Indian Institute of Technology Kanpur

National programme onTechnology enhanced Learning (NPTEL)

Course title
Basic cognitive processes

Lecture-19

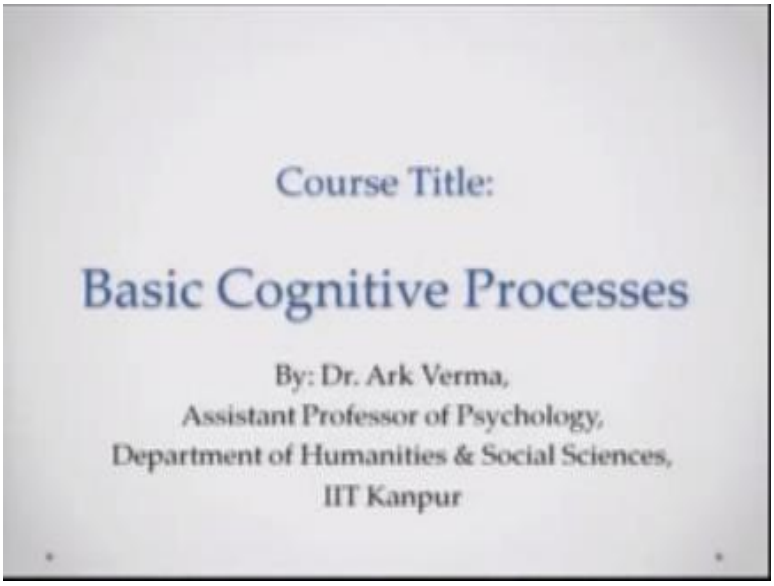
Approaches to visual perception-2

By
Prof, ARK.Verna
Department of humanities and social science

Indian Institute of technology Kanpur

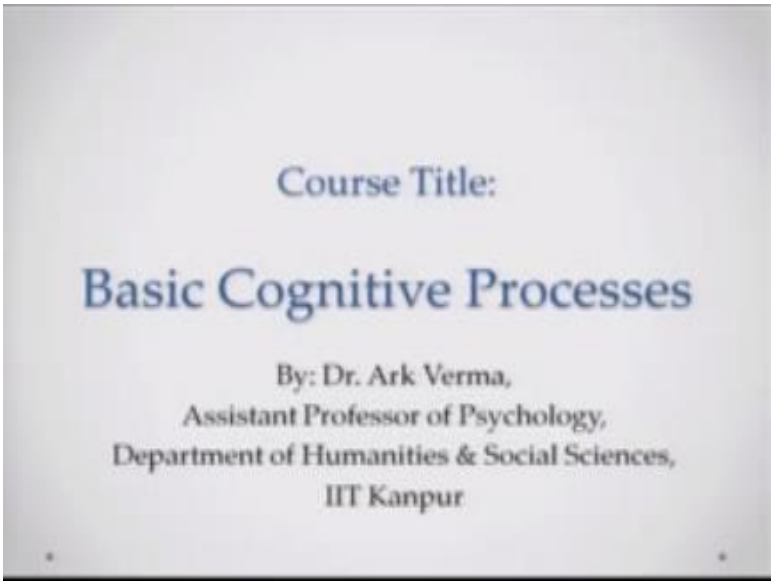
Hello and welcome to the course series basic cognitive processes.

(Refer Slide Time: 00:16)



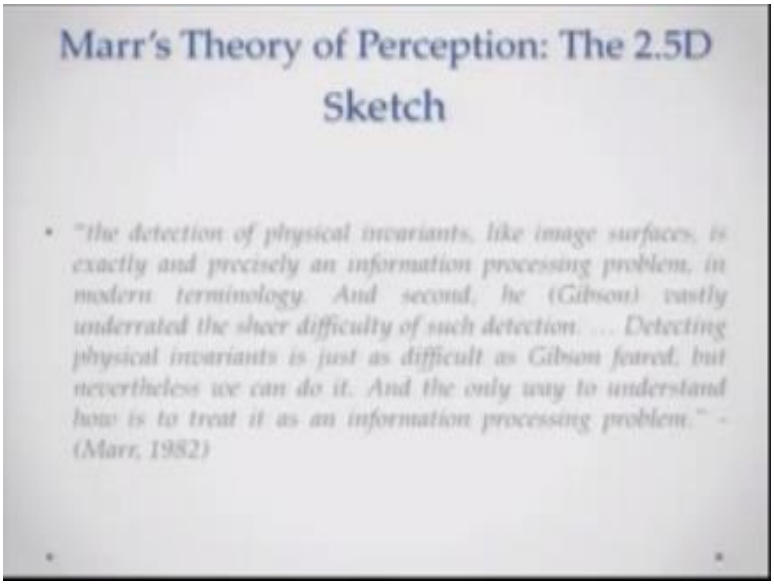
I am dr.Khanna from IIT Kanpur.

(Refer Slide Time: 00:19)



We are still talking about various approaches to visual perception today we will talk about David merge theory of perception which is more commonly known as the 2.5D sketch approach okay before we move further let me quote David Marr from one of his papers in 1982 and he says he was basically talking about how Gibson had approached perception and this quotations in that reference. So he says the detection of physical invariants like image.

