 **ĐẠI HỌC ĐÀ NẴNG**

**TRƯỜNG ĐẠI HỌC BÁCH KHOA**

**KHOA ĐIỆN TỬ VIỄN THÔNG**

**BÁO CÁO ĐỒ ÁN**

**CHUYÊN NGÀNH KĨ THUẬT MÁY TÍNH**

**Đề Tài : Phát hiện chuyển động**

**trong video và ứng dụng**

|  |  |  |
| --- | --- | --- |
| GVHD : | TS. Trần Thị Minh Hạnh  TS. Hoàng Lê Uyên Thục |  |
| SVTH : | Lý Huỳnh Hữu Trí  Nguyễn Ngọc Hiếu  Tô Đông Trung | 19DTCLC2  19DTCLC2  19DTCLC2 |

Đà Nẵng, tháng 12 năm 2022 **MỤC LỤC**

**LỜI NÓI ĐẦU**

Trong thời đại ngày nay công nghệ thông tin hầu như đã thâm nhập vào toàn bộ các lĩnh vực đời sống xã hội. Xã hội càng phát triển thì nhu cầu về công nghệ thông tin ngày càng cao, do vậy dữ liệu số hầu như không còn xa lạ đối với mỗi người chúng ta.Trong mọi lĩnh vực các ứng dụng công nghệ thông tin đã trợ giúp con người rất nhiều. Hiện nay, thông tin hình ảnh đóng vai trò rất quan trọng trong trao đổi thông tin, bởi phần lớn các thông tin mà con người thu nhận được đều thông qua thị giác. Trong các lĩnh vực công nghệ thông tin thì lĩnh vực giám sát tự động đã và đang thu hút được nhiều sự quan tâm của các nhóm nghiên cứu trong và ngoài nước. Cùng với sự phát triển của sức mạnh máy tính, các hệ thống giám sát tự động ngày càng tinh vi và hiện đại đã trợ giúp con người rất nhiều trong lĩnh vực an ninh, giám sát giao thông v.v.Thời gian qua, hệ thống giám sát bằng camera đã trở thành lĩnh vực nghiên cứurất phát triển. Trước các nhu cầu giám sát, đảm bảo an ninh mục tiêu, trụ sở, nhàriêng… của các tổ chức, cá nhân ngày càng cao đã thúc đẩy việc nghiên cứu, ứng dụngcác sản phẩm giám sát thông qua camera theo dõi. Bài toán đặt ra là xây dựng và đưavào sử dụng các hệ thống giám sát đủ mạnh, có độ chính xác cao, phù hợp với từngmục đích cũng như đối tượng sử dụng. Một số ứng dụng quan trọng của hệ thống giámsát camera trong các lĩnh vực như: Các thiết bị giám sát an ninh chuyên dụng sử dụngtrong lực lượng vũ trang; các thiết bị giám sát an ninh dân dụng; các thiết bị giám sát, phân luồng giao thông; công nghệ nghiên cứu, chế tạo Robot thông qua việc phân tích các cử động của con người…Ở nước ta hiện nay, lĩnh vực giám sát Video cũng có những bước phát triển đáng kể. Tuy nhiên nó chỉ mới dựa trên nền tảng phần cứng và cũng chưa áp dụng nhiềutrong thực tế. Việc giải quyết bài toán này theo hướng tiếp cận sử dụng phần mềm chưa được quan tâm phát triển.Một hệ thống giám sát camera bao gồm nhiều chức năng: Phát hiện chuyển động, phân loại, theo dấu, cảnh báo, phân tích hành động. Phát hiện chuyển động được xác định là một trong những chức năng đầu tiên và quan trọng nhất của một hệ thống giámsát camera. Nó không chỉ có chức năng trích xuất đối tượng chuyển động mà còn có ý nghĩa quan trọng đối với ứng dụng thị giác máy tính như mã hóa video dựa trên đối tượng, phân tích chuyển động của con người và tương tác người – máy.Xuất phát từ các nhu cầu thực tế, nhóm chúng em đã quyết định chọn đề tài là:

“Phát hiện chuyển động bằng phương pháp trừ nền”

**CHƯƠNG I . TỔNG QUAN**

**1.1 Giới thiệu đề tài**

Trong kỷ nguyên công nghệ thông tin hiện nay, với sự phát triển nhanh chóng của các công nghệ chế tạo thiết bị phần cứng ngày càng hiện đại, tinh vi thì ngành trí tuệ nhân tạo cũng không ngừng mở rộng để phù hợp với các yêu cầu của thực tế. Trong đó phải kể đến sự phát triển của các thiết bị thu nhận hình ảnh từ thế giới thực, chẳng hạn như các hệ thống giám sát bằng camera, song hành với nó là các vấn đề liên quan đến việc giám sát. Thách thức chính trong lĩnh vực này chính là việc xử lý các hình ảnh thu nhận được từ các hệ thống giám sát đó Giám sát là một vấn đề được rất nhiều nhà nghiên cứu quan tâm đặc biệt bởi những ứng dụng thiết thực của nó cho đời sống xã hội. Chẳng hạn như các hệ thống giám sát các hành vi khả nghi của tội phạm, khủng bố ở các địa điểm nhạy cảm của các chính phủ. Hệ thống giám sát trong các viện bảo tàng, lưu trữ để chống trộm cắp các di vật đang được trưng bày. Hệ thống giám sát các hiện tượng bất bình thường, vi phạm pháp luật, tai nạn ở các điểm giao thông. Hệ thống giám sát phòng chống hỏa hoạn. Các hệ thống giám sát trong các siêu thị, cửa hàng, công ty để chống trộm cắp,… Thách thức chính cho ngành công nghệ phần mềm là đưa ra các giải pháp nhằm xây dựng một hệ thống giám sát tối ưu nhất nhằm giúp con người phát hiện chính xác và kịp thời các hiện tượng bất thường để có biện pháp xử lý nhanh chóng nhằm tránh các thiệt hại đáng tiếc cho xã hội. Dữ liệu thu được từ hệ thống camera giám sát thường được lưu trữ dưới dạng video. Như vậy công việc hiện nay của chúng ta là nghiên cứu các vấn đề liên quan đến việc xử lý video.

* 1. **Mục tiêu của đề tài**

Khi xã hội phát triển càng mạnh, yêu cầu về các thiết bị công nghệ càng cao. Như vậy, xử lý video là một mảnh đất màu mỡ cho các trung tâm nghiên cứu, các công ty đầu tư vào. Nhất là trong giai đoạn hệ thống nhúng đang phát triển và mở ra một kỷ nguyên mới cho ngành công nghệ phần mềm như hiện nay. 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 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 2 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, chúng em thấy việc nghiên cứu và đưa ra các phương pháp để xử lý video là vô cùng thiết thực.

* 1. **Phương pháp nghiên cứu**

Nghiên cứu và sử dụng phương pháp trừ nền. Cơ sở của phương pháp này là phát hiện các đối tượng chuyển động từ sự khác biệt giữa khung hiện tại và khung tham chiếu

**CHƯƠNG 2 : CƠ SỞ LÝ THUYẾT**

**2.1 . Các khái niệm cơ bản về video**

**2.1.1 Khái niệm video :**

Video là tập hợp các khung hình, mỗi khung hình là một ảnh. dòng video gồm các chuỗi các khung hình liên tiếp, không thể chia nhỏ hơn, ứng với một thao tác camera đơn. Scene (cảnh) là các đơn vị logic của dòng video, một cảnh gồm các lia liên quan về không gian và liền kề về thời gian, cùng mô tả một nội dung ngữ nghĩa hoặc một tình tiết. Khi được chiếu, các khung hình lần lượt được hiển thị ở tốc độ nhất định. Tốc độ thường thấy ở các định dạng video khác nhau là 30 và 25 hình/s. Như vậy một giờ video sẽ có số khung hình tương ứng là 108000 hoặc là 90000. Dù là video ở định dạng nào thì nó cũng có dung lượng rất lớn và nếu xử lý với tất cả các khung hình thì thật không hiệu quả. Phân đoạn là quá trình phân tích và chia nội dung hình ảnh video thành các farme. Video có 4 đặc trưng chính đó là: Color (màu), Texture (kết cấu), Shape (hình dạng), Motion (chuyển động).

