễn ra càng gần thời điểm hiện tại.

#### 3.1.2. Image Moments

##### 3.1.1. Trích rút đặc trưng chuyển động

###### 1) Hướng của chuyển động

Với phương thức đo lường hướng tổng thể của chuyển động đã trình bày trong phần 11), mục 3.3.1 áp dụng cho tình huống ngã do vấp với đặc điểm chung là giai đoạn đầu chuyển động rất nhanh theo phương ngang, đến cuối hành trình mới chuyển động theo phương dọc thì hướng tổng thể vẫn theo phương ngang. Vì vậy luận văn này đề xuất một phương thức tính hướng chuyển động nhạy cảm hơn với trường hợp kể trên như sau:

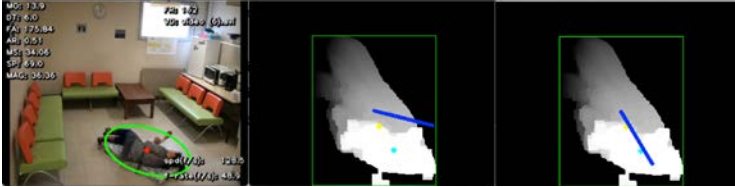
- Xác định trọng tâm của foreground ở khung hình hiện tại  $M_{recent}$
- Xác định trọng tâm của MHI  $M_{MHI}$
- Hướng chuyển động được xác định như sau:

$$\begin{aligned} dx &= |M_{recent}[0] - M_{MHI}[0]|; \\ dy &= |M_{recent}[1] - M_{MHI}[1]|; \\ angle &= \arctan \frac{dx}{dy} \end{aligned}$$

$$\text{Với độ lớn: } magn = \sqrt{dx^2 + dy^2}$$

Trong đó, trọng tâm được xác định dựa trên Image Moments.

Hình 3.11 mô tả thực nghiệm để khẳng định tính đúng đắn của quan sát đã được đề cập và cho thấy hiệu quả tốt hơn của phương thức đề xuất trong ngữ cảnh bài toán.



Hình 3.11. So sánh phương thức xác định hướng chuyển động. Từ trái qua phải: Ảnh 1 là khung hình từ chuỗi video; ảnh 2 thể hiện kết quả xác định hướng chuyển động bằng kỹ thuật truyền thống; ảnh 3 là kết quả của kỹ thuật được đề xuất

## 2) Độ lớn của chuyển động

Ảnh MHI thể hiện chuyển động ở thời điểm hiện tại bằng vùng điểm ảnh có giá trị cực đại (màu trắng trong ảnh), còn vùng chuyển động quá khứ là vùng điểm ảnh có giá trị nhỏ hơn giá trị cực đại và giảm dần về 0 (vùng màu xám trong ảnh). Vùng không xuất hiện chuyển động có giá trị điểm ảnh bằng 0. Chuyển động càng nhanh, vùng màu xám càng lớn. Vì vậy tác giả đề xuất phương thức ước lượng độ lớn của chuyển động từ MHI bởi:

$$M_{rate} = \frac{\sum N Z P_{MHI} - N Z P_{MotionMask}}{\sum N Z P_{MHI}} \times 100\%$$

Trong đó,  $\sum N Z P_{MHI}$  là tổng số điểm ảnh có giá trị lớn hơn 0 của MHI;  $N Z P_{MotionMask}$  là tổng số điểm ảnh có giá trị lớn hơn 0 của foreground phân tách được ở khung hình hiện tại.

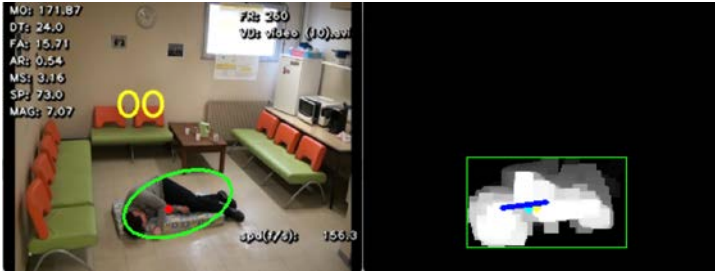
Giá trị của  $M_{rate}$  nằm trong khoảng từ 0% đến 100%, tương ứng với mức độ chuyển động từ nhỏ đến lớn.

