**Lời nói đầu**

**Mục Lục**

[Chương 1: Mở Đầu 5](#_Toc443572499)

[1.1. Lý do chọn đề tài 5](#_Toc443572500)

[1.2. Mục tiêu nghiên cứu 5](#_Toc443572501)

[1.3. Phương pháp nghiên cứu 6](#_Toc443572502)

[1.4. Đối tượng nghiên cứu 6](#_Toc443572503)

[Chương 2: Nội dung nghiên cứu 7](#_Toc443572504)

[2.1. Tổng quan về video và phát hiện đối tượng chuyển động 7](#_Toc443572505)

[2.1.1. Sơ lược về video 7](#_Toc443572506)

[2.2. Các phương pháp phát hiện chuyển động 10](#_Toc443572507)

[2.2.1. Tổng quan về các phương pháp phát hiện đối tượng chuyển động. 10](#_Toc443572508)

[2.2.2. Phương pháp sử dụng ngưỡng hình ảnh 12](#_Toc443572509)

[2.2.3. Phương pháp chênh lệch tạm thời 13](#_Toc443572510)

[2.2.4. Phương pháp trừ nền 14](#_Toc443572511)

[2.2.5. Phương pháp luồng quang học 16](#_Toc443572512)

[2.2.6. Phương pháp thống kê 19](#_Toc443572513)

[2.3. Một số đề xuất cài tiến 21](#_Toc443572514)

[2.4. Kết quả thực nghiệm 21](#_Toc443572515)

[Chương 3: Ứng dụng 21](#_Toc443572516)

[3.1. Môi trường cài đặt 21](#_Toc443572517)

[3.1.1. Môi trường thực nghiệm 21](#_Toc443572518)

[3.1.2. Thiết bị thực nghiệm 21](#_Toc443572519)

[3.1.3. Khoảng cách thực nghiệm 21](#_Toc443572520)

[3.2. Xây dựng ứng dụng 21](#_Toc443572521)

[Kết Luận 22](#_Toc443572522)

[Kết luận 22](#_Toc443572523)

[Kiến nghị 22](#_Toc443572524)

[Tài liệu tham khảo 23](#_Toc443572525)

# Mở Đầu

## Lý do chọn đề tài

Giám sát tự động là một hướng mới và có nhiều triển vọng trong sự phát triển tiếp theo của lĩnh vực nhận dạng và xử lý ảnh 2 chiều. Đồng thời, đó cũng là một hướng đi cho mảng phần mềm thiết kế chuyên dụng cho các thiết bị giám sát tự động. Việc phát hiện ra các đối tượng chuyển động trong video nhờ các kỹ thuật xử lý ảnh, trên cơ sở đó đoán nhận một số hành vi của đối tượng là một việc làm có ý nghĩa khoa học và thực tiễn. Nhất là trong hoàn cảnh Việt Nam chưa có nhiều những nghiên cứu và ứng dụng theo hướng này.

Xuất phát từ thực tế đó, việc nghiên cứu và đưa ra các phương pháp để xử lý video là vô cùng thiết thực. Được sự hướng dẫn của thầy Ngô Đức Vĩnh em đã tiến hành nghiên cứu đề tài “***Tìm hiểu một số phương pháp phát hiện đối tượng chuyển động trong video***” Các vấn đề cần giải quyết bao gồm:

- Nghiên cứu tổng quan về video và phát hiện đối tượng chuyển động.

- Nghiên cứu và đề xuất một số hướng khắc phục các nhược điểm trong việc phát hiện, đánh dấu, phân loại các đối tượng chuyển động và tạo kho cơ sở dữ liệu mẫu.

- Cài đặt một số phương pháp phát hiện, đánh dấu các đối tượng chuyển động dựa theo các phương pháp đã nêu ở trên.

## Mục tiêu nghiên cứu

* Tìm hiểu về phát hiện đối tượng
* Tìm hiểu về các kỹ thuật theo dõi đối tượng chuyển động .
* Phân tích đánh giá ưu nhược điểm của từng phương pháp.
* Tìm hiểu về các thuật toán liên quan đến lĩnh vực xử lý ảnh.
* Đánh giá hiệu quả của phương pháp cải tiến.

