LỜI CẢM ƠN

Sau thời gian thực tập và nghiên cứu tại trường ĐHDL Hải Phòng em đã hoàn thành việc tìm hiểu đề tài: Tìm hiểu một số phưcmg pháp phát hiện đối tượng đột nhập, trong thời gian làm đồ án tốt nghiệp em đã nhận được sự giúp đỡ hết sức nhiệt tình của Th.s Ngô Trường Giang và cùng với sự nỗ lực của bản thân nên em đã hoàn thành đề tài được giao.

Em xin chân thành cảm ơn sự giúp đỡ của các thầy cô khoa công nghệ thông tin, thầy giáo Th.s Ngô Trường Giang và toàn thể các bạn sinh viên khoa Công nghệ thông tin, trường Đại học Dân lập Hải Phòng cùng với gia đình đã ủng hộ giúp đỡ và tạo mọi điều kiện thuận lợi cho em trong suốt quá trình học tập và nghiên cứu để hoàn thành tốt đồ án tốt nghiệp này.

Hải Phòng, tháng 07 năm 2007  
Sinh viên

Nguyễn Quỳnh Nga

MỤC LỤC

[LỜI CẢM ƠN 1](#bookmark0)

[MỞ ĐẦU 4](#bookmark8)

[CHƯƠNG 1: TỔNG QUAN YỀ xử LÝ VIDEO 5](#bookmark9)

1. [Giới thiệu 5](#bookmark10)
2. [Các hiệu ứng biên tập video 7](#bookmark11)
3. [Các thuộc tính đặc trưng của video 9](#bookmark12)
4. [Color 9](#bookmark13)
5. [Texture 9](#bookmark15)
6. [Shape 10](#bookmark16)
7. [Motion 10](#bookmark17)
8. [Phân đoạn video 10](#bookmark18)
9. [Kỹ thuật trừ ảnh 13](#bookmark21)
10. [Dựa vào so sánh điểm ảnh 14](#bookmark22)
11. [Dựa vào khối 16](#bookmark29)
12. Dựa vào so sánh biểu đồ 18
13. [Dựa vào phưorng pháp thống kê 23](#bookmark42)

CHƯƠNG 2: MỘT SỐ PHƯƠNG PHÁP PHÁT HỆN CHUYÊN ĐỘNG 25

1. [Tổng quan phưomg pháp trừ nền 25](#bookmark61)
2. [Giới thiệu 25](#bookmark62)
3. [Giải thuật trừ nền 26](#bookmark63)
4. [Tiền xử lí 27](#bookmark64)
5. [Mô hình hóa nền 27](#bookmark65)
6. [Phát hiện đối tượng 32](#bookmark74)
7. [Hợp lí hoá dữ liệu 33](#bookmark78)
8. [Trừ nền sử dụng thông tin màu và gradiant 35](#bookmark79)
9. [Giới thiệu 35](#bookmark94)
10. Mô hình hóa nền 35
11. [Trừ nền 36](#bookmark82)
12. [Trừ nền dựa vào màu 37](#bookmark84)
13. [Trừ nền dựa vào biên 38](#bookmark85)
14. Kết hợp các kết quả trừ màu và trừ biên 39
15. Nhận xét 40
16. [Phát hiện đối tuợng chuyển động dựa vào kết cấu 40](#bookmark89)
17. Giới thiệu 40
18. [Toán tử mẫu nhị phân cục bộ 41](#bookmark90)
19. Phuomg pháp phát hiện đối tuợng chuyển động 42
20. Phát hiện đối tuợng chuyển động dựa vào sự biến thiên cục bộ của

véctơ kết cấu SP 45

1. Giới thiệu 45
2. Biểu diễn video với vectơ kết cấu SP 47
3. Phát hiện đối tuợng chuyển động dựa trên biến thiên cục bộ 48

CHƯƠNG 3: THỬ NGHIỆM PHÁT HỆN ĐỐI TƯỢNG ĐỘT NHẬP.... 50

* 1. [Mô tả bài toán 50](#bookmark100)
  2. [Môi truờng test 50](#bookmark101)
  3. [Một số giao diện 50](#bookmark102)
  4. [Đánh giá 52](#bookmark104)

[KẾT LUẬN 54](#bookmark105)

[TÀI LIỆU THAM KHẢO 55](#bookmark106)

MỞ ĐẦU

Ngày nay, sự phát triển của công nghệ thông tin đã xâm nhập vào hầu hết các mặt của đời sống xã hội với những ứng dụng rộng rãi hỗ trợ cho con người trên nhiều lĩnh vực, trong đó có lĩnh vực an toàn an ninh, chống xâm nhập.Việc đảm bảo an ninh là một vấn đề quan trọng và cần thiết (Vd : Trong các nút giao thông quan trọng, trong các siêu thị, trong các ngân hàng hay bảo tàng là những nơi mà vấn đề an ninh được đặt lên hàng đầu ). Có nhiều phương pháp nhằm đảm bảo an ninh sao cho đạt được hiệu quả cao, ưong đó có phương pháp áp dụng các hệ thống camera để quan sát sự chuyển động của các đối tượng trong khu vực. Ngày nay với sự tiến bộ vượt bậc của khoa học, ta có thể xây dựng một hệ thống cảnh báo tự động phát hiện đối tượng đột nhập mà không càn con người phải trực tiếp theo dõi camera, điều này giúp cho con người giảm thiểu được thời gian và sức lực đồng thời vẫn cho một kết quả chính xác.

Với bài toán phát hiện đối tượng đột nhập có 2 hướng tiếp cận để giải quyết đó là: dựa vào phần cứng và dựa vào các kĩ thuật xử lý ảnh. Trong đồ án tốt nghiệp này em xin trinh bày một số phương pháp phát hiện đối tượng chuyển động trong đoạn video dựa vào các kĩ thuật xử lí ảnh. cấu trúc của bài khóa luận gồm phàn mở đầu, phần kết luận, phàn phụ lục và 3 chương với nội dung:

Chương 1: Trinh bày tổng quan về video và các kĩ thuật trừ ảnh.

Chương 2: Một số phương pháp phát hiện chuyển động trong video.

Chương 3: Chương trình ứng dụng phát hiện đối tượng đột nhập.

CHƯƠNG 1: TỔNG QUAN VỀ xử LÝ VIDEO

1. Giói thiêu

Dữ liệu video số ngày càng được sử dụng rộng rãi. Dữ liệu video số bao gồm hình ảnh và âm thanh video được lưu trữ trong máy tính dưới dạng số. Trước tiên ta tìm hiểu video số là gì? Video số hay Video là tập hợp các khung hình được đặt liên tiếp nhau, mỗi khung hình là một ảnh số.

video

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Scene 1 |  | Scene 2 |  | Scene n |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Shot 1 |  | Shot2 |  | Shotn |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Frame 1 |  | Frame 2 |  | Frame n |

Hình 1.1: cấu trúc phân đoạn của video

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

Khung hình i khung hình i+1

Hình 1.2: Các khung hình

ĐỔ 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.

Shoti

Shot i+1

Shot i+2

Biên shot

Biên shot

Chuyển cảnh

Hì**nh** 1.3: Đoạn Ctf sở

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.

Ngoài ra có thể biểu diễn cấu trúc video dựa trên đối tượng:

video

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Shot 1 |  | Shot2 |  | Shot n |
|  | |  | |  |
| object 1 |  | object 2 |  | object n |
|  | |  | |  |
| Region 1 |  | Region 2 |  | Region n |

Color, texture, shap

* Motion, trajectoữy
* Spatial,temporal, stuctures

Feature

Hình 1.4: Biểu diễn video dựa trên đối tượng

1. Các hiệu ứng biên tập vỉdeo

Số các hiệu ứng video có thể là rất lớn. Các chương trình biên tập video nổi tiếng như Adobe Premiere hay Ulead MediaStudio có tới hàng trăm loại biên tập được tham số hóa. Tuy nhiên ừên thực tế, hơn 99% các biên tập đó đều nằm trong ba loại: cắt cứng, fade và chồng mờ; ngoài ra còn có wipe.

* Cắt cứng: là sự biến đổi đột ngột từ cảnh này sang cảnh khác, nó xảy ra giữa haỉ khung hình.
* Fade\ là sự bỉến đổỉ dần dần giữa một cảnh và một ảnh (fade out) hoặc là sự biến đổi dần giữa một ảnh và một cảnh phim (fade ỉn).
* Chồng mờ: là sự biến đồi dần dần từ cảnh cũ sang cảnh mới trong đó cảnh cũ được fade out và cảnh mới được fade in.
* Wipe: là hiện tượng cảnh chạy theo một đường trên màn hình, trong khỉ đó cảnh mới xuất hiện phía sau đường đó.

Các hiệu ứng đó tạo nên sự bỉến đổi cho các khung hình, tạo nên các chuyển cảnh, cắt cứng tạo nên chuyển cảnh đột ngột; fade, chồng mờ và wipe tạo nên chuyển cảnh dần dần.

Hình 1.5: Một sé loại Wipe cơ bản

1. Các thuộc tính đặc trưng của vỉdeo
2. Color

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image1.pngC:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image2.pngVới mỗi một ảnh đều có một biểu đồ màu để biểu diễn sự phân bố màu trong ảnh, biểu đồ màu không phụ thuộc vào việc quay ảnh, dịch chuyển ảnh, chiều nhìn ảnh mà phụ thuộc vào vào hệ màu và các phương pháp định lượng được dùng.

I

(a)

(b)



(c) (d)

Hình 1.6: Biểu đồ màu của 4 loại ảnh Ctf bản (a) ảnh tối (b) ảnh sáng (c)ảnh tương phản thấp (d) ảnh tương phản cao

1. Texture

Đây là một đặc trưng quan họng của bề mặt, nơi xảy ra việc lặp lại mẫu cơ bản. Có hai dạng biểu diễn Texture phổ biến: biểu diễn dạng ma trận đồng thời và biểu diễn Tamura.

- Ma trận đồng thời mô tả hướng và khoảng cách giữa các điểm ảnh, ta có thể trích chọn được các thống kê có ý nghĩa. Ngược lại, người ta thấy rằng entropi và mô-men chênh lệch nghịch đảo lại có khả năng phân biệt tốt nhất.

* Biểu diễn Tamura được thúc đẩy nhờ các nghiên cứu về tâm lí trong việc thu nhận trực giác của con người, nó bao gồm các thuộc tính đo tính thô, độ tưomg phản, hướng, tính tron, tính cân đối và độ thô ráp. Các đặc tính này rất quan trọng trong việc tìm hiểu nội dung ảnh vì nó biểu diễn rất trực quan.

1. Shape

Các đặc trưng hình dáng có thể được biểu diễn sử dụng phân tích hình dáng truyền thống như mô-men bất biến, mô tả Fourier, mô hình học tụ động quay lui và các thuộc tính hình học. Các đặc trưng này có thể được phân chia thành đặc trưng toàn cục và đặc trưng cục bộ.

* Đặc trưng toàn cục là đặc trưng thuộc tính thu được từ toàn bộ hình dáng ảnh (VD: chu vi, tính tròn, hướng trục chính...).
* Đặc trưng cục bộ là đặc trưng thu được từ việc thao tác với một phần của ảnh, không phụ thuộc vào toàn bộ ảnh.

