ComputerVision and Image Processing – Fundamentals and Applications
Professor Dr. M.K. Bhuyan
Department of Electronics and Electrical Engineering
Indian Institute of Technology, Guwahati
Lecture No. 39
Introduction to Video Surveillance
(Background Modelling and Motion Estimation)

Welcome to the NPTEL MOOC course on Computer Vision and Image Processing– Fundamentals and Applications. In this class I will briefly explain one important application of Computer Vision, that is, Video Surveillance. There are many applications of the Automated Video Surveillance System. So, I will explain about this, all these applications.

Also there are many research problems in developing an Automated Video Surveillance System. And mainly I will explain two important concepts. One is the Background Modelling and another one is Motion Estimation. In Motion Estimation I will explain one important algorithm, that is, Optical Flow. Let us see what are the important applications of a Video Surveillance System.

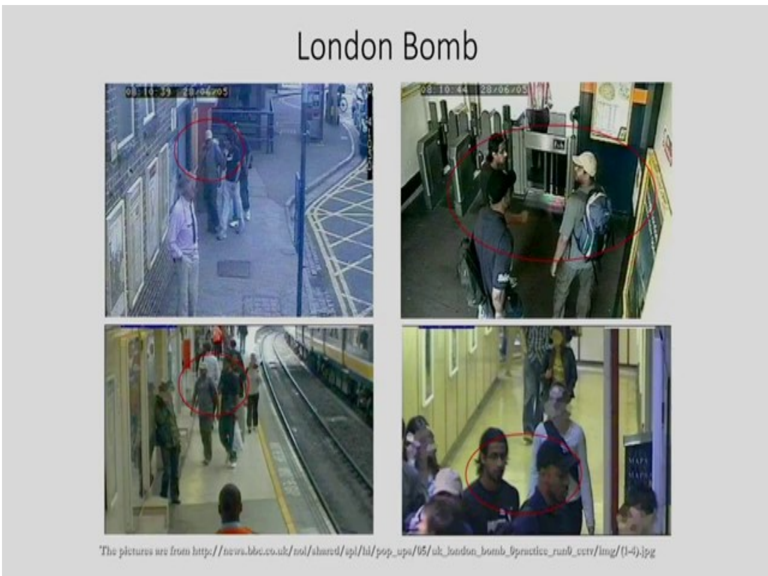
(Refer Slide Time: 01:24)

So, some of the important applications like, access control in special areas like in the airport, railways stations, maybe in the tunnels. So, this is one important application that is the access control in special areas. Anomaly detection and alarming, and another one is the Crowd flux statistics and congestion analysis. So, these are some applications of the Automated Video Surveillance system.

(Refer Slide Time: 01:53)

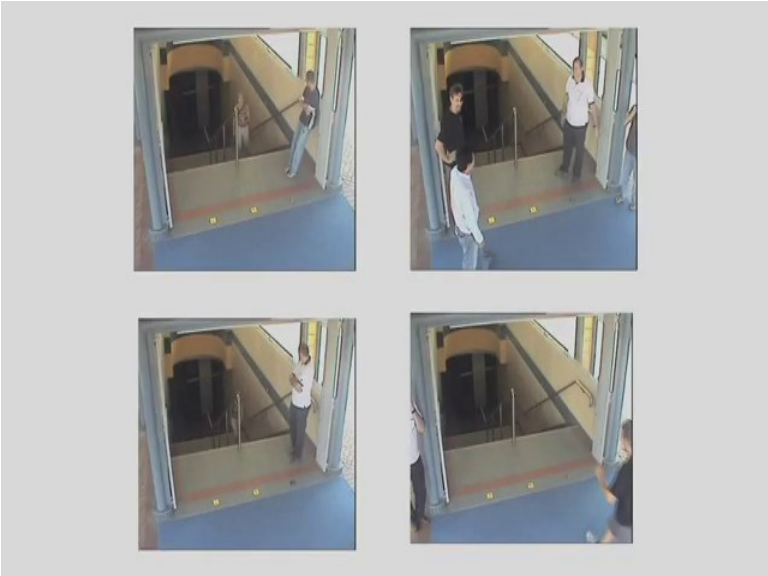
So, in the figure you can see, I have shown one example that is the anomaly detection and alarming. So, here I am considering, you can see the tracking that is people tracking. And from this I want to detect anomalies, and accordingly we can give some alarms based on the anomalies. So, this is anomaly detection and alarming.

(Refer Slide Time: 02:20)



And here I have shown one practical example that is the London bomb attack. Here in the figure you can see, I have shown some of the frames of the CCTV video sequence. And in this case you can see that the terrorists were detected by CCTV's. So, this is related to the London bomb attack. So, I have taken the pictures from this website and this is one application that is the anomaly detection.

(Refer Slide Time: 02:48)

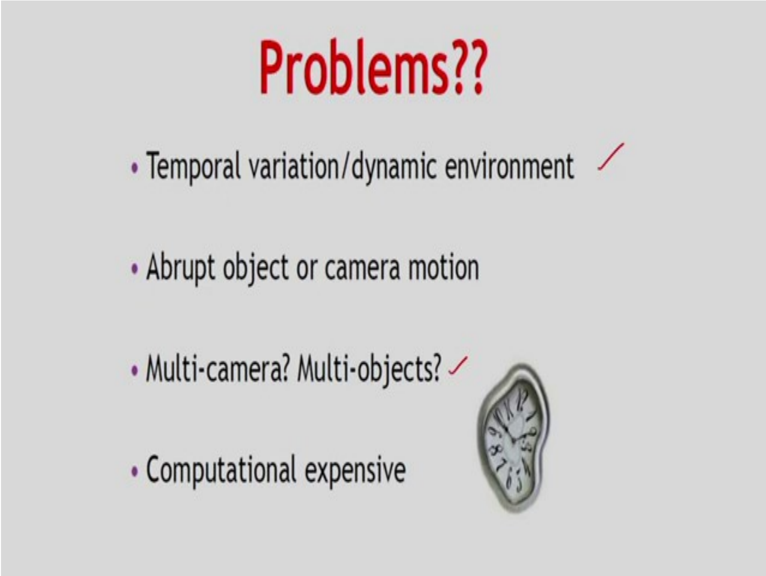


Next also I have shown some frames of a video sequence that is from the CCTV, and this is nothing but the anomaly detection.

(Refer Slide Time: 02:57)

And another application is the traffic information. So, I can track cars, I can track vehicles, and accordingly I can get the traffic information. So, this is also one important application.

(Refer Slide Time: 03:12)



But there are problems for this Automated Videos Surveillance system. The first problem is the temporal variation/dynamic environment. So, what is temporal variation? Suppose if I want to recognise activities, suppose the person is doing some activities then in this case we have to consider the spatio temporal variations, as I mention in the gesture recognition. In gesture recognition also we have to consider the spatio temporal variation. In this case also for human activity recognition or maybe the human action recognition we have to consider spatio temporal variations.

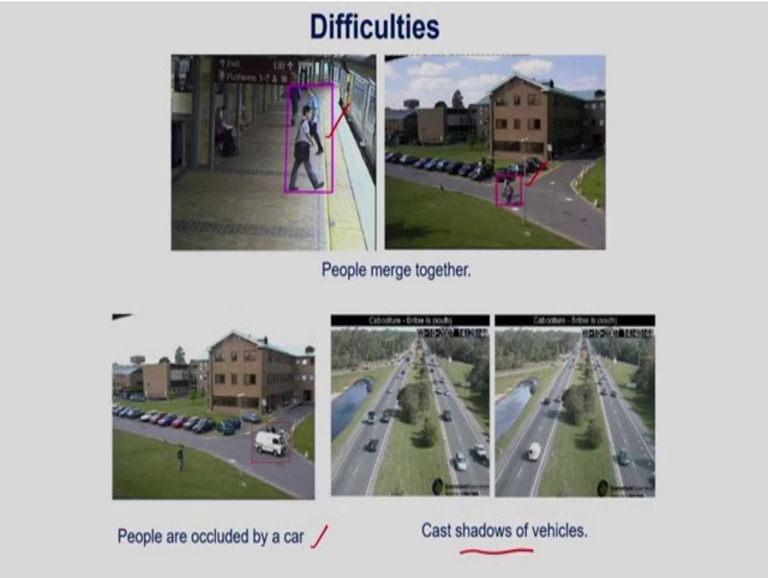
And also, whenever I want to detect a person or maybe a particular object, we have to do the segmentation, the segmentation of foreground from the background. But the problem is the background maybe cluttered background or maybe the illumination may changes or maybe the background maybe the dynamic background. Then in this case it is very difficult to do the background subtraction.

For all this conditions particularly the illumination variations, cluttered background, dynamic background it is very difficult to do the segmentation, the segmentation of the foreground objects from the background. So, that is the dynamic environment and also the temporal variations.

The next point is the abrupt object or the camera motion. So, there may be abrupt objects or the camera motions. Then in this case also we have to do the tracking, we have to do the background subtractions. So, all this things we have to do, but we have to consider this cases that is the abrupt object or the camera motion. And also we can employ multiple cameras for a video surveillance system. So, for multiple cameras the field of view will increase as compared to a single camera.

But the problem is I have to find the correspondence between all these cameras, the multiple cameras. So, that is one research problem, how to find the correspondence between all these cameras available in the video surveillance system. Also we need all the outputs in the real time so that is why one important issue is the computational complexity of the algorithm. So, these are the main problems for an automated video surveillance system.