## Phương pháp nghiên cứu

* Nghiên cứu dựa trên các bài báo, tạp chí khoa học về giám sát tự động, phát hiện đối tượng chuyển động, ranh giới ảo từ đó xây dựng góc nhìn tổng quan về lĩnh vực đang nghiên cứu.
* Từng bước nghiên cứu, đánh giá và đề xuất các kỹ thuật cải thiện cho nội dung nghiên cứu.
* Tham khảo các mã nguồn mở về xử lý ảnh, phát hiện chuyển động và theo dõi chuyển động trên các website như codeproject.com, codeplex.com, sourceforge.net, github.com, emgu.com, opencv.com, msdn.com… Nhằm mục đích tìm hiểu quá trình cài đặt các thuật toán xử lý ảnh cơ bản và nâng cao trên ngôn ngữ lập trình C#.
* Tìm hiểu kiến thức về xử lý ảnh thông qua các giáo trình Xử lý ảnh, các bài viết về xử lý ảnh trên các diễn đàn – song song với việc tham khảo các tài liệu do giáo viên hướng dẫn cung cấp.
* Tìm hiểu, cài đặt các thư viện mã nguồn mở về xử lý ảnh hỗ trợ cho ngôn ngữ C#, qua đó giảm quá thời gian trình xây dựng hệ thống.
* Thực nghiệm dựa trên xây dựng ứng dụng và đánh giá các kết quả nghiên cứu.

## Đối tượng nghiên cứu

* Kỹ thuật phát hiện đối tượng chuyển động trong video.

# Nội dung nghiên cứu

***Tóm tắt nội dung chính****: Trong chương này chúng ta sẽ nghiên cứu một số vấn đề cơ bản liên quan đến xử lý ảnh và video. Các vấn đề được nêu ra trong chương là những khái niệm, thuật toán, cách thức lưu trữ, xử lý ảnh, video. Thông qua đó, tạotiền đề cho việc phát triển, xử lý các vấn đề khó khăn liên quan đến đề tài đang xây dựng.*

## Tổng quan về video và phát hiện đối tượng chuyển động

### Sơ lược về video

Video là sự tái tạo ảnh tự nhiên theo không gian và thời gian hoặc cả hai, thực chất là một dãy ảnh liên tục theo thời gian nhằm mô phỏng sự chuyển động.

Video được thu nhận chủ yếu qua các thiết bị như camera. Camera sẽ thu thập âm thanh và hình ảnh để được video. Nghĩa là video= âm thanh + hình ảnh. Sau đó, video sẽ được nén hoặc định dạng tùy theo mục đích sử dụng. Việc lưu trữ video phụ thuộc khá nhiều vào định dạng của video. Với cùng 1 video nhưng định dạng khác nhau sẽ có kích thước khác nhau.

#### Video tương tự

**NTSC Video**: Đây là dạng Video tương tự với 525 dòng trên một khung hình, 30 khung hình trong một giây, quét cách dòng, chia làm hai trường (mỗi trường 262.5 dòng), có 20 dòng dự trữ cho thông tin điều khiển tại thời điểm bắt đầu mỗi trường.

**PAL Video**: Dạng Video này có 625 dòng trên một khung hình, 25 khung hình trong một giây, quét cách dòng. Khung gồm hai trườngchẵn lẽ, mỗi trường bao gồm 312.5 dòng.

#### Video số

Một số tiêu chuẩn của Video số lấy theo tiêu chuẩn CCIR (Consultative Committee for International Radio)

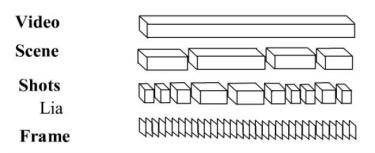
*Các tiêu chuẩn của Video số*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *CCIR*  *601525/6*  *NTSC* | *CCIR*  *601625/50*  *PAL/SECAM* | *CFI* | *QCFI* |
| *Độ phân giải*  *độ chói* |  |  |  |  |
| *Độ phân giải*  *màu sắc* |  |  |  |  |
| *Lấy mẫu màu* |  |  |  |  |
| *Số trường /s* |  |  |  |  |
| *Cách quét* |  |  |  |  |

#### Các dạng video

#### Cấu trúc của video

Video là tập hợp các khung hình, mỗi khung hình là một ảnh. Shot (lia) là đơn vị cơ sở của video. 1 lia là một đơn vị vật lý của dòng video, gồm các chuỗi các khung hình liên tiếp, không thể chia nhỏ hơn, ứng với một thao tác camera đơn. Sciene (cảnh) là đơn vị logic của dòng video, một cảnh gồm các lia liên quan về không gian và thời gian, cùng miêu tả một nội dung ngữ nghĩa hoặc một tình tiết.



*Cấu trúc phân đoạn video*

Khi video được chiếu, các khung hình lần lượt được hiển thị ở tốc độ nhất định. Tốc độ thường thấy là 24 frame / giây. Với càng nhiều khung hình trên 1 giây thì video sẽ được thể hiện càng mịn và chuyển động linh hoạt.

**Frame (khung):** Video được tạo nên bởi chuỗi các ảnh tĩnh. Một chuỗi các khung hình tĩnh đặt cạnh nhau tạo nên các cảnh phim chuyển động. Một khung đơn là một ảnh tĩnh.