* + 1. **Màu (Color)**

- Màu là một thuộc tính quan trọng của ảnh. Biểu đồ màu, biểu diến sự phân bố màu, là một đặc trưng màu phổ biến nhất hiện nay. Biểu đồ màu không phụ thuộc vào sự quay, dịch chuyển ảnh cũng như chiều nhìn ảnh. Tính hiệu quả của nó lại phụ thuộc vào hệ màu và phương pháp định lượng được dùng. Có một vấn đề với biểu đồ màu là nó không biểu diễn thông tin về không gian phân bố các điểm ảnh, do đó hai ảnh có cùng biểu đồ màu có thể có nội dung rất khác nhau.

**2.1.3 . Kết cấu (Texture)**

- Đây là một đặc trưng quan trọng 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, nhờ đó các thống kê có nghĩa có thể được trích chọn. Ngược lại, người ta thấy rằng entropy và mô men chênh lệch nghịch đảo lại có khả năng phân biệt tốt nhất.

**2.1.4. Hình dáng (Shape)**

- Các đặc trưng hình dáng có thể được biểu diễn sử dụng phân tích hình dáng truyền thống như bất biến mô men, mô tả Fourier, mô hình học tự 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, chẳng hạn như chu vi, tính tròn, mô men trung tâm, 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.

**2.1.5. Chuyển động (Motion)**

- 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 như mô men của trường chuyển động, biểu đồ chuyển động hoặc là 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 out), thu nhỏ (zoom in) cũng có thể được trích chọn.

* 1. **Giới thiệu về ngôn ngữ lập trình Python , thư viện OpenCV và các thư viện liên quan**
     1. **Ngôn ngữ Python**
* Python một ngôn ngữ lập trình bậc cao cho các mục đích lập trình đa năng, do Guido van Rossum tạo ra và lần đầu ra mắt vào năm 1991. Python được thiết kế với ưu điểm mạnh là dễ đọc, dễ học và dễ nhớ. Python là ngôn ngữ có hình thức rất sáng sủa, cấu trúc rõ ràng, thuận tiện cho người mới học lập trình và là ngôn ngữ lập trình dễ học; được dùng rộng rãi trong phát triển trí tuệ nhân tạo.
* Ngày nay, Python trở nên phổ biến nhanh vì tính đơn giản và khả năng đọc mã. So với C/C++, Python hoạt động chậm hơn do Python là ngôn ngữ thông dịch
  + 1. **OpenCV**
* OpenCV (Open Source Computer Vision Library) à một thư viện mã nguồn mở Python hỗ trợ các chức năng lập trình được sử dụng cho thị giác máy tính trong Trí tuệ nhân tạo, Học máy, nhận dạng khuôn mặt, v.v.
* Trong OpenCV, CV là dạng viết tắt của thị giác máy tính, được định nghĩa là lĩnh vực nghiên cứu giúp máy tính hiểu được nội dung của các hình ảnh kỹ thuật số như ảnh và video.
* Đặc biệt là về xử lý ảnh ứng dụng trong thị giác máy tính. OpenCV hỗ trợ nhiều ngôn ngữ khác nhau như: Python, C++, … và có sẵn trên các nền tảng như Windows, Linux, Mac OS, …
  + 1. **Numpy**
* NumPy là viết tắt của Numerical Python. NumPy được tạo ra vào năm 2005 bởi Travis Oliphant. Nó là một dự án mã nguồn mở cấu trúc dữ liệu mạnh mẽ, triển khai các mảng và ma trận đa chiều. Các cấu trúc dữ liệu này đảm bảo tính toán hiệu quả với ma trận và mảng. Việc triển khai thậm chí còn nhắm đến các ma trận và mảng khổng lồ. Bên cạnh đó, mô-đun cung cấp một thư viện lớn các hàm toán học cấp cao để hoạt động trên các ma trận và mảng này.
  + 1. **Một số Mô\_đun liên quan**
* Randint : tạo ra các số ngẫu nhiên
* Time , Date : cung cấp các lớp để thao tác ngày và giờ.
* Mixer : Mô-đun của PyGame để tải và phát âm thanh
  1. **Tìm hiểu về giải thuật trừ nền ứng dụng phát hiện chuyển động**

**2.3.1. Khái niệm :**

* Background subtraction hay còn gọi là trừ nền là một trong những giải thuật đơn giản và phổ biến trong lĩnh vực Thị giác máy tính (Computer Vision). Thuật toán được sử dụng nhằm xác định đối tượng chuyển động trong camera nền tĩnh.

**2.3.2 Các thuật toán trừ nền**

**2.3.2.1 Frame Difference(Khung khác biệt).**

**-** Phương pháp Frame Differences là một**thuật toán để xác định chuyển động của một đối tượng**. Sử dụng thuật toán này, chúng ta có thể phân biệt một đối tượng chuyển động trong môi trường. Phép trừ nền là một trong những phương pháp phù hợp để cải thiện hơn nữa sự khác biệt khung hình, do đó tăng hiệu quả và độ chính xác của nó.

**- Ưu điểm :**

- Đây là phương pháp đơn giản nhất của giải thuật trừ nền. Ta sẽ tính hiệu giá trị pixel của khung hình hiện tại và khung hình trước đó nếu hiệu này lớn hơn ngưỡng **T** thì pixel tại vị trí đó thuộc về Foreground.

- Ưu điểm của phương pháp này là tốc độ tính toán nhanh do việc khởi tạo background image chỉ đơn giản là việc lấy khung hình trước đó.

- **Nhược điểm :**

- Phương pháp này là nó chỉ tốt với những đối tượng di chuyển liên tục nhưng khi có một đối tượng đứng yên trong khung hình quá lâu thì đối tượng này sẽ bị cho là background cùng với đó phương pháp này phụ thuộc rất nhiều vào ngưỡng **T** do đó với mỗi video ta đều phải chọn 1 ngưỡng phù hợp.

**2.3.2.2 Mean Filter**

- Phương pháp Mean Filter sử dụng **N** khung hình phía trước khung hình đang xét để khởi tạo background image. Giả sử khung hình đang xét tại thời điểm **t** thì background tương ứng với nó sẽ được tính như sau:

Ảnh có chứa văn bản, đồng hồ

Mô tả được tạo tự động

* Trong đó **B** là background image tại thời điểm **t**, **N** là số khung hình trước thời điểm **t** để dùng cho việc tính toán ra background image, **I**(x, y, t) là khung hình tại thời điểm **t**.

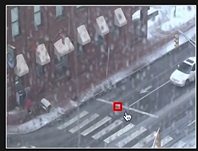
Sau khi khởi tạo background image thì phần còn lại của phương pháp này cũng giống với phương pháp frame difference.