(Refer Slide Time: 05:48)

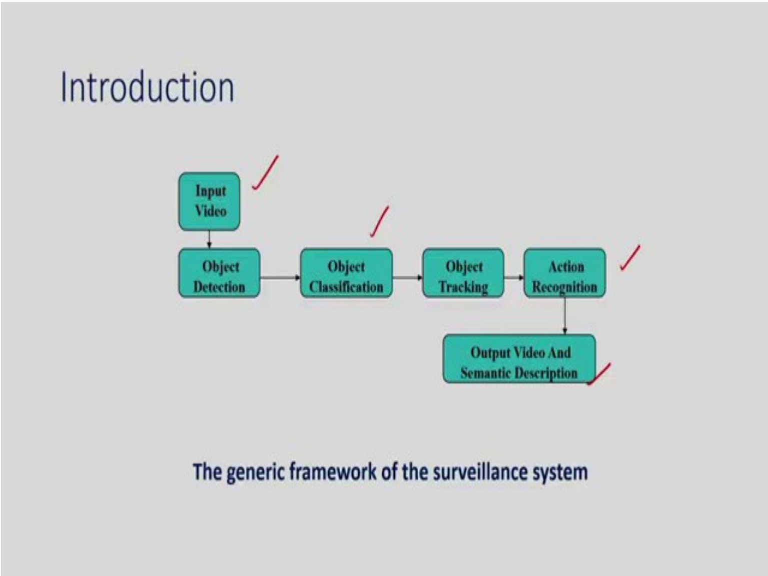


So, again I am showing some of the difficulties. In this case I am showing the concept of the tracking but in the image you can see the people merge together, that means there may be some occlusion. And because of this occlusion it is very difficult to track. So, here you can see that the people merge together. In this case also the people merge together. In this case we also have to do the tracking but it is a difficult situation. So, in this case we also have to do the tracking.

Next one is the people are occluded by a car and that is also another example of the occlusion. So, for a particular time maybe people are occluded by a car then also we have to do the tracking. So, this is one difficulty. And also if I consider the tracking of vehicles, then the problem will be the shadow. Even for tracking the people, the problem may be the shadow. So, the cast shadow that means the shadow of the vehicle, we have to consider. So, the vehicle will be moving at a particular speed and its shadow will be also moving.

So, in this case we have to do the segmentation, the segmentation of the vehicles from the background and after this we have to do the tracking. So, tracking in the presence of shadow is a big difficulty. So, we have to do this, we have to remove the shadow. So, that is also another research problem in video surveillance.

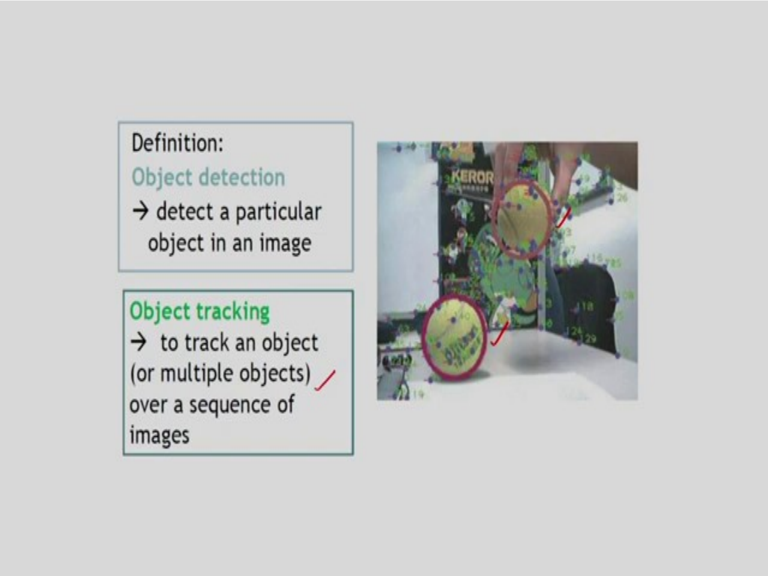
(Refer Slide Time: 07:18)



So, let us consider one generic framework of a video surveillance system. So, input is the video. After this we have to detect the objects present in an image, and after this we can employ some machine learning algorithms, the pattern classification algorithms for object classification. After this we can do the tracking, the tracking of the objects, and also we can mploy this one that is the action recognition, that is the activity recognition.

Suppose the people are walking or people are standing like this, so for this we can recognise different actions that are similar to the gate recognition, gate means the walking style. So, that is what we can do that is action recognition. The output is the output video and the semantic description. So, I will be getting the semantic description of the input video. So, this is the generic framework of a video surveillance system.

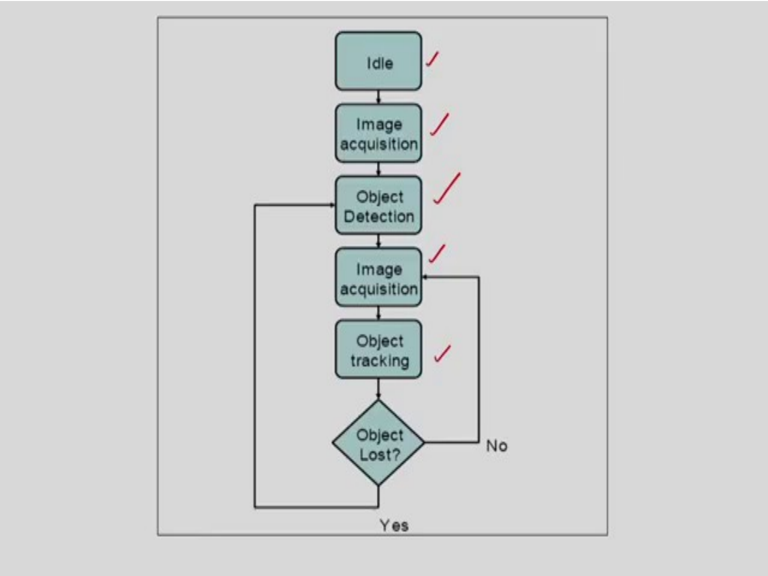
(Refer Slide Time: 08:20)



And you can see what is the difference between the object detection and the object tracking. So object detection means detect a particular object in an image. So, here you can see I am detecting this objects and that is nothing but the object detection. But, what is the object tracking? To track an object or maybe the multiple objects over a sequence of images that means, the sequence of images means the frames of a video.

So, in this case I have to find the correspondence between the frames of a video. That is to track an object, maybe a single object or maybe the multiple objects over a sequence of frames of a video that is the object tracking.

(Refer Slide Time: 09:07)

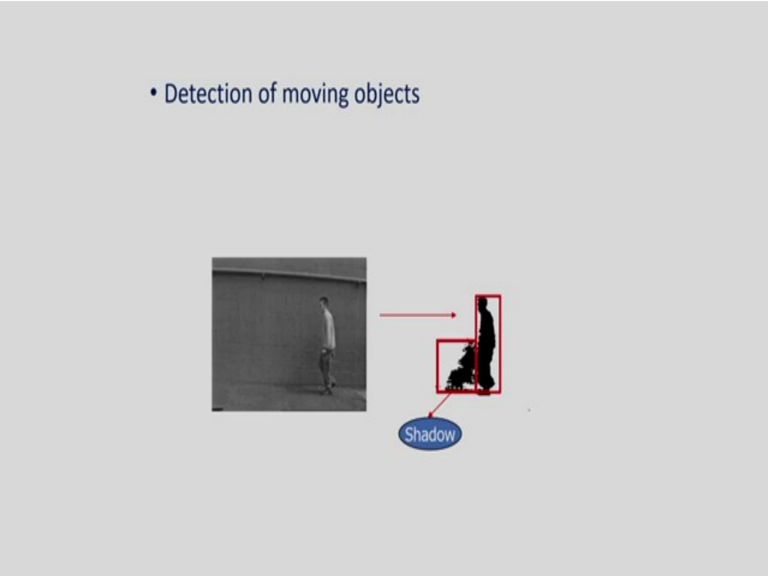


And how to detect the object? That I am showing the simple flowchart here. So, suppose the object is stationary, after this we are considering the image acquisition. After this we can detect object, so object detection algorithms are available. After this again the image acquisition from the video that is the frames I am collecting or the frames I am taking from the video.

And after this again I am doing the object tracking and suppose the object is lost in a particular frame, then again we have to detect the objects. And if the object is not lost then we can continue the tracking. So, you can see first I have to detect the object present in an image or present in a video, after this we have to do the tracking.

In some of the frames the object may get lost, then in this case again we have to acquire the frames of the video. And we can detect the object, and after this you can continue the tracking. Like this you can do, but the object, if it is not lost then we can continue tracking. So, this is a simple flowchart to detect objects and also do the tracking.