Để đoạn video có thể tạo cảm giác chuyển động, các khung hình phải được quay với tốc độ phù hợp. Mắt người chỉ có thể nhận được 24 hình/giây, nếu như lần lượt 24 hình hoặc nhiều hơn 24 hình được phát trong một giây thì mắt sẽ không nhận ra được sự rời rạc giữa những khung hình, mà chỉ thấy những cảnh liên tục. Có nhiều hệ video và mỗi hệ có tốc độ quay khác nhau như : NTSC 30 hình/giây, PAL 24 hình/giây, SECAM 29.99 hình/giây.

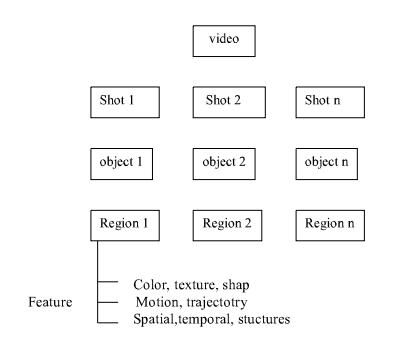
Khung hình là đơn vị cơ bản nhất của dữ liệu video. Theo chuẩn của hệ NTSC thì một giây có 30 khung hình, vậy một phút có 1800 khung hình, một giờ có 60x1800 = 108000 khung hình. Có thể thấy rằng số lượng khung hình cho một đoạn video thường là rất lớn, cần phải có một đơn vị cấp cao hơn cho video số.

Shot (lia): là đơn vị cơ sở của video. Một lia là một đơn vị vật lý của dòng video, gồm các chuỗi một hay nhiều khung hình liên tiếp, không thể chia nhỏ hơn, ứng với một thao tác camera đơn.

Những khung hình này liên tiếp theo thời gian mô tả một hành động liên tục, và được giới hạn bởi 2 chuyển cảnh.

Một đoạn video có thể có nhiều đoạn cơ sở, mà cũng có thể chỉ là một đoạn cơ sở. Những đoạn cơ sở đại diện cho toàn bộ đoạn video, và truy xuất đến chúng cũng coi như là truy xuất đến chính đoạn video. Tổ chức những đoạn video ở mức độ đoạn cơ sở là thích hợp nhất cho việc duyệt và truy tìm thông tin dựa vào nội dung.

**Scene (cảnh )** : là các đơn vị logic của dòng video, một cảnh gồm một hay nhiều shot liên quan đến nhau về không gian và liền kề về thời gian, cùng mô tả một nội dung ngữ nghĩa hoặc một tình tiết.

*Biểu diễn video dựa trên đối tượng*

## Các phương pháp phát hiện chuyển động

### Tổng quan về các phương pháp phát hiện đối tượng chuyển động.

Vấn đề phát hiện đối tượng đang được nghiên cứu và có nhiều ứng dụng trong cuộc sống. Các đối tượng được phát hiện nhờ những thông tin trong một frame ảnh. Có rất nhiều hướng tiếp cận để giải quyết vấn đề trên. Các tác giả Alper Yilmaz, Omar Javed và Mubarak Shah đã phân loại các hướng tiếp cận này được trình bày.

*Bảng phân loại các thuật toán phát hiện đối tượng*

|  |  |
| --- | --- |
| Loại | Những nghiên cứu có liên quan |
| Point detectors | 1. Moravec’s detector  2. Harris detector  3. Scale Invariant Feature Transform Affine  4. Invariant Point Detector |
| Segmentation | 1. Mean-shift  2. Graph-cut  3. Active contours |
| Background Modeling | 1. Mixture of Gaussians  2. Eigenbackground  3. Wall flower  4. Dynamic texture background |
| Supervised Classifier | 1. Support Vector Machines  2. Neural Networks  3. Adaptive Boosting |

Việc lựa chọn phương pháp áp dụng phải dựa vào tình huống cụ thể, đối với trường hợp có ảnh nền không thay đổi việc phát hiện đối tượng chuyển động có thể bằng các phương pháp trừ nền. Các giải thuật này sẽ được trình bày sau đây. Hướng giải quyết là xây dựng mô hình nền, sau đó sử dụng mô hình này cùng với frame hiện tại để rút ra được các foreground chuyển động. Để có thể tiếp cận cần phải xây dựng được mô hình background. Có nhiều phương pháp xây dựng mô hình background bởi các tác giả: Anurag Mittal dùng adaptive kernel density estimation được tính bằng[5]. Kết quả tốt tuy nhiên khó khăn về không gian lưu trữ, tính toán phức tạp, tốc độ không đáp ứng thời gian thực. Haritaoglu dùng giải thuật W4, Stauffer sử dụng Mixture of Gaussian để xây dựng mô hình nền… Nhằm phát hiện được các đối tượng chuyển động, xác định xem những đối tượng này có đúng là những đối tượng ta cần phát hiện hay không. Đây là các khó khăn cần khắc phục.