1. Motion

Là thuộc tính quan trọng của video, thông tin về chuyển động có thể được sinh ra bằng các kỹ thuật ghép khối hoặc luồng ánh sáng. Các đặc trưng chuyển động: mô-men của trường chuyển động, biểu đồ chuyển động, các tham số chuyển động toàn cục có thể được trích chọn từ vectơ chuyển động. Các đặc trưng mức cao phản ánh di chuyển camera như quét camera (pan), nghiêng (tilt), phóng to (zoom in), thu nhỏ (zoom out) cũng có thể được trích chọn.

1. Phân đoạn video

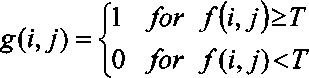
Là quá trình phân tích và chia luồng video dài thành các đom vị nhỏ như shot, quá trình này sẽ phân tích và phát hiện ranh giới các shot. Quá trinh này cũng được coi như là việc phát hiện những chuyển cảnh giữa các đoạn, xác định khung hình bắt đầu và khung hình kết thúc của một đoạn cơ sở.

Chuyển cảnh

Hì**n**h 1.7: Biên đoạn cơ sở

Trong những năm gàn đây, những nghiên cứu về việc phát hiện tự động đoạn cơ sở ngày càng phát triển, những ứng dụng ngày càng nhiều và càng có nhiều thuật toán được công bố để giải quyết vấn đề phân đoạn cơ sở cho những mức độ phức tạp khác nhau của dữ liệu thật.

Các kỹ thuật phân đoạn có thể được chia thành 3 loại: phân đoạn dựa vào ngưỡng, phân đoạn dựa trên phát hiện cạnh, phân đoạn bằng phương pháp nở vùng. Phân đoạn dựa vào ngưỡng sẽ biến đổi một ảnh đầu vào f thành một ảnh nhị phân đầu ra g dựa trên một ngưỡng T cho trước như sau:

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image5.pngC:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image6.png(1.1)

Đoạn cơ sở 1

Đoạn cơ sở 2

Đoạn cơ sở 3

Biên đoạn cơ sở

Biên đoạn cơ sở

Yới i và j là các tọa độ X và Y, g(i,j) =1 chỉ ra rằng điểm ảnh (i,j) thuộc về đối tượng và ngược lại g(i,j)= 0 chỉ ra rằng điểm ảnh thuộc về nền. T là giá trị ngưỡng trong không gian đặc trưng. Với một số ngưỡng cho trước ta có thể chia một ảnh thành các đối tượng.

Hoặc một phương pháp phân đoạn khác cũng dựa vào ngưỡng là: rút trích các đặc trưng của mỗi khung hình trong đoạn video, rồi tính toán sự khác biệt giữa đặc trưng của các khung hình liên tiếp, sau đó so sánh những khác biệt này với một ngưỡng cho trước. Mỗi khi sự khác biệt vượt quá ngưỡng thì có nghĩa là tìm thấy được một đoạn cơ sở. Từ vấn đề sử dụng đặc

trưng nào của khung hình, và sử dụng như thế nào đã nảy sinh ra nhiều phương pháp, hướng tiếp cận khác nhau.

Hình 1.8: Tính khác biệt đặc trưng giữa những khung hình liên tiếp

Hình 1.9: Đường sai biệt về đặc trưng lược đồ giữa các khung **h**ì**nh**

liên tiếp

Kỹ thuật phân đoạn dựa trên phát hiện cạnh dùng các điểm ảnh để tìm biên đối tượng. Cạnh là dữ liệu có tần số cao chỉ sự không liên tục của màu sắc trong một ảnh. Làm thế nào để phân biệt giữa các cạnh và nhiễu là một vấn đề khó của các thuật toán tìm cạnh. Một số lớn các phương pháp đã được giới thiệu để theo vết các biên đối tượng hoặc để nối các mô hình hình dạng mong muốn.

Hình 1.10: Các kết quả phát hiện cạnh của một ảnh

Dữ việc phát sinh nhiễu tự nhiên của phương pháp phát hiện cạnh, việc xây dựng nên các đối tượng bằng cách tìm kiếm biên của chúng gặp nhiều khó khăn. Các phương pháp nở vùng sinh ra các phân vùng một cách trực tiếp dựa trên các chuẩn đồng nhất (vd: màu sắc). Hướng tiếp cận dưới-lên gom các điểm ảnh vào các vùng nếu thòa mãn độ đo đồng nhất Trái lại, hướng tiếp cận trến-xuống chia ảnh thành các vùng cho đến khỉ thỏa mãn cẫc độ đo đồng nhất. Các hướng tiếp cận tổng hợp chia-và-trộn kể thừa, và phát triển những ưu điểm của hai phương pháp trên.

Hình 1.11: Một ví dụ phân đoạn ảnh bằng phương pháp nở vùng

Các kỹ thuật phân đoạn tiên tiến được phát triển trong cả ba loạỉ trên nhằm cho kết quả phân đoạn tất hơn. Một số ví dụ đã biết như các thuật toán watershed dùng các hình thái toán học để nở vùng, mạng neural học các thông tin ngữ cảnh từ dữ liệu huấn luyện, và gom nhóm dữ liệu dựa trên lý thuyết mờ.

1. Kỹ thuật trừ ảnh

Hiểu theo nghĩa hẹp, trừ hai ảnh có cùng kích thước là việc xây dựng ảnh mới từ sự khác biệt của hai ảnh. Theo nghĩa rộng hơn, trừ ảnh là việc tính toán độ chênh lệch giữa hai ảnh trên một đặc trưng ảnh nào đó như cường độ, màu sắc, texture (kết cấu), shape (hlnh dáng), chuyển động...

Có nhiều kỹ thuật trừ ảnh khác nhau, nhưng có thể chỉa thành 3 loại:

- Dựa vào đỉểm ảnh: So sánh các cặp điểm ảnh tương ứng trên hai ảnh.

* Dựa vào khối: Chia ảnh thảnh các miền và so sánh các miền tương

ứng.

* Dựa vào biểu đồ: So sánh sự phân bố của thuộc tính nào đó của ảnh.

Kết hợp các loại này với các thuộc tính so sánh của ảnh, ta sẽ có nhiều kỹ thuật trừ ảnh khác nhau.

1. Dựa vào so sánh điểm ảnh

Đây là phương pháp để tính toán sự sai khác giữa hai Ếrame bằng việc tính toán các giá trị, nó mô tả mọi thay đổi về cường độ điểm ảnh trong các ảnh. Có nhiều phương pháp để tính sự sai khác này, Nagasaka và Tanaka đã đưa ra một phương pháp tính tổng toàn bộ những thay đổi khác nhau về cường độ điểm ảnh giữa hai khung hình như là độ chênh lệch khung D(fi f2).

**D(fỉ,f2) = —X Ẽ l/i (\*» y) - /2 (\*» y)\ (**1**.**2**)**

^ x X x-0 J>=0

Sau khi tính được độ chênh lệch D, tiến hành so sánh D với ngưỡng chuyển cảnh T xác định xem có chuyển cảnh hay không.

Nhược điểm của phương pháp này là:

* Không thể phân biệt được thay đổi lớn cho vùng ảnh nhỏ và thay đổi nhỏ cho vùng ảnh lớn. Ví dụ như các cắt cảnh rất dễ bị bỏ sót khi một phần nhỏ của khung hình có sự thay đổi lớn hoặc nhanh.
* Nhạy với nhiễu và các di chuyển của camera.

Một bước phát triển hơn được Otsuji đề xuất đó là thay vì tính toán trực tiếp tổng những điểm khác biệt lớn về cường độ thực tế, tiến hành đếm các số điểm ảnh có thay đổi lớn hơn một ngưỡng nào đó, so sánh tổng đó với ngưỡng khác để phát hiện chuyển cảnh.

***DP(x*** ***v)*** = **í1 ’Nếu** 1/1(x’yì-f\*(x’yì>T' **(1.3)**

[0 , ngược lại

DƯì,Â) = 1^—ỵrỈDP(x,y) **(1.4)**

Nêu tỷ lệ số điểm ảnh thay đổi D(fi, f2) lớn hom ngưỡng T2 thì đã có chuyển cảnh do cắt. Tuy các thay đẩi không liên quan trong khung hình đã được loại bớt nhưng phưomg pháp này vẫn nhạy cảm với những di chuyển camera và di chuyển của đối tượng khi camera quay hướng theo đối tượng, rất nhiều đỉểm ảnh thay đổỉ dù chỉ một số ít điểm ảnh dịch chuyển.

Hình 1.12: Ảnh hưởng của chớp sáng

Một nhược điểm nữa của phương pháp phân biệt điểm ảnh là tính nhạy cảm những thay đổi về độ sảng của ảnh, ví dụ điền hình là các chớp sáng (đèn tlash).

Trên hình 1.12, giá trị độ xám nhảy lên mức cao khi chớp sáng xuất hiện. Đỉều này sẽ trở lại bình thường sau một số ữame do các thay đổi mỡ của camera. Nhưng với một cảnh thật, phân bố màu sẽ không trở lại mức ban đầu. Người ta dùng tỉ lệ khác biệt màu qua frame và khác biệt màu long tem để phát hiện ílash. Tỉ lệ này được định nghĩa:

***Frịị)*** = ***D(ỉ*,** i **-**1**)** ỉ ***CHẠ*** + ồJ-ĩ) **(**1**.**5**)**

Trong đó i là írame đang xét, và ổ là chiều dài trung bình của thay đổi mở của camera. Nếu Fr(i) nhỏ hom một ngưỡng cho trước thì một chớp sáng được phát hiện tại vị trí frame thứ i và ngược lại. Khi đó người ta điều chinh độ sai khác giá trị điểm ảnh bằng cách chia nó cho cường độ của điểm ảnh trên khung hình thứ hai.

(/1./2) XxYỵL fẢX<y) (1.6)

Phương pháp trừ giá trị điểm ảnh cơ bản là tính toán từ các giá trị cường độ, nhưng có thể mở rộng với các ảnh màu. Ví dụ với ảnh màu RGB, ta tính tổng có trọng số các sai khác của ba giá trị Red, Green và Blue của các điểm ảnh.

D{ẤJi)=ị,Ỷ, **z** wì\fu{x,y)~ f2i(x,y} **(1.7)**

x=ũ y=ữ *ìz{r,g,b}*

1. Dựa vào khối

Trái ngược với hướng tiếp cận sử dụng các đặc tính toàn cục của cả khung hình, hướng tiếp cận phân khối sử dụng các đặc tính cục bộ nhằm tăng tính độc lập với các di chuyển của camera và đối tượng. Mỗi khung hình được chia thành b khối. Các khối trên khung hình fi được so sánh với khối tương ứng trên khung hình f2. về cơ bản, độ chênh lệch giữa hai khung hình được tính như sau:

DƯí,f*2*)=ịlCt.DPỰt,fỉ,k) (18)

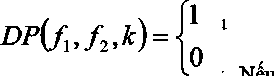
k-ỉ

Trong đó Ck là hệ số cho trước, DP(fi,f2,k) là độ chênh lệch giữa khối thứ k của hai khung hình fi và f2.

Kasturi so sánh các khối tương ứng áp dụng công thức:

ơ\k'ơ *2*k

(1.9)

Trong đó Mik ’ Pik là giá trị cường độ trung bình của khối thứ k, và Ơ1 k >ơ2k là độ chênh lệch tương ứng với hai khối đó.

Nếu \

Nếu ngược lại

(1.10)

Một cắt cảnh xảy ra khi số các khối thay đổi đủ lớn, nghĩa là D(f\,f2) > T2 và ck =1 cho tất cả các khối. Phương pháp này chậm đi theo độ phức tạp của hàm thống kê. Phương pháp này có một bất lợi là các chuyển shot sẽ bị bỏ qua trong trường hợp hai khối rất khác nhau có thể có cùng hàm mật độ. Tuy nhiên trường hợp đó cũng ít xảy ra.