(Refer Slide Time: 10:29)

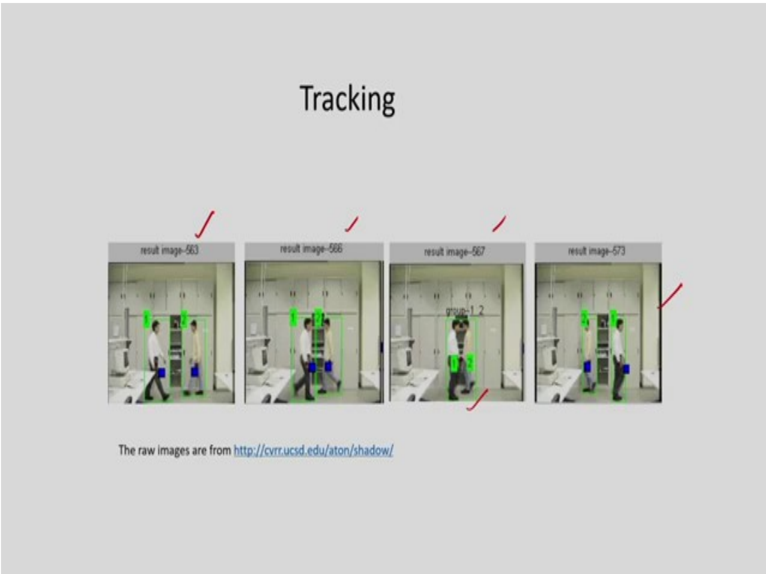


So, here you can see I am showing one example, the detection of the moving objects. And this person is moving, and we have to detect the moving objects based on some algorithms but the problem is the shadow. That is the car's shadow, we have to remove it. The person is moving and his shadow is also moving so that is why the problem is detecting the moving objects. So, that is why we have to develop some algorithms so that we can remove the shadow. That is the problem of detecting the moving objects in the presence of shadow.

(Refer Slide Time: 11:04)

After this object classification, so we can consider the classification of objects present in an image or maybe the video. So, here you can see we are classifying the objects, the car, the pedestrians, like this we can do the classification based on some algorithms. That is also very important. So, we can employ some machine learning algorithms for object classification.

(Refer Slide Time: 11:29)



And the tracking, already I have explained, it is nothing but finding the correspondence between the frames of a video. In this case I have shown some of the frames of a video, so here you can see the 563, that is the frame number, 566, 567, like this, these are the frames. And you can see the person one is moving and person two is also moving and there maybe occlusion in this case, you can see in the image 567 there is an occlusion. And in the presence of occlusion also we have to track the person, the person 1 and person 2.

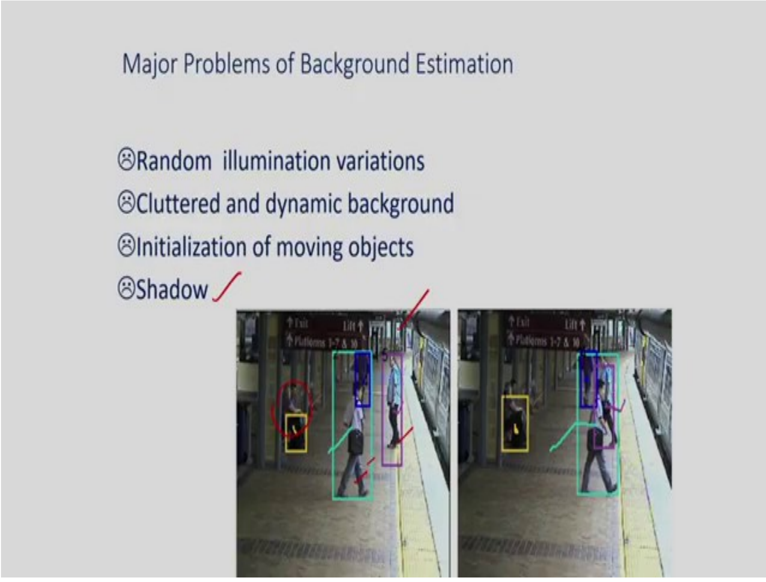
And you can see the output of this, the tracking algorithm. So, we can track person 1 and person 2 in this video. So, tracking means finding the correspondence between the frames of the video.

(Refer Slide Time: 12:21)

So, some of the traditional methods for all this activities, for moving object detection we can employ the Background subtraction algorithm and also we have to remove the shadow that is the Shadow removal algorithm we have to employ. For object classification simply we can employ the shape information that is the Shape-based classification we can do. And for tracking of the objects we can employ the Model-based tracking and also the Feature-based tracking.

Incase of the Model-based tracking, we can consider the model of the object, maybe the geometrical model we can consider, and based on this model we can do the tracking. In case of the Feature-based tracking, we can consider some of the features, maybe the color feature, maybe the texture. So, all this features we can consider and based on features we can do the tracking. So, these are the traditional methods employ in the video surveillance system.

(Refer Slide Time: 13:19)



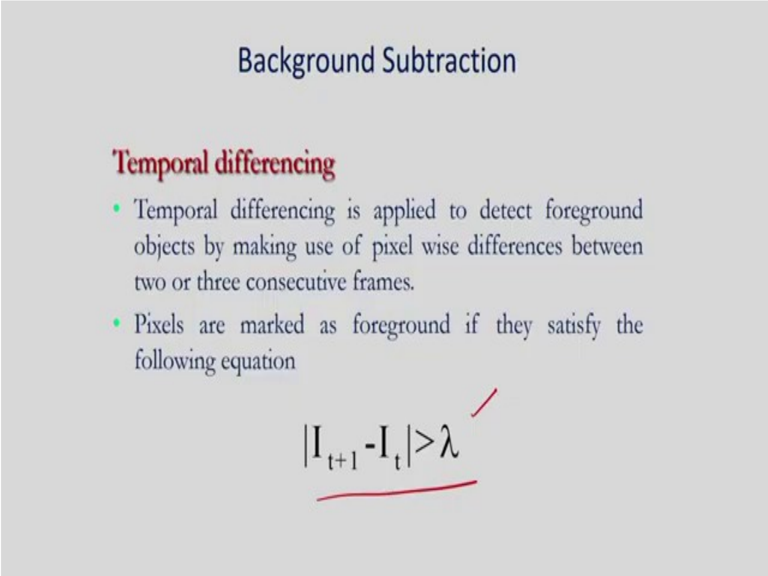
And the major problem of the background estimation, so one is the random illumination variation, so that we have to consider. So, illumination is not uniform. So, if there is random illumination variation then the problem of the background subtraction, the background estimation. And the background may be cluttered and also the dynamic background.

So, in this image you can see the background is very cluttered and one train is moving that means the background is dynamic. And we want to track these persons, you can see I am doing the tracking in the presence of the moving background and cluttered background, and also the illumination is not constant, not uniform. Random illumination variations are available in these images in the video. And also we have to do the initialization of the moving object.

In this video or in these images or in these frames of the video, you can see, you can see we have the moving objects, this person is moving, this person is moving. So, we have the moving objects, and also some persons they are sitting here that means they are stationary. So, we have to initialize the moving objects.

And also as explained earlier so we have to consider the shadow problem. And one difficult problem is the removal of the cast shadow. Because the shadow is also moving and the person is also moving, so based on the motion information I cannot do the separation between the moving objects and the shadow. So, we have to employ some algorithms so that we can remove the cast shadow.

(Refer Slide Time: 15:00)



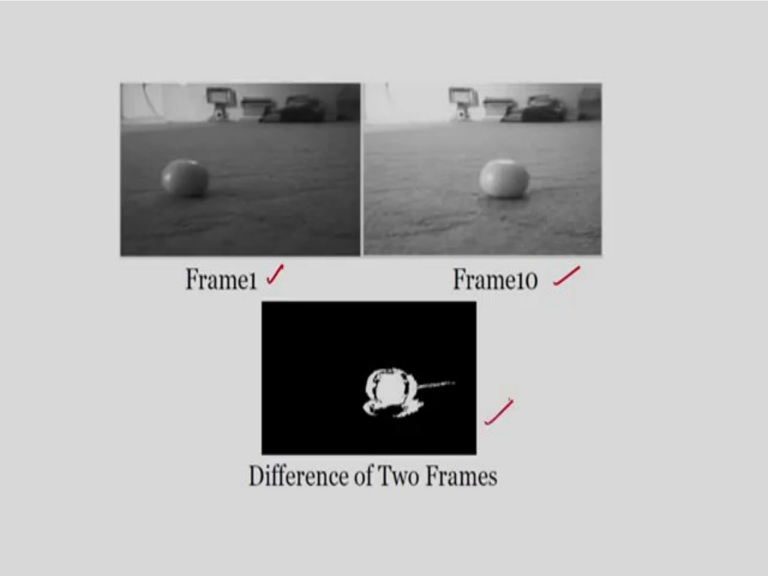
Now I will discuss the concept of the background subtraction. So, one very old algorithm is the temporal differencing. So, this algorithm I can apply to detect foreground objects by making use of pixelwise differences between 2 or 3 consecutive frames of the video. That means I have to find the differences between 2 or 3 consecutive frames.

And the pixels are marked as foreground if they satisfy the following equation. Suppose if I consider a particular pixel, the pixel at the time $t+1$ is $I_{t+1} - I_t$ that is that pixel at the time t. If the difference between these is greater than a particular threshold, then we can consider that pixel as a foreground pixel otherwise it is a background pixel. Like this we can determine the moving pixels.

(Refer Slide Time: 16:00)

In this case I am showing one example foreground extraction that is the background subtraction that is nothing but the sense detection algorithm. So, this is the current image and I am considering the background image. If I subtract the background image from the current image, I will be getting the foreground pixels. So, this is nothing but the background subtraction algorithm.