* + 1. **Tìm hiểu Phương pháp Mixture of Gaussian (MOG2)** 
       1. **Bài toán đặt ra**
* Giả sử rằng trong một video đầu vào thì nó có thể hoạt động ở một thời thiết bình thường có thể chuyển sang thời tiết xấu ví dụ như cảnh đường phố có xe ô tô chạy qua và chúng ta muốn đếm những chiếc xe , hoặc tìm ra những chiếc xe vi phạm giao thông nhưng những gì chúng ta có là những hạt mưa và tuyết trong video trên đó và hạt mưa (tuyết) có thể duy chuyển xung quanh (tăng thêm sự phức tạp của ảnh ) thì với phương pháp trừ nền cơ bản nó thật sự không giả quyết tốt các vấn đề trên

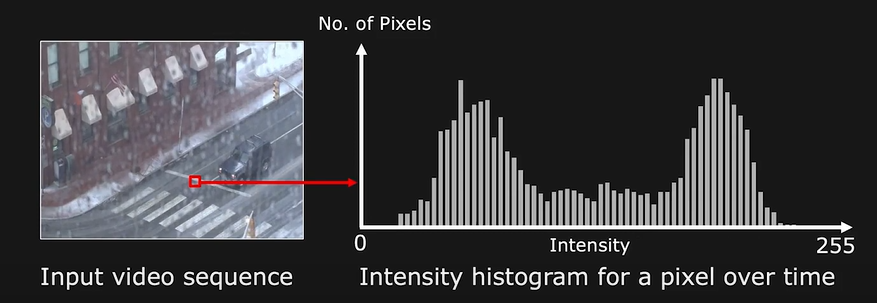
- Để thật sự đưa ra một giải pháp hữu ích , thì chúng ta cần có khả năng lập mô hình các biến đổi cường độ tại mỗi điểm trong ảnh vì vậy chúng ta xem các biến thể cường độ như một phân bố và sau đó chúng ta điều chỉnh một mô hình cho phù hợp với phân phối, bằng cách xem xét các thành phần khác nhau này của sự phân bố , chúng tôi đưa ra phán đoán rằng pixel sẽ thuộc về nền hay là tiền cảnh và mô hình này chúng ta nói đến là GMM (**Mixture of Gaussian**) .

* Bài toán đặt ra ở trên được giải quyết bằng GMM như sau :

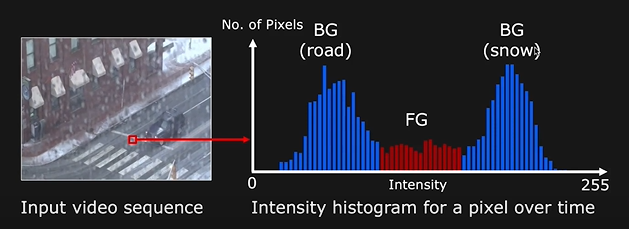
Hình dưới : đầu vào là một khung hình video . chúng ta giải sử xét một Pixel tại vòng màu đỏ . chúng ta tiến hành xem xét mà tính toán biểu đồ của các mẫu độ sáng bằng Histogram tương ứng với ảnh đó



* Biểu đồ ta thấy ở đây có giá trị từ 0-255 :

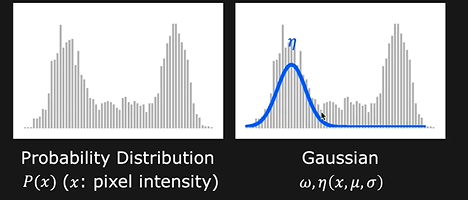


* Chi tiết do cường độ sáng về nền , đường . có thể thay đổi theo thời gian về độ chiếu sáng , trong trường hợp này có thể do ánh sáng hoặc giọt mưa , tuyết đi qua pixel và đôi khi đối tượng chuyển động đi ngang qua pixel ví dụ ở đây là xe ô tô nhưng hiếm khi xe đi qua pixel này mà thường thì do tuyết hay mưa đi qua
* Xem xét phân phối của của chúng ta
* Chúng ta có thể xem đây là một tập hợp các cường độ , mỗi hình chuông là một quần thể con (gồm con đường , tuyết , và đôi khi có xe oto chạy qua dẫn đến cường độ thay đổi được thể hiện bằng xanh lam và sau đó đối tượng hoặc phương tiện di chuyển không thường xuyên (Cụ thể là Ô Tô) được hiển thị bằng màu đỏ ở đây . vì vậy trực giác ở đây là pixel nền vì nếu 1 chiếc ô tô đi vào và nằm trên pixel và nó vẫn ở đó và tại một số điểm chúng ta phải nói rằng chiếc xe đó là một phần của background (Như hình dưới )

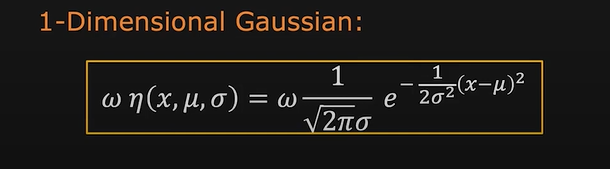


- Chúng ta sẽ thử và mô hình hóa sự phân bố này và sau đó phân loại pixel thuộc về background (Nền) hay là foreground (Tiền Cảnh )

- Chúng ta sẽ sử dụng mô hình gaussian này , chúng ta mô hình hóa tháp đầu tiên sử dụng gaussian nếu lắp một gaussian và dữ liệu đó sẽ có dạng như (Hình dưới )

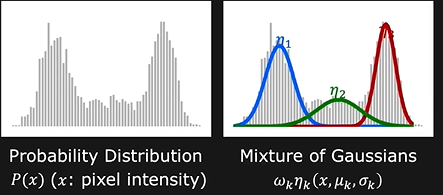


* Gaussian Dimensional :

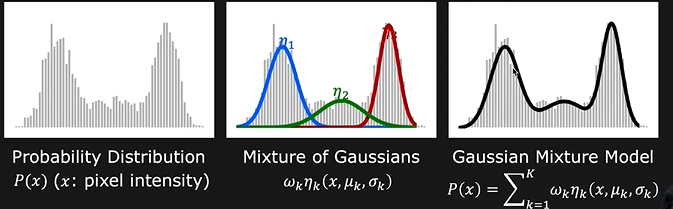


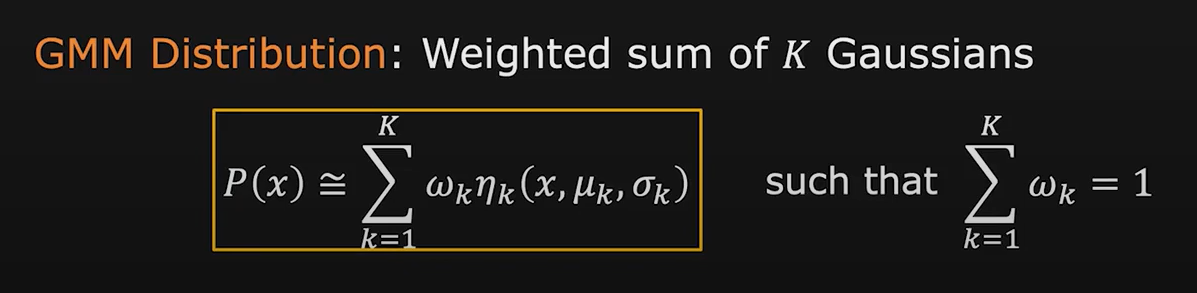
(Và gaussian như chúng ta đã biết chiều rộng được cho bởi sigma (σ ) và giá trị trung bình do Mean đưa ra kí hiệu (μ) , thường được gọi là quy mô liên quan đến gaussian )

* chúng ta sẽ có một tham số thứ ba đó là omega . Omega đc gọi là quy mô liên quan đến gausian này điều này có thể phụ thuộc vào thực thể của từng vật thể chuyển động qua pixel thì sẽ có những chiều cao của chuông khác nhau đó chính là gaussian => nên chúng ta có thể lập mô hình một trong những đỉnh với một gaussian tương ứng , vì chúng ta đang xem xét các giá trị độ sáng
* Với logic tương tự chúng ta có thể vẽ lên các gaussian và bây giờ chúng ta xem toàn bộ biểu đồ biến thiên cường độ như tập hợp các Gaussian



* một biểu diễn khá tốt bằng cách sử dụng hỗn hợp gaussian . khi trộn chúng với nhau và chúng ta sẽ nhận được một biểu đồ thực tế