Tuy nhiên ở giai đoạn cuối của hành động ngã, phần lớn cơ thể đã dừng chuyển động, vùng điểm ảnh có giá trị cực đại thu hẹp lại trong khi vùng màu xám còn rất rộng. Điều đó dẫn đến giá trị  $M_{rate}$  là rất lớn, không mô tả đúng thực tế là tốc độ chuyển động nhỏ trong trường hợp này. Thực tế này được minh họa như

ví dụ ở Hình 3.12. Vì vậy cần thiết phải sử dụng một ngưỡng cho độ dài vector chuyển động  $magn$  tính được ở phần 12) mục này trong xác định độ lớn chuyển động. Cụ thể như sau:

$$M_{rate} = \begin{cases} \frac{\sum NZP_{MHI} - NZP_{MotionMask}}{\sum NZP_{MHI}} \times 100\% & \text{nếu } magn > \tau \\ 0 & \text{nếu } magn \leq \tau \end{cases}$$

Trong đó,  $\tau$  là ngưỡng cho trước, được xác định qua thực nghiệm để chọn giá trị phù hợp nhất cho từng tập dữ liệu khác nhau.



Hình 3.12. Minh họa xác định  $M_{rate}$  lỗi trong thời điểm gần kết thúc chuyển động: Khi người ngã đã nằm trên sàn và vừa kết thúc chuyển động,  $M_{rate} = 73\%$  (rất lớn) nhưng  $magn$  chỉ là 7.07 (rất nhỏ)

Hạn chế lớn nhất của MHI khi áp dụng đo lường độ lớn chuyển động là, khi vùng chuyển động hiện tại chồng lấp phần chuyển động trong quá khứ, một phần thông tin về lịch sử chuyển động bị mất đi. Tương ứng với vùng màu xám bị nhỏ hơn thực tế phải có. Dẫn đến trong trường hợp xảy ra chồng lấp, độ lớn chuyển động đo được từ phương thức trên là nhỏ hơn thực tế. Để khắc phục vấn đề đó, tác giả đề xuất sử dụng thêm một đặc trưng về chuyển động nữa, đó là khoảng cách trọng tâm của cơ thể trong khung hình hiện tại so với vị trí trọng tâm ở trước đó 15 khung hình, gọi là mức độ chuyển động của trọng tâm. Nếu khoảng cách này lớn hơn ngưỡng chọn trước, rất có thể đã xuất hiện chuyển động nhanh bất thường. Trong trường hợp này, ngưỡng xác định chuyển động nhanh bất thường đối với  $M_{rate}$  và  $magn$  được điều chỉnh giảm xuống để không bỏ sót chuyển động nhanh trong trường hợp xảy ra chồng lấp.



Hình 3.13. Ví dụ cho ước lượng độ lớn chuyển động: Ảnh bên trái là khung hình trích ra từ chuỗi video; Ảnh bên phải là MHI. Khi ngã,  $M_{rate} = 61\%$ ,  $magn=28$ , độ chuyển động của trọng tâm = 37

Bằng các quan sát trong quá trình thực nghiệm, tác giả nhận thấy rằng trong điều kiện chiếu sáng phức tạp, foreground thu được từ quá trình phân tách vùng chuyển động là rất nhiều và không ổn định. Trong trường hợp camera được đặt quá thấp, ngang với chiều cao của người, khi đối tượng chuyển động đến gần sẽ khiến kích hoạt chức năng tự động đo sáng của camera, dẫn đến sự thay đổi đột ngột mức sáng trên toàn bộ khung hình. Điều đó dẫn đến sai lầm của giải thuật trừ nền. Để hạn chế bất lợi này, tác giả đề xuất sử dụng ngưỡng diện tích vùng chuyển động. Nếu diện tích vùng chuyển động quá lớn hoặc quá nhỏ, không phù hợp với vùng chuyển động tương ứng tạo ra bởi việc di chuyển của cơ thể người, chuyển động này sẽ được xem như là nhiễu và bị bỏ qua. Mặt khác, việc ứng dụng ngưỡng độ lớn diện tích trên còn giúp loại trừ các chuyển động nhỏ của một phần cơ thể trong các hoạt động thường nhật bình thường.

Cũng vì lý do nhiễu, vùng foreground có thể di chuyển đột ngột từ vị trí này tới vị trí khác. Trong trường hợp này, khoảng cách trọng tâm của foreground trong hai khung hình liên tiếp là lớn bất thường. Chuyển động này cũng được bỏ qua nếu khoảng cách kể trên lớn hơn một ngưỡng chọn trước.