(Refer Slide Time: 16:22)

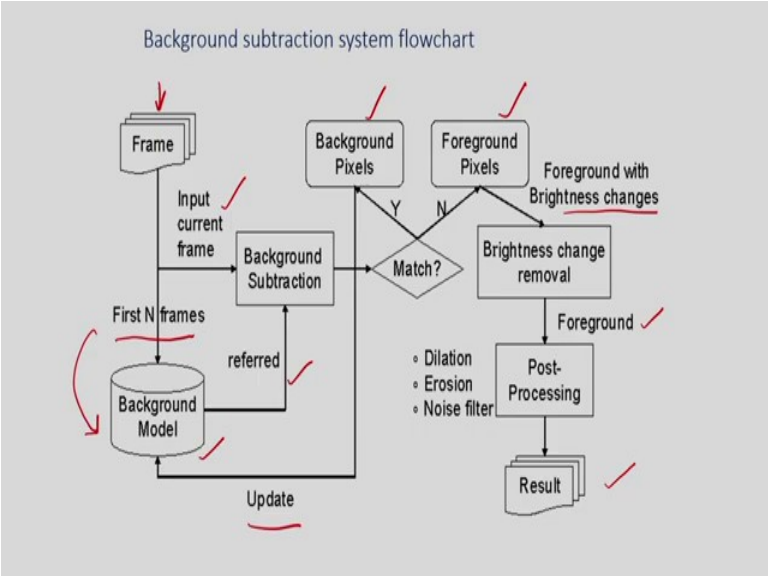


So, again I am showing you the same thing the background subtraction algorithm. So, this is frame one and another one is the fram ten. So, if I do the subtraction, then I will be getting the difference of the two frames that are nothing but the moving objects I can determine.

(Refer Slide Time: 16:38)

Similarly I am showing one example of the absolute differences. So, two frames I am considering. So, you can see in the first frame the moving objects are present, in the second frame the background that is nothing but the background. And if I subtract the second image from the first image, I will be getting the moving objects. That is the moving objects will be detected. That is nothing but the absolute differences.

(Refer Slide Time: 17:03)



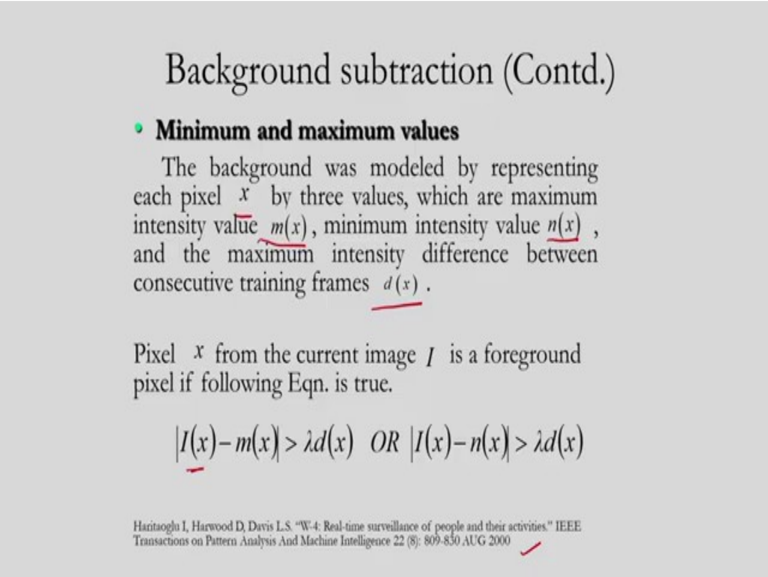
So, in this example, I have shown one algorithm that is a simple algorithm for background subtraction. So, background subtraction system flowchart. So, suppose I have the input video and I am considering the frames of the video. First I am considering the N frames…the first N number of frames I am considering and based on this I am determining the background model. So, the background model is determined from the first N number of frames.

After this suppose one input frame is coming that is the input current frame. And I can employ the background subtraction algorithm, so you can see I have the background image here, and I have the input current frame so I can employ the background subtraction algorithm and based on this I can determine the foreground pixels and the background pixels I can determine based on the matching.

And this, based on the background pixels I can update the background model that already I have developed from the first N number of frames. Also I have to consider the brightness since removal, so there may be changes of the brightness so that also we have to consider and I will be getting the foreground.

After this we have to do some post processings, so may be we can employ the morphological image processing operations like the dilation, erosion, noise filter like this we can employ. And ultimately I will be getting the output that is nothing but the foreground image I will be getting. So, I can separate foreground and the background by considering this background subtraction algorithm.

(Refer Slide Time: 18:41)

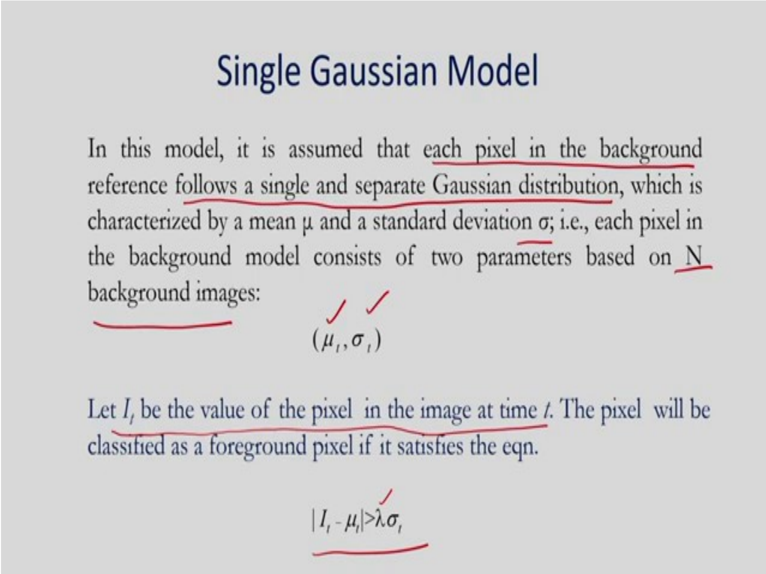


And this is also a very old technique; here it is published here that is IEEE Transactions on Pattern Analysis and Machine Intelligence. So, briefly the concept is like this, so the background is model by representing each pixel, the pixel is x by three values, so 3 values are considered. One is the maximum intensity value $m(x)$, one is the minimum intensity value $n(x)$and the maximum intensity difference between the two consecutive training frames that is the dx.

So, by using this three values, one is the maximum intensity value, another one is the minimum intensity value and another one is the maximum intensity difference between consecutive training frames, I can develop the background model. So, how to do this? So, pixel x from the current image I is considered as a foreground pixel, if the following equations or the following equation is true. That means $I_x$that is the current pixel – $m(x)$, what is $m(x)$ that is the maximum intensity value $m(x)$ is greater than lambda dx or maybe $I_x - n(x)$, $n(x)$ is nothing but the minimum intensity value, is greater than $\boldsymbol{\lambda} dx$.

So, based on this I can determine the foreground pixels. So, briefly this is about the algorithm, so based on this I can determine the foreground pixels.

(Refer Slide Time: 20:20)

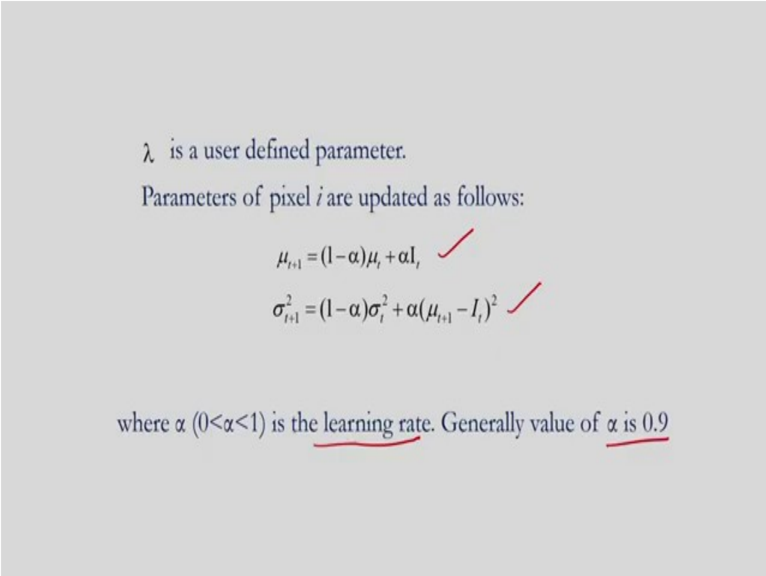


And this is also a very popular technique for background estimation, background modelling. So, we can employ a Single Gaussian model. So, in this case what we can consider, each pixel in the background reference follows a single and separate Gaussian distribution. So, that means each pixel in the background follows a single and separate Gaussian distribution. So, that is for a particular pixel we can consider Gaussian distribution and which is characterized by the mean and the standard deviation. So, that means each pixel in the background model consists of two parameters based on N number of background images.

So, we can consider N number of background images, so from N number of background images we can determine the mean and the standard deviation that we can determine. That is nothing but the background model of the image. And suppose $I_t$be the value of the pixels in the image at the time t, so suppose I am considering the pixel at the time t. The pixel will be assigned or may be classified as a foreground pixel if it satisfies the following equation. So, this equation should be satisfying.