(Refer Slide Time: 00:48)

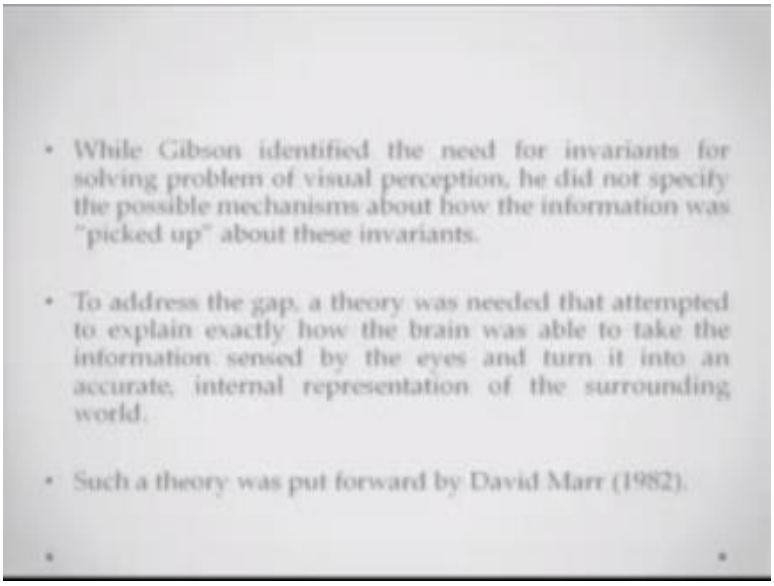


Surfaces is exactly in precisely an information processing problem in modern terminology and second Gibson vastly underrated the sheer difficulty of such detection detecting physical invariants is just as difficult as Gibson feared but nevertheless we can do it and only way to understand how is turret this as an information processing problem now you see David Marr is basically you know talking more about computational approach to perception he is usually talking about treating this problem of perception as a problem of information processing okay.

How do you really take up information from this external world how do you work upon intend how does that or say for example how does that information lead to the end product that is perception now this pretty much hold David Marr is talked about and we will discuss that in duty and enough details in today lecture.

Now why Gibson identifies the need for invariance you have seen horizontal ratio relation and using textures or surfaces you seen the thing of flow and all why Gibson identified the need for invariance for solving the problem of visual perception.

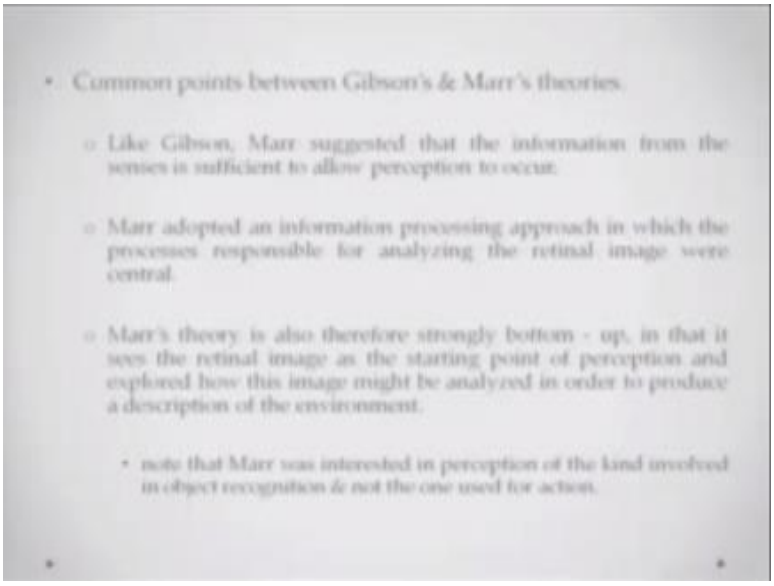
(Refer Slide Time: 02:00)



He does not really specify the possible mechanisms of how this information has to be picked up from these invariance he says if you remember the last lecture you know depending upon these sizes etc we can get information but he does not really say how are we doing that he says info you know you do a look at the surfaces whether the width is smaller or larger and you can make out whether the surface is receding or you know proceeding but how is this exactly done it not water this one really talks about.

So to address this gap about you know how this processor has to be done a particular theory was needed that attempted to explain how these processes will be done how does the brain take up information or sense by the eyes and turn it into an accurate internal representation of the surrounding world such a theory was put forward by David Marsh before we move into more detail about MA let me point out a couple of commonalities between Gibson's and Mars approaches.

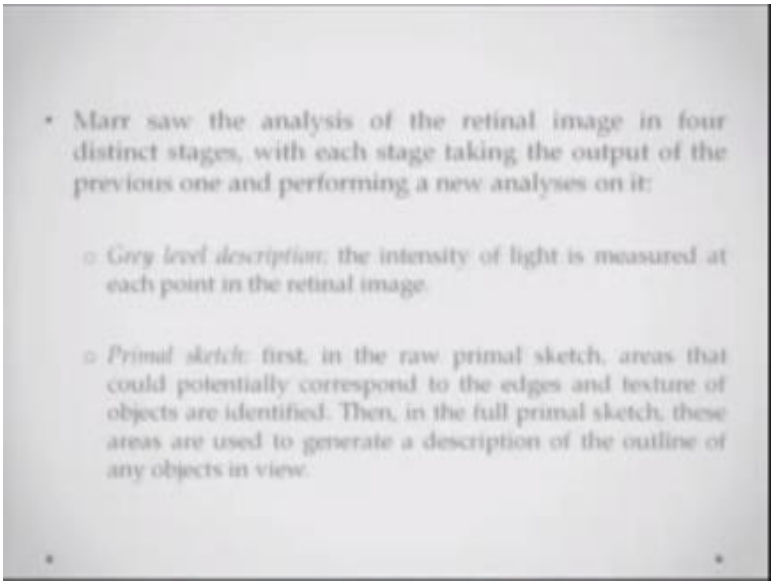
(Refer Slide Time: 03:00)



Like Gibson Mud also suggested that the information from the senses is sufficient to allow perception to occur more adapted an information processing approach in which the processes responsible for analyzing the retinal image were central so he is again saying light from the external world falling on the retina is the primary source of information and thesis the starting point of perception and this is how perception supposed to be built.