Một hướng tiếp cận khác với kỹ thuật trừ ảnh phân khối do Shahraray đưa ra. Shahraray đã chia khung hình thành 12 miền và tìm miền thích hợp nhất cho mỗi miền ở khung hình kia. Độ chênh lệch tính bằng kỹ thuật trừ ảnh dựa vào điểm ảnh của từng miền được sắp xếp. Tổng có trọng số của các chênh lệch đã sắp xếp cho ta kết quả D cuối cùng.

Xiong phát triển phưomg pháp trừ ảnh, gọi là so sánh thực, phát hiện chuyển cảnh do ngắt chỉ bằng việc so sánh một phàn của ảnh. Phương pháp này chỉ ra rằng, sai sót mắc phải hoàn toàn có thể bỏ qua nếu ít hơn một nửa số các cửa sổ cơ sở (các ô vuông không chồng nhau, hình 1.13) đều được kiểm tra. Trong trường giữa hai khung hình có sự biến đổi lớn thì kích thước các cửa sổ được chọn đủ lớn để bất biến với các thay đổi không làm vỡ và đủnhỏ để có thổ chứa thông tin về không gian nhiều chừng nào có thể. Các cửa sổ cơ sờ được so sánh và tính độ chênh lệch mức xám hoặc giá trị màu của các điểm ảnh. Khỉ giá trị chênh lệch lớn một ngưỡng nào đó thì xem như miền đang xét đã thay đổỉ. Khỉ số miền thay đổi lớn hơn một ngưỡng khác thì sự chuyển cảnh do ngắt đã xảy ra. Thực nghiệm đã chứng minh rằng hướng tiếp cận này cho tốc độ nhanh hơn phương pháp so sánh từng cặp điểm, thậm chí cả phương pháp biểu đầ xét dưới đây.

Hình 1.13: Các cửa sổ cơ sở trong thuật toán so sánh thực

1. Dựa vào so sánh bỉểu đồ

Phương pháp đo sự khác biệt giữa các írame dưới dạng giá trị màu không mạnh do chuyển động của camera và đối tượng có thể gây ra sự khác biết giá trị điểm ảnh quá lớn. Có thể dùng biểu đồ màu hoặc biểu đồ mức xám để tính toán sự sai khác giữa hai khung hình vì sự phân bố màu giữa các írame liên tục không bị ảnh hưởng nhiều bởi chuyển động của camera và chuyển động của đối tượng.

Biểu đà màu (mức xám) của khung hình i là một vectơ G chiều Hi = (Hi(l), Hi(2), ... , Hi(G)). Trong đó G là số màu (mức xám), Hi(j) là số điểm ảnh của khung hình ỉ có màu (mức xám ) j.

a) Bỉểu đề toàn cục

Phương pháp đơn gỉản nhất là tính tổng sự saỉ khác các cột của biểu đồ.

DƯi,f2) = Ễ\HiW-HM

k=0 v ' J

Trong đó Hn là biểu đồ mức xám của ảnh thứ n, k là giá trị hợp lí của mức xám G.

Có thể sử dụng thêm trọng số nếu có một số màu (mức xám) được xem xét quan trụng hơn với mục tiêu so sánh.

Dự,, f2) = ỵ w(k)\Hl (k) - H*2* (k)\

k=0 v ' J

Trong đó w(k) là trọng số ứng với giá trị màu (mức xám ) k.

Swain và Ballard lại sử dụng sự giao nhau của biểu đồ được so sánh:

Hình 1.14: So sánh haibỉễu đề

Vùng biểu đồ chung nhau, phần gạch chéo ừong hình 1, cho biết độ tương tự về nội dung hai ảnh có thể được định nghĩa như sau:

(1.13)

jfc=0

Độ tương tự còn có thể được định nghĩa như sau:

Zmin

S(Á,f2) = ỵ-

£maxiH^-H^k))

Như vậy có thể tính độ chênh lệch biểu đồ hai khung hình theo công

thức:

f>n(

fl«./2) = l-S(/l,/2) = l-g : H15)

**2>ax(**H,(k)-H2(k))

k-0

Phương pháp khác biệt về biểu đồ màu được sử dụng nhiều và thông dụng nhất vì nó tính toán nhanh, đơn giản và hiệu quả trong việc phát hiện chuyển cảnh đột ngột, chuyển cảnh rõ ràng, hoặc có sự dịch chuyển nhỏ của đối tượng và sự dịch chuyển nhỏ của camera. Nhưng với chuyển cảnh dần dần, các ảnh từ từ mờ đi, đan xen lẫn nhau, làm cho khác biệt về đặc trưng giữa các khung hình liên tiếp tương đối nhỏ, không đủ để vượt qua ngưỡng xác định đã đặt ra, dẫn đến khó phát hiện được chuyển cảnh. Nếu như ngưỡng xác định thấp quá thì sẽ tìm ra nhiều đoạn dư thừa, nếu đặt ngưỡng quá cao thì không phát hiện được chuyển cảnh. Mặt khác, trong trường hợp camera hay đối tượng chuyển động nhanh cũng tạo ra sự khác biệt tương đối lớn giữa các khung hình. Đổ giải quyết vấn đề này, một phương pháp được đề xuất là sử dụng hai ngưỡng để tăng mức độ tìm đúng và phát hiện được chuyển cảnh dàn dần, đó chính là phương pháp so sánh cặp.

Phương pháp so sánh cặp tính toán chênh lệch tích lũy giữa các khung hình sử dụng một ngưỡng lớn hơn cho chuyển cảnh trực tiếp Th và một ngưỡng nhỏ hơn cho chuyển cảnh dần dàn Ti.

Trước tiên sử dụng Th để phát hiện chuyển cảnh do cắt cảnh. Sau đó sử dụng Ti để phát hiện vị trí khung hình Fs có thể là khung hình đầu tiên của chuyển cảnh dần dần, khung Fs này được đem so sánh với các khung tiếp theo, công việc so sánh tích lũy vì trong suốt quá trình biến đổi dần dần độ chênh lệch sẽ tăng lến. Khung hình cuối cùng của chuyển cảnh dần dần sẽ được phát hiện khi độ chênh lệch giảm xuống thấp hơn mức ngưỡng Ti, trong đó so sánh tích lũy vượt mức ngưỡng Th.

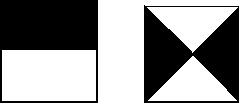
Hình 1.15: Phát hiện chuyển cảnh dần dần bằng phưomg pháp

so sánh cặp

Nêu độ chênh lệch giữa các khung hình liên tiếp giảm xuống dưới ngưỡng Ti mà so sánh tích lũy chưa vượt Th thì bỏ qua vị trí FB và việc tìm kiếm bắt đầu với một biến đổi dần dần khác. Tuy nhiên, có nhiều biến đổi dần dần mà chêch lệch giữa các khung hình liên tiếp đều nhô hơn ngưỡng bé Tị. Vấn đề này có thể giải quyết dễ dàng bằng cách đặt giá trị chấp nhận được cho phép chỉ một số lượng nhất định các khung hình liên tiếp có chêch lệch thấp trước khi loại trường hợp biến đổi đang xét. Như vậy, phương pháp so sánh cặp có thể phát hiện chuyển cảnh đột ngột và chuyên cảnh dần dần cùng một lúc. Qua kết quả thực nghiệm cài đặt thuật toán và so sánh các kỹ thuật phân đoạn khác nhau và thấy rằng so sánh cặp là phương pháp đơn giản và phân đoạn rất tết.

Nhược điểm của phương pháp này là:

- Biểu đồ chỉ mô tả sự phân bố các giá trị điểm ảnh (màu hay mức xám) chứ không chứa đựng các thông tin nào về không gian. Như vậy hai ảnh có cùng biểu đồ màu nhưng vẫn có thể có nội dung khác nhau do không gian phân bố khác nhau



Hì**n**h 1.16: Các ảnh có cùng biểu đồ màu nhưng nội dung khác nhau

- Với những vùng cảnh nhỏ, khi thay đổi vẫn gây ra chú ý nhưng lại không đóng vai trò quan trọng trong biểu đồ và như vậy thì rất dễ bị bỏ qua khi tiến hành kĩ thuật trừ ảnh để tìm ra sự sai khác.

b) Biểu đồ cục bộ

Biểu đồ cục bộ là biểu đồ mô tả sự phân phối các giá trị điểm ảnh trên một phần của khung hình. Như đã đề cập ở trên, phương pháp trừ ảnh dựa vào biểu đồ là phương pháp ít chịu ảnh hưởng của nhiễu và sự di chuyển đối tượng. Tuy nhiên với biểu đồ toàn cục thì vẫn gặp một số trở ngại, để khắc phục những nhược điểm của biểu đồ toàn cục, chúng ta sẽ kết hợp trừ ảnh dựa vào biểu đồ với kỹ thuật trừ ảnh phân khối. Trừ ảnh phân khối quan tâm đến thông tin về không gian, về cơ bản phương pháp này tốt hơn việc so sánh từng cặp điểm ảnh, nhưng nó vẫn chịu tác động của sự di chuyển của camera và đối tượng và cũng tốn kém. Còn kỹ thuật trừ ảnh dựa vào biểu đồ không chịu ảnh hưởng nhiều của camera hay sự di chuyển của đối tượng, nhưng lại có nhược điểm là không chứa đựng thông tin về không gian. Như vậy việc kết hợp hai phương pháp này sẽ bù đắp được những thiếu sót cho nhau. Bằng cách kết hợp hai ý tưởng, chúng ta vừa có thể giảm được tác động của sự dichuyển camera và đối tượng, vừa sử dựng thông tin về không gian ảnh, và cho kết quả phân đoạn tốt hơn.

Ý tưởng là, ta sẽ chia khung hình thảnh b khối, đánh số từ 1 đến b. So sánh biểu đồ của các khối tương ứng rồi tính tổng chênh lệch để có kết quả trừ ảnh cuối cùng.

D(A,fù = ỲJDPựl

(1.16)

\*=1

Với

**DPƯVÂ)** = ẳ|ff,(M)-7/2(M)| **(1.17)**

7=0

Trong đó H(j,k) là giá trị biểu đồ tại mức xám j ứng với khối thứ k.

Hướng tiếp cận khác trong kỹ thuật trừ ảnh dựa vào biểu đồ cục bộ được Swanberg đưa ra. Sự chênh lệch DP(fi, f2, k) giữa các khối được tính bằng cách so sánh biểu đồ màu RGB sử dụng công thức sau:

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image9.png***>J2> )* 2-1 *Hc(i kì* (l-18)**

*cz{r,G,b}j=* 0 *ĩ12*\J-)K)

1. Dựa vào phương pháp thống kê

Phương pháp sai khác thống kê dựa vào phương pháp trừ giá trị điểm ảnh, nhưng thay vì tính tổng sự sai khác của tất cả các điểm ảnh, ta chia ảnh thành các miền rồi so sánh các đại lượng thống kê điểm ảnh của các miền đó.

Ta sử dụng thống kê tỷ lệ số điểm ảnh thay đổi trên toàn bộ khung hình, sử dụng một giá trị d là ngưỡng sai khác được tính giữa hai điểm ảnh tương ứng.