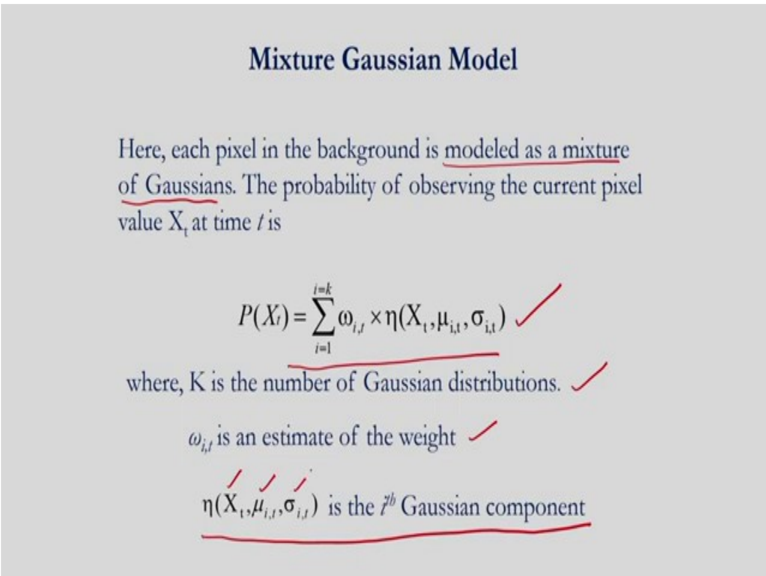
So, I_t means the value of the pixels in the image at time $t -$¿ the mean is greater than $\boldsymbol{\lambda}\sigma_t$. If this condition is satisfied then I can consider that particular pixel as a foreground pixel. So, lambda is a user defined parameter and I am considering the difference between the pixel and the mean that I am considering. $I_t - \mu_t$if it is greater than $\boldsymbol{\lambda}\sigma_t$ then we can consider that particular pixel as a foreground pixel.

(Refer Slide Time: 22:15)



So, lambda is a user defined parameter. And after this the parameters of the pixel I should be updated. Again we have to update the parameters corresponding to the pixel I so the mean is updated like this and also the variance I can update like this. So, these are the updating equations. And in this case the alpha is a parameter and that is used for the learning rate. So, I can control the learning rate based on the value of the parameter alpha. And generally the value of alpha is 0.9. So, this is the concept of the Single Gaussian model.

(Refer Slide Time: 22:54)

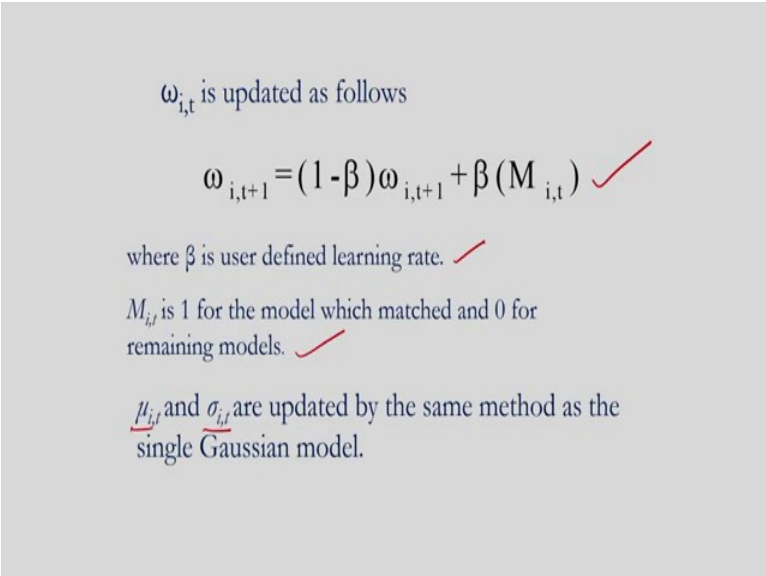


The next one is I can consider the mixture of the Gaussian model that is the Mixture Gaussian Model. So, what is the concept of this? So, each pixel in the background is modelled as a mixture of Gaussians. So, I can consider the pixel as a mixture of Gaussian because a Single

Gaussian model may not be appropriate in many cases, suppose the background is moving or may be the illumination is changing, then in this case the Single Gaussian model may not be appropriate. So, for this we can consider a mixture of Gaussians.

And the probability of observing the current pixel value of X t and the time t is given by this, so the probability of X t is given by this. And here the K is the number of Gaussian distributions. So, how many Gaussian distributions are employ to represent a particular pixel, or maybe the background. And w i t is an estimate of the weight. So, w it is nothing but the weight. And you can see I am considering the Ith Gaussian components. So, X t is nothing but the pixel value and I am considering the mean and the standard deviation.

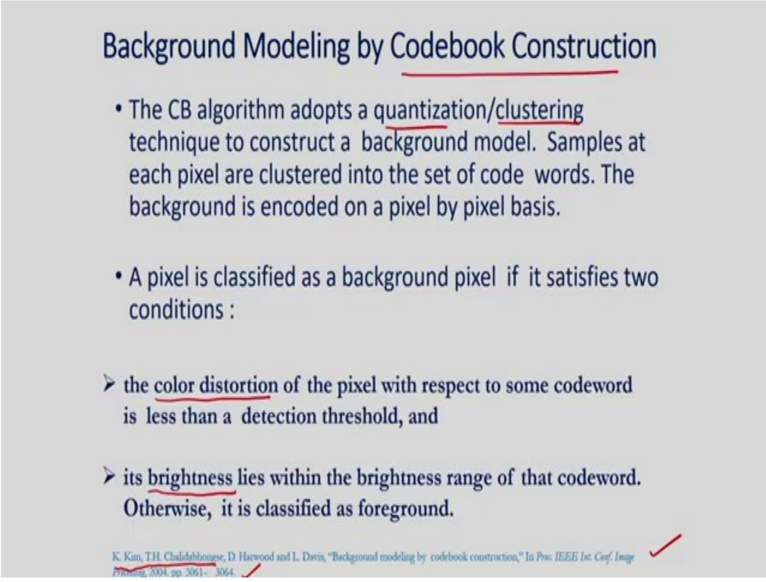
(Refer Slide Time: 24:04)



And after this we have to update the weights by using this equation and here the beta is the user defined parameter that is nothing but the learning rate and M_{it} is 1 for the model which is matched and 0 for the remaining models that is about the M_{it}. And this mean, mean and the standard deviation or the mean and the variance should be updated like we did in the Single Gaussian Model.

So, this is the concept of the Mixture Gaussian Model. So, briefly I have explained this concept the Mixture Gaussian model.

(Refer Slide Time: 24:39)



And this is also a very popular method, the Background Modelling by Codebook Construction. You can see this research paper by Kim. So, this research paper you can see and I think this is one of the popular methods but this is a very old method now. In this case the codebook algorithm is employ that is nothing but the codebook construction and in this case the concept of the quantization or the concept of clustering is employed to construct a background model.

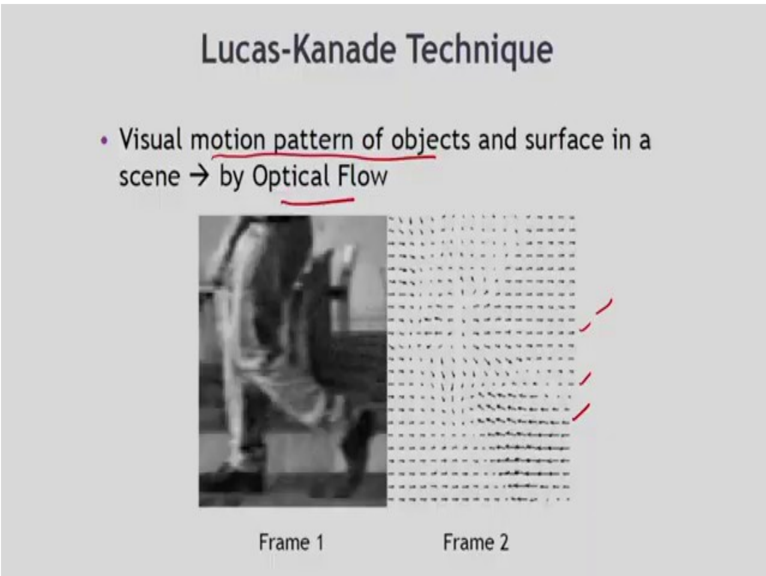
So, samples at each pixel are clustered into the set of code words. So, each pixel means if I consider samples of the pixel from the training image, they are clustered into the set of code words. And the background is encoded on a pixel by pixel basis. So, for each pixel I have the code word and the background is encoded on a pixel by pixel basis. A particular pixel can be classified as a background pixel if it satisfies the following two conditions.

So, first we have to consider the case of the color distortion, so based on the color distortion we can take a decision whether particular pixel is a foreground pixel or the background pixel. And also we have to consider the brightness difference. So, brightness lies within the brightness range of the code word. That also it is considered and based on this we can select the foreground and the background pixels. So, this is the briefly the concept of the Background Modelling by Codebook Construction. For more details you can read this paper.

(Refer Slide Time: 26:20)

So, these are some results we have done, Background Modelling by the Codebook technique. So, corresponding to this image you can see I am extracting the foregrounds. And similarly to this image I am extracting the foregrounds. Very difficult images in the second image because the background is you can see dynamic background is also there and also illumination variations are also there. And the background is very complex. So, these are some results of the background estimation that is the foreground separation from the background.

