

BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC ĐÀ NẴNG

TẠ THỊ ÁI NHI

PHÁT HIỆN VÀ THEO VẾT NGƯỜI
TỪ DỮ LIỆU VIDEO

Chuyên ngành : **KHOA HỌC MÁY TÍNH**

Mã số : **60.48.01**

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

Đà Nẵng - Năm 2012

Công trình được hoàn thành tại
ĐẠI HỌC ĐÀ NẴNG

Người hướng dẫn khoa học: **TS. Huỳnh Hữu Hưng**

Phản biện 1 : **TS. Huỳnh Công Pháp**

Phản biện 2 : **PGS.TS. Đoàn Văn Ban**

Luận văn đã được bảo vệ tại Hội đồng chấm Luận văn tốt nghiệp thạc sĩ kỹ thuật họp tại Đại học Đà Nẵng vào ngày 04 tháng 03 năm 2012

Có thể tìm hiểu luận văn tại:

- Trung tâm Thông tin - Học liệu, Đại học Đà Nẵng
- Trung tâm Học liệu, Đại học Đà Nẵng

MỞ ĐẦU

1. Lý do chọn đề tài

Với sự phát triển nhanh chóng của các loại máy móc hiện đại như máy ảnh số, máy quay phim kỹ thuật số, máy vi tính, điện thoại di động,...v.v thì lượng thông tin con người thu được dưới dạng hình ảnh là khá lớn. Để lượng thông tin này trở nên có ích hơn con người cần phải tiến hành xử lý nó và từ đó tạo điều kiện cho sự phát triển không ngừng của các kỹ thuật xử lý hình ảnh. Xử lý ảnh là một trong những công nghệ được ứng dụng rộng rãi hiện nay trong nhiều lĩnh vực của đời sống xã hội. Không chỉ dừng lại ở việc xử lý những vết nhòe, tái chế và phục hồi các ảnh cũ, ngày nay công nghệ xử lý ảnh đã mang lại những tiến bộ vượt bậc như nhận dạng vân tay, nhận dạng khuôn mặt, nhận dạng đối tượng khi nó kết hợp với lĩnh vực trí tuệ nhân tạo.

Bên cạnh đó, hệ thống camera quan sát ngày càng được sử dụng một cách rộng rãi với mục đích hỗ trợ an ninh cho các cửa hàng, công ty, ngân hàng,...v.v. Nhưng các camera này chỉ có khả năng để lưu giữ các dữ liệu video và chúng ta cần tốn thêm chi phí để thuê các nhân viên bảo vệ để quan sát và phát hiện khi có người nào đó xâm nhập.

Hiện nay trên thế giới cũng đã có một số ứng dụng cũng như nghiên cứu về việc giám sát hỗ trợ an ninh. Tuy nhiên những nghiên cứu này phát triển riêng lẻ, chưa tính hợp được với nhau để tạo thành một hệ thống giám sát hoàn chỉnh và giá thành cho các ứng dụng thì còn khá cao.

Từ yêu cầu thực tế được đặt ra, việc xây dựng một hệ thống phát hiện và theo vết (tracking) người từ dữ liệu video thu được thông qua hệ thống camera quan sát với sự hỗ trợ của kỹ thuật theo vết trong lĩnh vực thị giác máy tính (computer vision) là một điều cần thiết.

2. Mục tiêu và nhiệm vụ nghiên cứu

Mục đích của đề tài nhằm nghiên cứu các thuật toán phát hiện và theo vết một khuôn mặt người đang chuyển động từ dữ liệu video, từ đó xây dựng chương trình hỗ trợ với các chức năng sau:

- Phát hiện khuôn mặt người và thực hiện cảnh báo bằng âm thanh để thông báo khi có một hoặc nhiều người đi vào vị trí quan sát của camera đặc biệt tại những nơi bảo mật giới hạn người ra vào trong văn phòng công ty, ngân hàng,...v.v
- Theo vết khuôn mặt một người người từ dữ liệu video hoặc truyền từ camera thông qua thuật toán theo vết Meanshift.
- Thống kê số lượng người đi vào văn phòng, cửa hàng, ngân hàng,...v.v thông qua dữ liệu video hoặc từ camera quan sát.

Bên cạnh đó đề tài còn mong muốn giúp cho mọi người có một cái nhìn toàn diện hơn về vai trò và khả năng ứng dụng của công nghệ xử lý ảnh vào trong thực tế của đời sống xã hội.