Khi không xuất hiện chuyển động trong khung hình, phương thức được đề xuất yêu cầu phải xóa bỏ lịch sử chuyển động trước đó để không làm ảnh hưởng đến kết quả đo lường ở những khung hình tiếp theo.

## 3.2. Trích rút đặc trưng hình dạng cơ thể

Hình dạng cơ thể thay đổi rõ rệt trong quá trình ngã. Khi đứng thẳng hay đi lại, cơ thể người gần như vuông góc với phương nằm ngang, tỉ lệ chiều ngang so với chiều cao của người là nhỏ. Khi xảy ra ngã, người ngã nằm trên mặt sàn thường với tư thế cơ người lại và chân tay không ở sát người. Vì vậy tỉ lệ giữa chiều ngang và chiều cao của hình dạng tăng lên đáng kể. Sử dụng kỹ thuật fitting ellipse, xác định hình ellipse phù hợp nhất bao quanh hình dáng cơ thể rất phù hợp cho việc ước lượng, định lượng các đặc trưng về biến đổi hình dạng.

### 3.2.1. Kỹ thuật fitting ellipse

### 3.2.2. Áp dụng fitting ellipse đo lường đặc trưng hình dạng

Từ hình ellipse xác định được qua kỹ thuật đã mô tả, chuyển đổi sang dạng biểu diễn parametric ta có các tham số biểu diễn một ellipse gồm: hai trục ellipse  $a$ ,  $b$ ; tâm ellipse  $c(h,k)$ ; và góc quay  $t$ .

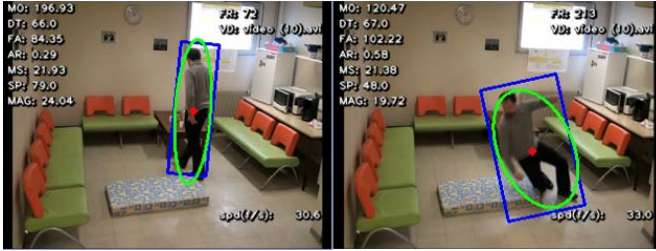
Tỉ lệ giữa bề ngang và chiều cao cơ thể AR tính bằng:

$$AR = \frac{b}{a}$$

Góc nghiêng cơ thể chính là góc nghiêng trục lớn của ellipse so với phương nằm ngang, hay chính là  $t$ .

Tuy nhiên, tác giả nhận thấy rằng vùng foreground thu được từ phân tách vùng chuyển động trong một số trường hợp bị ảnh hưởng rất nhiều bởi nhiễu. Một số trường hợp khác có thể không thu được foreground tại thời điểm người ngã vừa chấm dứt chuyển động, trong khi thông tin về hình dáng cơ thể tại ngay thời điểm này là rất quan trọng cho việc xác định ngã. Vì vậy luận văn này đề xuất sử dụng MHI cho xác định hình dáng cơ thể. Cụ thể như sau:

Từ MHI, sử dụng ngưỡng để tách lấy vùng điểm ảnh có giá trị lớn hơn ngưỡng. Ngưỡng được chọn qua thực nghiệm sao cho kết quả không quá sai khác so với hình dạng người chuyển động. Sau đó vùng điểm ảnh thu được sử dụng thay thế cho foreground trong quá trình xác định fitting ellipse. Kỹ thuật này giúp khai thác thông tin lịch sử chuyển động trong khoảng thời gian ngắn từ MHI, nhờ đó fitting ellipse thu được trở lên ổn định hơn rất nhiều.

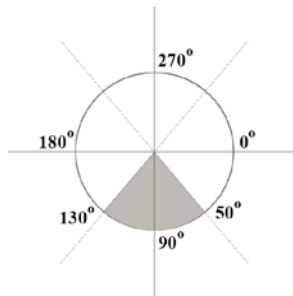


Hình 3.15. Minh họa sự thay đổi hình dạng cơ thể khi ngã: Ảnh bên trái,  $AR = 0.29$  khi đi lại thông thường; ảnh bên phải,  $AR = 0.58$  khi ngã

### 3.3. Phát hiện ngã

Việc phát hiện ngã phải trải qua phân tích trong suốt hành trình ngã, kể từ khi xuất hiện chuyển động nhanh bất thường cho đến khi không xuất hiện chuyển động sau khi ngã. Trình tự như sau:

Khi phát hiện chuyển động nhanh đột ngột với hướng chuyển động  $< 180^\circ$ , hệ thống tiếp tục theo dõi các khung hình tiếp theo trong khoảng 50 khung hình (tương đương với 2s với tốc độ 25fps). Nếu có chuyển động hướng xuống dưới sau đó (tương ứng với góc chuyển động nằm trong vùng màu xám minh họa ở Hình 3.16), tiếp theo là tỉ lệ  $AR$  vượt ngưỡng, góc nghiêng trục lớn vượt ngưỡng thì có thể xác định có thể là một tình huống ngã. Sau thời điểm đó, theo dõi tiếp trong 25 khung hình tiếp theo, nếu không phát hiện chuyển động hoặc chuyển động nhỏ thì kết luận xảy ra ngã. Nếu một trong các điều kiện kể trên không được thỏa mãn sau 50 khung hình thì không cảnh báo và tiếp tục theo dõi.



Hình 3.16. Quy ước về góc trong xác định hướng chuyển động và góc nghiêng cơ thể



## CHƯƠNG 4.

# THÍ NGHIỆM VÀ ĐÁNH GIÁ

### 4.1. Tập dữ liệu và phương pháp đánh giá hiệu quả thuật toán

#### 4.1.1. Tập dữ liệu thực nghiệm

Tập dữ liệu thực nghiệm cho các nghiên cứu phát hiện ngã do phòng thí nghiệm Le2i cung cấp [28] bao gồm 221 videos, được quay ở môi trường trong nhà, kích thước 320x240 với tốc độ khung hình 25fps, bao gồm 4 phần riêng biệt, tương ứng với 4 môi trường hoàn toàn khác nhau. Bảng 4.1 dưới sẽ đây mô tả về 4 tập đó.

*Bảng 4.1. Bảng mô tả các tập dữ liệu thực nghiệm*

Datasets	Videos	Falls
Lecture-room	27	14
Coffee-room	70	62
Home	60	33
Office	64	17

#### 4.1.2. Phương pháp đánh giá độ hiệu quả của giải thuật

Tính hiệu quả của giải thuật được đánh giá bởi chỉ số khôi phục (Recall), tỉ lệ chính xác (Precision), và hệ số điều hòa F-measure.

### 4.2. Cài đặt thí nghiệm

Để xác thực hiệu quả của các kỹ thuật đề xuất, tác giả thực hiện 4 thí nghiệm riêng biệt, trong đó 3 thí nghiệm đầu tiên lần lượt không áp dụng một kỹ thuật được giới thiệu. Ở thí nghiệm cuối cùng, tất cả các kỹ thuật được sử dụng. Cụ thể như sau:

- Thí nghiệm thứ nhất: Không sử dụng MHI trong đo lường hình dáng cơ thể mà đã được mô tả trong phần 3.4.2
- Thí nghiệm thứ 2: Sử dụng phương thức tính hướng gradient thông thường được đề xuất tại [33] thay cho phương thức giới thiệu tại phần 12) của mục 3.3.1

- Không áp dụng thông tin chuyển động trọng tâm để hạn chế lỗi trong ước lượng độ lớn chuyển động sử dụng MHI, mô tả tại phần 13) của mục 3.3.1
- Áp dụng tất cả các kỹ thuật được đề xuất.

### 4.3. Kết quả và thảo luận

Mỗi bộ dữ liệu cần đặt các ngưỡng giá trị cho các đặc trưng khác nhau do các đặc điểm môi trường, khoảng cách và vị trí đặt camera khác nhau. Các ngưỡng này được xác định thông qua thực nghiệm để đạt được kết quả tốt nhất.

Bảng 4.2 dưới đây trình bày các kết quả thí nghiệm đối với từng tập dữ liệu. Các dòng từ 1 đến 4 tương ứng với các thí nghiệm 1 đến thí nghiệm 4 đã được mô tả ở phần trước của chương này.