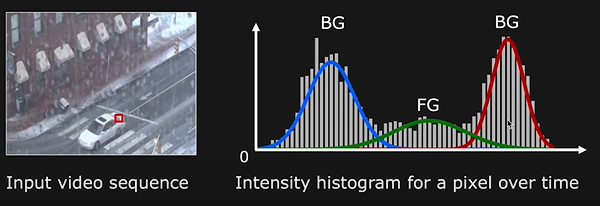


* Lưu ý : chúng ta chỉ xét 1 pixel ở vòng tròn màu đỏ ( 1 video có thể đọc vào nhiều frame nhưng chúng ta chỉ ví dụ là ngay tại pixel đó chỉ có 3 trường hợp sảy ra với pixel đang xét)
* Vậy đó là ý tưởng mà chúng ta điều chỉnh một mô hình hỗn hợp gaussian vào biểu đồ , và chúng ta sẽ có mô hình là tổng có trọng số của các gaussian chính
* Mỗi gaussian đều có giá trị trung bình riêng và chiều rộng của nó

- Cách để sử dụng mô hình GMM

Lưu ý : Chọn K . nếu K quá lớn có thể tưởng tượng dữ liệu sẽ không ổn định , nếu K quá nhỏ sự phù hợp sẽ không chính sác và ko đủ linh hoạt trong một mô hình

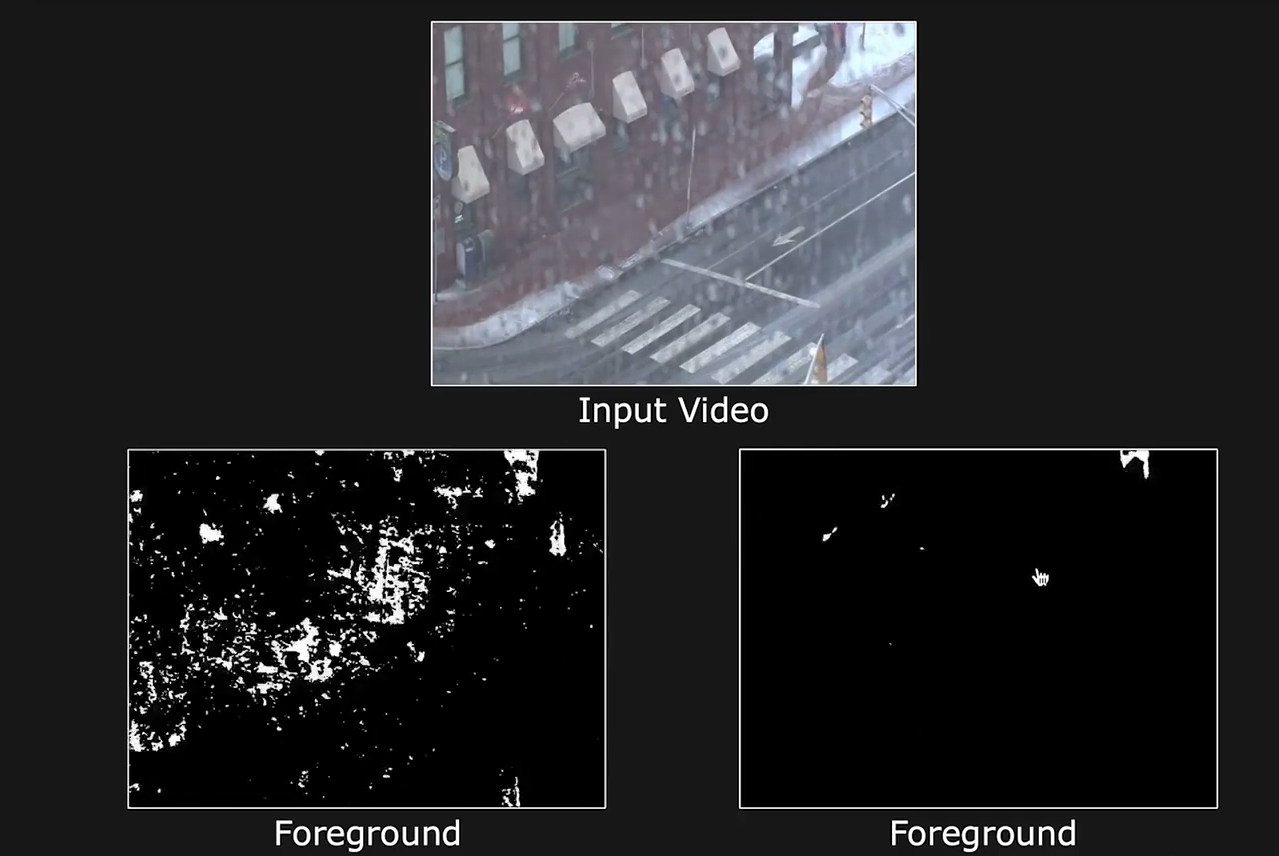
* Để đưa ra một mô hình hỗn hợp gaussian cho cường độ biến đổi màu sắc của pixel theo thời gian chúng ta sẽ có các vấn đề giải quyết đó là chúng ta phân loại đối tượng trong số guassian hình là tiền cảnh hay hậu cảnh
* Trước tiên muốn tìm ra thì chúng ta phải có một mô hình và tìm ra gaussian đại diện cho Foreground hoặc background để chúng ta nhìn thấy được các pixel trong tương lai (Frame tiếp theo), và phân loại chúng thành các thành phần khác nhau
* Hình dưới chúng ta thấy ba gaussian được liên kết với phân phối cụ thể này (khi omega lớn và sigma (σ ) nhỏ => sẽ là là gaussian nền )



* Nếu chia omega và sigma có độ lệch chuẩn lớn thì chúng ta gọi đó là nền . và bất cứ khi nào có omega chia chi sigma nhỏ sẻ đc gọi là tiền cảnh
* Công thức : σ  lớn => background , σ  nhỏ => foreground
* **Một điều gặp phải là cập nhập lại biểu đồ do quá trình xử lý video là đọc vào từng fame**
* Nên chúng ta không muốn là điều này cho mỗi khung hình mà chúng ta sẽ tính toán lại các mô Hình hỗn hợp gaussian và có thể là hàng triệu pixel cho mỗi khung hình ta thấy quá rắc rối và tính toán phức tạp nên ta làm như sau :
* B1. khung hình đầu tiên của video chúng ta tiến hành chuẩn hóa biểu đồ để đảm bảo nó được phân phối
* B2. bây giờ chúng ta lắp một GMM , và chúng ta chọn K bằng 3-5 để phân loại pixel và đối với mỗi khung hình tiếp theo
* B3. chúng ta sẽ lấy giá trị pixel X và một pixel nhất định cho cùng 1 pixel ở đây , giá trị X và ta tìm Gaussian nào trong mô hình của chúng ta gần nhất về giá trị trung bình μk gần nhất với X. đó là Gausian mà giá trị pixel thuộc về muốn biết Gaussian đó là gì chúng ta nhìn vào omega chia cho sigma để phân biệt



* B4. Và bây giờ có một điều là chúng ta cần cập nhật lại biểu đồ : nghĩa là chúng ta không làm điều này cho mọi khung hình mà chúng ta chỉ cần tính toán lại các mô hình hỗn hợp gaussian và có thể là hàng triệu pixel cho mỗi khung hình
* B5. Vì vậy chúng ta chỉ cần cập nhật những biểu đồ cho pixel , nếu biểu đồ chuẩn hóa khác với biểu đồ hiện tại mà nếu sự khác biệt đó lớn thì chúng ta sẽ thay biểu đồ mới và sửa lại mô hình GMM (nhưng điều này hiếm khi sảy ra vì 1 pixel duy nhất sẽ không thay đổi đáng kể biểu đồ trên hàng triệu biểu đồ pixel nhưng theo thời gian thì có thể thay đổi biểu đồ đó (có thể là độ sáng thay đổi , hoặc thởi tiết thay đổi đột ngột có thể sảy ra))
* **Kết quả cuối cùng :**