Gọi s là tập các điểm ảnh có độ sai khác lớn hơn d:

s = {(x,y) \ |fi(x,y) - f2(x,y)| > d} (1.19)

Độ sai khác giữa hai khung hình được tính bằng tỷ lệ các điểm ảnh có độ chênh lệch lớn hơn d.

Dựl9f2 ) =

***S.count***

***X\*Y***

(1.20)

Chúng ta có thể sử dụng cách khác là dùng các đại lượng thống kê cho từng miền, như biểu đồ chẳng hạn. Phương pháp này có khá nhiều sai sót trong phát hiện cảnh phim.

CHƯƠNG 2:

**MỘT SỐ PHƯƠNG PHÁP** PHÁT HIỆN CHUYỂN ĐỘNG

1. Tổng quan phương pháp trừ nền
2. Giới thiệu

Nhận dạng đối tượng chuyển động từ chuỗi video là nền tảng và tác vụ quan họng trong giám sát video, kiểm tra và phân tích giao thông, dò tìm và theo dõi con người, nhận dạng cử chỉ trong giao diện người máy. Phưcmg pháp chung để nhận dạng đối tượng chuyển động là trừ nền, ở đây mỗi khung video được so sánh với mô hình tham chiếu hoặc mô hình nền. Các pixel trong frame hiện thời mà lệch đáng kể so với nền sẽ được xem như là đối tượng chuyển động. Những pixel này được xử lí tiếp cho định vị và theo dõi đối tượng. Trừ nền là bước đầu tiên và quan trọng trong nhiều ứng dụng thị giác máy, nó được sử dụng để trích trọn ra những điểm tương ứng đúng với những đối tượng chuyển động quan tâm. Mặc dù nhiều giải thuật trừ nền được đề xuất, nhưng 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 họp. Xét một chuỗi video từ camera theo dõi cố định ở một nút giao thông (tại nơi giao nhau của giao thông). Đây là một môi trường ngoài trời, do đó một giải thuật trừ nền phải thích nghi với nhiều mức khác nhau của độ sáng tại những thời điểm khác nhau trong ngày và xử lí 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 lí tiếp theo. Luồng giao thông phức tạp tại chỗ giao nhau luôn đặt ra những 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 tốt phải xử lý những đối tượng chuyển động màban đầu nó hòa trộn với nền và sau đó nó không phải là 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.

Trong phần này sẽ tập trung giới thiệu vấn đề của trừ nền và tổng họp lại các phương pháp hiện tại, trên cơ sở đó đưa ra sự so sánh giữa chúng dựa vào 4 bước chức năng khác nhau: Tiền xử lí, mô hình hóa nền, phát hiện đối tượng và họp lệ hóa dữ liệu.

1. Giải thuật trừ nền.

Mặc dù tồn tại vô số các giải thuật trừ nền trong các tài liệu, nhưng hầu hết chúng đều theo một lưu đồ đơn giản như trong hình 2.1. Bốn bước chính trong giải thuật trừ nền là: Tiền xử lý, mô hình hóa nền, phát hiện đối tượng và hợp lệ hóa dữ liệu. Tiền xử lí bao gồm những tác vụ xử lí ảnh đơn giản cho video đầu vào tạo điều kiện cho xử lí trong những bước tiếp theo. Mô hình hóa nền sử dụng những Ếrame video mới để tính toán và cập nhật một mô hình nền. Mô hình nền này cung cấp phàn mô tả thống kê của toàn bộ cảnh nền. Dò tìm đối tượng là xác định những pixel trong ữame video không tương ứng với mô hình nền, và đưa ra mặt nạ nhị phân tương ứng với các đối tượng. Cuối cùng, họp lệ dữ liệu khảo sát mặt nạ thích hợp này, loại bỏ những pixel không tương ứng với những đối tượng chuyển động thật, và đưa ra mặt nạ cuối cùng.

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image10.png Delay

Preprocessing

Background

Modeling

Foreground

Detection

Data

Validation

Foreground

Masks

Background Subtraction

Hình 2.1: Biểu đồ luồng chung của các giải thuật trừ nền

1. Tiền xử lí

Trong đa số hệ thống thị giác máy tính, những bộ lọc đom giản thường được sử dụng trong trong gian đoạn đầu của xử lý để giảm bớt nhiễu camera. Những bộ lọc có thể được sử dụng để loại bỏ những nhiễu môi trường nhất thời như mưa và tuyết... Với hệ thống thời gian thực, giảm bớt kích thước Ếrame và tốc độ frame thường được dùng để làm giảm tốc độ xử lý dữ liệu. Nếu camera chuyển động hoặc nhiều camera được sử dụng tại các vị trí khác nhau, đăng kí ảnh giữa các frame liên tục và những camera khác nhau là cần thiết trước khi mô hình hóa nền. Một vấn đề khác trong tiền xử lý là định dạng dữ liệu được sử dụng bởi các giải thuật trừ nền khác nhau. Hầu hết các giải thuật đều xử lý cường độ chiếu sáng. Mặc dù vậy, ảnh màu, trong không gian màu RGB hoặc HSV, cũng được sử dụng khá phổ biến trong trừ nền [6,7]. Người ta cũng đã chỉ ra rằng khi xác định đối tượng trong vùng độ tưomg phản thấp và khử bóng sinh ra do đối tượng di chuyển thì sử dụng màu sắc tốt hom sử dụng độ sáng. Ngoài ra, những đặc trưng ảnh dựa trên pixel cũng thường được sử dụng để kết họp các biên và thông tin chuyển động. Ví dụ, các giá trị cường độ và dẫn xuất không gian có thể kết họp để hình thành một không gian trạng thái đom cho kiểm tra nền với bộ lọc Kalman. Trường họp khác kết họp cả dẫn xuất không gian và thời gian để tạo thành mô hình nền vận tốc đều cho phát hiện tốc độ xe. Mặt hạn chế chính của phưomg pháp này là làm tăng sự phức tạp cho ước lượng tham số mẫu. Độ phức tạp thường tăng đáng kể khi đa số các kỹ thuật mô hình hóa nền sử dụng mẫu độc lập cho mỗi pixel.

1. Mô hình hóa nền

Mô hình hóa nền là vấn đề trọng tâm của bất kỳ giải thuật trừ nền nào. Nhiều nghiên cứu đã tập trung để phát triển mô hình hóa nền linh hoạt chống lại sự thay đổi của môi trường, nhưng đủ nhạy để xác định mọi chuyển động

Cần quan tâm. Các kĩ thuật mô hình hóa nền được phân thành 2 loại - không đệ quy và đệ quy.

a) Kỹ thuật không đệ quy

Kỹ thuật không đệ quy sử dụng phương pháp cửa sổ trượt để ước lượng nền. Nó lưu trữ một bộ đệm cho L frame video trước và ước lượng ảnh nền dựa vào sự biến thiên thời gian của mỗi pixel trong bộ đệm. Kỹ thuật không đệ quy có khả năng thích ứng cao như là chúng không phụ thuộc vào thời điểm những frame lưu trữ trong bộ đệm. Mặt khác, yêu cầu vùng lưu trữ có thể lớn nếu bộ đệm lớn cần thiết để đối phó với những chuyển động chậm của giao thông. Cho một bộ đệm với kích thước cố định, vấn đề này có thể giảm bớt từng phần bằng việc lưu trữ các frame video tại một tốc độ frame chậm r. Dưới đây là một số kỹ thuật không đệ quy thường được sử dụng:

* Sự khác biệt khung

Đây là kỹ thuật mô hình hoá nền được coi là đơn giản nhất, nó sử dụng Ếrame video tại thời điểm t -1 làm mô hình nền cho khung tại thời điểm t. Khi nó chỉ sử dụng trame đơn phía trước thì sự phân biệt khung có thể sẽ không xác định được những pixel bên trong một phạm vi lớn mà đối tượng chuyển động có màu đồng dạng. Đây chính là một vấn đề hạn chế.

* Lọc trung bình

Lọc trung bình là một trong những kỹ thuật mô hình hoá nền hay được sử dụng nhất. Ước lượng nền được định nghĩa là trung bình tại mỗi vị trí pixel của tất cả các frame trong bộ đệm. Độ phức tạp của tính toán số trung bình là 0(L log L) cho mỗi pixel.

* Bộ lọc ước đoán tuyến tính

Toyama tính ước lượng nền hiện thời bằng cách áp dụng bộ lọc ước đoán tuyến tính trên các pixel trong bộ đệm. Hệ số lọc được ước lượng tại mỗi thời điểm của khung dựa vào hiệp phưomg sai mẫu, do đó kỹ thuật này khó áp dụng trong xử lý thời gian thực.

* Mô hình không có tham số

Khác với các kỹ thuật trước sử dụng ước lượng nền đon cho mỗi vị trí pixel, Elgammal đã sử dụng toàn bộ It\_L,It\_L+1,...,It\_ị để tạo một ước lượng không có tham số của hàm mật độ pixel f(It = u):

/(/,=«) = ị !>(»-/,) (2.1)

L *i=t-L*

K(.) là hàm ước lượng, pixel hiện thời It được coi như không phải nền nếu nó không giống như sự phân bố này, nghĩa là f(It) nhỏ hon so với ngưỡng giới hạn. Ưu điểm của việc sử dụng hàm mật độ đầy đủ cho ước lượng đon là khả năng quản lý phân bố nền phức tạp. Những ví dụ của nền phức tạp bao gồm các điểm từ cây đung đưa hoặc các biên có độ tưong phản cao mà chúng bị nhòe do sự chuyển động của camera.

b) Các kỹ thuật đệ quy

Kỹ thuật đệ quy không thể duy trì bộ đệm cho việc ước lượng nền. Thay vào đó, người ta sử dụng một phưong pháp đệ quy mới, đó là cập nhật các mô hình nền đơn dựa vào các írame đầu vào. So với các phương pháp không có đệ quy thì phương pháp sử dụng đệ quy cần bộ lưu trữ ít hơn rất nhiều, tuy vậy một số lỗi trong mô hình nền vẫn còn tồn tại trong bất kỳ thời điểm nào.

* Lọc trung bình xấp xỉ

Đổ có được sự thành công của lọc trung bình không đệ quy, McFarlane và Schotĩeld đề xuất một phương pháp lọc đệ quy đơn giản. Phương pháp này đã được sử dụng trong mô hình mô hình hoá nền để theo dõi giao thông đô thị. Trong mô hình này, giá trị trung bình sẽ được tăng một đơn vị nếu điểm ảnh đầu vào lớn hơn ước lượng ban đầu, và giảm đi một đơn vị nếu cácđiểm ảnh này nhỏ hơn. Giá trị này sẽ dàn hội tụ tại một ngưỡng gọi là trung điểm được ngăn cách bởi một bên là những điểm ảnh có giá trị lớn, và một bên là những điểm ảnh có giá trị nhỏ.

> LọcKalman

Đây là bộ lọc phổ biến sử dụng kỹ thuật đệ quy cho những hệ thống theo dõi chuyển động dưới nhiễu Gaussian. Có rất nhiều phiên bản đã được đề xuất cho mô hình hoá nền, sự khác biệt chính trong không gian trạng thái được sử dụng cho việc theo dõi. Phiên bản đơn giản nhất chỉ sử sụng cường độ chiếu sáng. Karmann và Brandt sử dựng cả cường độ và trạng thái thời gian, trong khi Koller, Weber và Mailk lại sử dụng cường độ và trạng thái không gian. Phương trình sau mô tả trạng thái của hệ thống, với Bt là cường độ, và Bt’ là trạng thái thời gian:

*Bt*

***b:***

*= A*

*Bt-1*

Ki.