But this therefore also strongly bottom-up if you remember the distinction between top-down and bottom-up approaches mentioned in the last lecture in that it sees the retinal image as a starting point of perception and also explores how this image might be analyzed in order to produce adscription of the environment note that Mar is not really concerned about perception for action rather he is concerned about perception for recognition or perception that is meant to build the description of the world.

(Refer Slide Time: 04:03)

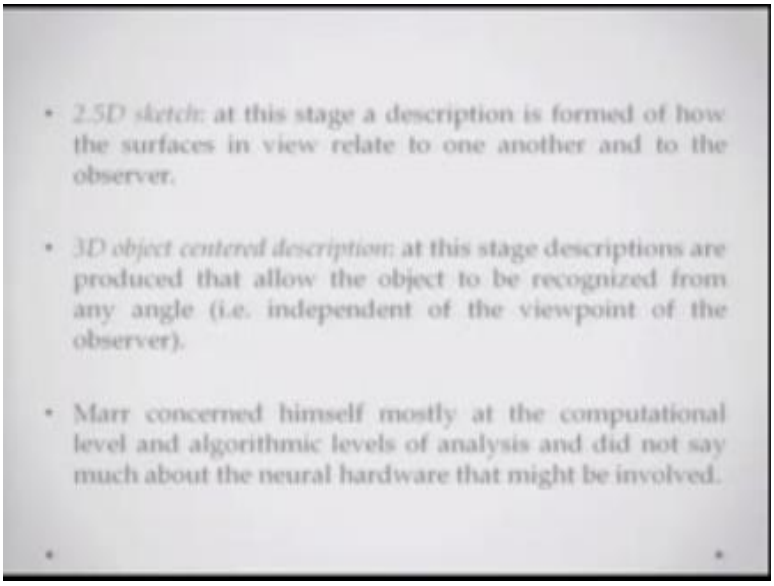


More easily saw the analysis of the retinal image in four distinct stages within each stage taking the output of the previous one and performing a new analysis so itis basically an incremental process of perception that matter we talked about you see here something happens in the first date and where the output is processed further in the second stage and so on and so fore.

Let us have a brief look at these stages before we go and elaborate upon them so the first stage is level description is basically Varian you are measuring the intensity of light layer o at each point in the retinal image the second is the generation of the primal sketch variant first there are two phases here so first in the raw primal sketch areas could potentially correspond to the edges and textures of the objects that are identified then you move on to constructing.

A full primal sketch which is basically you use the areas to generate a description of the outline of any objects we will actually elaborate on these in much more detail as we go ahead and just giving an overview of what really man was talking about the third stage in Mass approach is 2.5 D sketch it is at this stage adscription of the is form of how the surfaces in view relate to each other.

(Refer Slide Time: 05:11)



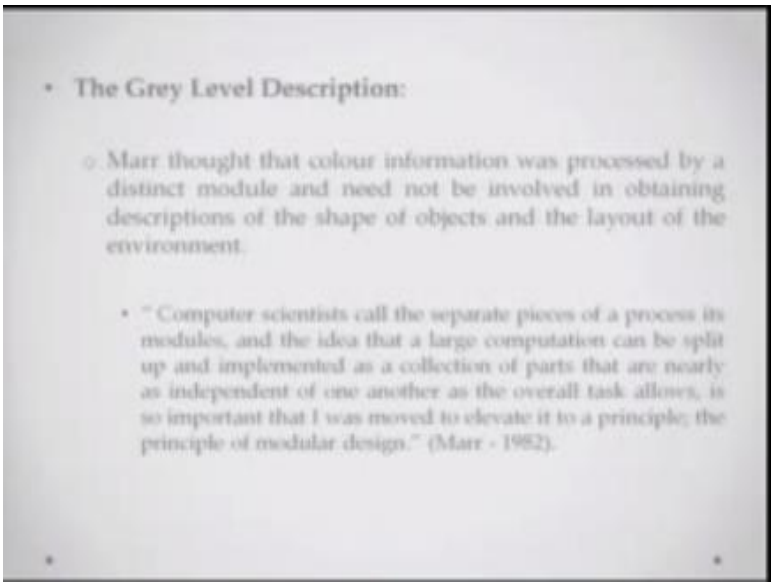
And also to the observer the final stage of Mar is basically objects centered is a3d or the centered description in this stage descriptions are produced that allow the object we recognize from any angle so you have the object even move around you can still recognize the object false basically every stable object centered representation Marking son himself mostly at the computational algorithmic levels of analysis and it do not really very much about the neural hardware that might actually be implicated in doing all of these computations that he was talking about.

Now let us sign elaborate these stages you will have to kind of follow this in more detail in order to understand this entire sequence of events so the first step in this sequence of events is building up of a gray level description what is the gray level description my thought that color information was processed by a distinct module he says that the process of perception is handled by different module components of processing say for example if you have to see the shape.