Trong các lĩnh vực về phát hiện phần đầu của người thì Wei Qu, Nidhal Bouaynaya and Dan Schonfeld đề ra hướng tiếp cận bằng cách kết hợp mô hình màu da cùng với mô hình màu tóc (skin and hair color model). Những màu này được phát hiện dựa vào mô hình Gauss. Sau đó bằng cách áp dụng phương pháp so khớp mẫu (template matching) để đạt được mục đích phát hiện phần đầu người đáp ứng thời gian thực. Khó khăn trong hướng tiếp cận này thường gặp ở việc thu thập dữ liệu huấn luyện màu da và màu tóc, độ chính xác dể bị ảnh hưởng bởi độ sáng của môi trường.

Việc phát hiện đối tượng có thể được thực hiện bằng các phương pháp máy học. Các phương pháp này có thể kể đến như: mạng neural, adaptive boosting, cây quyết định, support vector machines. Điểm chung của các phương pháp này đều phải trải qua giai đoạn huấn luyện trên một tập dữ liệu. Tập dữ liệu này phải đủ lớn, bao quát hết được các trạng thái của đối tượng. Sau đó các đặc trưng sẽ được rút trích ra trên bộ dữ liệu huấn luyện này. Việc lựa chọn đặc trưng sử dụng đóng vai trò quan trọng ảnh hưởng đến hiệu quả của các phương pháp máy học. Một số đặc trưng thường được sử dụng như: đặc trưng về màu sắc, đặc trưng về góc cạnh, đặc trưng histogram… Sau khi đã có được đặc trưng, ta sẽ đánh nhãn lớp cụ thể cho các đặc trưng đó để sử dụng trong việc huấn luyện. Trong quá trình huấn luyện, các phương pháp máy học sẽ sinh ra một hàm để ánh xạ những đặc trưng đầu vào tương ứng với nhãn lớp cụ thể. Sau khi đã huấn luyện xong thì các phương pháp máy học trên sẽ được dùng để phân lớp cho những đặc trưng mới. Đặc điểm của phương pháp này là độ chính xác cao. Tuy nhiên nó gặp phải khó khăn trong việc thu thập dữ liệu huấn luyện ban đầu, tốn thời gian và chi phí cho quá trình học máy.

### Phương pháp sử dụng ngưỡng hình ảnh

Phương pháp sử dụng ngưỡng hình ảnh là phương pháp phát hiện đối tượng đơn giản. Trong kỹ thuật này, người ta cho rằng các đối tượng chuyển động và nền có màu sắc khác nhau. Ritter và Wilson đưa ra phương pháp ngưỡng hình ảnh để phân loại các điểm ảnh là đối tượng hoặc nền. Mỗi điểm ảnh trong khung hình được phân loại là đối tượng hoặc nền. Nếu giá trị màu của một điểm ảnh nằm trong giá trị ngưỡng màu đưa ra, thì sẽ được gán giá trị nhị phân 1 trái lại nó được coi như một điểm ảnh nền và được gán giá trị 0.

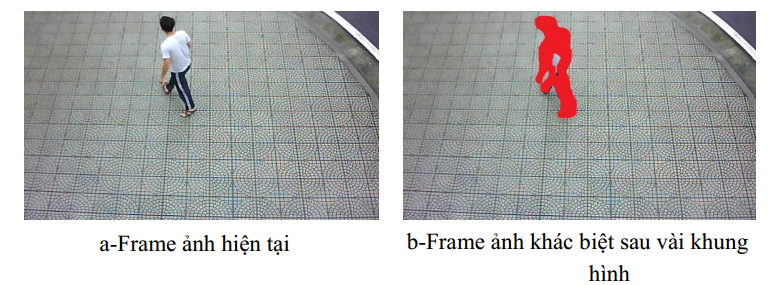
Thay vì sử dụng một giá trị ngưỡng toàn cục, thì phương pháp tích lũy, sử dụng ngưỡng cục bộ hoặc ngưỡng động thường được sử dụng.

Trong hầu hết các tình huống thực tế, các đối tượng và hình nền được chia sẻ nhiều màu sắc phổ biến, điều này gây ra khó khăn để chọn giá trị ngưỡng tối ưu để tách đối tượng từ nền. Su và Amer tập trung vào hai loại ngưỡng (ước lượng sự phân bố các vùng ảnh thay đổi do vị trí không gian màu) và đề xuất một thuật toán phi tham số để tính toán ngưỡng toàn cục. Phương pháp này cải thiện được việc phát hiện đối tượng. Petrou và Bosdogianni sử dụng biểu đồ màu nhằm tìm ra các giá trị ngưỡng để tách các đối tượng từ nền.