**Bảng 4.2. Kết quả thực nghiệm**

Datasets	Thí nghiệm	Recall(%)	Precision(%)	F-measure
Lecture-room	1	92.86	81.25	0.867
	2	92.86	92.86	0.929
	3	100	82.35	0.903
	<b>4</b>	<b>100</b>	<b>93.33</b>	<b>0.965</b>
Coffee-room	1	88.71	94.83	0.917
	2	82.26	100	0.903
	3	70.97	100	0.83
	<b>4</b>	<b>90.32</b>	<b>94.92</b>	<b>0.926</b>
Home	1	87.88	93.55	0.906
	2	87.88	100	0.935
	3	72.73	96	0.828
	<b>4</b>	<b>93.94</b>	<b>96.88</b>	<b>0.954</b>
Office	1	82.35	77.78	0.8
	2	64.71	73.33	0.688
	3	52.94	52.94	0.529
	<b>4</b>	<b>82.35</b>	<b>87.50</b>	<b>0.848</b>

Trong các thí nghiệm, thí nghiệm thứ 4 luôn cho giá trị Recall và F-measure cao hơn hẳn các thí nghiệm khác. Tốc độ xử lý của giải thuật đạt 90fps.

**CHƯƠNG 5.****KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN**

Qua quá trình thực hiện luận văn, tác giả đã tiến hành tìm hiểu lý thuyết tổng quan về lĩnh vực xử lý ảnh số và thị giác máy, có cơ hội tiếp cận với một số giải thuật học máy. Tác giả cũng đi sâu vào tìm hiểu các giải thuật thị giác máy quan trọng như giải thuật trừ nền, giải thuật trích chọn một số đặc trưng quan trọng như góc, điểm bất biến, đặc trưng SIFT, SUFT, v.v..., có hiểu biết cơ bản về các dạng bài toán thuộc ngành thị giác máy, qua đó làm nền tảng cho quá trình học tập nghiên cứu tiếp theo của bản thân trong lĩnh vực này.

Về mặt thực tiễn, luận văn này đã giới thiệu một phương thức tự động phát hiện ngã dựa trên việc kết hợp các đặc trưng chuyển động như hướng và độ lớn, với các đặc trưng về sự thay đổi hình dáng cơ thể. Luận văn cũng đã đề xuất một số cải tiến trong giải thuật MHI, đề xuất sử dụng đặc điểm chuyển động trọng tâm cơ thể để cải thiện kết quả đo lường chuyển động dựa trên MHI. Điểm mấu chốt của phương thức được đề xuất là thông qua phân tích đặc điểm thực tế của quá trình ngã dựa trên quan sát kỹ lưỡng, từ đó có thể khai thác hợp lý các đặc trưng đã trích rút được. Việc đưa ra dự đoán về việc ngã không dựa trên giá trị của các đặc trưng trong cùng một thời điểm, mà dựa trên quan sát giá trị các đặc trưng này trên toàn bộ khoảng thời gian tương ứng với hành động ngã, từ khi bắt đầu xuất hiện chuyển động nhanh bất thường đến khi không xuất hiện chuyển động sau ngã.

Các ngưỡng được xác định thủ công dựa trên suy luận từ các đặc điểm của việc ngã và quá trình quan sát các video thử nghiệm. Với việc lựa chọn tập dữ liệu thực nghiệm với nhiều môi trường khác nhau, điều kiện ánh sáng khác nhau, vị trí và góc độ camera được đặt khác nhau, kịch bản ngã phong phú và được xen giữa bởi các hoạt động thông thường hằng ngày, kết quả đạt được của luận văn là hết sức khả quan.

Các trường hợp nhận diện sai lầm chủ yếu là do nhiễu, do thay đổi ánh sáng đột ngột hoặc do người di chuyển quá gần ống kính camera, khiến kích hoạt chức năng tự động đo sáng của camera, ảnh hưởng đến giải thuật phân tách vùng chuyển động. Một số trường hợp nhận diện nhầm các hành động như nằm, ngồi

dứt khoát. Để giải quyết các vấn đề trên, tác giả dự kiến tìm hiểu các giải pháp trừ nền phù hợp hơn nữa, giúp loại trừ trường hợp camera điều chỉnh độ sáng, bổ xung các kỹ thuật phát hiện vùng đầu người (head detection) và kỹ thuật giới hạn vùng quan tâm (inactivity zone) trong các nghiên cứu tiếp theo. Ngoài ra, để mở rộng phạm vi của bài toán trong trường hợp bối cảnh có nhiều hơn một người, tác giả dự định tìm hiểu các kỹ thuật theo vết đối tượng (object tracking) cho việc cải tiến phương thức đã đề xuất.

Kết quả nghiên cứu và kỹ thuật được đề xuất đồng thời cũng được trình bày trong bài báo [51] gửi hội thảo quốc tế SoICT và đã được chấp nhận.