It is a different module if you see the color is different module if you see a depth another things they are carried over by different modules you might remember in the physiology part we are talking about different areas in the occipital loop areas v 1 v 2 v 3 4 etc which are all doing different stuff so mark end of you know is talking about something similar he says he was

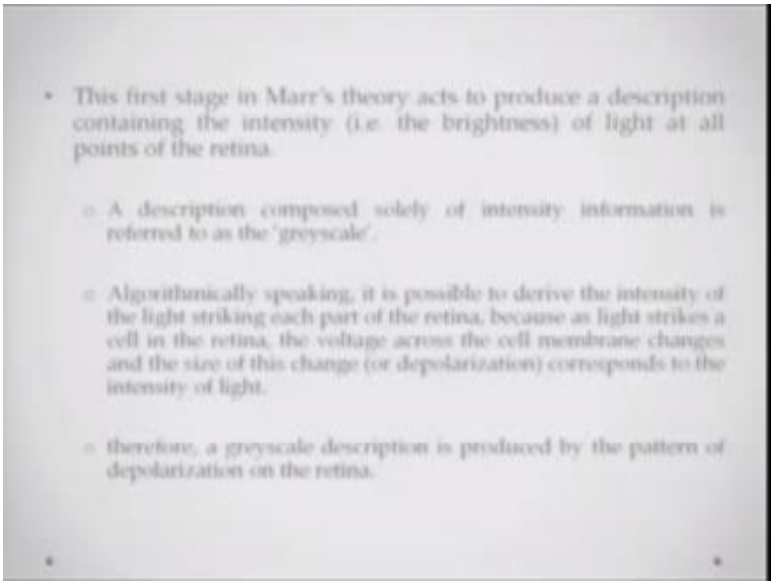
actually fascinated by you know this idea in computer science that a particular larger process can be actually split into modules.

(Refer Slide Time: 06:58)



So a large computation he says can be split into aspirin of and implemented as a collection of parts that are nearly as independent of one another as this entire task is and you are so moved by this that it actually elevated this to a particular principle which was later referred to as the principle of modular design if you remember one of our most you know earlier lectures we have been talking about this principle of modularity in great detail.

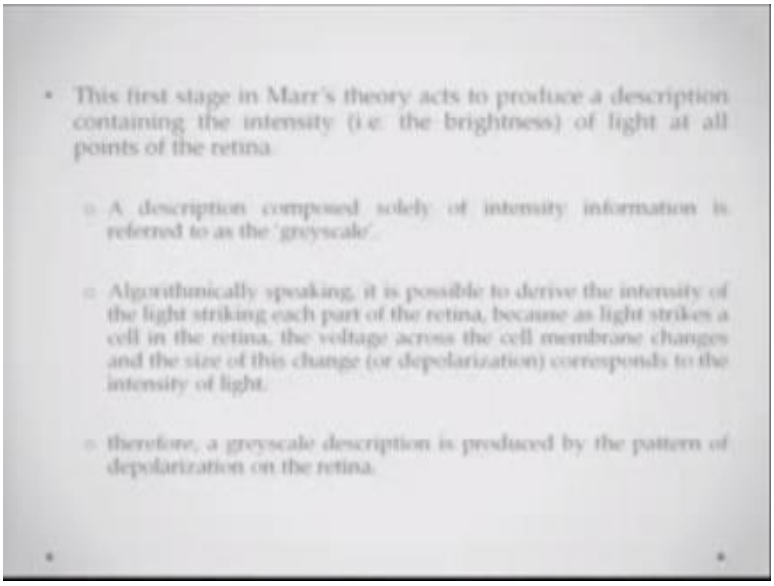
(Refer Slide Time: 07:20)



Now the first stage in mastering is basically to produce a description containing the intensity of the intensity distributions of light at different points in the retina how is this done this basically is say for example the way it is done is that it is possible to derive the intensity of the light striking at each part of the retina because as light strikes a cell in a retina the voltage across the cell membrane changes and the size of this change how much this voltage is changing corresponds to the intensity of the light.

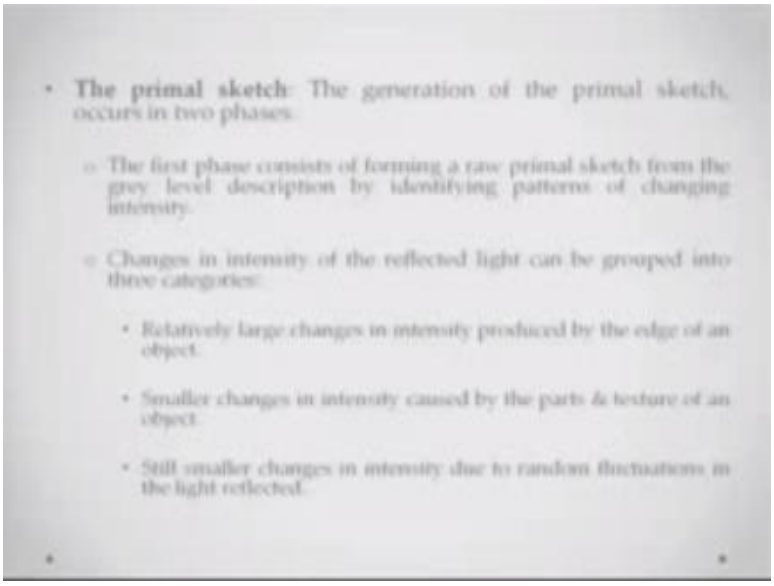
If you remember the chapter on physiology we will been talking about neural impulses and how the polarization happens it is a bit about that but what you need to understand for the moment is just that the amount of change that happens in a neuron in the retina is corresponding to the intensity of the light that is falling upon it so using this information you can actually construct a description of you know any surface with respect to whatever you know the intensity of light has fallen upon it.