Ngưỡng hình ảnh là một phương pháp hữu ích cho phát hiện đối tượng chuyển động, khi các đối tượng và màu nền là rất khác nhau.

### Phương pháp chênh lệch tạm thời

Phương pháp chênh lệch tạm thời hoặc khác biệt khung hình thực hiện việc phát hiện vùng chuyển động bằng cách tính toán sự khác nhau ở mức điểm ảnh giữa hai hoặc ba khung hình liên tiếp trong một chuỗi các khung hình video, chỉ những điểm ảnh có biến động đáng kể (lớn hơn một ngưỡng quy định) được đánh dấu là điểm ảnh của đối tượng chuyển động và điểm ảnh còn lại được phân loại là ảnh nền. Phương pháp này khá tốt đối với khung cảnh động, có nhiều thay đổi song nó lại có nhược điểm trong việc phát hiện ra các điểm ảnh có liên quan đến một số kiểu chuyển động đặc biệt của đối tượng trong cảnh quay.



Phương pháp chênh lệch tạm thời được thực hiện qua công thức

|I\_t (x,y)-I\_(t-1) (x,y)| > τ

Trong đó I\_t (x,y) là điểm ảnh có tọa độ (x,y) của khung hình tại thời điểm t, I\_(t-1) (x,y) là điểm ảnh có tọa độ (x,y) của khung hình tại thời điểm t-1. τ là ngưỡng được định nghĩa trước đó.

Verbeke và Vincent thay vì tính toán sự khác biệt ngay ở khung hình trước đó, họ tính toán sự khác biệt với mười khung hình đã tích lũy trước đó, kết hợp với phương pháp phân tích thành phần chính để giảm kích thước dữ liệu, kỹ thuật họ đề xuất được đánh giá tốt hơn so với kỹ thuật khác biệt khung hình đơn giản. Tuy nhiên với các điều kiện ánh sáng thay đổi, nhiễu và đặc biệt là hiện tượng dừng tạm thời của đối tượng thì kỹ thuật vẫn chưa giải quyết được. Như các kỹ thuật khác biệt khung hình khác, kỹ thuật này cũng không tách được đối tượng trên nền chuyển động.

### Phương pháp trừ nền

Nhận dạng đối tượng chuyển động từ chuỗi video là nên tảng và tác vụ quan trọng trong giám sát video, kiểm tra và phân tích giao thông, dò tìm và theo dõi mỗi con người, nhận dạng cử chỉ trong giao diện người máy . Phương pháp chung để nhận dạng đôi tượng chuyển động là trừ nền, ở đây mỗi khung hình video được so sánh với mỗi mo hình tham chiếu hoặc mô hình nền. Các pixel trong frame hiện thời mà lệch đang kể so với nền sẽ được xem xét như là đối tượng chuyển động. Những pixel này được xử lý tiếp cho định vị và theo dõi đối tượng. Từ nền là bước đầu tiên và quan trọng trong nhiều ứng dụng thị giác máy tính, nó được sử dụng để trích chọn ra những điểm tương ứng đúng với những đối tượng chuyển động quan tâm. Mặc dù nhiều giải thuật trừ nền được đề xuất, những vấn đề của nhận dạng đối tượng chuyển động trong những môi trường phức tạp vẫn còn chưa được giải quyết hoàn toàn.

Có một vài vấn đề mà một giải thuật trừ nền tốt phải giải quyết phù hợp. Xét mội chuỗi video từ camera theo dõi côs định ở một nút giao thông ( tại nơi giao nhau của giao thông). Đây là một môi trường ngoài trời, dó đó một giải thuật trừ nền phải thích nghi với nhiều mức khác nhau với độ sáng tại những thời điểm khác nhau trong ngày và xử lý điều kiện thời tiết bất lợi như sương mù hoặc tuyết làm thay đổi nền. Việc thay đổi bóng, đổ bóng bởi sự di chuyển của đối tượng cần phải được loại bỏ để những đặc trưng thích hợp có thể được trích chọn từ những đối tượng trong quá trình xử lý tiếp theo. Luồng giao thông phức tạp tại chỗ giao nhau luôn đặt ra thách thức mới cho giải thuật trừ nền. Xe di chuyển tốc độ bình thường khi đèn xanh, nhưng dừng lại khi đèn bật đỏ. Xe cộ đứng yên không chuyển động cho đến khi đèn xanh bật lại. Một giải thuật trừ nền tố nhưng phải xử lý những đối tượng chuyển động mà ban đầu nó hòa trộn với nền và sau đó nó không phải nền. Ngoài ra, nhiều ứng dụng cần phải xử lý thời gian thực, nên giải thuật trừ nền phải tính toán nhanh và có yêu cầu bộ nhớ ít, trong khi vẫn có khả năng xác định những đối tượng chính xác trong video.