3. Đối tượng và phạm vi nghiên cứu

❖ Đối tượng nghiên cứu

Trong bài toán này, dữ liệu được xử lý là đoạn video được quay từ một camera tĩnh và ghi lại với chuẩn AVI (Audio Video Interleave).

❖ Phạm vi nghiên cứu

Phạm vi nghiên cứu của đề tài liên quan đến lĩnh vực xử lý ảnh số thông qua việc sử dụng bộ thư viện mở OpenCV (Open Computer Vision).

4. Phương pháp nghiên cứu

❖ *Phương pháp nghiên cứu lý thuyết*

- Tìm hiểu cách lập trình với thư viện OpenCV.
- Tìm hiểu phương pháp AdaBoost kết hợp với các đặc trưng Haar-like để phát hiện khuôn mặt.
- Tìm hiểu phương pháp trừ nền (background subtraction) để phát hiện chuyển động,
- Tìm hiểu thuật toán theo vết Meanshift và Camshift.

❖ *Phương pháp nghiên cứu thực nghiệm*

- Tiến hành phân tích và cài đặt: thuật toán trừ nền để phát hiện chuyển động, thuật toán phát hiện và theo vết khuôn mặt người từ dữ liệu video.
- So sánh và đánh giá kết quả đạt được.

5. Kết quả dự kiến

- Nắm vững và cài đặt thành công các thuật toán: phát hiện chuyển động bằng phương pháp trừ nền, phát hiện khuôn mặt, theo dõi khuôn mặt trong video
- Xây dựng được chương trình để hỗ trợ 3 chức năng: cảnh báo người bằng âm thanh, theo vết khuôn mặt bằng thuật toán Meanshift và đếm số người ra vào tòa nhà từ dữ liệu video.
- Tạo được bảng so sánh kết quả, độ chính xác của các phương pháp phát hiện và theo vết đối tượng chuyển động dựa trên tập các video có sẵn.

6. Ý nghĩa khoa học và thực tiễn của luận văn

❖ Về mặt lý thuyết

- Ứng dụng thành công công nghệ xử lý ảnh vào trong thực tế.
- Tạo tiền đề cho những nghiên cứu tiếp theo trong tương lai.

❖ Về mặt thực tiễn

- Cung cấp các thuật toán hỗ trợ cho điều khiển rôbốt, các thiết bị phát hiện và nhận dạng khuôn mặt người.
- Giúp giảm giá thành cho công tác bảo vệ và bảo mật.
- Thống kê được số lượng khách hàng mỗi ngày để từ đó có những chiến lược thích hợp trong việc phát triển các chiến lược kinh doanh trong tương lai.
- Tạo tiền đề cho việc xây dựng cơ sở dữ liệu nhận dạng khách hàng một cách tự động thông qua khuôn mặt.

7. Bố cục luận văn

Nội dung của luận văn được chia thành các phần như sau:

❖ Mở đầu

❖ Chương 1: Nghiên cứu tổng quan

❖ Chương 2: Đánh giá các thuật toán và xây dựng các chức năng chính

❖ Chương 3: Cài đặt và kết quả

❖ Kết luận và hướng phát triển

❖ Phụ lục

CHƯƠNG 1: NGHIÊN CỨU TỔNG QUAN

1.1. TỔNG QUAN VỀ CAMERA SỐ

1.1.1. Khái niệm về camera số

Camera là một thiết bị ghi hình có thể ghi lại được những hình ảnh trong một khoảng thời gian nào đó và lưu trữ các dữ liệu hình ảnh đó.

1.1.2. Phân loại camera số

1.1.2.1. Phân loại theo kỹ thuật hình ảnh

1.1.2.2. Phân loại theo đường truyền

1.1.2.3. Phân loại theo tính năng sử dụng

1.1.3. Hệ thống camera quan sát

1.2. TỔNG QUAN VỀ VIDEO

1.2.1. Khái niệm về video

Thuật ngữ video dùng để chỉ nguồn thông tin hình ảnh trực quan (pictorial visual information), bao gồm một chuỗi các ảnh tĩnh (still image) liên tiếp nhau, được sắp xếp theo chiều thời gian

1.2.2. Video số

1.2.2.1. Tín hiệu video số

1.2.2.2. Ưu và nhược điểm của video số

1.2.2.3. Chuẩn video số AVI

Theo wikipedia, Audio Video Interleave (AVI) là một đa phương tiện định dạng container của Microsoft được giới thiệu vào tháng 11 năm 1992. AVI là tập tin có thể chứa cả âm thanh và video dữ liệu trong một container cho phép đồng bộ tập tin âm thanh với video.