(Refer Slide Time: 08:26)



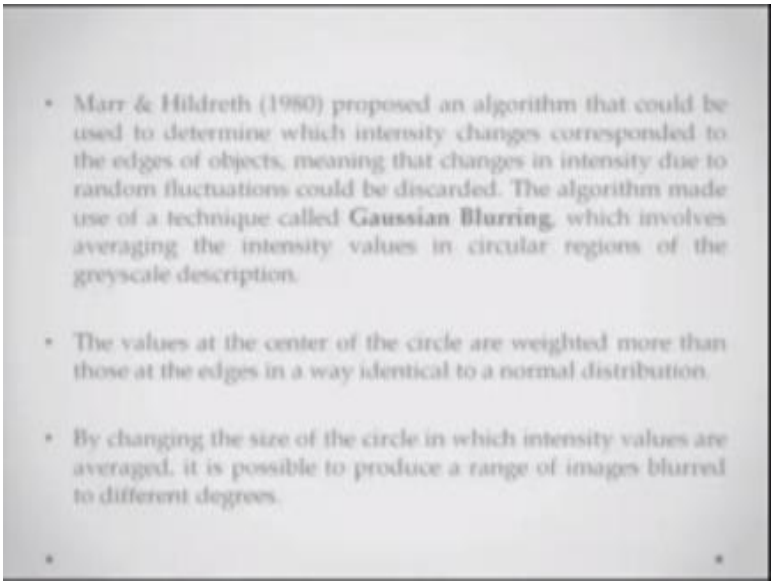
So a grayscale description basically is produced by a pattern of depolarization on the retina so you get a description which basically contains only the intensity information of the scene or of the surface that you are looking at now moving to the second part the primal sketch the generation of the preemie sketch basically occurs in two parts in the first part basically the forming of a raw primal sketch happens from the gray level description by identifying the patterns of changing intensity.

(Refer Slide Time: 09:11)



So in the first case you have you had a description of what are the different levels of intensity in the surface or in the scene in the second level you are actually taking into account how this intensity is changing throughout the surface or throughout the scene changes in intensity of the reflected light can be grouped into three categories so selectively large changes in intensity prod can be produced by the edge of an object smaller changes in intensity can be caused by the parts and the texture of the object.

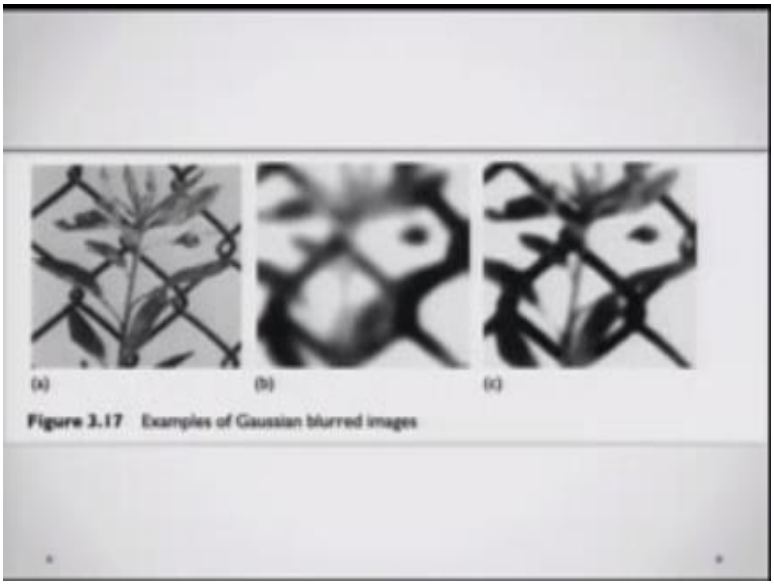
(Refer Slide Time: 09:36)



Even smaller changes in intensity might be just because of random fluctuations in the light etc modern Hildreth actually they work together and propose an algorithm that could be used to determine which intensity changes correspond to the edges of the objects meaning that changes in intensity due to the random fluctuation could be discarded this algorithm made use of a technique called Gaussian learning which involves averaging the intensity values in circular regions of the grayscale description.

So basically it's like you know the algorithm works upon the intensity values that it gets from the description and this is basically done over circular regions of this agrayscaleare you sure figure which talks about this the values at the center of the circle are weight mode are weighted more than those at the edges of the circle I exist in a way identical to a normal distribution so if you look at if you have the concept of how normal distribution looks is basically like bell-shaped curve the highest intensity is at the center the lower intensity is kind of arc towards outside by changing the size of the circle in which intensity values are averaged it is possible to produce a range of images blurred to different degrees so let us have a look at this figure here.

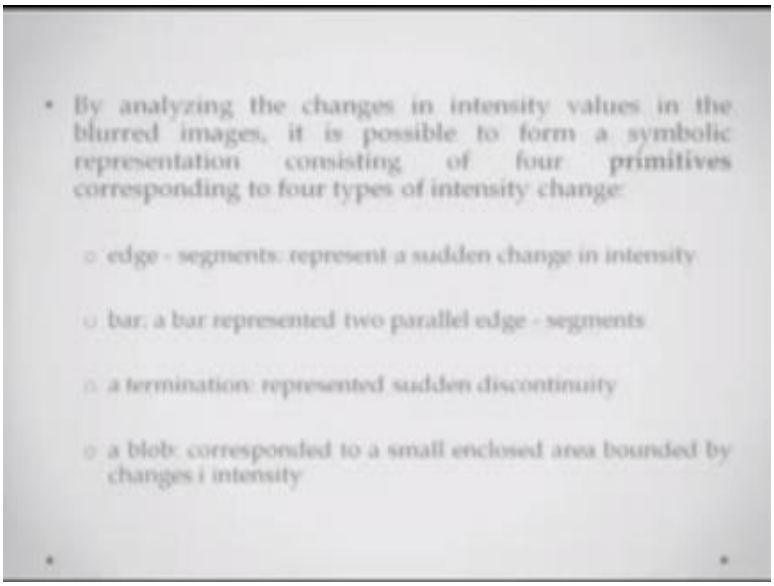
(Refer Slide Time: 10:44)



You can see the main picture is the figure 8 the figure B and figure C are blurred to slightly different degrees now kirimodern Hilliard's algorithm really basically works by comparing the images that have been blurred to different degrees now if an intensity change in these different you know figures invisible at two or adjacent levels of learning so for example you blow something up to 10% and20% very well if these images you know if an intensity change is visible at two or more adjacent levels of learning then it is assumed that it cannot correspond to random fluctuations and must relate to an edge of the objects.