+ Kt (It - H ‘A

*Bt-1*

K1.

(**2**.**2**)

Ma trận A mô tả các nền động, và H là phép đo ma trận:

A =

10.7

0 0.7] ’

H = [l **0]**

(**2**.**3**)

Kalman đã khuếch đại ma trận Kt tạo những biến đổi nằm giữa hai mức: thích nghi chậm (Xi và thích nghi nhanh 0,2.

Kt =

a,

nếu It-1 là đối tượng, và

*a.*

*a.*

ngược lại

(**2**.**4**)

> Mixture of Gaussians(MoG)

Không giống như bộ lọc Kalman là chỉ theo dõi sự biến đổi của một hàm Gaussian, phương pháp MoG đồng thời theo dõi sự biến đổi của nhiều hàm Gaussian. Tương tự như mô hình không có tham số được trình bày trong mục a), MoG duy trì một hàm mật độ cho mỗi điểm ảnh. Như vậy, nó có khả năngxử lý tốt trên những mô hình có nền phức tạp. Mặt khác, MoG là một tham số, những mô hình tham số có khả năng thích nghi và cập nhật nhưng lại không thể lưu trữ những vùng đệm lớn của video. Hàm f(It = u) là sự pha trộn của K Gaussian:

*K*

/(/, = u) = YJC0ÍJ-r){u-,ạil,ơit)

(2.5)

i=\

Ở đây rj(u;juinơịt)là thành phần Gaussian thứ i với cường độ trung bình Mi t ĩ ị- M?■ơ-và độ lệch chuẩn ơit, Cữit là phàn dữ liệu được

tính cho thành phàn thứ i, phạm vi của K là từ 3 đến 5 phụ thuộc các giá trị lưu trữ. Với mỗi pixel đầu vào It bước đầu tiên là xác định thành phần ĩ. Thành phần ĩ được mô tả như là thành phần được đối sánh nếu

<D • ƠỊ

*i,t*-1

Ở đây D là ngưỡng với giá trị xác định. Những tham số của các thành phàn được đối sánh được cập nhật lại như sau:

= (1 - «)<»=,,\_, + a Mh = (1 - /»)/<?,\_, + pỉ,

(**2**.**6**)

**<,=(!-**-Pt,,**)2**

Ở đây a là tỉ lệ nằm bong khoảng 0 < a < 1 do người sử dụng định nghĩa, p bằng xấp xỉ tỉ lệ của những tham số:

a

Cờ,

(2.7)

i,t

Nếu không tìm thấy thành phàn tương ứng, thành phàn có trọng số nhỏ nhất sẽ được thay thế bởi một thành phần mới với giá trị trung bình lị, độ biến

thiên ban đầu lớn ỡ"o và trọng số nhỏ ữ>0 .

Phần còn lại có giá trị trung bình và độ biến thiên tương tự nhưng trọng số của nó thấp hơn được tính theo:

- (1 a)(ởi,t-l (2.8)

Cuối cùng tất cả các trọng số đều được chuẩn hóa lại với tổng là 1. Đe xác định It là một pixel đối tượng trước tiên chúng ta phải xếp hạng tất cả các

thành phần bởi các giá trị ®i// ơ i t của chúng. Thành phần có hạng cao hơn sẽ có độ biến thiên thấp và tỉ lệ xác suất cao, đây chính là đặc trưng cơ

bản của nền. Nếu là thứ tự các thành phần sau khi sắp xếp, thì M

thành phàn đầu tiên thoả mãn các điều kiện được khai báo là những thành phàn nền.

*i M*

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image11.png

k = ỉ'j

*a>k,t \** r

**(2.9)**

Ở đây r là trọng số ngưỡng, It là những điểm khác nền.

1. Phát hiện đối tượng

Phát hiện đối tượng là việc so sánh khung video hiện thời với mô hình nền và xác định các pixel đối tượng từ trame đầu vào đó. Ngoại trừ mô hình không có tham biến và mô hình MoG, còn lại tất cả các kỹ thuật được giới thiệu ở mục 2.1.2.2 đều sử dụng một ảnh đơn làm mô hình nền cho chúng. Những phương pháp phổ biến nhất được sử dụng cho phát hiện đối tượng là kiểm tra sự khác biệt của các pixel đầu so với ước lượng nền tương ứng.

I,(x,y)- Bt(x,y)>T

(2.10)

Một cách khác hay được sử dụng là phân ngưỡng dựa trên các tiêu chuẩn thống kê:

(2.11)

Trong đó Pd và ơd là giá trị trung bình và độ lệch tiêu chuẩn của It (x, y) - Bt (x, y) cho mọi vị trí không gian (x, y). Việc xác định T và Ts là dựa vào thực nghiệm.

Một hướng tiếp cận khác là sử dụng độ khác nhau tưcmg đối hem là độ khác nhau tuyệt đối để nổi bật độ tưcmg phản trong những vùng tối như là đổ bóng

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image12.png(**2**.**12**)

Tuy nhiên, kĩ thuật này không thể sử dụng để tăng độ tương phản trong các ảnh sáng như cảnh ngoài trời dưới sưcmg mù dày đặc. Một hướng tiếp cận khác để đưa ra độ biến thiên không gian là sử dụng hai ngưỡng với hiện tượng trễ. Ý tưởng cơ bản là ban đầu xác định các pixel không thuộc nền mà chúng khác biệt tuyệt đối với các ước lượng nền vượt quá một ngưỡng. Sau đó vùng đối tượng được phát triển từ các pixel đã được xác định, có thể được thực hiện bởi việc sử dụng giải thuật kết nối các nhóm thành phần.

1. Họp lí hoá dữ liệu

Chúng ta định nghĩa hợp lí hoá dữ liệu là quá trình xử lí, cải tiến mặt nạ xác định đối tượng dựa vào thông tin thu được từ bên ngoài mô hình nền. Tất cả các mô hình nền trong phàn 2.1.2.2 có ba hạn chế chính: thứ nhất, chúng bỏ qua tất cả các mối liên quan giữa các pixel lân cận; thứ hai, tốc độ thích nghi có thể không phù họp với tốc độ chuyển động của các đối tượng; thứ ba, những pixel động do đung đưa của các lá cây, hoặc đổ bóng do các đối tượng chuyển động dễ gây hiểu lầm là đối tượng thực sự. vấn đề điển hình đầu tiên

dẫn đến khẳng định sai hoặc phủ định sai các vùng phân bồ ngẫu nhiên ngang qua mặt nạ ứng cử. Hướng tiếp cận chung để giải quyết vấn đề đầu tiên là kết hợp lọc hỉnh thái và nhóm thành phàn liên kết để loại bỏ những vùng ngẫu nhiên do hạn chế thứ nhất sinh ra. Áp dụng bộ lọc hình tháỉ trên các mặt nạ đối tượng loại bỏ những pixel đối tượng đứng riêng biệt và hợp nhất các vùng liền kề. Nhóm thành phần kết nối có thể sử dụng để xác định tất cả các vùng đối tượng, và loại bỏ những vùng quá nhỏ để tương ứng với các đối tượng chuyển động thực. Một hướng tiếp cận đơn giản để làm gỉảm bớt các vấn đề trên là sử dụng nhiều các mô hình nền chạy tại các tốc độ thích ứng khác nhau, ngoài ra phân đoạn màu có thể được sử dụng để phát triển các vùng đối tượng bằng giả thuyết tổng hợp các màu tương tự khắp toàn bộ đối tượng... Nếu nhiều camera cùng thu được những cảnh tương tự tại các góc khác nhau, thông tin chênh lệch giữa các camera có thể sử dụng để ước lượng chiều sâu. Thông tin chiều sâu rất hữu ích khi đối tượng không phải là nền. vấn đề cuối cùng có thể giải quyết bởi việc sử dụng những lã thuật mô hình hoá nền phức tạp như MoG và áp dụng bộ lọc hình thái để làm sạch dữ liệu. Mặt khác, khử bóng chuyển động là vấn đề khó giải quyết, đặc biệt với video có độ chiếu sảng cao. Những nghiên cứu và so sánh của các giải thuật khử bóng được trình bày trong [7].

1. b) c) d)

Hình 2.2: Khung mẫu và các nền tương ứng từ bến cảnh: a)Cảnh sáng,

1. Cảnh có sương mù, c)Cảnh có tuyết, d)Cảnh có xe cộ đông đúc.
2. Trừ nền sử dụng thông tin màu và gradỉant
3. Giới thiệu

Nhiều phương pháp đã được phát triển để phát hiện chuyển động trong một chuỗi các ảnh. Các giải thuật này sử dụng trừ nền liên quan đến những sự thay đổi màu hay độ chiếu sáng. Nó được thực hiện bằng cách so sánh màu hoặc cường độ các pixel của ảnh đầu vào với ảnh tham chiếu, nếu có sự khác nhau đáng kể so với ảnh tham chiếu thì có nghĩa là có sự chuyển động. Phương pháp này có thể cho kết quả sai lệch khi mà có các điều kiện không lí tưởng. Một phương pháp trừ nền khác sử dụng thông tin màu và gradient để nâng cao hiệu quả phát hiện, được đề xuất bởi Jabri sau đó được thay đổi bởi Javed, Shatlque và Shah. Hướng tiếp cận này là xây dựng mô hình nền sử dụng thông tin màu và thông tin gradient, sau đó thực hiện trừ nền dùng các mô hình đó. Mô hình này luôn được cập nhật để thích ứng với những thay đổi chậm trong chiếu sáng.

1. Mô **h**ì**nh** hóa nền

Mô hình nền được xây dựng từ hai phàn, mô hình màu và mô hình gradient. Mô hình màu được xây dựng cho mỗi một kênh màu. Nó bao gồm hai ảnh đại diện cho giá trị trung bình và độ lệch tiêu chuẩn cho thành phàn màu đó. Mỗi pixel trong ảnh trung bình được tính toán sử dụng.

**ut = axt + (l - a** )wí\_1 (2.13)

Trong đó Ut là trung bình được tính trên frame t, a là tốc độ học của mô hình, và xt là cường độ thành phần màu trong frame t. Trừ ảnh đầu vào với ảnh trung bình sẽ xác định những pixel bị thay đổi cường độ. Trong quá trình trừ nền, ảnh độ lệch chuẩn ơt được sử dụng để chuẩn hóa ảnh tin cậy và nó được tính như sau:

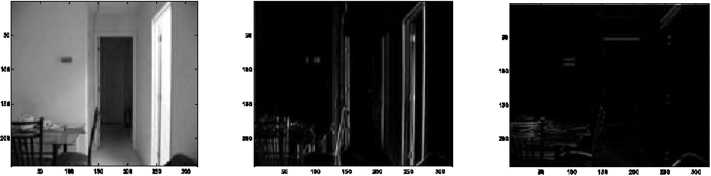
ơỉ = a(xt - ut Ỵ + (l - aRi (2.14)

Mô hình hiên bao gầm 2 ảnh trung bình và 2 ảnh độ lệch chuẩn, được tính bằng cách áp dụng dò tìm biên sobel phưomg ngang và thẳng đứng với ảnh mức xám. Kêt quả gradỉent phương ngang là H, gradient phương thẳng đứng là V. Ảnh trung bình được tính như sau:

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image13.jpeg(2.15)

Trong đó p là tốc độ học của mô hình. Cảc ảnh lệch chuẩn ơH,t và ơv,t được tính tương tự như mô hình màu. Mô hình biên được dùng để xảc định những thay đểỉ trong cấu trúc một ảnh.