1.3. CÁC PHƯƠNG PHÁP TRỪ NỀN PHÁT HIỆN ĐỐI TƯỢNG

1.3.1. Tổng quan về phát hiện đối tượng

1.3.2. Giới thiệu về phương pháp trừ nền

Ý tưởng chung của các phương pháp trừ nền là: Để phát hiện ra được các đối tượng chuyển động trong video chúng ta phải có được mô hình nền (background). Mô hình nền này có thể được học qua nhiều frame ảnh nếu nền bị thay đổi, ngược lại ta có thể chọn một nền cố sẵn nếu nền không bị thay đổi. Sau đó, ta sẽ dùng mô hình nền này để so sánh với frame ảnh hiện tại và kết quả là ta sẽ nhận biết được đâu là vùng phân nền, đâu là các phần chuyển động.

1.3.3. Frame Difference

1.3.3.1. Thuật toán cài đặt

1.3.3.2. Ưu điểm

1.3.3.3. Nhược điểm

1.3.4. Running Gaussian Average

1.3.4.1. Thuật toán cài đặt

1.3.4.2. Ưu điểm

1.3.4.3. Nhược điểm

1.3.5. Codebook

1.3.5.1. Thuật toán cài đặt

1.3.5.2. Ưu điểm

1.3.5.3. Nhược điểm

1.4. PHƯƠNG PHÁP PHÁT HIỆN KHUÔN MẶT NGƯỜI

1.4.1. Tổng quan về phát hiện người

1.4.2. Phát hiện khuôn mặt người

1.4.2.1. Giới thiệu

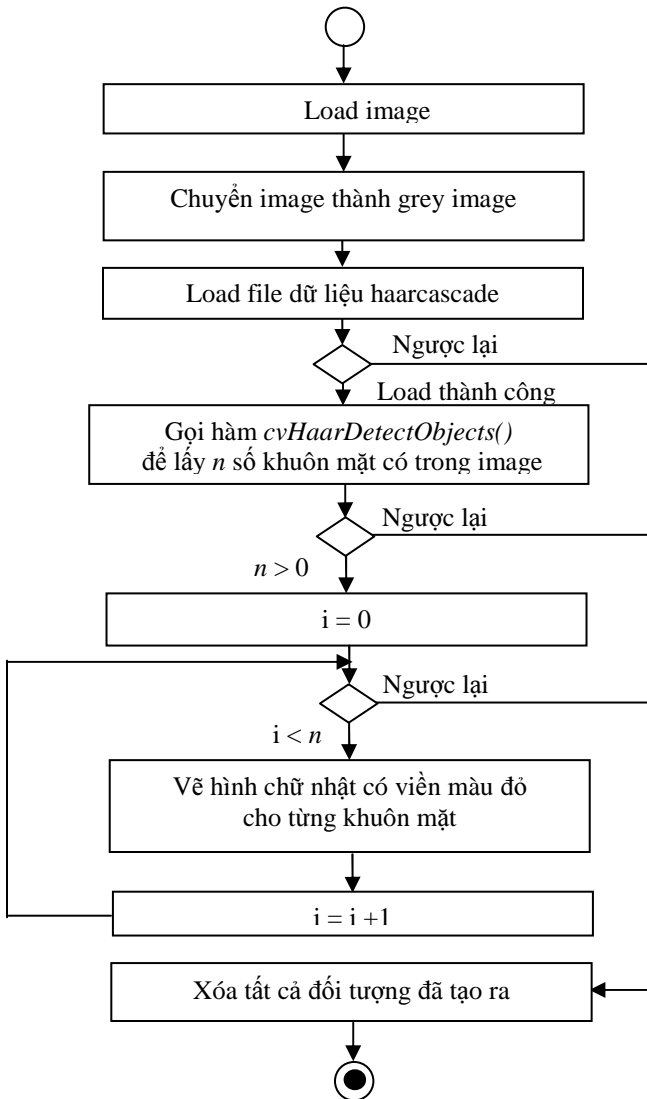
1.4.2.2. Các phương pháp phát hiện khuôn mặt người

1.4.2.3. Đặc trưng Haar-like

1.4.2.4. AdaBoost

1.4.2.5. Hệ thống xác định vị trí khuôn mặt người

1.4.2.6. Thuật toán xác định khuôn mặt người



Hình 1.14 Sơ đồ thuật toán xác định khuôn mặt người

1.5. CÁC PHƯƠNG PHÁP THEO VẾT ĐỐI TƯỢNG

1.5.1. Giới thiệu

Theo vết đối tượng là giám sát các thay đổi theo không gian và thời gian của đối tượng trong suốt chuỗi video như vị trí, kích thước hoặc hình dáng của đối tượng.