* 1. **Tổng quan về Mixture of Gaussian**

**2.4.1 Cơ sở lý thuyết :**

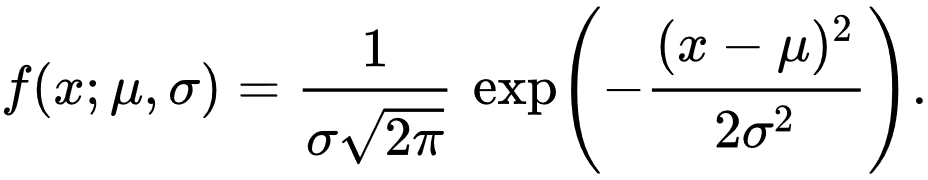
**2.4.1.1 Phân phối chuẩn :**

* Phân phối chuẩn, còn gọi là sự phân bố dữ liệu hay phân phối Gauss (cách gọi khác là Hình chuông Gauss), là một phân phối xác suất cực kì quan trọng trong nhiều lĩnh vực , Mà ở đó giá trị tập trung nhiều nhất ở khoảng giữa và các giá trị còn lại rải đều đối xứng về phía các điểm cực trị
* Nó là họ phân phối có dạng tổng quát giống nhau, chỉ khác tham số vị trí (giá trị trung bình **μ**) và tỉ lệ (phương sai **σ**).

Chart, histogram

Description automatically generated

Hàm mật độ xác suất của phân phối chuẩn với trung bình **μ** và phương sai **σ2** (hay độ lệch chuẩn **σ** ) là một ví dụ của một hàm Gauss có công thức :



Trong đó : μ= Giá trị trung bình

σ = Độ lệch chuẩn

π≈3.14159π≈3.14159

e≈2.71828

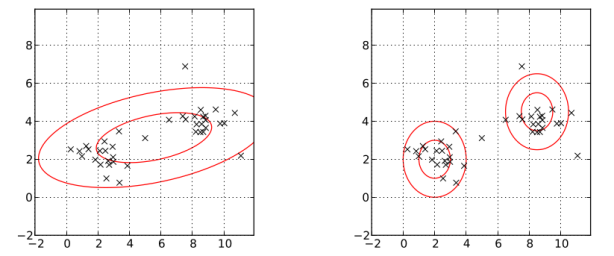
**2.4.1.2 Mô hình Gaussian hỗn hợp (GMM)**

**+** Mô hình hỗn hợp Gaussian là một mô hình xác suất để đại diện cho các quần thể con phân bố bình thường trong một tổng thể tổng thể (Như ví về bài toán đặt ra ở trên xét trên 1 pixel thì 1 pixel có thể có nhiều quần thể con về cường độ màu sắc cụ thể như (đường , xe ô tô , và có khi là có tuyết hoặc giọt mưa đi qua một pixel))

+ Là một khái niệm xác suất được sử dụng để mô hình hóa các tập dữ liệu trong thế giới thực

+ Là một hàm tham số mật độ xác suất được biểu diễn như là một tổng trọng số của các mật độ Gaussian thành phần. GMM được sử dụng rộng rãi như là một mô hình tham số của phân phối xác suất

+ Là một hình thức học tập không có giám sát .

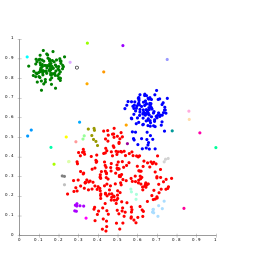
*Hình trên. ( Bên trái) Phù hợp với một phân phối Gaussian ( Biên phải) Phù hợp với mô hình hỗn hợp Gauss với hai thành phần*

Một gợi ý rằng dữ liệu có thể tuân theo mô hình hỗn hợp là dữ liệu trông [đa phương thức](https://brilliant.org/wiki/data-mode/) , tức là có nhiều hơn một "đỉnh" trong phân phối dữ liệu. Việc cố gắng điều chỉnh phân phối đa phương thức với một mô hình [đơn](https://brilliant.org/wiki/data-mode/) phương thức (một "đỉnh") nói chung sẽ cho kết quả kém phù hợp, như thể hiện trong ví dụ dưới đây. Vì nhiều phân phối đơn giản là đơn phương thức, một cách rõ ràng để mô hình hóa phân phối đa phương thức là giả định rằng nó được tạo ra bởi nhiều phân phối đơn phương thức. Vì một số [lý do lý thuyết](https://brilliant.org/wiki/central-limit-theorem/), phân phối được sử dụng phổ biến nhất trong mô hình hóa dữ liệu đơn phương thức trong thế giới thực là phân phối Gauss. Do đó, mô hình hóa dữ liệu đa phương thức như một hỗn hợp của nhiều phân phối Gaussian đơn phương thức có ý nghĩa trực quan. Hơn nữa, GMMs duy trì nhiều lợi ích về mặt lý thuyết và tính toán của các mô hình Gaussian, khiến chúng trở nên thiết thực để lập mô hình hiệu quả các tập dữ liệu rất lớn.

Mô hình hỗn hợp Gaussian được tham số hóa bởi hai loại giá trị, **trọng số thành phần** hỗn hợp và giá trị trung bình thành phần và **phương sai / hiệp phương sai** . Đối với một mô hình hỗn hợp Gaussian với Kcác thành phần,kthứ tự thành phần có nghĩa là *μk*​. và phương sai của *σk* ​đối với [trường hợp đơn biến](https://brilliant.org/wiki/normal-distribution/) và trung bình của vector *μ*​*k*​.và ma trận hiệp phương sai của Σ*k* ​đối với [trường hợp đa biến](https://brilliant.org/wiki/multivariate-normal-distribution/) . Trọng lượng thành phần hỗn hợp được định nghĩa là *ϕk* ​cho thành phần *Ck*​, với ràng buộc rằng =1để tổng phân phối xác suất chuẩn hóa thành11. Nếu trọng số thành phần không được học, chúng có thể được xem như một phân phối ưu tiên so với các thành phần như *p*( *x* được tạo bởi thành phần  *Ck*​)=*ϕk*​. Nếu chúng được học thay vào đó, chúng là các ước tính [hậu](https://brilliant.org/wiki/a-posterior/) kỳ của các xác suất thành phần được cung cấp dữ liệu

**2.4.2 Học tập không giám sát**

Hai hình thức suy luận phổ biến nhất được thực hiện trên GMM là [**ước tính mật độ**](https://brilliant.org/wiki/density-estimation/) và [**phân cụm**](https://brilliant.org/wiki/clustering/) .



Phân cụm sử dụng mô hình hỗn hợp Gaussian. Mỗi màu đại diện cho một cụm khác nhau theo mô hình.

**2.4.2.1 Ước tính mật độ**

Vì GMM hoàn toàn được xác định bởi các tham số của các thành phần riêng lẻ của nó, một GMM phù hợp có thể đưa ra ước tính xác suất của cả điểm dữ liệu trong mẫu và ngoài mẫu, được gọi là ước tính mật độ. Hơn nữa, vì có thể lấy mẫu số từ một phân phối Gaussian riêng lẻ, nên người ta có thể dễ dàng lấy mẫu từ GMM để tạo tập dữ liệu tổng hợp.

Lấy mẫu từ GMM bao gồm các bước sau:

B1: Lấy mẫu thành phần Gaussian theo phân bố được xác định bởi p ( CS )=ϕS.

B2: Mẫu x từ phân phối cho thành phần CS, theo phân phối được xác định bởi N ( x∣μS,σS) .

