

IT4090

Xử lý ảnh

(Digital Image Processing)

PGS.TS. Nguyễn Thị Hoàng Lan
Bộ môn Truyền thông và Mạng máy tính
Viện CNTT&TT - ĐHBK Hà nội
lan.nguyenthihuang@hust.edu.vn

Chương 9

Phát hiện và theo dõi chuyển động

(Motion Detection and Tracking)

- Tổng quan về phát hiện và theo vết đối tượng chuyển động trong video
- Các kỹ thuật phát hiện chuyển động trong video
 - Kỹ thuật trừ nền
 - Kỹ thuật phát hiện chuyển động khác
- Kỹ thuật theo vết (theo dõi) chuyển động
- Phát hiện và theo dõi người chuyển động trong video
(Human Motion Detection & Tracking for Video Surveillance)

Tổng quan – Giới thiệu chung

- **Phát hiện và theo dõi chuyển động** (Motion Detection and Tracking):
Mục đích nhằm phát hiện đối tượng chuyển động và theo vết chuyển động trong cảnh một cách độc lập với phông nền của cảnh (background)
- **Hai ngũ cảnh của bài toán:**
 - Cảnh được quan sát bằng camera cố định (Static Camera)
 - Cảnh được quan sát bởi camera chuyển động (Moving Camera)
- **Các vấn đề phát hiện chuyển động** (MotionDetection):
 - Phát hiện hành động chuyển động vật lý của đối tượng trong cảnh
 - Chuyển động có thể được phát hiện khi thay đổi thang đo tốc độ hay hướng của vector chuyển động của đối tượng
- **Tùy theo mục đích ứng dụng các bài toán** về phát hiện và theo dõi chuyển động được xác định, có các giải pháp kỹ thuật phù hợp.

Giới thiệu các áp dụng phát hiện và theo dõi chuyển động

- Các ứng dụng giám sát/ kiểm soát
(Surveillance/Monitoring Applications)
 - Security Cameras
 - Traffic Monitoring
 - People Counting
- Các ứng dụng điều khiển (Control Applications)
 - Object Avoidance
 - Automatic Guidance
 - Head Tracking for Video Conferencing....
- Many intelligent video analysis systems are based on motion detection and tracking

Phát hiện chuyển động trong cảnh tĩnh với camera cố định

- Ngữ cảnh: Quan sát đối tượng chuyển động trong cảnh tĩnh với camera cố định
- Các kỹ thuật phát hiện chuyển động (Motion Object Detection)
Phát hiện và phân vùng chuyển động là bước đầu tiên và có ảnh hưởng quyết định tới sự chính xác của hệ thống. Có rất nhiều hướng tiếp cận để phát hiện chuyển động:
 - Trừ nền (Background Subtraction)
 - Tìm sự khác biệt khung hình theo thời gian (frame differencing)
 - Phương pháp luồng thông lượng quang ảnh (Optical Flow)
 - Phương pháp tương quan...
- Các phương pháp khác nhau đem lại hiệu quả khác nhau. Không có phương pháp nào là tối ưu với tất cả các trường hợp ảnh đầu vào

Phát hiện chuyển động dùng Kỹ thuật trừ nền

- **Nguyên tắc:**

$$\text{Ảnh trừ nền} = \text{Ảnh đầu vào} - \text{Ảnh nền}$$



- Phát hiện đối tượng chuyển động bằng cách tính toán sự khác biệt giữa khung hình hiện tại với ảnh nền, dựa trên kỹ thuật xử lý mỗi khung hình (frame) video được so sánh với mô hình tham chiếu hoặc mô hình nền và áp dụng ngưỡng để phát hiện các vùng tiền cảnh chuyển động.
- Các pixel của vùng chuyển động sẽ được định vị và theo dõi (theo vết quỹ đạo) chuyển động của đối tượng.
- Có nhiều phương pháp để mô hình hóa nền.

