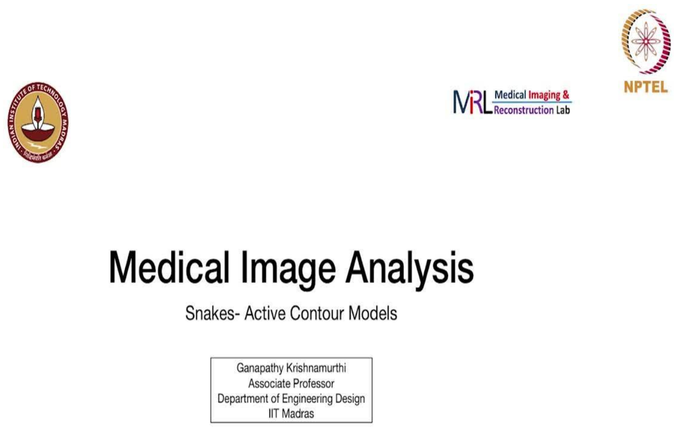
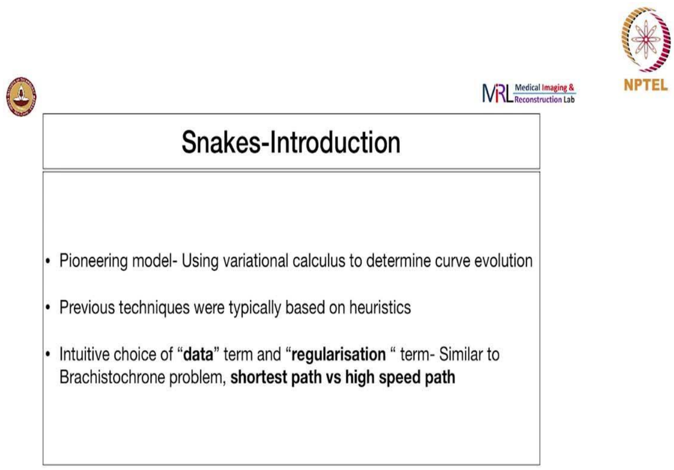
Medical Image Analysis
Professor Ganapathy Krishnamurthi
Biomedical Engineering Design
Indian Institute of Technology, Madras
Lecture 29
Snakes - Active Contour Models

(Refer Slide Time: 00:15)



Hello, and welcome back. So, in this video, we will look at the so-called active contour models or snakes. So, one of the systematic kind of PDE based models to come out of calculus of variations, and one of the pioneering methods in that class of methods to pioneering work.

(Refer Slide Time: 00:37)

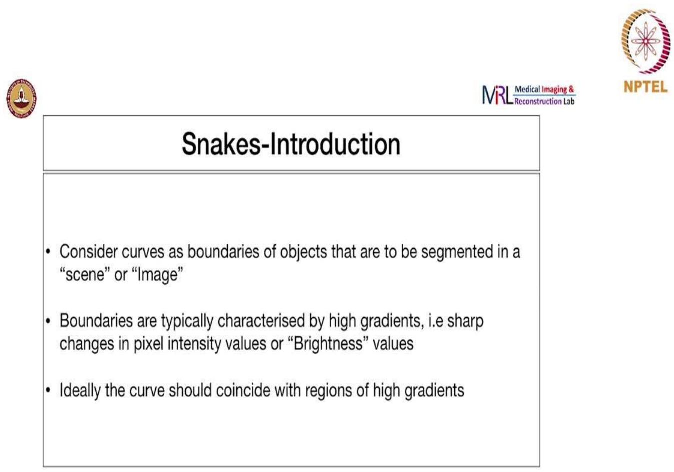


So, like I said before, it is a pioneering model using calculus, variational calculus to determine curve evolution. So, why is this so important? Because previous techniques or techniques that came before this were all typically based on heuristics, like a bunch of rules that you put together and hope that it worked.

But in this one, this is of course used for segmentation. It actually, the rules actually came out of variational calculus by putting together a correct choice of data terms, loss function and regularization terms, in the functional, we arrived at a curve propagation equation, without having to explicitly specify rules.