Trừ nền là một kỹ thuật thường được sử dụng cho việc phân đoạn chuyểnđộng trong những cảnh tĩnh. Mục đích của phương pháp là để phát hiện các khu vực di chuyển bằng cách trừ từng điểm ảnh của hình ảnh hiện tại với một hình nền tham chiếu được tạo ra bằng cách trung bình hình ảnh theo thời gian trong một khoảng thời gian. Các điểm ảnh có sự khác biệt lớn hơn một ngưỡng được phân loại thành foreground. Sau khi tạo ra một ma trận điểm ảnh foreground, một số phương pháp hình thái trong xử lý ảnh chẳng hạn như phép giãn nở, phép xói mòn và phép đóng được thực hiện để làm giảm ảnh hưởng của nhiễu và cải thiện các khu vực được phát hiện. Các nền tham chiếu được cập nhật mới theo thời gian để thích ứng với các đoạn cảnh chuyển động.

Có nhiều phương pháp khác nhau để thực hiện trừ nền. Heikkila và Silven sử dụng phương pháp đơn giản của phương pháp này đó là một điểm ảnh tại vị trí (x, y) trong hình ảnh hiện tại It thuộc về foreground nếu: |It(x,y) – Bt(x-y)| > τ được thỏa mãn trong đó τ là một ngưỡng được xác định trước. Hình nền Bt được cập nhật bằng cách sử dụng bộ lọc Infinite Impulse Response (IIR) như sau:

Bt+1= αIt + (1- α)Bt

Sau đó thực hiện tạo ma trận các điểm ảnh foreground bằng các phép hình thái và loại bỏ các vùng có kích thước nhỏ. Ta có thể hình dung phương pháp này như hình ảnh bên dưới. Từ trái qua phải:

- Ảnh nền

- Ảnh chứa đối tượng

- Khoanh vùng đối tượng sau khi trừ nền

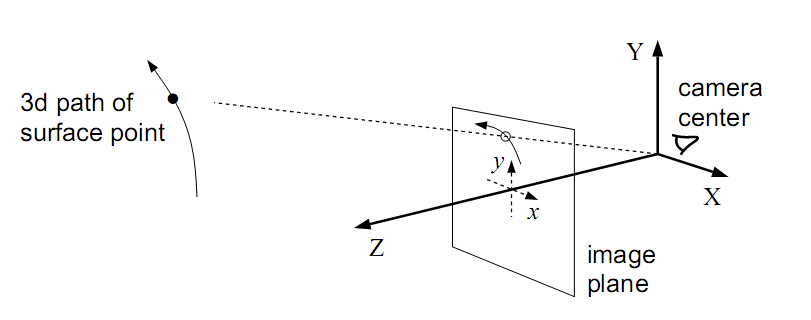
*Mô hình hoạt động của phương pháp trừ nền*

Vì kỹ thuật này khá đơn giản, dễ cài đặt và có thể thực hiện trong thời gian thực nên được sử dụng rất phổ biến. Mặc dù kỹ thuật trừ thực hiện rất tốt việc tác được hầu hết của các điểm ảnh có nằm trong vùng di chuyển ngay cả khi đối tượn dừng lại, nhưng lại thường nhạy cảm với những thay đổi đột ngột, ví dụ, đối tượn đứng yên ra khỏi nền (ví dụ như một chiếc xe đậu di chuyển ra khỏi bãi đậu xe hoặc ánh sáng bị thay đổi đột ngột.

### Phương pháp luồng quang học

Phương pháp dựa vào luồng quang học (optical flow) thực hiện bằng cách sử dụng các vector có hướng của các đối tượng chuyển động theo thời gian để phát hiện các vùng chuyển động trong một ảnh. Chúng có thể phát hiện chuyển động trong các dãy video ngay cả các video thu được từ camera di chuyển.

Luồng quang học là khái niệm chỉ sự chuyển động tương đối của các điểm trên bề mặt một đối tượng, vật thể nào đó gây ra, dưới góc quan sát của một điểm mốc (mắt, camera…). Sự chuyển động của các vật thể (mà thực tế có thể coi là sự chuyển động của các điểm trên bề mặt của vật thể ấy) trong không gian 3 chiều, khi được chiếu lên một mặt phẳng quan sát 2D được gọi là trường chuyển động (motion field). Nói chung, mục đích của các phương pháp luồng quang học là để xác định (xấp xỉ) trường chuyển động từ một tập các khung hình thay đổi theo thời gian, chúng được sử dụng rộng rãi trong các bài toán phân đoạn đối tượng, phát hiện đối tượng chuyển động, theo vết đối tượng…



*Luồng quang học*

Có nhiều kĩ thuật xác định luồng quang học khác nhau, luận án sẽ tập trung vào phương pháp dựa trên gradient, chi tiết các phương pháp khác có thể tham khảo khảo sát của Raveshiya .