So you just have intensity data you have a data about how this pattern of intensity is changing and you can see how these computations are cumulatively telling us important information about what is out there what does the object or the scene look like although this algorithm was implemented on a computer there is also evidence that shows that retinal processing delivers descriptions that have been actually blurred to different degrees so it could be kind of an evidence in support of you know the way Maris stipulating this process is happening.

(Refer Slide Time: 12:04)



Now by analyzing this changes intensity values in the blurred images it is actually possible to form a symbolic representation consisting of four primitives corresponding to four types of intensity change so how do you know really construct this kind of an image say for example if you want to look at edge segments they should represent a really sudden change in intensity if you look at a bar kind of a thing it basically should represent two parallel edge segments.

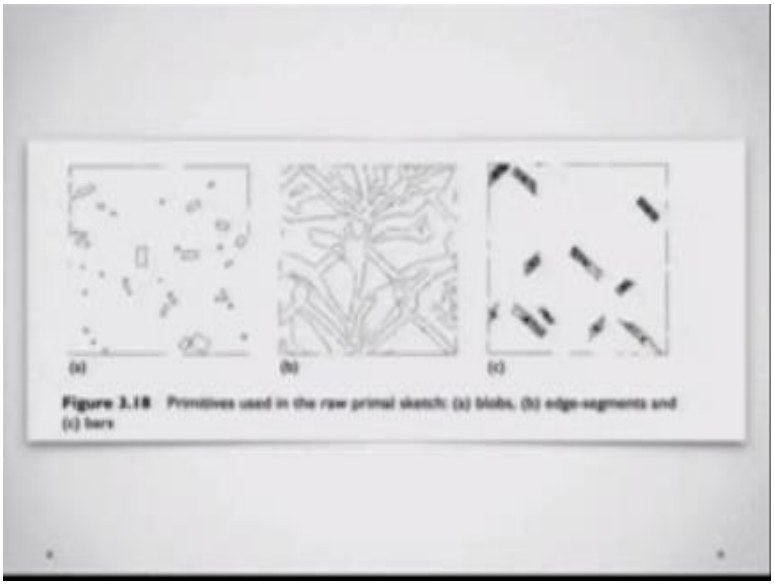
So two sudden changes in intensity a termination of any surface say for example the edge of this table here can represent sudden discontinuity also if you're talking about a particular object which can correspond to a small enclosed are abound by changes in intensity say for example you are looking at a at a face or you're looking at say for example face still has a lot when you can do if you looking at a picture of an orange.

So it has continuous inside it is just looking like a blob now here again you can see you can see the flower is basically more looking like a blob by the steel buyer sat the back of it are actually more looking like edges you can see this in more clearly here you can see here picture a has a lot of blobs picture B has a lot of edge segments and picture C which is basically those wires have altar's you can see how this can be used on the straight lines or curved lines in that sense now the

next step after you are got this intensity description into transform a raw primal sketch into adscription known as the full primal sketch

Just to recap what we have done is we have done a grayscale description then we have come to a raw primal sketch which is basically formed by analyzing the patterns of changes in this intensity now what we need to do is weave to move towards forming a full primal sketch what is full when we sketch it contains information about how the image is organized particularly about the location the shape texture and internal parts of any objects that are on view basically the idea is that you will have some things called you know place tokens what a place to fill their assigned.

(Refer Slide Time: 13:13)

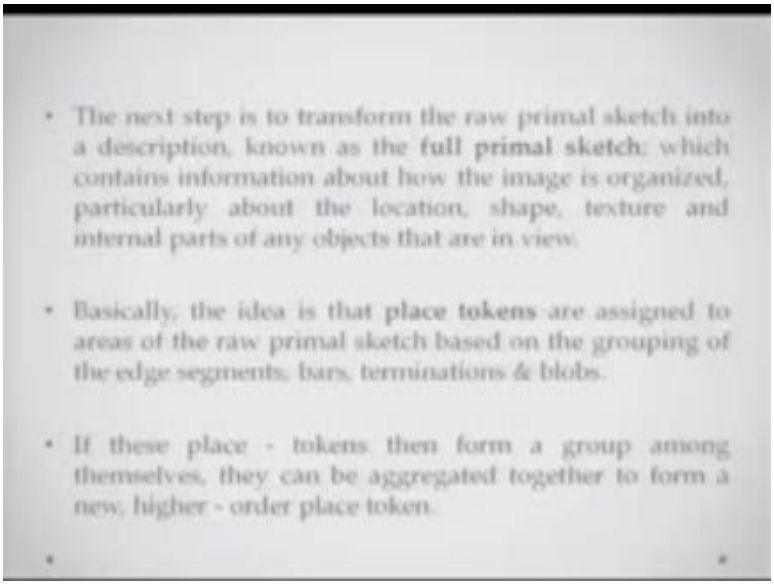


The areas of raw primal sketch based on grouping of edge segments grouping of bias terminations and blobs so once you have done the analysis like say for example in this figure here now would what you would want to do is you would want to group things together that is what one needs to do in by forming the full primal sketch so basically so if these place tokens then form a group among themselves they can be aggregated together to form a new higher order place to be.