So sánh hiệu năng các giải thuật trừ nền

Giải thuật trừ nền	Tốc độ	Yêu cầu bộ nhớ	Độ chính xác
Trung bình (Average)	Nhanh	Cao	Chấp nhận được
Running Average	Nhanh	Thấp	Chấp nhận được
Trung vị	Nhanh	Cao	Chấp nhận được
MOG	Trung bình	Trung bình	Trung bình
KDE	Trung bình	Cao	Tốt hơn
Không gian nền riêng	Trung bình	Trung bình	Trung bình
SKDA	Trung bình	Trung bình	Tốt hơn
Mean-shift chuẩn	Chậm	Cao	Tốt hơn
Mean-shift tối ưu	Trung bình	Trung bình	Tốt hơn

Phát hiện chuyển động dựa trên sự khác biệt khung hình theo thời gian

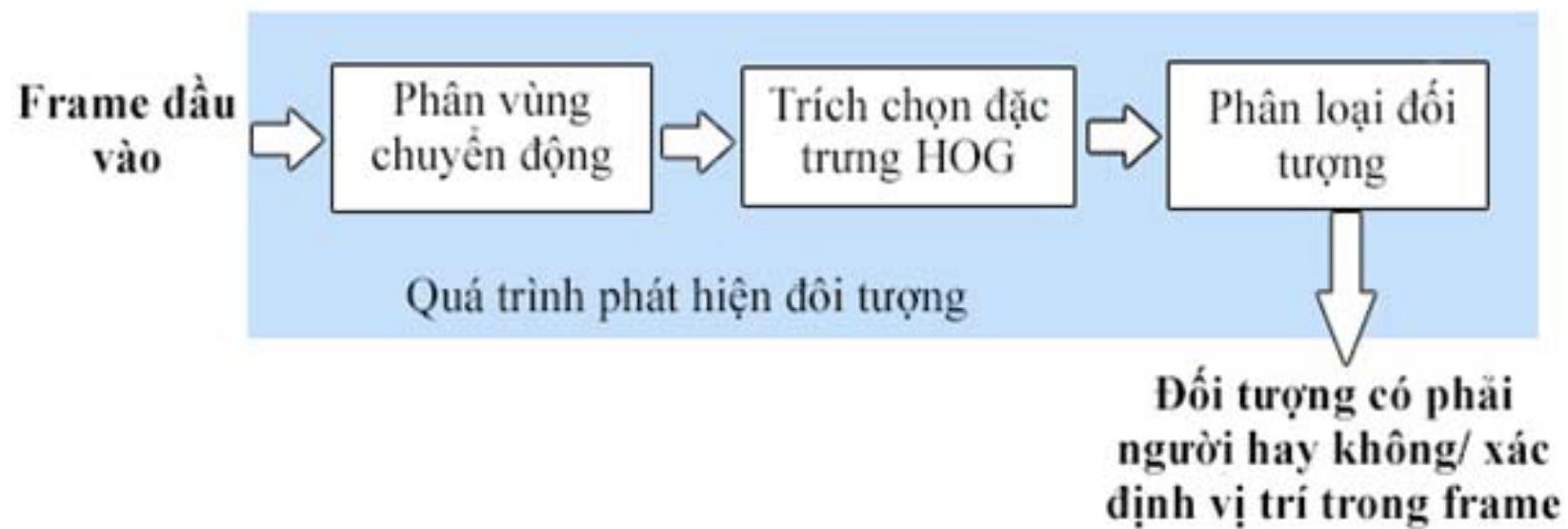
- **Nguyên tắc:** Sự khác biệt 2 khung hình theo thời gian tại mỗi thời điểm t:

$$D(t) = \text{frame}(t) - \text{frame}(t-1)$$

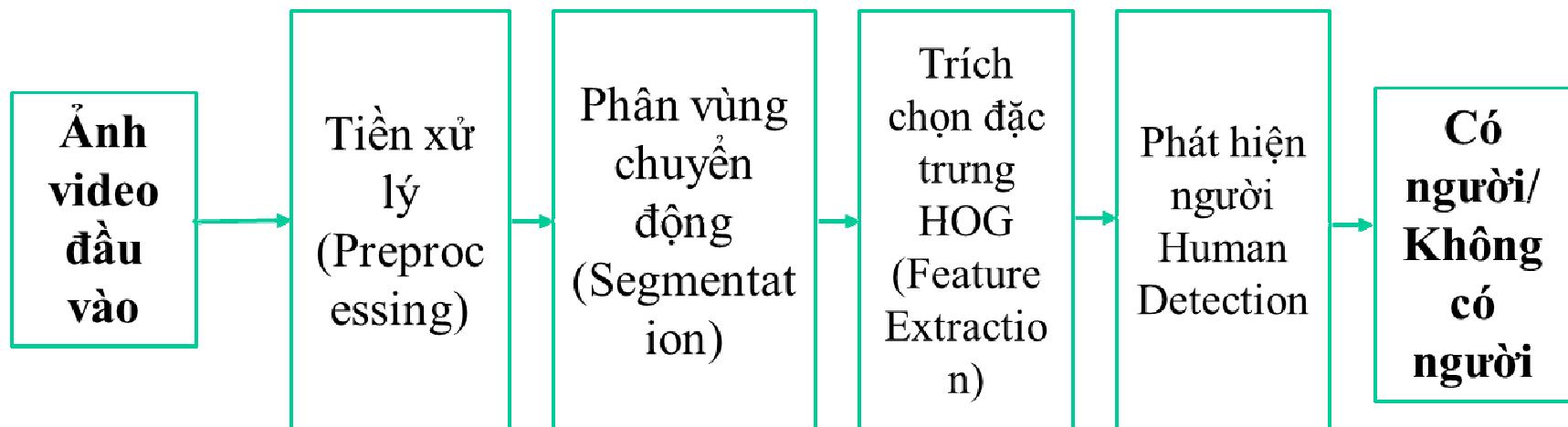


- Dựa trên $D(t)$ sẽ phát hiện và phân tách vùng chuyển động (Motion segmentation)

Hệ thống phát hiện người trong video (Human Detection in video)



Sơ đồ xử lý dữ liệu ảnh phát hiện người trong video

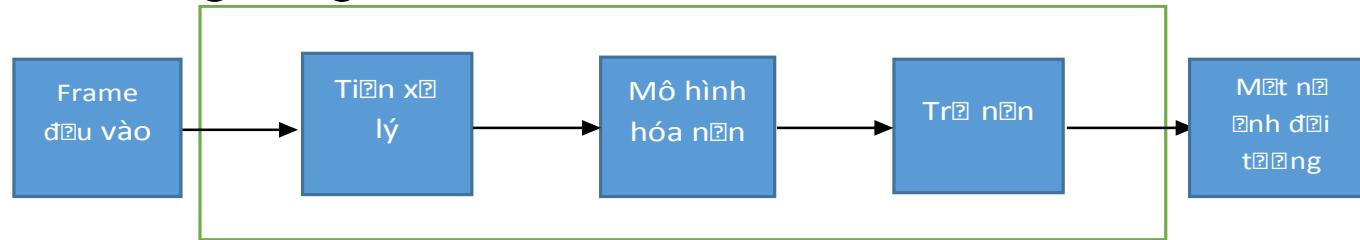


Ảnh video đầu vào:

- Ảnh video từ các bộ dữ liệu
- Ảnh video thu nhận từ camera

Thuật toán phân vùng và phát hiện chuyển động trù nền MOG

- **Motion segmentation:** MOG (Mixture of Gaussian) là một thuật toán trù nền sử dụng mô hình hóa nền có đệ quy để phân vùng các đối tượng chuyển động trong chuỗi ảnh video
- Sơ đồ hoạt động của giải thuật MOG



- MOG là một mô hình tham số, các mô hình tham số có khả năng thích nghi và cập nhật nhưng cần lưu trữ những vùng đệm lớn của video. Hàm mô hình nền frame I_t : $f(I_t=u)$ là sự pha trộn của K Gauss
- **Motion object detection:** Sau tìm được mô hình nền theo MOG, sẽ thực hiện trừ frame hiện tại với ảnh nền MOG và kết quả có được vùng chuyển động - mặt nạ ảnh các đối tượng chuyển động. Các đối tượng riêng biệt sẽ được phân tách từ mặt nạ theo phép xử lý liên kết các thành phần liên thông.

Sơ đồ hoạt động hệ thống

Hoạt động của hệ thống gồm 2 quá trình: Quá trình huấn luyện và quá trình phát hiện