Các ảnh độ lệch tiêu chuẩn được tính toản chỉ với chuỗi các ảnh nền tình trong khỉ những ảnh trung bình được cập nhật liên tục. Điều này cho phép mô hình nền được đỉều chỉnh với những thay đểỉ dần dần trong độ chiểu sáng.



Hình 2.3: Ảnh trung bình cho kênh màu đỏ, biên ngang, và biên thẳng

đứng tại ỉrame thứ 25.

1. Trừ nền

Trừ nền được thực hiện bằng cách thực hiện trừ dựa vào màu và trừ dựa vào biên riêng biệt sau đó kết hợp các kết quả với nhau. Hình 2.4 biểu dỉẫn các ảnh mà phép trừ sẽ được thực hiện

Hình 2.4: Khung 65 và 70 từ chuỗi 96 khung.

150

IM

200

**Mí**

«0 IM 1S0 Mù 250 Mb

M IM IM MO 2S0 MO

1. Trừ nền dựa vào màu

Phép trừ dựa vào màu được thực hiện bằng cảch trừ ảnh hiện thời với ảnh trung bình trong mẫỉ một kênh màu. Kêt quả của phép trừ này là đưa ra 3 ảnh khác nhau được sử dụng để tạo nên 3 ảnh tin cậy được chuẩn hóa. Điều này được thực hiện bằng cách so sánh sự khảc biệt vởỉ 2 ngưỡng nicơ và Mcơ, được lấy từ các ảnh độ lệch chuần. Với mỗi pixel, độ tin cậy được tỉnh toán như sau:

0

CM=<

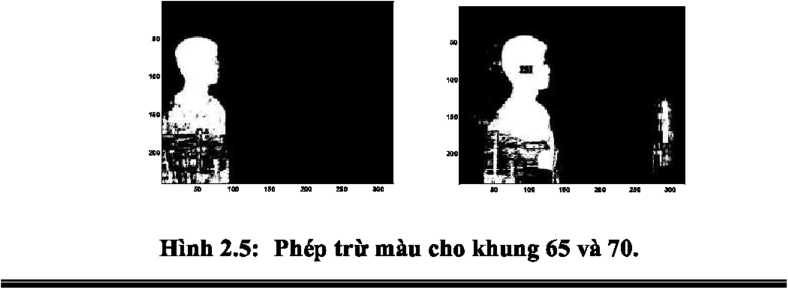
*Đ — m„ơ*

*Mơ - m\_a* 100

xioo

D < mcơ mcơ <D< Mcơ D > Mcơ

(2.16)

Khỉ có một sự thay đểỉ lớn trong một vàỉ màu thì đó là miền có chứa đối tượng. Môi một ảnh tin cậy Cc có thể được tạo bằng cách lấy giả trị tin cậy lớn nhất cho mẫỉ pixel. Hình 2.5 biểu diễn những ảnh tin cậy cho ữarne 65 và 70.

1. Trừ nền dựa vào biên

Phép trừ dựa vào biên được thực hiện bằng cách lấy các ảnh phương ngang hiện thời trừ đi ảnh trung bình Ht và lấy các ảnh phương thẳng đứng hiện thời trừ đi ảnh trung bình vt:

*AH = \H-Ht\, AV = \V-Vt\* (2.17)

Biên gradient của ảnh là:

AG = AH + AV(2.18)

Ảnh tin cậy được tính toán bằng việc nhân AG với một hệ số nhân tin cậy R và so sánh két quả với 2 ngưỡng meơ và Meơ. Ở đây, ơ là tổng độ lệch chuẩn phương ngang và độ lệch chuẩn phương thẳng đứng. Với mỗi pixel, ta giả sử:

G = ịỉỉị + ịvị, Gt = \Ht I + ịvt |, G\* = max{ơ, Gt} (2.19)

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image16.pngSau đó

*R =*

(**2**.**20**)

Độ tin cậy của mỗi pixel được tính toán như sau:

CB

(2.21)

#### 'o

*RÁG - meơ*

#### < — xioo

*Meơ-meơ*

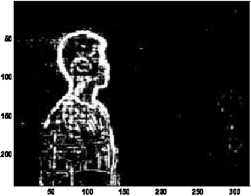
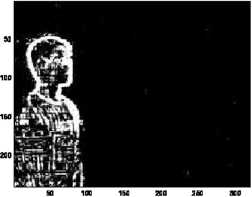
100

*RÁG < mơ*

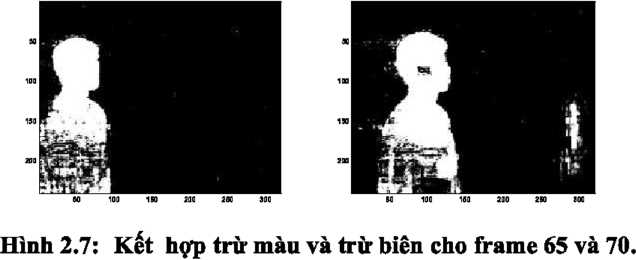
*e*

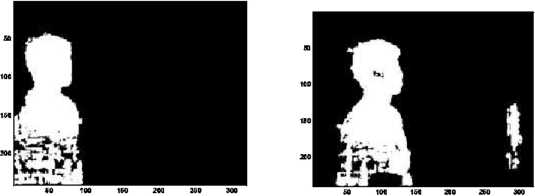
*meơ < RAG < MecT RAG > Meơ*

Biên có thể được phân loại tại điểm này. Nếu có một sự khác nhau lớn và có một biên quan trọng trong ảnh hiện thời, thì biên xem xét đang được khép kín. Nếu có một sự khác nhau lớn và không có một biên quan trọng trong ảnh hiện thời, thì nó đã được khép kín.

Hình 2.6: Phép trừ biên cho frame 65 và 70.

1. Kết họp các kết quả trừ màu và trừ bỉên

Việc trừ màu và trừ biên được kết hợp bằng cảch lấy giả trị lớn nhất giữa hãi ảnh tin cậy tại mỗi pixel. Kết quả kết hợp được thể hiện ở hình 2.7. Nhỉễu được loại bỏ bằng cách sử dụng bộ lọc trung bình. Những vùng không được nối 100% tới miền tin cậy, được coi là những đại lượng dương sai và bị loại bỏ bởi ngưỡng trễ. Hình 2.7 biểu diễn đốỉ tượng màu trắng và nền màu đen trong hai ữame. Với phương pháp này, đối tượng không thuộc nền cuếỉ cùng được định nghĩa tất hơn hai phương pháp trừ biền và trừ màu. Trong ữarne 70, người ta cho một cái bóng lên tường và cái bóng đổ được xác định như là một đối tượng không thuộc nền khác. Đó là một nhận định saỉ.



Hình 2.8: Kết quả kết hợp sau khi áp dụng bộ lọc trung bình và ngưỡng trễ cho frame 60 và 70.

1. Nhận xét

Sử dụng thông tin màu và thông tin gradient để thực hiện trừ nền đã xảc định được đếỉ tượng không thuộc nền một cách tết hơn. Gỉảỉ thuật vẫn thực hiện tốt khi cổ nhiễu và những thay đổi nhỏ về cường độ chiếu sáng. Tuy nhiên, do giải thuật này dựa trên pỉxel nên sự thay đểỉ lớn trong màu hoặc biên sẽ ảnh hưởng đến đối tượng. Như vậy nó không thành công khỉ có sự thay đổi đột ngột trong chiếu sáng.

23. Phát hiện đếỉ tượng chuyển động dựa vào kết cấu

1. Giới thiệu

Trừ nền thường là các thao tác đầu tiên trong cảc ứng dụng thị giác máy, là một phần then chốt của hệ thống. Kết quả đầu ra của trừ nền là đầu vào của công việc xử lí ở mức cao hơn cố thể thực hiện, ví dụ như: theo dõi đối tượng. Sự thực thỉ của trừ nền phụ thuộc chủ yếu vào kỹ thuật mô hình hóa nền được sử dụng. Đặc biệt, các cảnh thiên nhỉên đặt ra nhiều thách thức cho mô hình hóa nền do các cảnh này trong tự nhiên thường có những thay đẩi độ sảng, cây cỏ đung đưa, nước gợn sóng.. .Một giải thuật trừ nền tốt là phải xử lí được những vị trí mà tại đó xuất hiện một đối tượng tĩnh hoặc là một đối tượng cũ dỉ chuyển từ cảnh. Hơn nữa, bóng của chuyển động và cảc đối tượng cảnh có thể là nguyên nhân của vấn đề. Thậm chí trong một cảnh tĩnh nhưng vẫn có sự thay đểỉ giữa hai frame do nhỉễu và sự di chuyển của camera. Ngoài ra, giảithuật mô hình hóa nền phải thực thi trong thời gian thực. Trong phàn này, một hướng tiếp cận mới dựa vào kết cấu cho trừ nền sẽ được giới thiệu. Mục đích của phương pháp mới này giải quyết tất cả các khó khăn ở trên ngoại trừ việc xử lí bóng vì đây là vấn đề rất khó trong mô hình hóa nền.

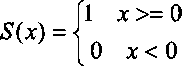
1. Toán tử mẫu nhị phân cục bộ

Phương pháp dựa vào kết cấu được đề xuất cho trừ nền dựa vào độ đo kết cấu LBP (Local Binary Pattem). LBP là phương tiện hữu hiệu cho mô tả kết cấu. Toán tử gán nhãn các pixel của khối ảnh bằng tạo ngưỡng cho lân cận của mỗi pixel với giá trị trung tâm và nhận kết quả là số nhị phân (mã LBP):

LBP(xc,yc) = ỵ^S(gp- gc)\*2p **(2.22)**

*p=0*

Trong đó gc tương ứng với giá trị mức xám của điểm ảnh trung tâm (xc,yc), và g p là các giá trị mức xám của p điểm ảnh lân cận. Hàm s(x) được định nghĩa:

(2.23)

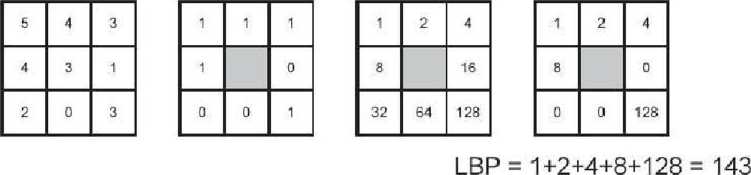
Toán tử LBP gốc được thực hiện với 3\*3 lân cận của một điểm ảnh. Hình 2.9 là một minh họa cho toán tử. Dạng tổng quát của toán tử LBP sử dụng các lân cận đối xứng thành vòng tròn như trong hình 2.10. Trong trường hợp này gp tương ứng vói các giá trị mức xám của p pixel trong vòng tròn với bán kính R. Biểu đồ của mã LBP tính toán trên các khối ảnh và có thể sử dụng để mô tả kết cấu cho khối. Từ (2.22) có thể thấy, LBP bất biến với các thay đổi đơn trong mức xám.