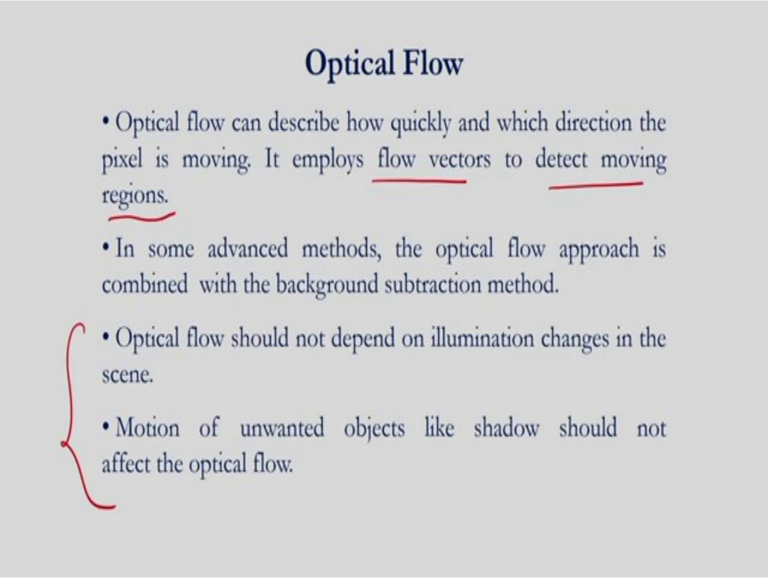
(Refer Slide Time: 26:54)



And one important technique is the Lucas- Kanade technique. That is the concept of the optical flow. So, we can determine the visual motion pattern of the moving objects and that is called the optical flow. So, here you can see this person is moving and corresponding to this motion. I can determine the visual motion pattern I can determine. That means I can determine the motion factors. So, here you can see these are the motion factor that is nothing but the optical flow.

So, based on the optical flow determination, I can determine the moving objects that are nothing but the motion estimation. So, this algorithm I am going to explain is the motion estimation by optical flow.

(Refer Slide Time: 27:37)



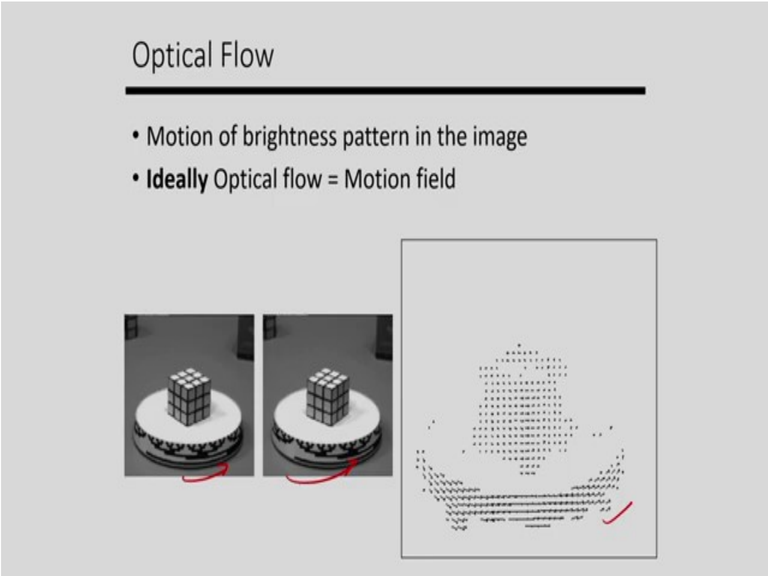
So, what is optical flow? This concept is very important. So, optical flow can describe how quickly and which direction the pixel is moving. So, that means I am determining the motion pattern that is the visual motion pattern can be determined by the optical flow algorithm.And we can determine the motion factors. That is nothing but the flow factors to detect moving regions.

And the optical flow approach can be combined with other background subtraction methods. So, by combining this optical flow method and the other background subtraction method we can determine the moving regions that motion estimation we can do, and we can separate the foreground from the background. And I have to consider these assumptions.

One is the optical flow should not depend on illumination changes in the scene. That means the illumination should be almost constant for optical flow determination. That is one assumption. The second assumption is the motions of the unwanted object like shadow should not affect the optical flow. So, that also we have to consider that means we have to avoid or we should not consider the unwanted objects and is the shadow is present that will be a big problem for optical flow algorithm.

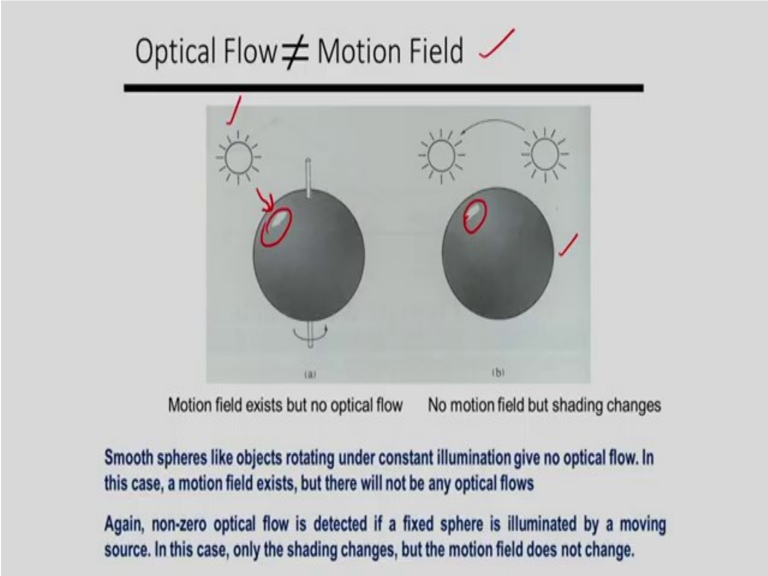
So, that we are not considering the presence of the shadow or the presence of the unwanted objects. So, these two assumptions we are considering for determination of the optical flow. Now let us see how to determine the optical flow.

(Refer Slide Time: 29:13)



Here you can see in this example I have shown this object is rotating, it is rotating like this in the counterclockwise direction and corresponding to this I will be getting the optical flow that is the motion pattern I will be getting. And ideally, the optical flow is equal to motion field, because the object is also moving and corresponding to this I will be getting the motion factors that is the motion field I will be getting and also the optical flow I will be getting.

(Refer Slide Time: 29:42)



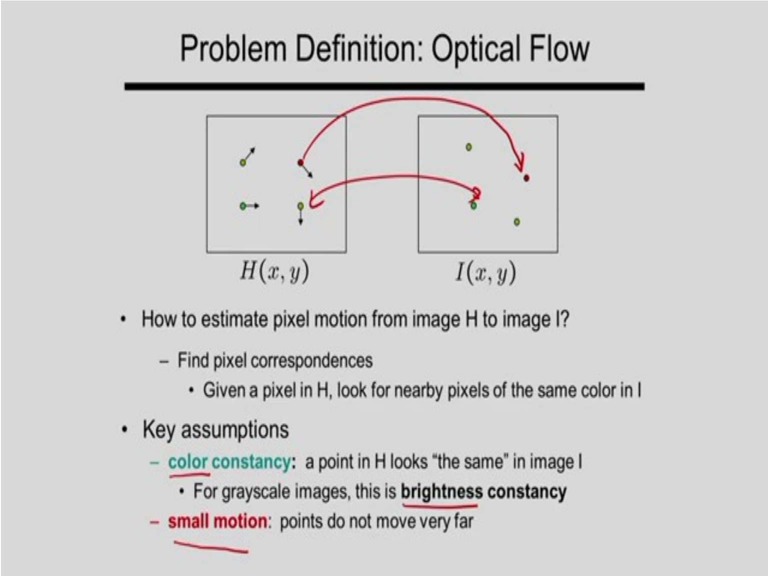
But sometimes, it may not be the optical flow may not be equal to motion field. Here I am giving two example, in the example a, you can see, this object is rotating this is rotating and I have the source, one source is available. So, corresponding to these paths, suppose these paths I am considering, the illumination will be almost constant that means if the illumination is constant then corresponding to this there will not be any optical flow.

Because the optical flow, how to determine? Optical flow is determined based on the illumination change, because the object is moving and based on this the illumination at different points will be changing. But in this case, the illumination at this particular point will be almost constant.

In the second case, there is no motion you can see. The object is in stationary that is the sphere is in the stationary, but I have the moving sources, the source is moving. So, corresponding to this path suppose, I have different brightness. The brightness will be changing. So, whenever the brightness is changing, then I will be getting the optical flow. But I will not be getting the motion because there is no motion field here, the motion field is 0, so inspite of having the motion is 0, the optical flow is present.

So, that means no motion field, but shading changes that means the brightness corresponding to that particular paths, it will be changing because the sources are moving. So, in the first case motion field is there, but there is no optical flow. In the second case, motion field is not there, but there is a change of the brightness and because of this I will be getting the optical flow. So, that means I can say the optical flow is not equal to motion field.