So here, when you say data term and regularization term, I want you to think about the Brachistochrone problem where we looked at whether we go fast, by dropping vertically down a bit, so that time is reduced, or we reduce time by going that shortest path, so that was a tension there, here, there is a tension between data and regularization terms. We will look at this in more detail in the next few slides.

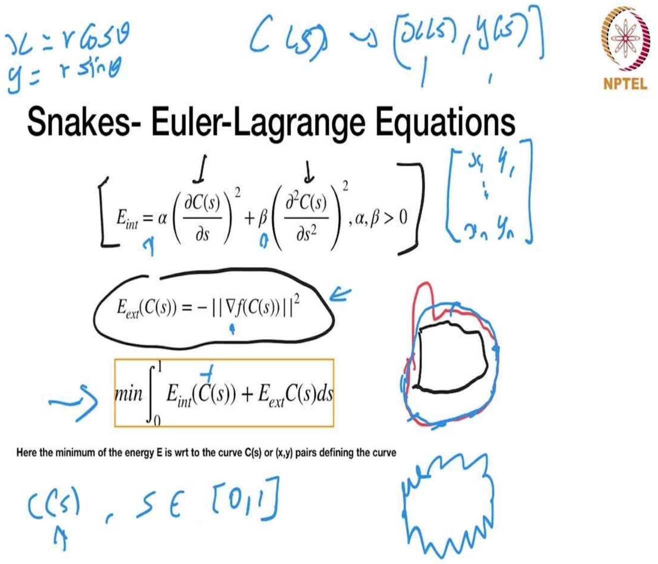
(Refer Slide Time: 01:50)



So why are they called Snakes? So, if you consider any object that is to be segmented, in this case, an organ to be segmented in a scene or an image in the medical image, the boundaries are basically the curves, the boundaries of objects, or the curves. And that is the one that so, we seek to segment objects by propagating closed curves, so that they can go and latch on to the boundaries of objects.

Now, what are the other characteristic of boundaries objects? Typically, in a medical imaging setting, the boundaries are characterized by high gradients, the sharp changes in pixel intensity values or brightness values. So, ideally, the curve that we are propagating that is we are trying to latch on to the boundary, we should also coincide with regions of high gradients. So, that is one of the important criteria that is considered in this method.

(Refer Slide Time: 02:52)



So, we have, right away we are looking at the various terms that go into the functional. So, first we discussed that the gradient is an important term. So, the gradient will tell you were sharp transitions in pixel intensity occurs, and we surmise that wherever there are sharp transition pixel intensity means that there is a change in the regions, that is the edge of the object that we are looking for.

So, we do not know the object, but our hypothesis is that the edges of those objects correspond regions of high intensity, and we want our curves to go and rest, or latch onto regions of high intensity. So that is one term so, that is the term that you see here. Negative gradients squared. So that means that you have, you are looking for regions of, why is it negative?

So there the gradient is high, which means that it is a very high positive value, so you put in a negative sign there, so that it becomes a very small number. So, you minimize, so it becomes a very small number. So, you are trying to minimize our functional. So, one of the terms in the integrand is this becomes a very small number when the gradient is very high.

The other term that we want really small are this, so here, this is where the tension is because we want them to latch on to the gradients, but we want them to be smooth, or they will not be smooth curves. So, this is your regularization term, which makes your curve smooth. Of course, there are various explanations for this. But I will just give you a couple here. So, what this term does, the delta C or delta s I will tell you what these variables mean.

This is the first derivative, there is a square there to make sure it is always dealing with positive numbers. The square, so the delta C or delta s make sure that you do not have very sharp discontinuities in your curve. And what this also ensure is that you have a very tight curve or basically your curve is as small as possible. So, it does not have sharp discontinuities and it is generally very tight, very short as possible, the length of the curve is shortened, is the link, shortening term. That is the first derivative.

This second term, this just to make sure that the curve is not too wriggly. So do you do not want some so, there if you have an object, let us say this object that you are trying to segment and your curve you do not want some curves going like this. This is what we saw, you do not want a lengthy curve. So, this is a very long curve. And this term makes sure that the sharp discontinuous do not occur rather, you have a smooth curve, fairly shortened curve.