## **Danh mục công trình khoa học của tác giả liên quan đến luận văn**

1. Viet Anh Nguyen, Thanh Ha Le and Thuy Thi Nguyen. Single camera based Fall detection using Motion and Human shape Features. In *The Seventh International Symposium on Information and Communication Technology (SoICT 2016)*, accepted.

## TÀI LIỆU THAM KHẢO

### Tài liệu tiếng Việt

- [1] Bộ Y tế, "Ngày quốc tế người cao tuổi (IDOP) 2015: Già hóa dân số do nâng cao chất lượng cuộc sống," 1 tháng 10 2015. Bản điện tử: <http://moh.gov.vn:8086/news/Pages/ChuongTrinhMucTieuQuocGiaYTe.aspx?temID=4110>. [Truy cập 3 tháng 5, 2016]
- [2] Quỹ Dân số Liên hợp quốc, "Già hóa dân số và người cao tuổi ở Việt Nam: Thực trạng, dự báo và một số khuyến nghị chính sách," Hà Nội, 2011.

### Tài liệu tiếng Anh

- [3] A. Bourke, J. O'Brien, G. Lyons: Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait Posture* 2007, 26:194–199.
- [4] A. F. Bobick and J. W. Davis. The recognition of human movement using temporal templates. *IEEE transactions on pattern analysis and machine intelligence*, 23(3):257–267, March 2001.
- [5] A. H. Nasution and S. Emmanuel. Intelligent video surveillance for monitoring elderly in home environments. In *Multimedia Signal Processing, 2007. MMSP 2007. IEEE 9th Workshop* on pages 203–206. IEEE, 2007.
- [6] A. Leone, G. Diraco, and P. Siciliano. Detecting falls with 3d range camera in ambient assisted living applications: A preliminary study. *Medical Engineering & Physics*, 33:770–781, July 2011.
- [7] A. W. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least squares fitting of ellipses. *Technical Report DAIRP-794*, January 1996.
- [8] A. Zweng, S. Zambanini, and M. Kampel. Introducing a statistical behavior model into camera-based fall detection. In *International conference on Advances in visual computing*, pages 163–172, 2010.
- [9] B. Mirmahboub, S. Samavi, N. Karimi, and S. Shirani. Automatic monocular system for human fall detection based on variations in silhouette area. *IEEE Transactions on Biomedical Engineering*, 60:427–436, February 2013.
- [10] C. Doukas, I. Maglogiannis, F. Tragkas, D. Liapis, G. Yovanof: Patient Fall Detection using Support Vector Machines. *Int Fed Inf Process* 2007, 247:147–156.
- [11] C. F. Lai, S. Y. Chang, H. C. Chao, Y. M. Huang: Detection of cognitive injured body region using multiple triaxial accelerometers for elderly falling. *IEEE Sensors J* 2011, 11:763–770.
- [12] C. Rougier, E. Auvinet, J. Rousseau, M. Mignotte, and J. Meunier. Fall detection from depth map video sequences. In *International Conference on Smart Homes and Health Telematics*, pages 121–128, June 2011.