1. Phương pháp phát hỉện đối tượng chuyển động

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image23.jpegMultiply

Threshold

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image22.jpeg



Hình 2.9: Ví dụ cho tính toán mã LBP gếc

Mục đích chính của phương pháp là xác định những vùng củã khung video có chứa đối tượng chuyển động. Phương pháp này chia khung video thành các khối có kích thước đều nhau bằng cách sử dụng cấu trúc lưới chông cục bộ (hình 2.11). Bằng việc sử dụng các khối chằng lấp cục bộ chúng ta có thề trích chọn hình dáng đối tượng chuyền động một cách chính xảc hơn trường hợp sử dụng các khối không chồng lấp. Chúng ta coi những vectơ đặc tính của khấỉ ảnh cục bộ theo thời gian như một qui trình xử lý khối. Khỉ sử dụng biếu đồ LBP là vectơ đặc trưng, thì xử lý khôi được định nghĩa là chuỗi thời gian của các biểu đồ LBP. Khi đó biểu đồ khối tại thời điểm t sẽ được biểu diễn là Xt

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | c |  | ơ | J | k |  |  |
|  |  |  |  |  |  |  |  |  |  |  | < |  | o |  |  |  |  |  | r |  | y |  | V. | í1 |  |  |
|  |  |  |  | o |  |  |  |  |  |  | c | p |  | c | ) |  |  |  |  | ) |  |  |  |  | J |  |
|  |  |  | o | • | o |  |  |  |  |  | o |  | m |  | o |  |  |  | o |  |  | • |  |  | o |  |
|  |  |  |  | 0 |  |  |  |  |  |  | c | s |  | c | ) |  |  |  | r |  |  |  |  | ị |  |  |
|  |  |  |  |  |  |  |  |  |  |  | V |  | ỡ |  |  |  |  |  |  | ) |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | k | ? | ô |  | r |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

p - 4. R. — 1 P-8. R-2 p = 12 R = 3

Hình 2.10: Lân cận cân đối vòng tròn được đặt vói những giá trị khác

nhau của p và R

Như đã đề cập trong các mục trước, LBP bất biến với những thay đổi đơn trong mức xám, vì vậy nó ít chịu sự ảnh hường do thay đổi của ánh sáng. Do biểu đồ LBP không bao gồm các thông tin về các vị trí các mã LBP riêng được tính toán, nên nó hỗ ừợ mô hình hóa nền đa mức. LBP tính toán rất nhanh, đây là đặc tính quan trọng dưới góc độ thi hành cục bộ và đặc biệt trong trường hợp này sù dụng nhiều biểu đồ để mô hình mỗi khối.

Mỗi khối {xi X2,...,Xt}, được mô hình hóa bởi tập hợp K trọng số các biểu đồ LBP. Các bước thực hiện cập nhật mô hình nền cho một khối:

Bước 1: Sử dụng một độ đo khoảng cách để so sánh biểu đồ khối mới Xt với K biểu đồ mô hình đã có. Trong trường hợp này có thể sử dụng độ đo biểu đồ giao nhau để đo khoảng cách. Biểu đồ giao nhau của hai biểu đồ đã được chuẩn hóa Xi và x2 được định nghĩa như sau:

C:\Users\Dung.VT\AppData\Local\Temp\FineReader12.00\media\image25.png(2.24)

Trong đó i là chỉ số cột của biểu đồ. Người dùng định nghĩa giá trị ngưỡng TD cho sự giao nhau biểu đồ là một tham số phương pháp. Chú ý rằng nó có thể sử dụng với các độ đo khác như Chi-square hoặc Log- likelihood.

Hình 2.11: Giải thuật sử dụng cấu trúc lưới chềng cục bộ

Nếu không có cái nào trong số các biểu đồ mô hình trùng khớp với biểu đồ mới, thì biểu đồ mô hình với trọng số thấp nhất sẽ được thay thế bởi biểu đà mới và biểu đồ này sẽ được coi là có giá trị trọng số ban đầu thấp. Sau đó,các trọng số sẽ được chuẩn hóa để chúng có tổng bằng một. Nếu có một biểu đồ mô hình trùng khớp với biểu đồ mới được tìm thấy, thì số cột của biểu đồ sẽ được cập nhật như sau:

*Xk,tƯÌ =abxtụ\ +(l-ab)Xkt\_l\i] ,ab* e[0,l] (2.25)

Trong đó ab là tốc độ lấy mẫu được người dùng định nghĩa. Ngoài ra, các trọng số Wkt được cập nhật như sau:

&k,t = ơ-O^í-1 +awMkít,aw e [0,1] (2.26)

Với là tốc lấy mẫu được người dùng định nghĩa và Mkjt =1 với biểu đồ phù hợp và = 0 các trường hợp khác.

Bước 2: càn quyết định những biểu đồ nào của mô hình được tạo ra thích hợp nhất bằng xử lí nền. M. Heikkila, M. Pietikainen and J. Heikkila sử dụng “độ lưu” của biểu đồ như là một bằng chứng cho điều này. Bởi vì “độ lưu ” của biểu đồ thứ k có liên hệ trực tiếp với trọng số Wk,t, những biểu đồ này được phân loại theo độ giảm của trọng số. Kết quả là đa số các biểu đồ nền đều ở trên đỉnh của danh sách.

Bước 3: B biểu đồ đầu tiên được lựa chọn trở thành mô hình nền như

sau:

+ + ^B.t >TBĩTb e [0,1] (2.27)

Trong đó Tb là ngưỡng lựa chọn do người dùng định nghĩa. Nếu mô hình hóa một nền, thì chỉ cần sử dụng một giá trị nhỏ TB. Trong trường hợp đa mô hình nền, thì một giá trị TB lớn được sử dụng. Giá trị nhỏ TB lựa chọn một biểu đồ chắc chắn nhất làm mô hình nền, ngược lại giá trị lớn cho phép biểu đồ xt lấy nhiều cấu hình cột từ mô hình nền bao gồm nhiều biểu đồ.

Phát hiện đối tượng không phải là nền có được do sự so sánh biểu đồ mới xt với B biểu đồ nền được lựa chọn tại thời điểm tức thời trước. Nếu không có giá trị thích hợp, khối đó sẽ thuộc về đối tượng. Ngược lại, khối đó được đánh dấu là nền.

1. Phát hiện đối tượng chuyển động dựa vào sự biến thiên cục bô của véctơ kết cấu SP
2. Giới thiệu

Chúng ta hãy tập trung vào một vị trí cố định trong một ảnh video và quan sát chuỗi các vectơ trực quan ( ví dụ: mức xám, tia hồng ngoại, màu hay các vectơ kết cấu) tại mỗi vị trí này trong một video. Giả thiết rằng, camera không chuyển động. Nếu quan sát nền cảnh tại vị trí này, thì ta sẽ chỉ thấy sự thay đổi rất nhỏ của các vectơ trực quan do có những thay đổi nhỏ về ánh sáng và các lỗi của thiết bị thu video. Mặt khác, nếu có một đối tượng chuyển động qua vùng này, thì chúng ta sẽ thấy sự khác nhau giữa các vùng của đối tượng, những vùng này sẽ cho những kết cấu khác nhau. Bởi vậy, tính kết cấu tại một vị trí đã cho rất phù hợp để phát hiện những thay đổi lớn.

Nếu tại vị trí đang quan sát trong video chỉ xem xét giá trị của một điểm thì sẽ không đủ để có thể phát hiện có sự chuyển động. Chẳng hạn, xét một đối tượng màu trắng di chuyển qua vị trí đang quan sát. Tại đường biên của đối tượng, sự bién đổi sẽ rất cao, nhưng sự biến đổi này có thể rất thấp ở trong thân đối tượng. Như vậy, bằng cách phát hiện sự biến đổi cao, chúng ta sẽ xác định đường biên của đối tượng chứ không phải một phần bên trong của đối tượng.

Trong phàn này, thay vì chỉ xem xét các giá trị màu sắc hoặc các giá trị hồng ngoại tại một điểm, chúng ta xem xét tất cả các điểm trong vùng spatiotemporal được biểu diễn là khối 3D. Đe mô tả cho các giá trị và giảm nhiễu, Ronand Mezianko và Jan Latecki đề xuất một kết cấu spatiotemporal mô tả cho những khối 3D. Mô tả kết cấu này là đầu vào cho các kỹ thuật phát hiện chuyển động dựa vào sự biến đổi cục bộ.

Trước tiên đoạn video được chia thành các khối spatiotemporal 3D (chẳng hạn: 8x8x3) và áp dụng một kỹ thuật giảm số chiều để thu được biểu diễn cô đọng về màu sắc, hồng ngoại hay giá trị mức xám ở mỗi khối. Vectơ kết cấu SP thu được cung cấp một biểu diễn mối quan hệ giữa kết cấu và mẫu chuyển động trong video và được sử dụng là dữ liệu đầu vào cho giải thuật phân tích video.

Ưu điểm của công nghệ giảm số chiều cho việc biểu diễn cô đọng đã được thừa nhận trong nén video. Ở đây, biến đổi cosine rời rạc 3D, và những biến đổi wavelet 3D cũng được sử dụng để rút gọn những giá trị màu hoặc giá trị mức xám của phần lớn các điểm ảnh trong khối cho một vài vectơ đã lượng tử hóa. Tuy nhiên, kỹ thuật này không hoàn toàn phù hợp để phát hiện những đối tượng chuyển động, vì những thành phần thu được không cung cấp những đặc điểm cần thiết để phân biệt các khối. Ngược lại, phưong pháp được Ronand Mezianko và DragoỊịub Pokrajac [10] đề xuất ở đây có thể thu được sự khác nhau thích hợp nhất cho tập các khối 3D đã cho. Để làm được vấn đề này, cần có một sự tác động từ ngoài vào và một biến đổi cảm ngữ cảnh sao cho sự thể hiện của khối đã cho phụ thuộc vào ngữ cảnh của nó. Việc ứng dụng phưomg pháp phân tích thành phàn chính (Principal Component Analysis - PCA) đã đáp ứng được yêu cầu này.

Như đã trình bày ở trên, kết cấu tại một vị trí đã cho trên video rất thích hợp cho những thay đổi lớn khi mà đối tượng chuyển động qua vị trí quan sát. Bởi vậy cần phải có một phưomg pháp phát hiện chuyển động dựa trên cơ sở những biến đổi cục bộ của vectơ kết cấu SP. Ronand Mezianko và DragoỊjub Pokrajac xác định là cần đo trong một cửa sổ giới hạn và càng ngắn càng tốt. Họ định nghĩa biển đổi cục bộ là giá trị riêng lớn nhất của vectơ kết cấu SP trong một khoản thời gian nhỏ. Nó được tính toán bằng cách áp dụng phương pháp PCA vào ma trận hiệp phương sai của vectơ kết cấu SP trong một khoảng thời gian nhỏ. Như vậy trong cách tiếp cận này, ta sử dụng phương

pháp PCA hai lần, lần đầu dùng để tính toán vectơ kết cấu SP, và làn thứ hai dùng để tính toán sự biến đổi của vectơ kết cấu. Sau đó sử dụng một ngưỡng động để quyết định xem đây là đối tượng chuyển động hay là nền tĩnh tại vị trí không gian đã cho.

1. Biểu diễn vỉdeo vói vectơ kết cấu SP

Các vectơ được mô tả là những mảng 3D của mức xám hoặc các giá trị pixel hồng ngoại đon sắc gi,j,t tại thời điểm t, vị trí pixel (i,j). Một video được đặc trưng bởi chiều thời gian là z tương ứng với số khung, và hai chiều không gian đặc trưng số vectơ trong phương ngang và phương thẳng đứng của mỗi írame. Các ảnh trong video được chia thành chuỗi những hình vuông rời rạc Nblockx Nblock ( chẳng hạn hình vuông 8x8), những hình vuông này bao trùm toàn bộ ảnh. Những khối spatíotemporal 3D có được là do sự kết hợp của nhiều hình vuông liên tiếp trên các írame tại cùng một vị trí của video. Ví dụ sử dụng 8x8x3 khối rời rạc trong không gian nhưng chồng lấp về thời gian, khi đó, hai khối ở cùng một vị trí và tại thời điểm t và t+1 có chung một hình vuông. Thực tế là những khối 3D nối tiếp nhau cho phép chúng ta phát hiện được chuyển động trong video với tàn số thời gian thấp. Những khối 3D thu được được biểu diễn thành các vectơ 192 chiều của mức xám hoặc là những giá trị điểm ảnh hồng ngoại đơn sắc.