Ý tưởng quan trọng của phương pháp tính toán luồng quang học dựa trên gradient là giả định sau:

Bề ngoài của đối tượng không có nhiều thay đổi (về cường độ sáng) khi xét từ frame thứ n sang frame n+1. Nghĩa là:

|  |  |  |
| --- | --- | --- |
|  | (1.2) |  |

Trong đó I(x ⃗,t) là hàm trả về cường độ sáng của điểm ảnh tại thời điểm t (khung hình thứ t). là tọa độ của điểm ảnh trên bề mặt (2D), là vector vận tốc, thể hiện sự thay đổi vị trí của điểm ảnh từ frame thứ t sang frame t+1. Mặc dù có vẻ không thực tế, nhưng nhiều công trình cho thấy giả sử (1.2) làm việc khá tốt trong thực tế.

Sử dụng khai triển Taylor, ta xấp xỉ:

|  |  |  |
| --- | --- | --- |
|  | (1.3) |  |

với ∇I(x ⃗,t), và I\_t (x ⃗,t) là đạo hàm tức thời của I tại thời điểm t.

Bỏ qua các phần tử có bậc đạo hàm cao hơn trong (1.3), so sánh (1.2) và (1.3), đồng thời xấp xỉ bằng , ta nhận được phương trình sau, gọi là phương trình ràng buộc gradient (gọi là phương trình luồng quang học):

|  |  |  |
| --- | --- | --- |
|  | (1.4) |  |

Phương trình (1.4) còn có thể đạt được bằng cách xuất phát ngay từ giả thiết ban đầu là giá trị cường độ sáng của điểm ảnh không thay đổi từ frame t sang t+1, như vậy:

|  |  |  |
| --- | --- | --- |
|  | (1.5) |  |

Phương trình (1.4) là phương trình cơ bản của luồng quang học, trong đó chính là luồng cần tìm. Tuy nhiên phương trình này có dạng với 2 ẩn là u\_1 và u\_2 nên không giải được. Ta có nhận xét rằng phương trình (4), với ẩn u ⃗=(u\_1,u\_2 ), là phương trình đường thẳng trong 2D với vector pháp tuyến là ∇I.

Mọi phương pháp ước lượng luồng quang học đều xoay quanh việc cố gắng tìm thêm một phương trình nữa để tính u ⃗.

Để tìm thêm một ràng buộc nữa cho u ⃗, ta có thể dùng đến các điểm ảnh lân cận với điểm ảnh đang xét, với giả sử là các điểm ảnh lân cận này cũng nằm trên vật thể và chuyển động với cùng vận tốc của điểm ảnh đang xét. Trong thực tế, các điểm ảnh lân cận có thể không chuyển động cùng vận tốc với điểm ảnh đang xét, do đó ta tìm vector vận tốc mà làm cực tiểu bình phương lỗi:

|  |  |  |
| --- | --- | --- |
|  | (1.6) |  |

Phương pháp này gọi là ước lượng bình phương tối thiểu (Least-squares estimation), trong đó g(u ⃗ ) xác định “độ rộng” mà hàm ước lượng này bao phủ. Thông thường người ta chọn hàm Gaussian. Vận tốc cần tìm là giá trị u ⃗ làm cực tiểu E(u ⃗ ).

Người ta chứng minh được E(u ⃗ ) đạt cực tiểu tại điểm đạo hàm riêng của nó bằng zero, nghĩa là:

|  |  |
| --- | --- |
|  | (1.7) |
|  | (1.8) |

Hai phương trình này có thể viết ở dạng ma trận:

|  |  |
| --- | --- |
|  | (1.9) |

Với :

|  |  |
| --- | --- |
|  | (1.10) |

Nghiệm của (1.9) là 

Các phương pháp luồng quang học rất phù hợp với tình huống camera di chuyển, tuy nhiên các phương pháp này luôn đòi hỏi sự tính toán khá phức tạp và khá nhạy với nhiễu.

### Phương pháp thống kê

Một phương pháp tiên tiến hơn là việc sử dụng các đặc tính thống kê của các điểm ảnh riêng lẻ đã được phát triển để khắc phục những thiếu sót của phương phá trừ nền cơ bản. Những phương pháp thống kê chủ yếu được lấy tư tưởng từ các phương pháp trừ nền trong việc lưu giữ và tự động cập nhật các số liệu thống kê củ các điểm ảnh thuộc quá trình xử lý nền. Các điểm ảnh foreground được xác định bằng cách so sánh số liệu thống kê của mỗi điểm ảnh với mô hình nền. Cách tiếp cận này đang trở nên phổ biến hơn do độ tin cậy của nó trong những cảnh chứa cácthay đổi tiếng ồn chiếu sáng, và bóng tối.