So there would be different levels at which this place tokens would really appear we can understand this better by looking at this example say for your for example you're looking at tiger and how is this an else is happening so the raw primal sketch of the tiger would contain information both about the edges of the Tigers body that is the contours but also there will be information about the edges of the tiger stripes so tiger is basically entirely covered in stripes but then there is also at a higher level the contour of the tiger.

So in the full primal sketch place tokens will be produced by the grouping of individual hairs into each of the stripes so we'll have information about the stripes then say for example the place tokens for each stripe would also be grouped at a higher level in higher order place token meaning that there will be at least two levels of place tokens making up the tigers one level which is just making up the stripe the other level based on the stripes making up the entire contour of the tiger now various mechanisms have been postulated various mechanisms exist for grouping the raw primal sketch components into place tokens.

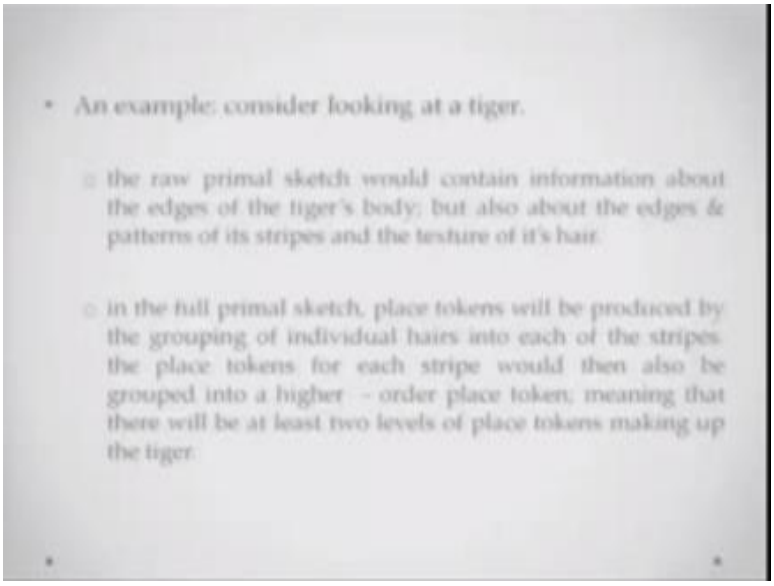
(Refer Slide Time: 14:00)



And then grouping the placed organs together to form the full primal sketch some of these could include things like clustering wherein tokens that are close to one another are grouped in a verisimilar in a very similar to the digital principle of proximity basically means that you have you know you put things that are close together in one object okay another thing would-be curvilinear aggregation which is basically that tokens which are related to each other which have similar alignments and also grouped or clustered together okay.

say for example if you're seeing a particular line you're more likely to see that line continuing here than you know breaking at different places we will talk about these Kestrel principles in more detail as we move it in the later lectures now the third stage the third stage is the generation of the 2.5 D sketch now much modular approach to perception basically means that while the full.

(Refer Slide Time: 15:04)

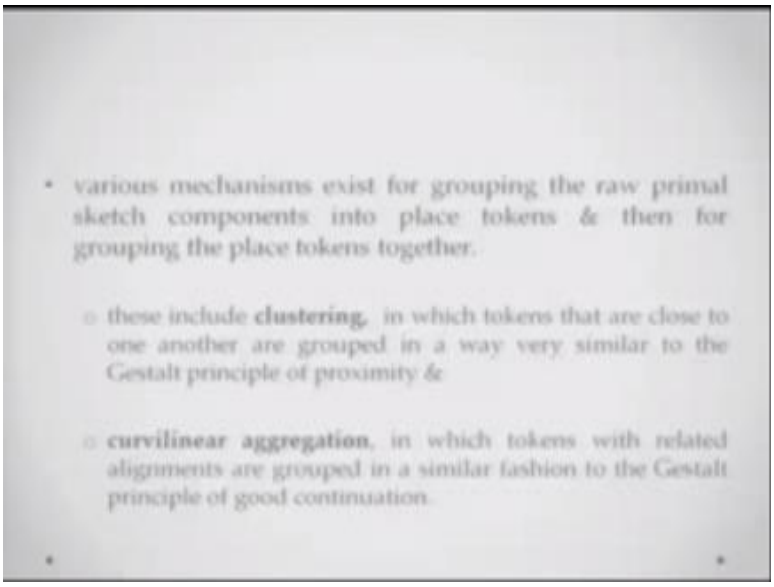


Primal sketch is being produced other piece of visual information are also being organized or analyzed simultaneously for example things like depth relations distance information between a surface and the observer information about whether an object is moving or you are moving in those kinds of things now more visually propose that the information from all of such modules

distance from shape color motion all of these videos will be combined together to form what is called 2.5 dimension sketch it is also called the two point five dimensions.

Case because the specification of the position and the depth of the surfaces and objects is done in relation to the observer now this is the view of the object in relation to the observer that is why this is also called a viewer centered representation an object as it is looking to me okay the image of the object that is falling on my retina this will not contain any information about the object that is not present in the retinal image okay.