1.5.2. Phương pháp Meanshift

1.5.2.1. Mô tả

Meanshift là một thuật toán dịch chuyển đệ quy một điểm dữ liệu đến trung bình của các điểm dữ liệu tại vùng lân cận của nó, tương tự như việc gom các điểm dữ liệu lại tạo thành một nhóm.

1.5.2.2. Hàm `cvMeanshift` trong `OpenCV`

1.5.3. Phương pháp Camshift

1.5.3.1. Mô tả

Một thuật toán theo vết khác được phát triển từ Meanshift là Camshift (Continue Adaptive Mean-shift). Về cơ bản, thuật toán này giống với Meanshift tuy nhiên với Camshift, *window* tìm kiếm có thể tự điều chỉnh kích thước.

1.5.3.2. Hàm `cvCamshift` trong `OpenCV`

CHƯƠNG 2: ĐÁNH GIÁ CÁC THUẬT TOÁN VÀ XÂY DỰNG CÁC CHỨC NĂNG CHÍNH

2.1. ĐÁNH GIÁ CÁC THUẬT TOÁN

2.1.1. Các thuật toán trừ nền phát hiện đối tượng



Hình 2.1 Kết quả với Frame Difference trên Video1



Hình 2.2 Kết quả với Running Gaussian Average trên Video1

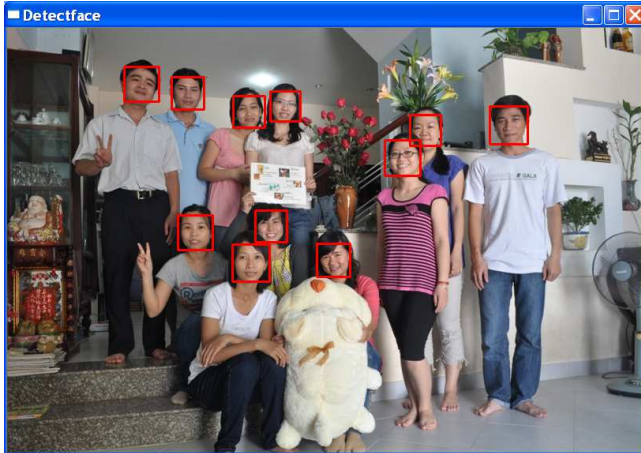


Hình 2.3 Kết quả với Codebook trên Video1

2.1.2. Thuật toán phát hiện khuôn mặt người

Bảng 2.2 Kết quả chạy thuật toán xác định khuôn mặt

Tên ảnh	Kích thước ảnh	Số người detect	Số người thực	Độ chính xác
Image 1.jpg	768x510	14	13	93%
Image 2.jpg	810 x 538	11	11	100%
Image 3.jpg	640x480	0	0	100%
Image 4.jpg	525x410	10	9	90%
Image 5.jpg	768x511	21	17	80%



Hình 2.14 Kết quả xác định khuôn mặt với Image2

2.1.3. Các phương pháp theo vết đối tượng

Đoạn Video1.avi có 1220 frame kích thước 320x240, có 2 người đang đi qua lại trong phòng và có khuôn mặt hướng về phía camera với ánh sáng vừa phải của đèn điện. Kết quả của theo vết được thể hiện như Hình 2.18 và Hình 2.19. Vì video này được quay đúng với điều kiện đặt ra của bài toán (camera cách đối tượng di chuyển một khoảng cách cố định, ánh sáng trong phòng vừa đủ) nên độ chính xác của việc theo vết của 2 phương pháp hoàn toàn giống nhau.



Hình 2.18 Kết quả theo vết với Meanshift trên Video1



Hình 2.19 Kết quả theo vết với Camshift trên Video1

2.2. XÂY DỰNG CÁC CHỨC NĂNG CHÍNH

Trong chương trình mà ta thiết kế, dữ liệu được xử lý sẽ là đoạn video được quay từ một camera tĩnh, ghi lại với chuẩn AVI với điều kiện ánh sáng tốt và nền không thay đổi. Nếu dữ liệu được dùng cho việc cảnh báo người hoặc theo vết khuôn mặt thì camera sẽ được đặt trước cửa ra vào sao cho có thể thu được các khuôn mặt từ phía trước (khoảng cách từ camera đến khuôn mặt khoảng 2-5m). Nếu dữ liệu được dùng cho việc đếm số người thì camera được đặt với một khoảng cách cố định (khoảng 2m) so với nền nhà tại cửa ra vào của một tòa nhà. Từ dữ liệu video, chương trình cung cấp một các chức năng sau:

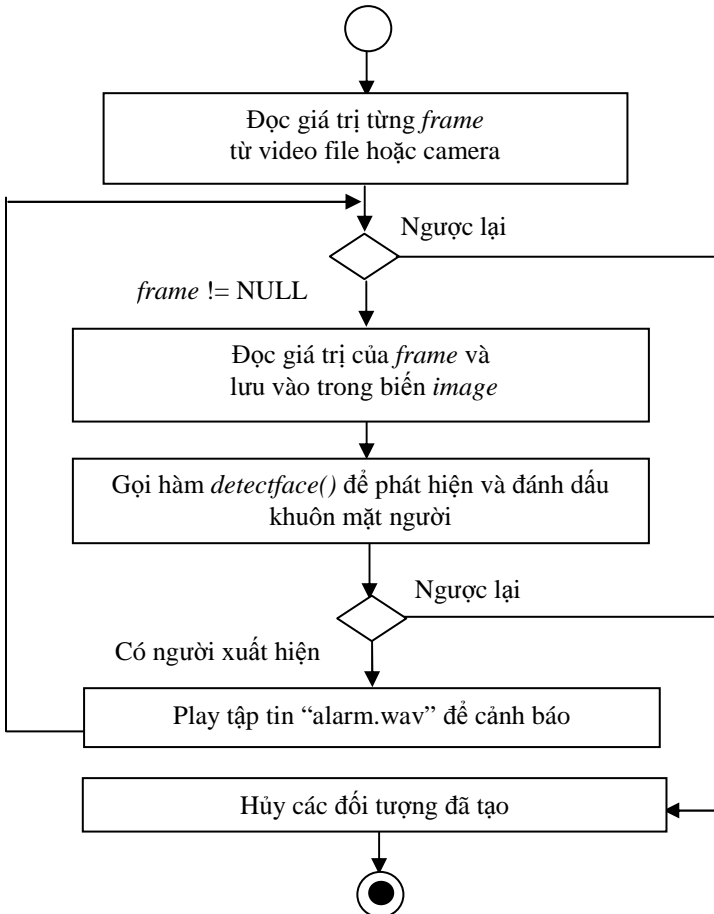
- Thực hiện cảnh báo bằng âm thanh để thông báo cho người dùng khi phát hiện khuôn mặt người trong video.
- Phát hiện và theo vết một khuôn mặt người trong video.
- Thống kê số lượng người đi vào và ra khỏi toà nhà trong video.

2.2.1. Chức năng cảnh báo bằng âm thanh

2.2.1.1. Mô tả chức năng

Hệ thống cung cấp chức năng phát hiện khuôn mặt người và cảnh báo bằng âm thanh giúp cho người dùng có thể phát hiện được có kẻ lạ đột nhập vào tòa nhà.

2.2.1.2. Thuật toán xử lý



Hình 2.25 Thuật toán xử lý cho chức năng cảnh báo

2.2.1.3. Kết quả

Đoạn video Video1.avi có 200 frame kích thước 320x240 và có 2 người đang nói chuyện, 1 người có khuôn mặt hướng về phía camera với ánh sáng trong phòng của đèn điện. Chương trình có thể phát hiện ra khuôn mặt của cô gái đang hướng về phía camera như Hình 2.26.