Các hệ thống W4 sử dụng một mô hình nền thống kê mà mỗi điểm ảnh được biểu diễn bởi giá trị cường độ tối thiểu (M) và giá trị cường độ tối đa (N) của nó và sự khác biệt cường độ tối đa (D) giữa bất kỳ khung hình liên tiếp quan sát trong suốt quá trình huấn luyện ban đầu mà cảnh không có các đối tượng chuyển động. Một điểm ảnh trong It thuộc về foreground nếu thỏa mãn:

|M(x,y) – It(x,y)| > D(x,y) or |N(x,y) – It(x,y)| > D(x,y)

Sau khi tách ngưỡng, thực hiện một số lần phép toán hình thái xói mòn cho ảnh foreground để loại bỏ các điểm ảnh nhiễu. Để phát triển các khu vực bị xói mòn về kích thước ban đầu, nối tiếp sau phép xói mòn là thực hiện phép giãn nở trên ma trận các điểm ảnh foreground. Ngoài ra, khu vực có kích thước nhỏ được loại bỏ. Sau đó áp dụng dán nhãn các thành phần được ghép nối để phát hiện các vùng. Số liệu thống kê của các điểm ảnh nền thuộc các khu vực không chuyển động của hình ảnh hiện tại được cập nhật với các dữ liệu hình ảnh mới.

Một ví dụ về phương pháp thống kê, Stauffer và Grimson mô tả một mô hình nền hỗn hợp thích ứng để theo dõi trong thời gian thực. Trong cong việc của họ, mỗi điểm ảnh riêng biệt được mô hình hóa bằng một hỗn hợp của Gaussians được cập nhật liên tục từ dữ liệu hình ảnh hiện tại. Để phát hiện xem một pixel thuộc foreground hay background, sẽ dựa vào đánh giá phân phối Gaussian của mô hình hỗn hợp cho điểm ảnh đó.

## Một số đề xuất cài tiến

## Kết quả thực nghiệm

# Ứng dụng

## Môi trường cài đặt

### Môi trường thực nghiệm

* Các môi trường thông thường như: Nhà ở, phòng học, thang máy với các yếu tố
* Ánh sáng vừa đủ
* Không có khói bụi nhiều
* Người đi lại không mang đồ đạc kích thước lớn
* Không có cây cối trong khu vực giám sát
* Ảnh nền không có nhiều vật chuyển động mạnh như quạt trần.
* Màu nền không có quá nhiều màu sắc và vùng màu.

### Thiết bị thực nghiệm

**Thiết bị ghi hình hoặc file video:**

* Thiết bị ghi hình: độ phân giải > = 3.0 Mps
* File video : Kích thước lớn hơn 480x600 pixel

**Cấu hình máy**

* CPU tối thiểu: 2.0 GHz
* RAM tối thiểu 1Gb

### Khoảng cách thực nghiệm

* Thiết bị ghi hình cách mặt đất không quá 4m
* Không dưới 2m
* Góc quay từ 45-90 độ. (tùy thuộc vào tiêu cự của thiết bị ghi hình – Khoảng cách đề nghị trên thông dụng cho các camera an ninh tại trường học)

## Xây dựng ứng dụng

# Kết Luận

## Kết luận

Phát hiện chuyển động của đối tượng trong đoạn video là một đề tài với nhiều ứng dụng trong các lĩnh vực đời sống đặc biệt là ngành đảm bảo trật tự an ninh. Trong báo cáo này em đã trình bày một số phương pháp phát hiện đối tượng chuyển động cùng các ưu nhược điểm của từng phương pháp khi áp dụng chúng. Tùy từng yêu cầu cụ thể về thời gian, về địa điểm và chất lượng của video cần xử lý mà chúng ta chon sử dụng phương pháp phát hiện thích hợp.

Bên cạnh việc đưa ra một số cải tiến cho phương pháp phát hiện đối tượng chuyển động, báo cáo cũng đã cài đặt thửu nghiệm thuật toán trừ nền và sự khác biệt thời gian cho bài toán phát hiện đối tượng chuyển động.

Những mục tiêu đạt được:

* Tìm hiểu các phương pháp phát hiện đối tượng chuyển động trong video.
* Cài đặt thử nghiệm một số phương pháp.
* Phân tích ưu nhược điểm của từng phương pháp.

Những mục tiêu chưa đạt được:

Hướng phát triển trong tương lai:

Tuy nhiên do hạn chế về điều kiện và thời gian , báo cáo sẽ không tránh khỏi nhữn thiếu sót. Kính mong được sự đóng góp ý kiến của thầy cô và các bạn để đề tài nghiên cứu của em sẽ được hoàn thiện hơn.

## Kiến nghị

# Tài liệu tham khảo