(Refer Slide Time: 31:38)

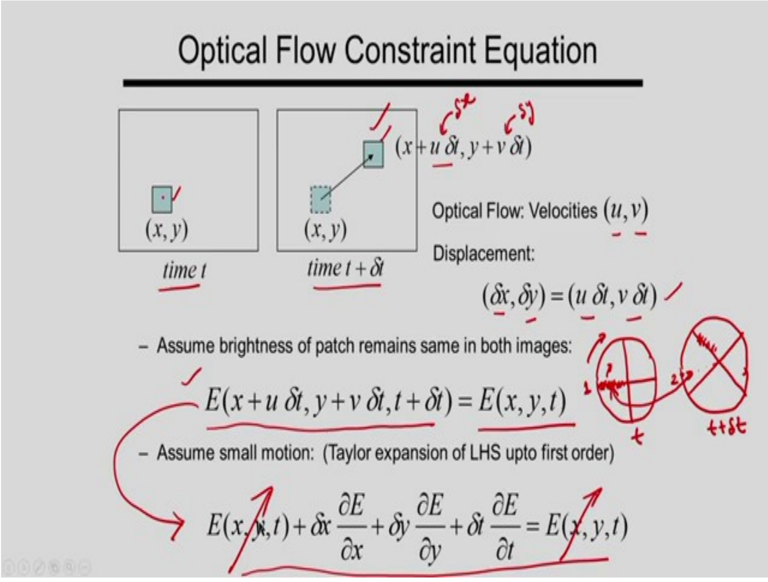


And the problem definition of the optical flow, so here I have shown two images, one is H another one is I. And I have shown some pixels, the red pixels, the green pixels like this I have shown. And the pixels are moving, so I have to determine the motion of the pixels that is nothing but the motion estimation.

So, for this you can see, find the pixel correspondence. So, I have to find the correspondence between the pixels and based on this I can determine the optical flow, I can determine the motions. So, suppose this red pixel is corresponding to this red pixel and similarly this pixel is corresponding to this pixel like this. So, I have to find the correspondence between the pixels.

For this I can consider some assumptions like the color property I have to consider or may be the brightness property we can consider for the gray scale image. And also we have to consider this assumption that is the small motion assumption we have to consider. Points do not move very far. So, for this we have to find the correspondence between the pixels of the images. So, the actual concept of the optical flow, I will be explaining in the next slide.

(Refer Slide Time: 32:51)



Here I have shown two frames of a video at time t another one is time $t+\delta t$. Let us consider a paths at the point x, y. So, these paths I am considering. And other time $t+\delta t$, this path moves to this position, second position. So, these paths will be moving suppose. Now what will be the position of the paths in the second image? The position of the paths in the second image $x+u\delta t, y+v\delta t$

So, in this case, what is $x+u\delta t$? $u\delta t$ is nothing but δx. So, that means it is nothing but δx. And similarly, the $v\delta t$ is nothing but δy. So, what will be the position of the paths in the second frame? The position of the paths in the second frame will $x+\delta x$, $y+\delta y$. And I am considering the velocity that is the optical flow velocities u and v. So, u is the velocity along the x direction and v is the velocity along the $y$direction.

So, from u and v I can determine the displacement. So, δx is nothing but $u\delta t$ and δy is nothing but $v\delta t$ that I can determine. So, assume that the brightness of a path remains same in both the images, that is true here. You can see this path is moving to this position; the second position. So, that means the brightness of this path will remain same in both the images.

Only the position is changing, but the brightness of the path remains the same in both the images. So, here you can see that is the brightness at the point x comma y at the time t. So, E is the brightness of the paths at the point x comma y and at the time, the time is t. And since the paths is moving to the another position in the time t plus delta t, the same paths is moving, so there will not be any brightness since.

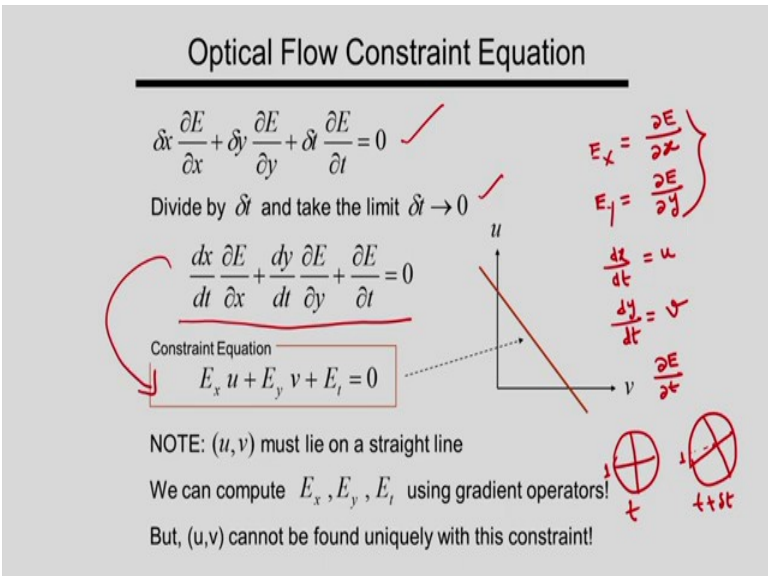
So, corresponding to this…what will be the brightness? The brightness will be $E(x, u\delta t)$ because that is the new position $x+\delta x, y+\delta y$ and what is the time, the time is $t+\delta t$. So, brightness will remain same. So, I can show another diagram, suppose I am considering one object and it is rotating suppose. So, this object is rotating like this. So, this is at time t and this is at $t+\delta t$.

So, if I consider the brightness at this point suppose, will be same as that of this, because this point is moving and this is moving to this point, so brightness of this will be same as that of this. And if I consider the same position, so brightness at this position, the position number 1 and the position number suppose 2, the same position; it will be changing because of the rotation.

So, from this I can determine the optical flow. Here you can see corresponding to the position 1 at the time t, and corresponding to the position 2 at the time t plus delta t, the brightness will be different because of the rotation. So, based on this concept I can determine the optical flow because the brightness at the same point, because this point and this point is same point 1 and 2 is the same point so brightness will be different because of the rotations. So, that gives the optical flow. So, now you can see the assume the brightness of the paths remains same in both the images, then I will be getting this equation.

And this I can expand by considering the Taylor series expansion. So, higher order terms I am neglecting so I will be getting this one. So, this equation I will be getting and I can cancel out this, this and this will be cancelled out.

(Refer Slide Time: 37:04)



So, from this what I will be getting? I will be getting this equation.

$$\frac{dx}{dt}\frac{\partial E}{\partial x}+¿\frac{dy}{dt}\frac{\partial E}{\partial y}+\frac{dx}{dt}\frac{\partial E}{\partial x}+\frac{\partial E}{\partial t}=0.$$

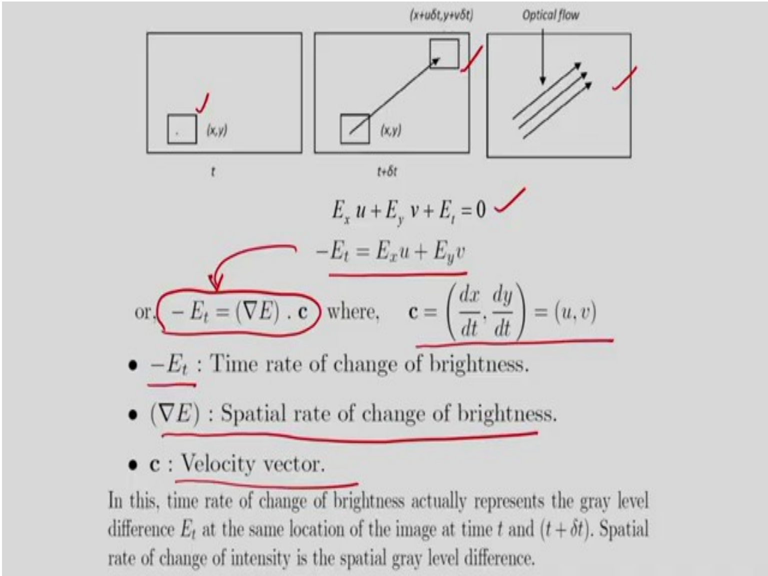
So, from this I will be getting this equation. So, I have to divide by delta t so I will be getting this equation. And this equation can be represented like this. So, I will be getting this one the $E_x u+E_y v+E_t=0$.

So, what is E_x? What is E_x in this case? E_x is nothing but $\frac{\partial E}{\partial x}$ that is the change of brightness with respect to the coordinate, the coordinate is x. That is the spatial change of the brightness. And similarly, if I considerE_y, that is nothing but $\frac{\partial E}{\partial y}$. This corresponds to the spatial change of the brightness, because of the motion. And what is the $\frac{dx}{dt}$? $\frac{dx}{dt}$ is nothing but$u$ that is the velocity along the acceleration. And what is the d$\frac{dy}{dt}$? So, $\frac{dy}{dt}$is nothing but velocity along the y direction that is v.

And what is delta E delta t that is the time rate of change of brightness. The brightness is changing with respect to time. Because already I have explained, suppose this object is moving like this at time t and at the time $t+\delta t$. So, this is the $t+\delta t$, so you can see the brightness at this point, so brightness at this point suppose one and one, it will be changing. So, change of the brightness with respect to time, so that is $\frac{\partial E}{\partial t}$.

So, I will be getting this equation, this is the equation of a straight line so the velocity u and v must lie on a straight line and based on this we can compute E_x, E_y and $E_t$using the gradient operators. But the problem is it is not a simple problem, u and v cannot be found uniquely with this constraint. So, with this constraint cannot determine the velocity u and v.