**2.4.2.2 Phân cụm (Thuật toán** **K-mean clustering)**

1. **Định nghĩa và nguyên lý của K-mean clustering :**
2. **Định nghĩa :**

+ K-mean clustering là một**thuật toán phân cụm đơn giản thuộc loại học không giám sát** (tức là dữ liệu không có nhãn) và được sử dụng để giải quyết bài toán phân cụm. Ý tưởng của thuật toán phân cụm K-mean

// Bổ sung

**2.5. Tìm hiểu về xử lý ảnh**

**2.5.1 Giới thiệu**

**2.5.1.1 Tổng quan**

- Xử lý ảnh là đối tượng nghiên cứu của lĩnh vực thị giác máy tính, là quá trình biến đổi từ một ảnh ban đầu sang một ảnh mới và tuân theo ý muốn của người sử dụng. . Đây là một phân ngành khoa học mới rất phát triển trong những năm gần đây

- Xử lý ảnh gồm quá trình phân tích, phân lớp đối tượng, làm tăng chất lượng phân đoạn, tìm biên, gán nhãn cho vùng. Xử lý ảnh được áp dụng rộng rãi trong trong đời sống và các lĩnh vực khác: xử lý nâng cao chất lượng ảnh, nhận dạng ảnh , Photoshop, Nhận dạng khuôn mặt, nhận dạng chữ viết, nhận dạng biển số xe , xử lý ảnh thiên văn, ảnh y tế, …Sự phát triển của xử lý ảnh đem lại rất nhiều lợi ích cho cuộc sống của con người.

**2.5.1.2 Một số khái niệm cơ bản trong xử lý ảnh**

**\* Ảnh và điểm ảnh :** Điểm ảnh được xem như là dấu hiệu hay cường độ sáng tại 1 toạ độ trong không gian của đối tượng và ảnh được xem như là 1 tập hợp các điểm ảnh

\* **Pixel :** Ảnh gồm nhiều điểm ảnh được gọi là pixel . mỗi pixel gồm một cặp tọa độ x, y và màu . cặp tạo độ x,y tạo nên độ phân giải . khi được số hóa thì thường được biểu diễn dưới dạng ma trận 2 chiều có kích thước N\*M ( với N là hàng , M là cột ) , và ảnh gồm N\*M = Số Pixel

\* **Mức xám, màu :** Là số các giá trị có thể có của các điểm ảnh của ảnh , mức xám là cường độ sáng của điểm ảnh , màu là màu của điểm ảnh (R,G,B).

**2.5.2 Phép hình thái học trong xử lý ảnh**

**2.5.2.1 Giới thiệu :**

**- Phép toán hình thái học** là tập hợp các phép toán dùng để phân tích hình ảnh và xử lý tín hiệu số đa chiều theo mẫu , theo cấu trúc được chọn

- Yếu tố cấu trúc : Là yếu tố do chúng ta quy định , có thể là bất kì hình dạng nào hoặc kích thước nào

- Fit : Tất cả các pixel trong yếu tố cấu trúc sẽ phủ hết các pixel trên ảnh

- Hit : Bất kì yếu tố cấu trúc nào phủ lên trên 1 pixel ảnh

-Miss : trường hợp còn lại ( các pixel trong cấu trúc sẽ không phủ lên trên 1 pixel ảnh nào )

**- Điều kiện :**

+ Đối tượng là X : ảnh xám hoặc ảnh nhị phân

+ Phần tử cấu trúc S : là mặt nạ bất kì mà các phần tử của nó tạo nên mô típ . tiến hành rê mặt nạ đi khắp ảnh và tính gài trị điểm ảnh cới các điểm lân cận với mô tip của mặt nạ theo cách lấy hội (Phép AND ) và tuyển (Phép OR)

+ Bx là phép dịch chuyển vị trị B đến vị trí x

**- Mục đích :**

+ Nghiên cứu cấu trúc hình học của đối tượng ảnh

+ Đơn giản hóa bằng cách loại bỏ các thông tin không cần thiết

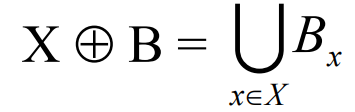
**2.5.3 Các phép hình thái học cơ bản :**

**2.5.3.1. Phép giãn (Nở) ảnh** :

**Khái niệm :**

* Phép toán này có tác dụng làm cho đối tượng ban đầu trong ảnh tăng lên về kích thước (giản nở ra).

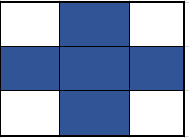
**Công thức :**



Trong đó : + X : Ma trận điểm ảnh của ảnh nhị phân.

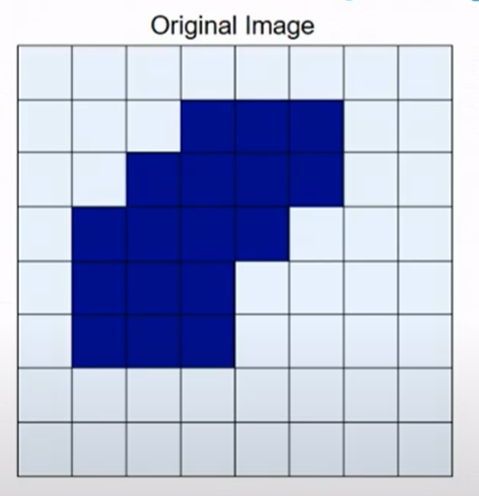
+ B: Là phần tử cấu trúc. Phép giãn nở (Dilation) ảnh sẽ cho ra hợp các Bx với x thuộc X , dễ dàng thấy rằng đây là một phép tổng giữa X và B.

**Ví dụ :**

**+ Phần tử cấu trúc S :**  

**+ Cách hoạt động :di chuyển đến từng pixel của ảnh nếu trùng 1 phần , hoặc trùng hoàn toàn thì Tâm phần tử cấu trúc chuyển thành pixel thuộc ảnh**

**TRƯỚC**  **=>** **SAU**

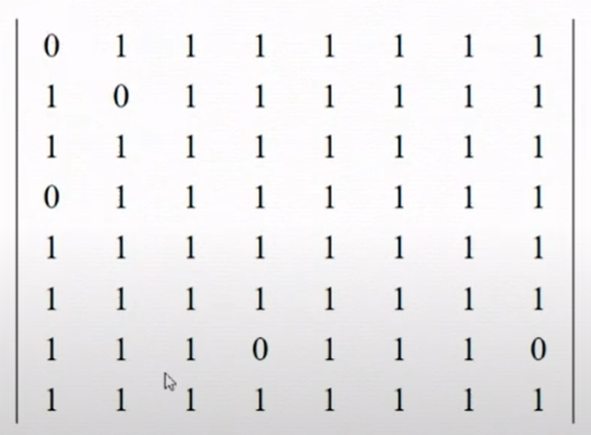
 

**Mô phỏng Tính toán ảnh dưới dạng ma trận :**

**Giãn ảnh với ma trận Phần tử cấu trúc**

****

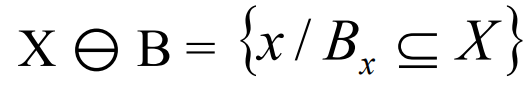
**Kết quả sau khi giãn ảnh**

****

**2.5.3.2. Phép co ảnh**

**Khái niệm :**

* Phép toán co (Erosion) là một trong hai hoạt động cơ bản (khác phép giãn nở) trong hình thái học có ứng dụng trong việc giảm kích thước của đối tượng, co hẹp đối tượng, tách rời các đối tượng gần nhau
* **Công thức :**

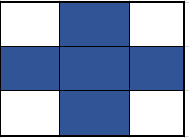
****

**Trong đó :**

+ X : Ma trận điểm ảnh của ảnh nhị phân.