- [13] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau. Monocular 3d head tracking to detect falls of elderly people. In *Engineering in Medicine and Biology Society*, 2006. EMBS'06. 28th Annual International Conference of the IEEE, pages 6384–6387. IEEE, 2006.
- [14] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau. Robust video surveillance for fall detection based on human shape deformation. *IEEE Transactions on Circuits and Systems for Video Technology*, 21:611–622, May 2011.
- [15] C. Stauffer and W.E.L. Grimson. Adaptive background mixture models for real-time tracking. *International conference on Computer Vision and Pattern Recognition*, 2, 1999.
- [16] D. Anderson, R. H. Luke, J. M. Keller, M. Skubic, M. J. Rantz, and M. A. Aud. Modeling human activity from voxel person using fuzzy logic. *Fuzzy Systems, IEEE Transactions on*, 17(1):39–49, 2009
- [17] D. Anderson, R. Luke, J. Keller, M. Skubic, M. Rantz, and M. Aud. Linguistic summarization of video for fall detection using voxel person and fuzzy logic. *Computer Vision and Image Understanding*, 113(1):80–89, Jan 2009.
- [18] E. Auvinet, F. Multon, A. St-Arnaud, J. Rousseau, and J. Meunier. Fall detection using body volume reconstruction and vertical repartition analysis. In *International conference on Image and signal processing*, pages 376–383, 2010.
- [19] E. Auvinet, F. Multon, A. St-Arnaud, J. Rousseau, and J. Meunier. Fall detection with multiple cameras: An occlusion-resistant method based on 3-d silhouette vertical distribution. *IEEE Transactions on Information Technology in Biomedicine*, 15:290–300, March 2011.
- [20] E. Auvinet, L. Reveret, A. St-Arnaud, J. Rousseau, and J. Meunier. Fall detection using multiple cameras. In *Engineering in Medicine and Biology Society*, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, pages 2554–2557. IEEE, 2008.
- [21] F. Bagalà, C. Becker, A. Cappello, L. Chiari, K. Aminian and H. Jeffrey, "Evaluation of Accelerometer-Based Fall Detection Algorithms on Real-World Falls," vol. 7, no. 5, May 2012.
- [22] F. Bianchi, S. J. Redmond, M. R. Narayanan, S. Cerutti, N. H. Lovell: Barometric pressure and triaxial accelerometry-based falls event detection. *IEEE Trans Neural Syst Rehabil Eng* 2010, 18:619–627.
- [23] F. Sposaro, G. Tyson: iFall: an Android application for fall monitoring and response. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Minneapolis: Institute of Electrical and Electronics Engineers; 2009:6119–6122. doi:10.1109/IEMBS.2009.5334912.

- [24] G. Diraco, A. Leone, and P. Siciliano. An active vision system for fall detection and posture recognition in elderly healthcare. In *Design, Automation & Test in Europe Conference & Exhibition*, pages 1536–1541, March 2010.
- [25] H. Foroughi, A. Naseri, A. Saberi, and H. S. Yazdi. An eigenspace-based approach for human fall detection using integrated time motion image and neural network. In *Signal Processing*, 2008. ICSP 2008. 9th International Conference on, pages 1499–1503. IEEE, 2008.
- [26] H. Kerdegari, K. Samsudin, A. R. Ramli, S. Mokaram: Evaluation of fall detection classification approaches. In *Proceedings of the 4th International Conference on Intelligent and Advanced Systems*. Kuala Lumpur: Institute of Electrical and Electronics Engineers; 2012:131–136. doi:10.1109/ICIAS.2012.6306174.
- [27] I. C. Lopes, B. Vaidya, J. Rodrigues: Towards an autonomous fall detection and alerting system on a mobile and pervasive environment. *Telecommun Syst* 2011, 48:1–12.
- [28] I. Charfi, J. Miteran, J. Dubois, M. Atri, and R. Tourki. Definition and performance evaluation of a robust svm based fall detection solution. *SITIS'12*, page 218–224, 2012.
- [29] J. Chen, K. Kwong, D. Chang, J. Luk and R. Bajcsy, "Wearable Sensors for Reliable Fall Detection," in *Proceedings of the IEEE Engineering in Medicine and Biology 27th Annual Conference*.
- [30] J. Cheng, X. Chen, M. Shen: A framework for daily activity monitoring and fall detection based on surface electromyography and accelerometer signals. *IEEE J Biomed and Health Inform* 2013, 17(1):38–45.
- [31] J. Dai, X. Bai, Z. Yang, Z. Shen, D. Xuan: Mobile phone-based pervasive fall detection. *Pers Ubiquitous Comput* 2010, 14:633–643.
- [32] J. Tao, M. Turjo, M. Wong, M. Wang, and Y. Tan. Fall incidents detection for intelligent video surveillance. In *IEEE Conference on Information, Communications and Signal Processing*, pages 1590–1594, 2005.
- [33] J. W. Davis. Recognizing movement using motion histograms. Technical report, 1999.
- [34] K. Kim, T.H. Chalidabhongse, D. Harwood, and L. Davis. Real-time foreground-background segmentation using codebook model. *Real-Time Imaging*, 11:172–185, 2005
- [35] K. Tra and T. Pham. Human fall detection based on adaptive background mixture model and hmm. In *International Conference on Advanced Technologies for Communications*, pages 95–100, Oct 2013.
- [36] M. Kangas, A. Konttila, P. Lindgren, I. Winblad, T. Jms: Comparison of low-complexity fall detection algorithms for body attached accelerometers. *Gait Posture* 2008, 28:285–291.