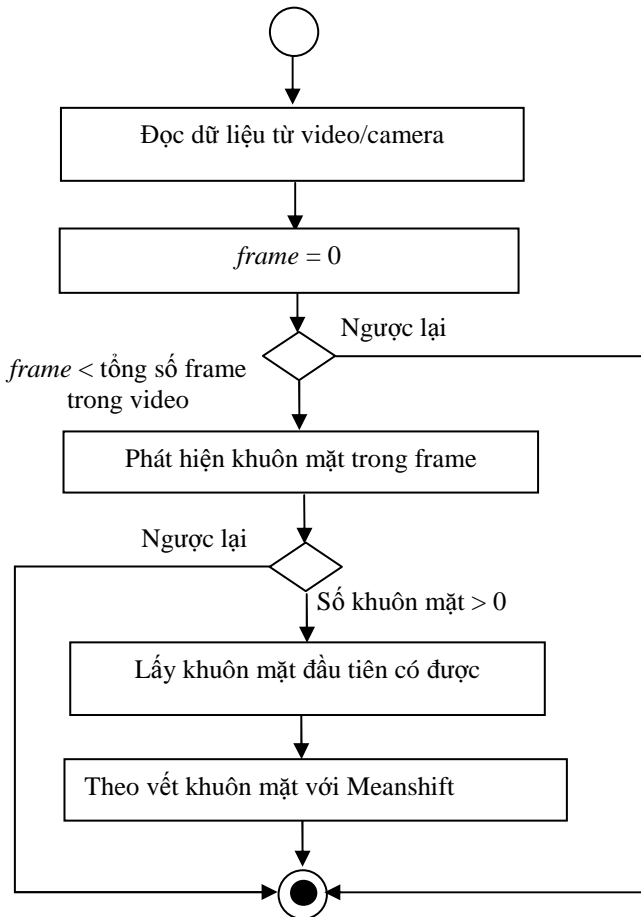
Hình 2.26 Kết quả cảnh báo với Video1

2.2.2. Chức năng theo vết khuôn mặt người

2.2.2.1. Mô tả chức năng

Một trong các chức năng tiếp theo của chương trình là chức năng theo vết khuôn mặt người trong video. Chức năng này giúp cho người sử dụng có thể theo dõi sự xuất hiện của khuôn mặt một người nào đó trong toàn bộ đoạn video được quay lại hoặc theo dõi trực tiếp thông qua camera.

2.2.2.2. Thuật toán xử lý



Hình 2.30. Thuật toán theo vết một khuôn mặt

2.2.2.3. Kết quả

Đoạn video Video1.avi có 1083 frame kích thước 640x360 và có 1 người đang nói khuôn mặt hướng về phía camera với ánh sáng của đèn điện. Chương trình có thể phát hiện ra khuôn mặt của người đó và thực hiện việc theo dõi khuôn mặt trong toàn bộ video như Hình 2.31.



Hình 2.31 Kết quả theo vết 1 khuôn mặt với Video1

2.2.3. Chức năng đếm số người vào tòa nhà

2.2.3.1. Mô tả chức năng

Việc thống kê được số lượng khách hàng từ dữ liệu video là một điều cần thiết cho chiến lược kinh doanh của mỗi công ty. Vì thế, chương trình cung cấp thêm chức năng này giúp cho người dùng có thể thống kê gần đúng số lượng khách tới văn phòng để giao dịch mỗi ngày. Điều này sẽ làm cơ sở để công ty có thể đề ra những chiến lược thích hợp trong việc phát triển cho tương lai.

2.2.3.2. Thuật toán xử lý

Thuật toán xử lý bao gồm việc lặp lại 3 bước sau với mỗi frame trong video.

Bước 1:

- Dùng phương pháp trừ nền để lấy những phần chuyển động bằng cách so sánh frame hiện tại với frame đầu tiên
- Chuyển ảnh thành ảnh xám và thực hiện khử nhiễu bằng cách loại bỏ bóng.

Bước 2:

- Tính toán các thành phần của ảnh để xác định hướng di chuyển của đối tượng là vào hay ra bằng việc sử dụng hàm `cvCalcGlobalOrientation()`. Nếu giá trị của góc dịch chuyển từ $0-180^0$ thì xác định đối tượng đang đi vào, ngược lại thì đối tượng đang đi ra.

- Với mỗi đối tượng di chuyển (connected component) sẽ được đánh dấu bởi một hình chữ nhật màu vàng và một dòng màu đỏ để xác định hướng di chuyển của đối tượng

Bước 3: Dựa vào kết quả hướng di chuyển của đối tượng trong Bước 2, ta có thể tính tổng được số lượng đi vào và ra khỏi tòa nhà.

2.2.3.3. Kết quả

Đoạn video Video1.avi có 2946 frame với kích thước 384x288 và có 34 người lần lượt đi qua cửa ra vào với ánh sáng bình thường. Kết quả của việc đếm số người được thể hiện như Hình 2.36.