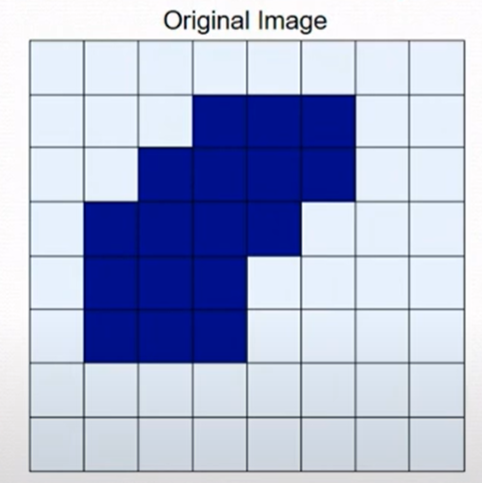
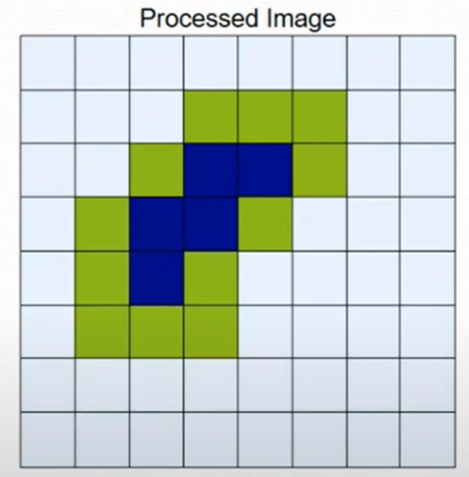
+ B: Là phần tử cấu trúc. Phép giãn nở (Dilation) ảnh sẽ cho ra hợp các điểm x sao cho Bx thuộc X , dễ dàng thấy rằng đây là một phép tổng giữa X và B.

**Ví dụ :**

**+ Phần tử cấu trúc S :**  

**+ Cách thực hiện : Bỏ các vị trí trùng 1 phần**

**TRƯỚC**  **=>** **SAU**

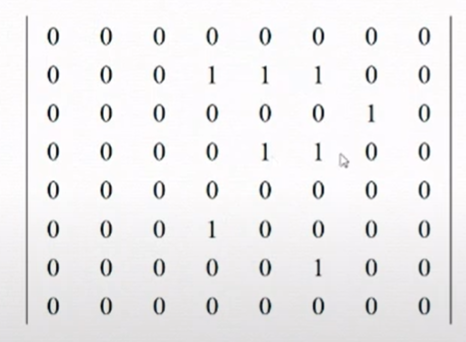
**** ****

**Mô phỏng Tính toán ảnh dưới dạng ma trận :**

**Co ảnh với ma trận Phần tử cấu trúc**

****

**Kết quả sau khi giãn ảnh**

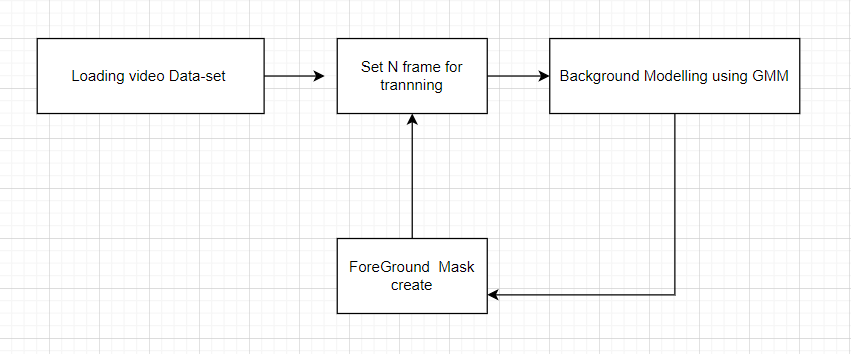
****

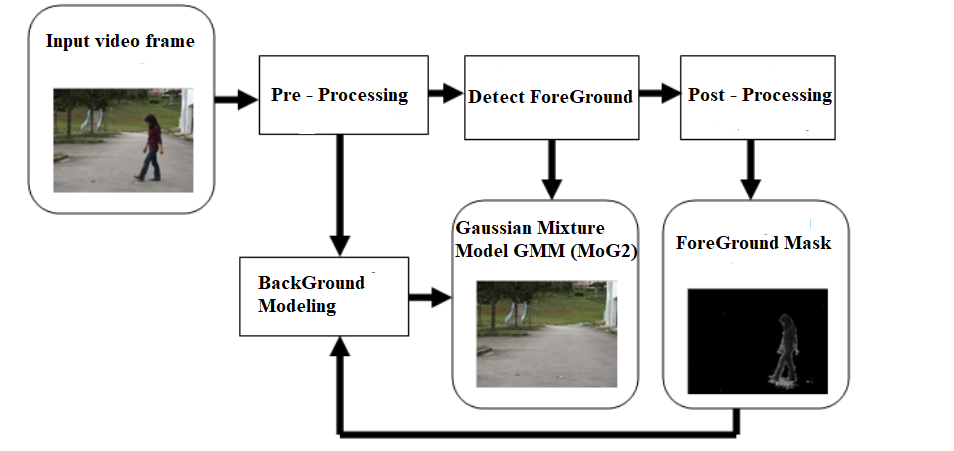
**2.6. Phương pháp đánh giá mô hình**

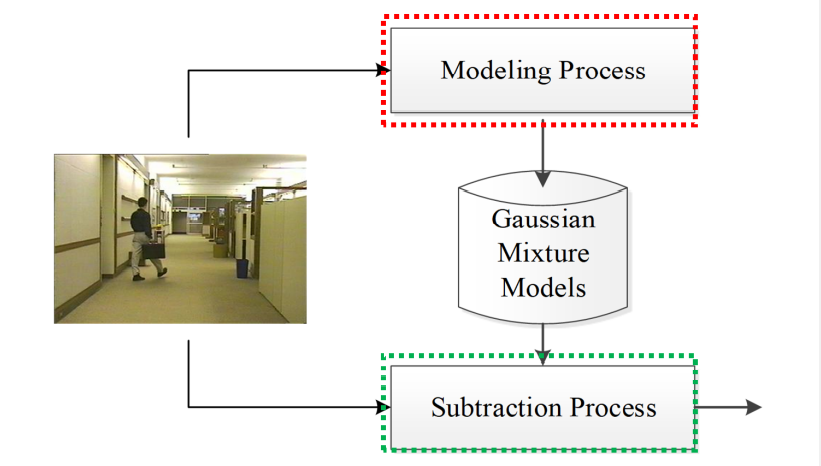
**CHƯƠNG 3 : THỬ NGHIỆM VÀ ĐÁNH GIÁ**

**3.1** **Sơ đồ khối và hoạt động**

**a. Sơ đồ khối của quy trình trừ nền sử dụng GMM**

****

****

****

**+ Loading video data-set** : đọc vào dữ liệu của video

**+ Set N frame for Trainning :** Set N frame để thực hiện mô hình nền

**+ Background Modelling Using GMM :** sử dụng mô hình GMM

**+ ForeGround Mask create :** tạo ra mặt nạ tiền cảnh

**Lưu Ý** : khối **Background Modelling Using GMM và khối Set N frame for Trainning** được tích hợp sẵn trong thư viện openCV với : **CV2.CreateBacgroundSubtractorMOG2(Set N frame for Trainning , varThreshold, detShadows )**

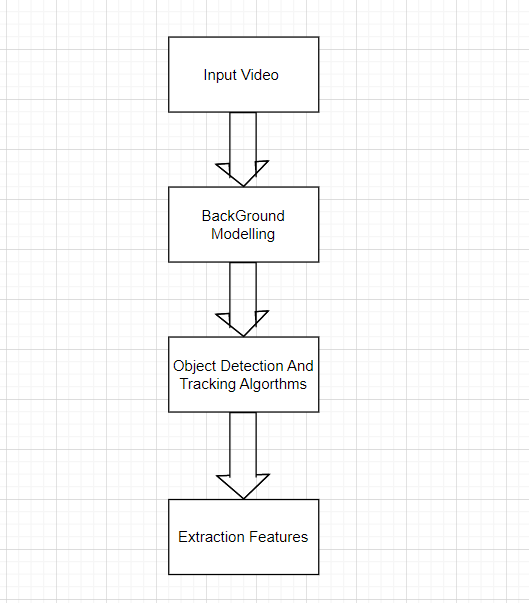
Ngoài ra :