The second term, what it does is make sure that no, you can also have curves like this. So, you are abrupt changes in the first derivative. So, in essence, what this happens is make sure that your curve is not super wriggly, it is more smooth. So, what are these C is the curve, obviously, just a notation, what is C of s represent? This is just some parameterization. So, s is a parameter that goes from 0 to 1. I am sorry, I am just writing it in a very poor way. Let me just wipe this off. So, s belongs to 0 1. So, if there is s very smoothly from 0 to 1.

So, it is a parameter which generates the curve. So, s is a parametric curve by s. C of s basically, it just means you can think of it as generating x of s, $C(s)$ is nothing but a pair of points x of s and y of s. So, x and y are basically points on the curve. So, you think of a curve it is in a plane, x and y are points on the curve. And you can generate the curve by wearing s smooth as there are many parameterizations available, we will not go into that.

So, for instance, you can think of s as maybe cosine theta and sine and, for instance, x can be, let us say,$rcos(\theta)$, and y can be $rsin(\theta)$, this is one parameters you can put r to be 1, so you get only $cos(\theta)$ $sin(\theta)$ and you can actually vary θ. So, you can of course normalize it. So that data varies from 0 to 1, and then you multiply accordingly. So those things can be done.

So that s is just one parameterization like that. So, lots of parameterization available, do not have to worry about that now, we will cross that bridge when you get there. So, $C(s)$ is nothing but it is a vector of these points, x and y. It is a collection of x and y, which make up the curve. So that is what we always when you say C they are only looking at know this bunch of points x come of y.

Once second, I will just rub this out and so you have this points x, y so you can generate them. So, you have x1, y1 up to xn, yn if you discretize it, if you choose to discretize otherwise, it is a continuous parameterization. So, s, smoothly varies from 0 to n so that you can generate all the x and y's possible. So, infinite number of x and y's possible, which make up the curve.

So, there are these two terms, one is, like I said, the negative of the gradient squared, the gradient of f here f, f refers to the image. Here, we can use i, but typically, it is bad choice in this kind of work, f is refers to the image $C(s)$ here, it basically we are evaluating the gradient only at the curve. So, at all the points that make up the curve, we are evaluating f of $C(s)$.

So, the gradient so for the curve to be optimal so, if you want to the curve to latch on to the object boundary, we stipulate that at the points wherever the curve is, you want a very high gradient. Which means that if you take the negative of the gradient, which means a very large, very large negative value, which means a very small number there, so minimum that you are trying to minimize that.

So large gradients, and we want the curve to be smooth, so which means that is why they are adding the first and second derivatives. These are called regularization terms. So, these are the two competing forces, forces that we are talked about, like the shortest path, highest speed path, we are looking at it has to latch onto strong gradients. At the same time, it should not be too big a curve or too long a curve and it should also not be too wiggly a curve.

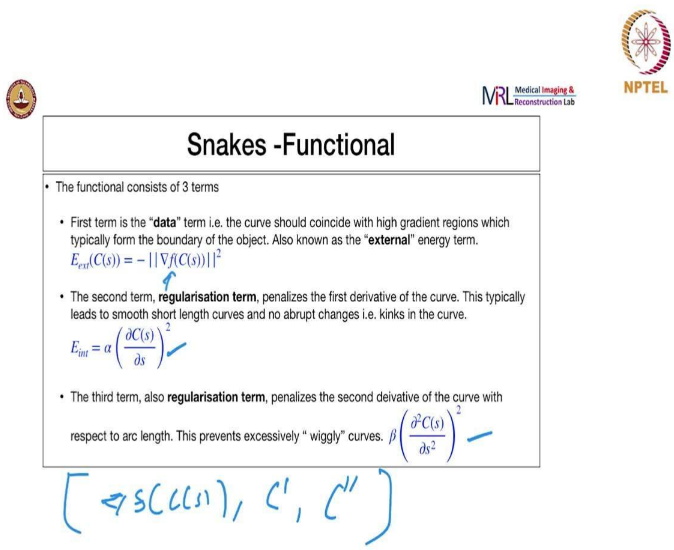
So, that is the tension here. And of course, we are trying to minimize this following function functional, where E internal is this energy, which corresponds to the regularization term. And this I call the data term because f is the image, that is the data that is given to you. And you are trying to estimate the gradient from the data. So that is the data term. So, and here in the paper, they have referred to as an external term.