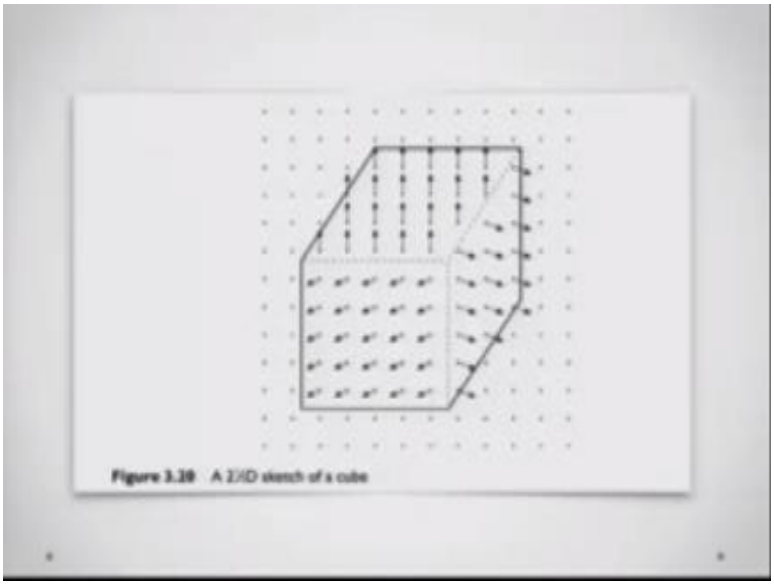
(Refer Slide Time: 15:43)



He do centered image is later turned into a fully 3d object Center different representation which we discussed in one of the later lectures Mara series saw the 2.5 g sketch as consisting of a series of primitives that contained vectors showing the orientations of each surface so once you have these air you know surfaces you actually you know can combine all of them together and the 2.5e sketch will appear is a series of primitives that contain these vectors you can actually understand this by looking here so if you're looking at it here you're looking at the 2.5 D sketch of a particular cube.

So you can seethes vectors pointing in the 3directions are actually telling us the direction of the surface or the orientation of the surface let us try and evaluate this approach let us look back and see what man was talking about a lot of research has followed davidmuntz theory some of it actually confirming his proposed mechanisms well some of them have found out some a few shortcomings modern Hildreth's proposal of a premium sketch being formed by looking for changes in intensity worked very well with computer simulations computer simulations but it could not really be guaranteed.

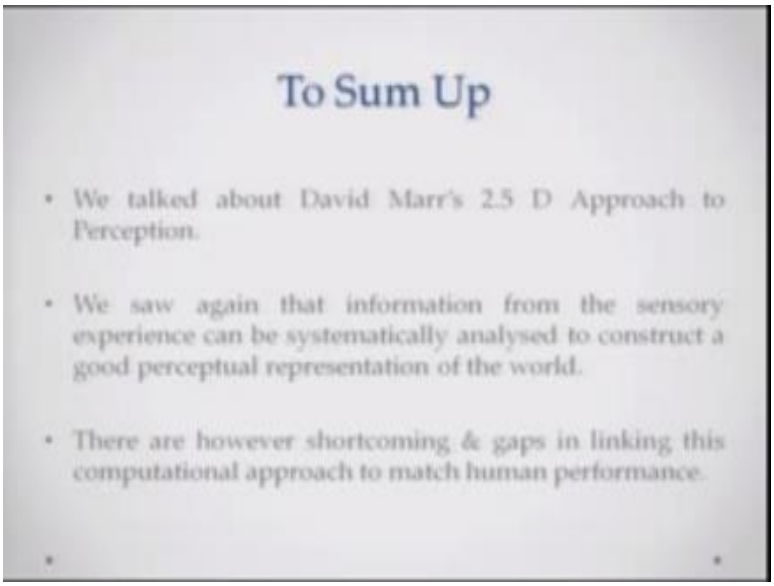
(Refer Slide Time: 17:51)



That this is the same process followed by the human visual system you would have seen that we have been discussing in one of the earlier lectures about there could be two possible descriptions of the same outcome ensign listening 1990 they showed that the participants of this study could also use three-dimensional information instead of only the two-dimensional information that is needed to form a full previous case so probably in some sense the underestimated the efficiency of our perceptual systems our mark proposal of the integration of depth cues in 2.5 D sketch was actually supported by experiments.

That were done by young increase in 1993 who basically reported that the perceptual system does process this cues separately so and we also make use selective use of them depending on how noisy they are so we are actually using these you know the perceptual system is actually using the excuse these invariant features or these depthcuesin order to you know develop a really rich representation of the external world around us and of the objects etc.

(Refer Slide Time: 20:29)



So this was all about David mask theory of perception let us try on some this up we have talked about David mass 2.5diablos to perception we have seen that information from the sensory experience can actually we systematically analyze to construct a good perceptual representation though by good I mean rich perceptual representation of the world that has all the knowledge that is needed for you to first construct a visual representation of the world and because at least Marv was not really concerned.

With about perception fractions we not talked about that in this part however they were indeed shortcomings and gaps in linking this kind of a computational approach to mass human performance that is what has been summoned out after you know a lot of research conducted on mass principals but nonetheless this was good and whendrownded computational approach about

visual perception in one's the next pictures we will talk about some other kinds of approaches to how visual perception is achieved thank you.

Acknowledgement

Ministry of Human Resources & Development

Prof. Satyaki Roy
Co – ordinator, NPTEL IIT Kanpur

NPTEL Team
Sanjay Pal
Ashish Singh
Badal Pradhan
Tapobrata Das
Ram Chandra
Dilip Tripathi
Manoj Shrivastava
Padam Shukla
Sanjay Mishra
Shubham Rawat
Shikha Gupta
K.K Mishra
Aradhana Singh
Sweta
Ashutosh Gairola
Dilip Katiyar
Sharwan
Hari Ram
Bhadra Rao
Puneet Kumar Bajpai

Lalty Dutta

Ajay Kanaujia

Shivendra Kumar Tiwari

an IIT Kanpur Production