**+ Tham số varThreshold :** thể hiện sự tương quan giá trị trọng lượng của các pixel trên khung hình hiện tại với các giá trị trên mô hình nền

**+** **Tham số detShadows :** đây là tham số bật hoặc tắt tính năng phát hiện bóng

1. **Sơ đồ khối của quy trình dectect (Phát hiện đối tượng)**

**c. sơ đồ khối của của chương trình**

****

Hình …. Sơ đồ khối hoạt động tổng quát của hệ thống

Các khối báo gồm :

+ **Input video** : là đọc video vào

**+ BackGround Modelling** : sử dụng phương pháp trừ nền

+ **Object dectect and tracking algorrthms :** Sử dụng thuật toán phát hiện và theo dõi đối tượng

+ **Extraction Features** : thêm chức năng

**3.2 các bài toán thực hiện**

**3.2.1 phát hiện xe ô tô chuyển động trên đường và thực hiện tracking và đếm số lượng xe trên đường sử dụng mô hình GMM**

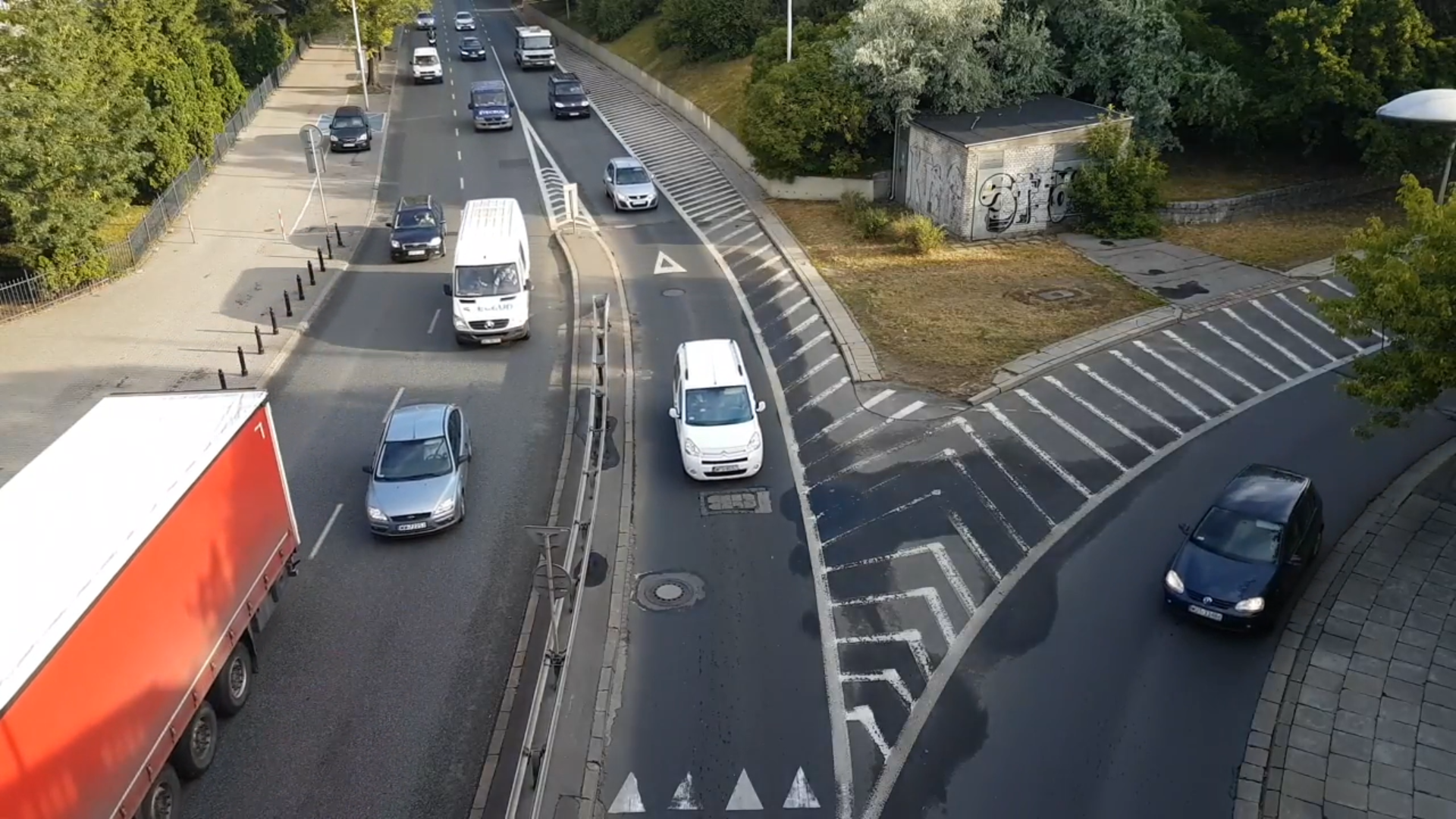
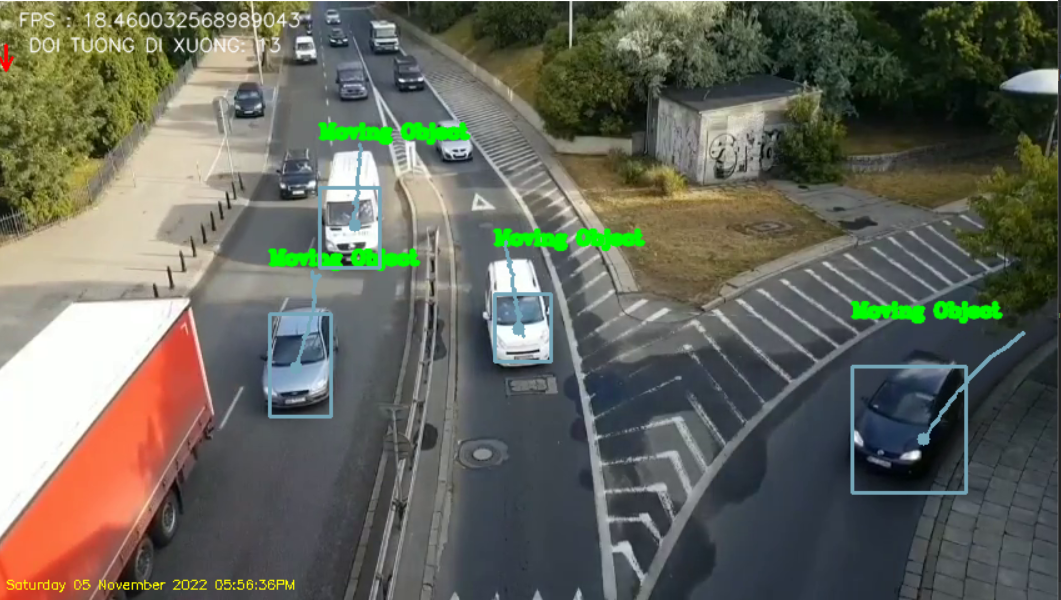
- Bài toán phát hiện xe chuyển động là bài toán được áp dụng mô hình GMM để xử lý . Đồng thời sử dụng một số thuật toán để tiến hành detect được xe tiến hành tracking và đếm số lượngxe

**3.2.1.1**  Bài toán detect :

**3.2.1.2** Bài toán tracking :

**3.2.1.3** Bài toán đếm số lượng xe :

**Kết quả mô hình :**

** **

Hình … Kết quả hoạt động của mô hình GMM

**3.2.1 phát hiện trộm di chuyển vào trong nhà và thực hiện báo động**

- Bài toán phát hiện trộm chuyển động vào ban đêm là bài toán được áp dụng mô hình GMM để xử lý . Đồng thời sử dụng một số thuật toán để tiến hành detect được đối tượng và tiến hành phát âm thanh cảnh báo

- Bài toán detect :

- Bài toán phát hiện cảnh báo :

**Kết quả mô hình :**

** **