Các khối được biểu diễn bằng các vectơ N chiều bjj t với chỉ số không gian ( i,j )và tại thời điểm t. Các vectơ bij,t bao gồm các giá trị mức xám gij;t của các pixel trong khối 3D tương ứng. Như vậy, những khối đã cho với chỉ số không gian (i j) và tại thời điểm t, tương ứng với vectơ khối bij,t chứa đựng những giá trị điểm ảnh gi,i;t từ những tọa độ không gian:

(Nblock-1) X (I-1)+1,...,Nblockx I (Nblock-1) X (J-l)+l,..., Nblock X J và từ khung t-T...t+T. Với T là tổng số khung.

Như vậy chiều dài N của vectơ khối sẽ bằng Nblockx Nblockx (2T+1).

Đe rút gọn chiều của bij,t trong khi vẫn duy trì được thông tin trong phạm vi lớn nhất có thể, ta tính một phép chiếu của vetor khối được chuẩn hóa thành một vectơ có độ dài thấp hơn K (K«N) bởi việc sử dụng ma trận chiếu PCA PKij để tính cho tất cả bij t tại vị trí không gian (i,j) của video. Kết quả vectơ kết cấu không gian b\*i,j t= PyK cung cấp mối quan hệ giữa biểu diễn kết cấu và mô hình chuyển động trong video, và chúng được sử dụng là phần đầu vào của giải thuật phát hiện chuyển động và sự di chuyển của đối tượng.

Đe tính toán PKịj, các tác giả đã sử dụng các giá trị phân tích trong các tài liệu [11,12]. Ma trận của tất cả vectơ khối được chuẩn hóa bi,j,t trong vị trí (i,j) được sử dụng để tính ma trận hiệp phương sai NxN chiều Sij. Ma trận chiếu PCA Pịj cho vị trí không gian (i,j) được tính từ ma trận hiệp phương Si,j. Ma trận chiếu P(i,j) kích thước NxN biểu diễn cho N thành phàn chính. Bằng việc chỉ đưa ra những thành phần chính tương ứng với K giá trị riêng lớn nhất, ta thu được PKij.

1. Phát hiện đổi tưọng chuyển động dựa trên biến thiên cục bộ

Ý tưởng của phương pháp này là dựa vào sự biến đổi của vectơ cục bộ. Trong thực tế, với mỗi vị trí (x,y) chúng ta xét các vectơ:

VX,y,t-Ws Vx,y,t-w+lJ• ">yx,y,tj..., Vx,y,t+w

tương ứng với cửa sổ có kích thước 2W+1 xung quanh thời gian t, Ở đây, Vi,j t = b\*ij,t là vectơ kết cấu SP. Với những vectơ này, chúng ta tính ma trận hiệp phương sai CXy,t. Chúng ta gán giá trị tại vị trí video spatiotemporal đã cho bởi giá trị biến thiên cục bộ. Độ đo biến thiên này sẽ được gọi là độ đo chuyển động:

mm(x,y,t) = À,,, (2.28)

Trong đó Ajc yt là giá trị riêng lửn nhất của cxy!t. Khi có sự hiện diện

của một đối tượng chuyển động tại vị trí (x,y,t), giá trị độ đo mm(x,y,t) chuyển động sẽ lớn.

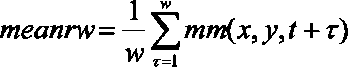
Cuối cùng, việc gán những vị trí trên video là chuyển động hoặc là nền sẽ phụ thuộc vào sự độ đo chuyển động là lớn hay nhỏ hơn so với ngưỡng chuyển động, ở đây các tác giả đã sử dụng giải thuật phân ngưỡng tự động để xác định giá trị ngưỡng tại vị trí (x,y,t) dựa vào các giá trị mm(x,y,s) với các giá trị thời gian (s=l,...,t-l).

Đầu tiên, tính giá trị trung bình meanl và độ lệch chuẩn stdl (sử dụng giá trị trung bình động) cho mọi mm(x,y,s) với s=l,...,t-l đã được gán nhãn ữnh. Một đối tượng chuyển động được phát hiện nếu :

meanrw-meanl>cỊ \* stdỉ

(2.29)

với Cị là hằng số và

(2.30)

Một chuyển động được phát hiện được chuyển về trạng thái tĩnh nếu:

meanrw-meanl<c*2* \* stdỉ

(2.31)

với C2<Ci là hằng số thứ hai.

Hình 2.12: Đồ thị của biến thiên cục bộ **mm** qua thởi gian

**CHƯƠNG 3:**

**THỬ NGHIỆM PHÁT HIỆN ĐỐI TƯỢNG ĐỘT NHẬP**

99999

* 1. **Mô tả bài toán**

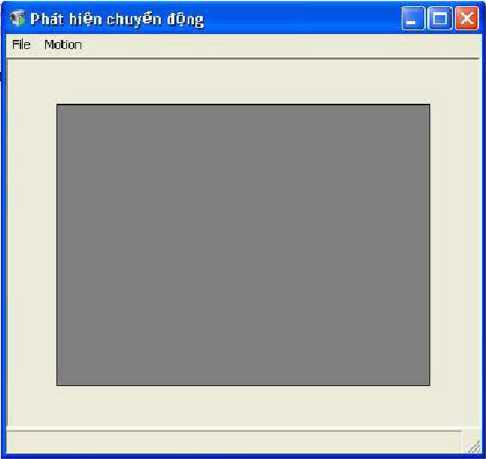
Dữ liệu đầu vào dùng cho bàỉ toán phát hiện đếỉ tượng đột nhập là một đoạn video. Nêu trong đoạn vỉdeo này có sự chuyển động củâ đếi tượng thì khoanh vùng các đối tượng chuyển động đó.

* 1. **Môi trường test**

ứng dụng được cài đặt thử nghiệm bằng ngôn ngữ lập trình c# cho Microsoft.NET plaíòrm phát triển trên hệ điều hành windows XP, bộ xử lý Pentium 2.8GHz, 512 MB bộ nhớ.

**33. Một số giao diện**

Khi bắt đầu chạy chương trình có giao diện như sau:



Hình 3.1 Giao diện chính của chưomg trình



Các kết quả:

(a)

(b)



(c)

Hình 3.2 (a), (b), (c) Kết quả sau khỉ thực hiện thuật toán vái video có

dạng \* AVI

**3\*4. Đánh giá**

Giải thuật sử dụng để phát hiện chuyển động thực hiện tương đối tốt trong chương trình thử nghiệm.

Chương trình thử nghiệm đã phát hiện và đóng khung những đối tượng chuyển động trong những vỉdeo định dạng \*.AVI có sẵn, hoặc qua các thiết bị thu (camera, webcam...). Tuy nhiên khỉ thực hiện phảt hiện chuyển động trong các đoạn video có nhiều đối tượng, tốc độ chuyển động của các đối tượng tương đối lớn thì việc đánh dấu những đếi tượng chuyển động gặp nhiều khố khăn (hình 3.3). Kêt quả kiểm tra vẫn bị ảnh hưởng bởi sự di chuyển của camera và bóng của cảc đếỉ tượng (hình 3.4). Để chương trình có thể ứng dụng vào thực tế, cần phải phát triển thêm để chương trình có thể phát hiện những đối tượng chuyển động với tấc độ lởn, khử nhiễu và giảm tác động của môi trường tớỉ kết quả phát hiện.



Hình 3.3 Phát hiện đối tưựng chuyển động nhanh



Hình 3.4 Camera bị rung

**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 bảo đảm trật tự an ninh. Trong đồ án tốt nghiệp này em đã trình bày một số phương pháp phát hiện đối tượng chuyển động cùng các ưu và nhược điểm của từng phương pháp khi áp dụng chúng. Tùy từng yêu cầu cụ thể về thời gian, về chất lượng và đặc điểm của video cần xử lý mà chúng ta chọn sử dụng phương pháp phát hiện thích họrp.

Bên cạnh việc đưa ra một cải tiến cho phương pháp phát hiện đối tượng chuyển động, khóa luận cũng đã cài đặt thử nghiệm thuật toán so sánh nền cho bài toán phát hiện đối tượng đột nhập với các file video \*.AVI. Đây là chương trình mang ý nghĩa demo kỹ thuật. Ngoài ra ta có thể phát triển chương trình tiếp để có thể phát hiện đối tượng chuyển động trên tất cả các fĩle video khác nhằm mang lại nhiều hiệu quả rộng rãi hơn.

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

**TÀI LIỆU THAM KHẢO**

1. . Cao Chánh Nguyên Hiển, Khương Trường Giang, “ Nghiên cứu

*phương phát phát hiện đổi tượng chuyển động trong video và ứng dụng* ” -Trường ĐH KHTN TP Hồ Chí Minh.

1. . Kỹ thuật trừ ảnh và ứng dụng trong phim cảnh- Đồ án tốt nghiệp

Nguyễn Thị Thúy Hòa - TS Đỗ Năng Toàn hướng dẫn [3 ]. Motion Detection Algorithms - Andrew Kirillov.

1. . Nguyễn Quốc Tuấn, Nguyễn Văn Kỷ Cang, Phan Vĩnh Phước,

“ *Phân tích tự động dữ liệu video sẻ hỗ trợ truy tìm thông tin thị giác dựa vào nội dung* ” - Trường ĐH KHTN TP Hồ Chí Minh.

1. . Sen-Ching s. Cheung and Chandrika Kamath, “ Robust techniques

for background subtraction in urban traffic video ”Center for Applied Scientitĩc Computing Lavvrence Livermore National Laboratory

1. . R.Cucchiara, M. Piccardi, and A. Prati, “ Detectỉng moving objects,

ghosts, and shadows in video streams ” IEEE Transactions on Pattem Analysis and Machine Intelligence 25, pp. 1337-1342, Oct 2003.

1. . A. Prati, I. Mikic, M. Trivedi, and R. Cucchiara, “ Detecting

moving shadows: algorithms and evaỉuation ” IEEE Transactions on Pattem Analysis and Maching Intelligence 25, pp. 918-923, July 2003.

1. . Thuan D. Vong, “*Background Subtraction Using Coỉor and*

Gradient Informatìorí\ Department of Electrical and Computer Engineering Clemson University Clemson.

1. . M. Heikkila, M. Pietikainen and J. Heikkila, “A Texture-based

*Method for Detectìng Moving Obịects*”, Infotech Oulu and

Department of Electrical and Iníormation Engineering p.o. Box 4500 FIN-90014 University of Oulu, Finland.

1. . Roland Miezianko1 , DragoỊjub Pokrajac2, “ Motion Detectỉon

Based on Local Variation of Spatiotemporal Texture ”, lTemple University, CIS Dept., Philadelphia, PA, 2 Delaware State University, CIS Dept., Dover, DE.

1. . Duda, R., p. Hart, and D. Stork, Pattem Classỉfication, 2nd edn.,

John Wiley & Sons, 2001.

1. . Flury, B. A *First Course in Muỉtivarỉate Statỉstỉcs,* springer Verlag.