- [37] M. Kangas, I. Vikman, J. Wiklander, P. Lindgren, L. Nyberg, T. Jämsä: Sensitivity and specificity of fall detection in people aged 40 years and over. *Gait Posture* 2009, 29:571–574.
- [38] M. Kepski and B. Kwolek. Fall detection using ceiling-mounted 3d depth camera. In *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pages 640–647, 2014.
- [39] M. V. Albert, K. Kording, M. Herrmann, A. Jayaraman: Fall classification by machine learning using mobile phones. *PLoS One* 2012, 7:e36556.
- [40] M. Yuwono, B. Moulton, S. Su, B. Celler, H. Nguyen: Unsupervised machine-learning method for improving the performance of ambulatory fall-detection systems. *Biomed Eng Online* 2012, 11:1–11.
- [41] N. Thome, S. Miguet, and S. Ambellouis. A real-time, multiview fall detection system: A lhm-based approach. *Circuits and Systems for Video Technology, IEEE Transactions on*, 18(11):1522–1532, 2008.
- [42] R. Cucchiara, A. Prati, and R. Vezzani. A multi-camera vision system for fall detection and alarm generation. *Expert Systems*, 24(5):334–345, Nov 2007.
- [43] R. M. Haralick and L. G. Shapiro. *Computer and Robot Vision*, volume 1. Addition Wesley, 1993.
- [44] R. Y. W. Lee, A. J. Carlisle: Detection of falls using accelerometers and mobile phone technology. *Age Ageing* 2011, 0:1–7.
- [45] S. Abbate, M. Avvenuti, F. Bonatesta, G. Cola, P. Corsini, A. Vecchio: A smartphone-based fall detection system. *Pervasive Mob Comput* 2012, 8:883–899.
- [46] S. H. Fang, Y. C. Liang, K. M. Chiu: Developing a mobile phone-based fall detection system on android platform. In *Proceedings of the Conference on Computing, Communications and Applications*. Hong Kong: Institute of Electrical and Electronics Engineers; 2012:143–146. doi:10.1109/ComComAp.2012.6154019.
- [47] S. Miaou, P. Sung, and C. Huang. A customized human fall detection system using omni-camera images and personal information. In *Transdisciplinary Conference on Distributed Diagnosis and Home Healthcare*, pages 39–42, Apr 2006.
- [48] S. Shan, T. Yuan: A wearable pre-impact fall detector using feature selection and support vector machine. In *Proceedings of the IEEE 10th International Conference on Signal Processing*. Beijing: Institute of Electrical and Electronics Engineers; 2010:1686–1689. doi:10.1109/ICOSP.2010.5656840.
- [49] T. Zhang, J. Wang, L. Xu, P. Liu: Fall detection by wearable sensor and one-class SVM algorithm. In *Lecture Notes in Control and Information Science*, Volume 345. Edited by Huang DS, Li K, Irwin GW. Berlin Heidelberg: Springer; 2006:858–863.

- [50] U. Lindemann, A. Hock, M. Stuber, W. Keck and C. Becker, "Evaluation of a fall detector based on accelerometers: A pilot study," *Medical and Biological Engineering and Computing*, vol. 43, no. 5, pp. 548-551, 2005.
- [51] Viet Anh Nguyen, Thanh Ha Le and Thuy Thi Nguyen. Single camera based Fall detection using Motion and Human shape Features. In *The Seventh International Symposium on Information and Communication Technology (SoICT 2016)*, accepted.
- [52] W. Feng, R. Liu, and M. Zhu. Fall detection for elderly person care in a vision-based home surveillance environment using a monocular camera. *Signal, Image and Video Processing*, 8(6):1129–1138, 2014.
- [53] World Health Organization, WHO Global Report on Falls Prevention in Older Age, 2007.
- [54] Y. T. Liao, C.-L. Huang, and S.-C. Hsu. Slip and fall event detection using bayesian belief network. *Pattern recognition*, 45(1):24–32, 2012.
- [55] Yannick Benezeth, Pierre-Marc Jodoin, Bruno Emile, H el ene Laurent, Christophe Rosenberger. Comparative study of background subtraction algorithms. *Journal of Electronic Imaging*, Society of Photo-optical Instrumentation Engineers, 2010, 19, <10.1117/1.3456695>. <inria-00545478>
- [56] Z. Zhang, E. Becker, R. Arora, and V. Athitsos. Experiments with computer vision methods for fall detection. In *International Conference on Pervasive Technologies Related to Assistive Environments*, pages 25:1–25:4, 2010.