Hình 2.36 Kết quả đếm người với Video1

CHƯƠNG 3: CÀI ĐẶT VÀ KẾT QUẢ

3.1. CÀI ĐẶT CHƯƠNG TRÌNH

3.1.1. Yêu cầu hệ thống

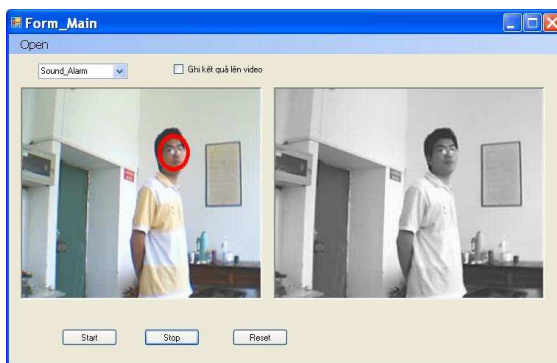
3.1.2. Cài đặt

3.2. KẾT QUẢ MINH HỌA

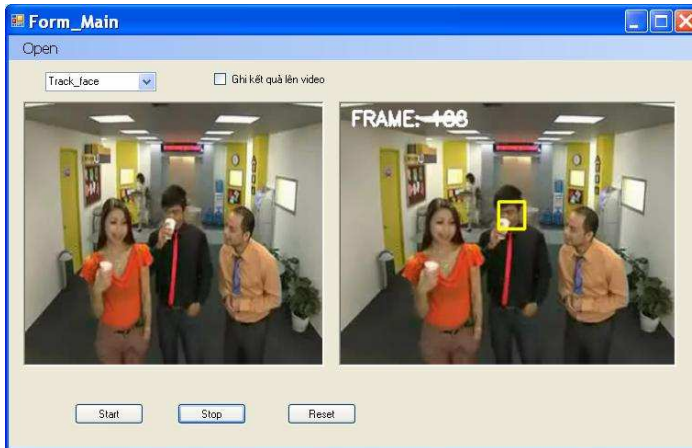
Sau khi chương trình khởi động, màn hình chính của chương trình hiện ra như Hình 3.1.



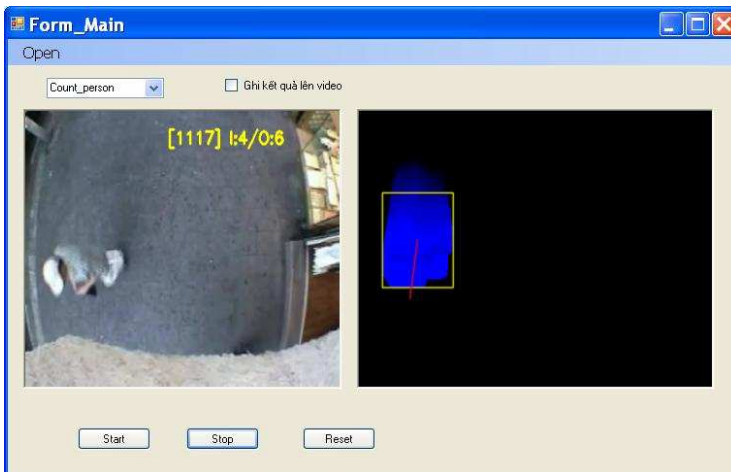
Hình 3.1. Giao diện chính



Hình 3.2. Chức năng cảnh báo bằng âm thanh



Hình 3.3. Chức năng theo vết khuôn mặt người



Hình 3.4. Chức năng đếm số người đi vào tòa nhà

3.3. ĐIỀU KIỆN RÀNG BUỘC CỦA CHƯƠNG TRÌNH

Để chương trình có thể hoạt động và cho kết quả chính xác thì cần phải thỏa mãn một số điều kiện sau:

- Nền của video không được thay đổi

- Camera được dùng để quay video phải đặt ở vị trí cố định, ở phía trên cửa ra vào và khoảng cách với người duy chuyển khoảng 1.5m đến 2m đối với chức năng đếm người ra vào tòa nhà
- Camera được dùng để quay video phải đặt ở vị trí cố định và ở phía trước khuôn mặt của người ra vào với khoảng cách khoảng 0.5m đến 1m đối với chức năng cảnh báo người bằng âm thanh
- Ánh sáng trong video phải ổn định, không quá tối hoặc quá sáng thì các chức năng sẽ làm việc chính xác hơn.
- Chương trình sẽ làm việc chính xác hơn với các video màu.

3.4. ĐÁNH GIÁ KẾT QUẢ CHƯƠNG TRÌNH

Sau khi thực nghiệm trên khoảng 40 video mẫu, nếu chương trình hoạt động như các điều kiện ràng buộc ở chương 3.3 thì nó sẽ mang lại độ chính xác khá cao.

- Độ chính xác của chức năng cảnh báo bằng âm thanh khoảng 90%
- Độ chính xác của chức năng theo vết khuôn mặt khoảng 80%
- Độ chính xác của chức năng đếm số người ra vào tòa nhà khoảng 88%

Từ kết quả này, ta có thể thấy rằng chương trình có khả năng để triển khai trong thực tế. Tuy nhiên, trong tương lai chương trình cần có một số cải tiến thêm để tăng độ chính xác của các chức năng hơn nữa.