(Refer Slide Time: 39:28)



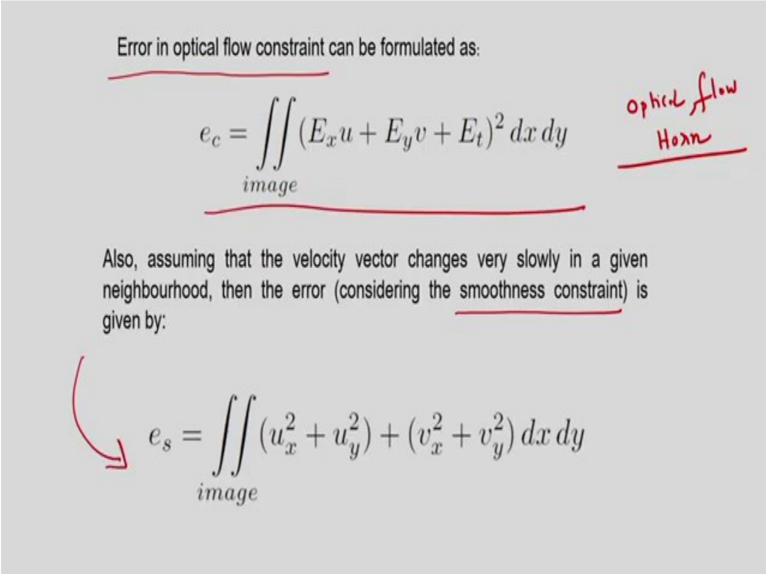
So, the same concept I am showing here, so this path is moving to this position and corresponding to this I will be getting the optical flow. The optical flow it depends on the change of the brightness because of the motion. So, already I have this equation, so this equation I can write like this. So, what is the interpretation of this equation?

You can see I will be getting this equation from this minus E_t is equal to that is the ∇E that is the gradient of E dot c. So, c is nothing but the velocity factor. So, c is the velocity factor, so I have two components, one is the u and another one is v, velocity along the x direction and the velocity along the y direction.

And what is Et? Et means the time rate of change of brightness, so with respect to time the brightness will be changing because of the motion. That is the time rate of change of brightness. And what is this ∇ E? $\nabla E$is nothing but the spatial rate of change of brightness.

So, one is the velocity factor, so our objective is to determine the velocity factor from this equation. So, this is the equation, this is the constraint equation or the optical flow equation, so from this equation, we can determine the velocity factor. The velocity factor is c, so I have two components u and v.

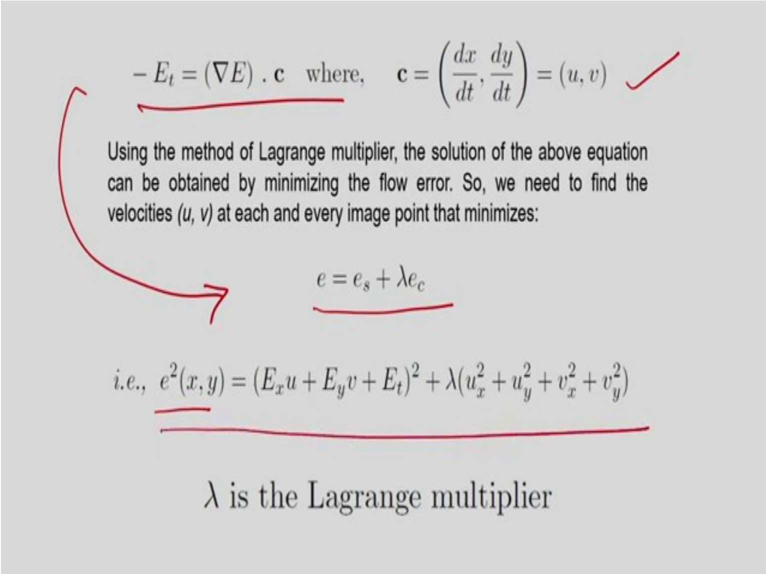
(Refer Slide Time: 40:53)



And for optical flow solution, there are many methods, so simply I am showing one method but if you want to see the solution of this method, so you can see the optical flow paper; research paper by Horn, the original paper is by Horn, so you can see this research paper, the optical flow algorithm you can see by Horn, but briefly I have explained the solution of this equation and that is the optical flow equation.

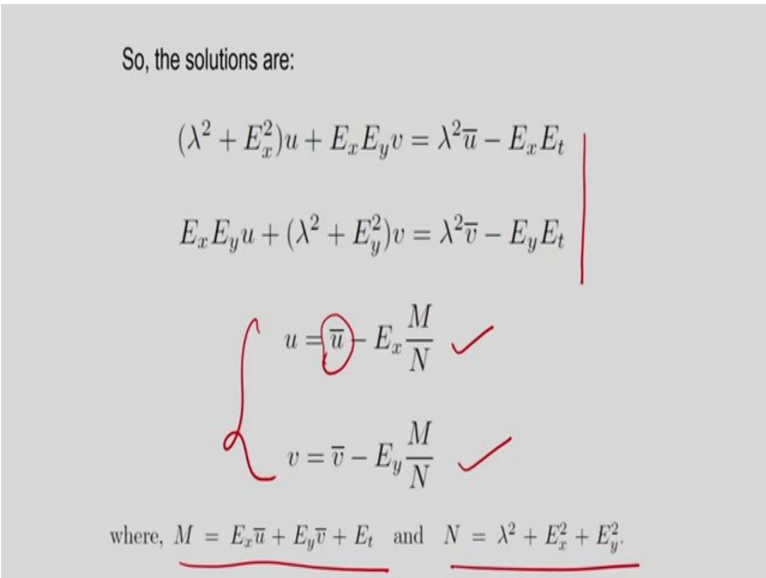
So, for this we have to consider error in optical flow constraint. So, this error I am considering and also we have to consider another assumption that is the velocity factor changes very slowly in a given neighbourhood. That means the motion is very slow. So, this assumption I have to consider that is nothing but the smoothness constraint we have to consider and based on this I am having this equation that is nothing but the smoothness constraint.

(Refer Slide Time: 41:51)



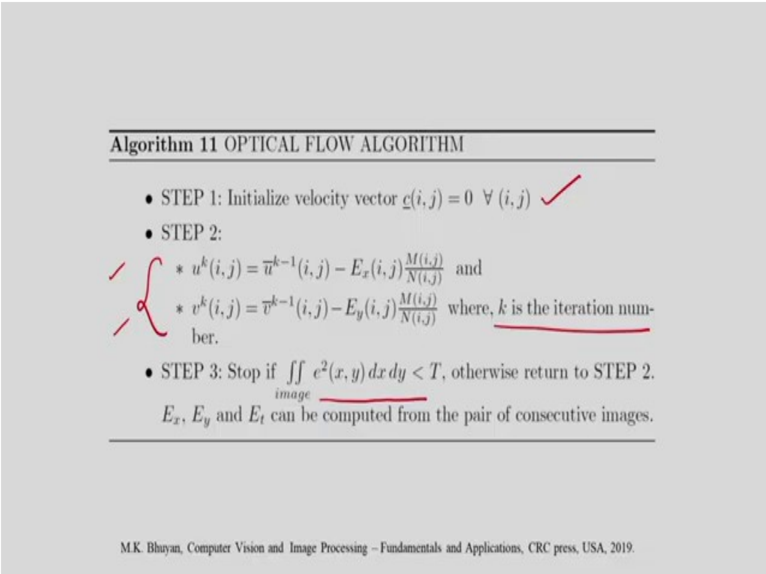
So, already you ave this quation that is optical flow equation and this equation we can solve by considering the Lagrange's multiplier method. So, $e = e_s + \lambda e_c$ that you can consider. So, lambda is nothing but the Lagrange's multiplier. And based on this error, the flow error already we have calculated, because e is equal to $e_s + \lambda e_c$, we have already defined e_s and also we have defined e_c. The solution of this equation we are considering like this we have to minimise the error and we will be getting this one. So, we have to do some mathematics and I will be getting this one.

(Refer Slide Time: 42:33)



And after this the solution of this equation is nothing but, I have to do the solutions and finally I will be getting the velocity along the x direction that is the u I can determine and velocity along the y direction I have to determine. u bar is nothing but the mean velocity. These are the mean velocity. And M is defined like this, M and N is defined like this. So, from these two equations I can determine the u and v that is the velocity factors I can determine.

(Refer Slide Time: 42:59)



So, I can show the simple algorithm from my book Optical Flow Algorithm. So, initialize velocity factor, first step to initialize the velocity factor. After this I have to consider the iterations, so k number of iterations I am considering. So, by using the previous equations you can see I am calculating the value u and the v, and we have to stop these iterations so first we have to consider this error. The error should be less then a particular threshold and based on this we can stop the iterations, other wise I can, I have to continue the iteration and I can determine the u and v.

So, this algorithm I can employ in a video, so in a video I have the frames, number of frames. So, in that case the k is the frame numbers. So, all the frames I have to read one by one, one number frame, number frame two, like this. So, k number of frames of a video I have to read and from this I can determine the velocity along the x direction and velocity along the y direction. So, briefly I have explained the concept of the optical flow, how to determine the optical flow.

In this class, I briefly explained the concept of a video surveillance system and also I have shown different applications of a video surveillance system. The two main concepts, I have briefly explained. One is the concept of the background modelling and another one is the concept of the motion estimation. For motion estimation I have explained the concept of optical flow that is very important. In my next class I will be explaining some tracking algorithms for object tracking, so let me stop here today. Thank you.