E internal energy and the external energy, I like to call it regularization and the data term. So, there is just a play interplay between these two terms, typically, you will have these parameters alpha and beta. Now alpha and beta can depend on s also, or they can be independent of s either way is fine. And they are the terms that gives you the ability to control how much importance you give to each of those terms.

So, by making the alpha and beta small or 0, you can make their influence smaller or larger. So that is the idea. So that is the entire loss function so that you put them into the loss functional. You are trying to minimize the loss functional with respect to what, you are trying to estimate $C(s)$ like we had for the curve, the brackets problem. Here, we are trying to estimate the segmenting curve $C(s)$, which will form the boundary of the object.

So, once you have the boundary of the object, we can find the interior also. So, we are trying to estimate $C(s)$, or alternatively since $C(s)$ is nothing but collection of these points x, y, which make up the curve where we are trying to estimate a bunch of x, y, which form the boundary of the object by minimizing this functional. So, and of course, remember, s is the parameterization of the curve, s varies continuously from 0 to 1. And as the series continues from 0 to 1, you can generate x and y as well.

(Refer Slide Time: 11:33)



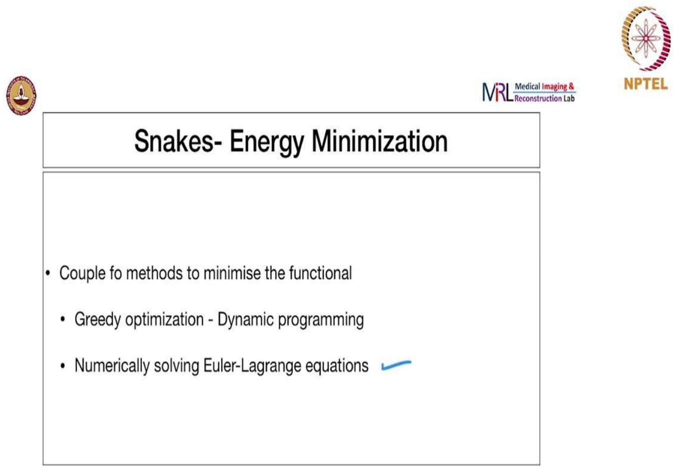
So, likewise, just to get a better understanding of this Snakes functional, consisting of three terms, like I said, the first term, the data term, the curve should coincide with high gradient regions, which typically form the boundary of the object. So, in medical imaging, also, in lots of situations, the boundary does form, the boundary does have high gradient, it is sometimes the paper they refer to as the external energy.

The second term is regularization term penalizes the first derivative. So, it means this typically leads to a smooth short length curves, that is without kinks, suddenly there will be going to making, so it is smooth short length curves so this is a curve shortening term like, so it will just collapse the curve into a point, if there is nothing stopping it. The third term is also the regularization term, again penalizes the second derivative of the curve with respect to arc length but s it is basically some parameterization. This prevents excessively wriggly curves.

So, there are two terms, the first derivative term, second derivative term, so you have some function. You will see that there is an energy term I call it energy term, there will be this f, gradient of f depends on $C(s)$, so it is a function of that. You also have your function, the integrand, inside the functional is has the C' term, and the C'' term. Here, the C is implicit in f. So, these three terms are there and so then you have the performance equation you know to use to our Euler Lagrange equation derive.