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

1. Kết luận

❖ Về mặt lý thuyết

- Tìm hiểu và sử dụng thành thạo thư viện OpenCV trên nền Microsoft Visual Studio 2008.
- Tìm hiểu được các khái niệm cơ bản về video và camera số.
- Tìm hiểu được các thuật toán về phát hiện chuyển động bằng phương pháp trừ nền: Frame Difference, Running Gaussian Average và Codebook.
- Tìm hiểu được phương pháp phát hiện khuôn mặt người bằng kỹ thuật Haar kết hợp AdaBoost.
- Tìm hiểu được các phương pháp theo vết: Meanshift, Camshift.
- Tìm hiểu được cách thức đếm được số lượng người ra vào trong một tòa nhà.

❖ Về mặt thực tiễn

Về cơ bản, đề tài đã hoàn thành được các chức năng được đặt ra như yêu cầu ban đầu như:

- Phát hiện được phần chuyển động qua từng frame của video.
- Phát hiện và đánh dấu thành công tất cả các khuôn mặt người trong ảnh tĩnh bằng việc kết hợp giữa các đặc trưng Haar-like với bộ phân loại mạnh AdaBoost.
- Hoàn thành chức năng cảnh báo người bằng cách sử dụng âm thanh.
- Hoàn thành được chức năng theo vết khuôn mặt của một người trong video giúp ích cho việc theo dõi cũng như

lưu dữ liệu khuôn mặt nhằm tạo cơ sở dữ liệu phục vụ cho việc nhận dạng người tự động sau này.

- Hoàn thành được chức năng đếm số người ra vào tòa nhà từ dữ liệu video thu được thông qua camera, giúp ích cho việc thống kê số khách hàng vào giao dịch tại cửa hàng.
- Sử dụng thành công các công cụ phục vụ cho việc xử lý video như Motion Video, FLV Converter để chuyển đổi các định dạng video khác nhau về định dạng AVI để dễ dàng cho việc thao tác bằng thư viện OpenCV.

2. Phạm vi ứng dụng

Đề tài có thể được ứng dụng trong công tác bảo mật tại các cửa hàng công ty, giúp tiết kiệm được chi phí thuê bảo vệ và hiệu quả khá cao trong việc phát hiện chuyển động, theo dõi và nhận dạng người chuyển động từ dữ liệu video.

Bên cạnh đó đề tài còn tạo nền tảng cho những nghiên cứu về cảm quan máy tính nói riêng và xử lý ảnh nói chung trong tương lai.

3. Hạn chế

Bên cạnh những kết quả đã đạt được, đề tài vẫn còn có một số hạn chế cần phải được khắc phục như:

- Tốc độ xử lý của chương trình còn chậm khi xử lý những video có độ phân giải cao và kích thước lớn.
- Việc phát hiện khuôn mặt người có thể bị sai trong trường hợp một số đối tượng trong video có hình dáng gần giống với khuôn mặt người.

- Chức năng đếm số người ra vào chỉ đếm dựa trên chiều di chuyển của các đối tượng chuyển động ra vào, không phân biệt đó là người hay là vật.

4. Hướng phát triển

Trong quá trình thực hiện đề tài, do những hạn chế về trình độ và thời gian thực hiện đề tài có hạn, chương trình được xây dựng chỉ là phần demo các thuật toán phát hiện chuyển động, phát hiện khuôn mặt và theo vết đối tượng dựa vào video. Để triển khai trong thực tế nó đòi hỏi cần phải cải tiến hơn nữa. Hy vọng trong tương lai, những phát triển dưới đây sẽ giúp đề tài hoàn thiện hơn.

- Kết hợp việc phát hiện khuôn mặt với việc phát hiện mắt, phát hiện hình dáng của con người.
- Xây dựng được thuật toán cải thiện chất lượng của video như loại trừ nhiễu, loại trừ bong và tối ưu hóa các thuật toán để tăng tốc độ của chương trình.
- Hỗ trợ các tùy chọn giúp người dùng có thể điều chỉnh khoảng cách giữa camera và đối tượng tùy thuộc vào cấu trúc của tòa nhà mà họ sử dụng để đặt camera quan sát.
- Xử lý được việc che lấp đối tượng nhằm tạo kết quả chính xác cao hơn cho chức năng theo dõi khuôn mặt người.
- Mở rộng thêm thuật toán Camshift để có thể theo dõi được nhiều đối tượng trong video cùng lúc và thực hiện đếm được tổng số người xuất hiện trong video thông qua chức năng này.
- Xây dựng cơ sở dữ liệu giúp nhận dạng khách hàng một cách tự động thông qua khuôn mặt.