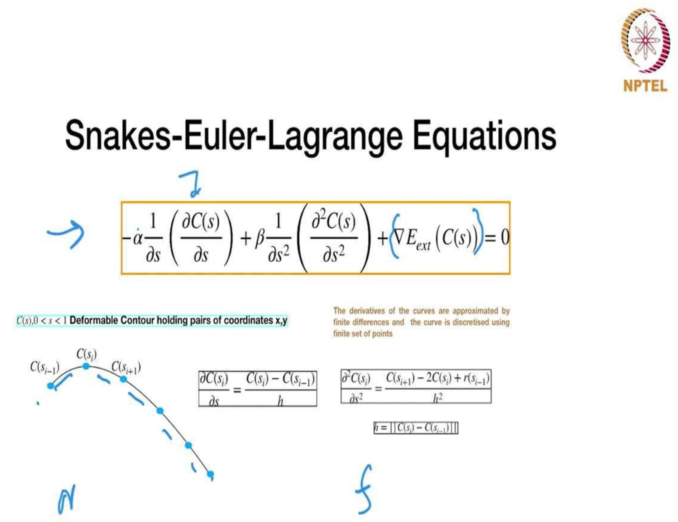
So, I had used f of Cs. Ideally think of just do not use f, maybe you just use E. This f is basically because I have used this f inside the functional when you are trying to derive the Euler-Lagrange equations. So, then that can be a bit confusing. So, feel free to instead of f use some other symbol. So that you do not get confused. So, I am not.

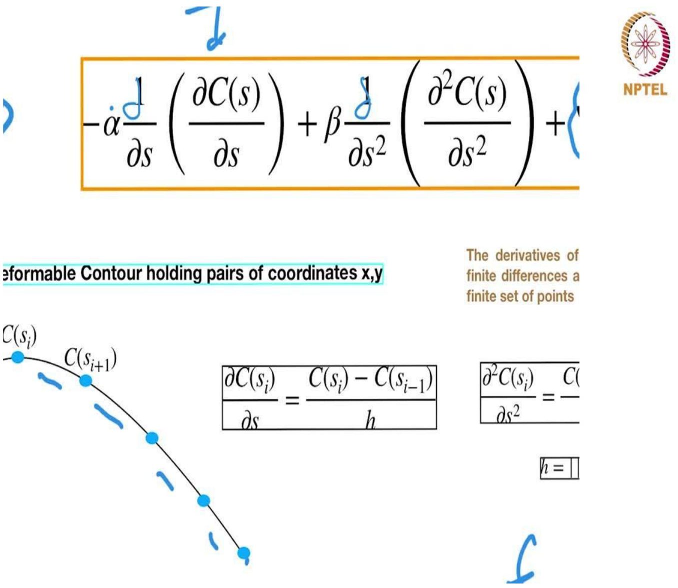
(Refer Slide Time: 13:47)

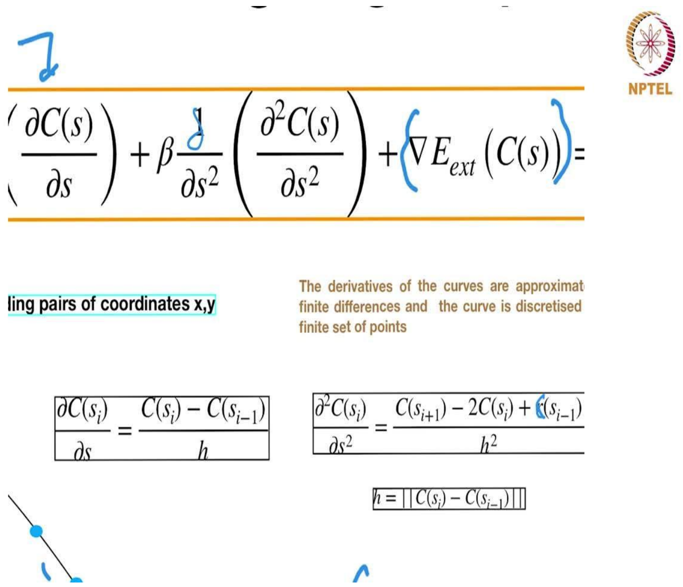


So now, sorry about this spelling mistakes, couple of methods, not couple fo methods. So, we are going to be only looking at the second one. This year, a numerically solving Euler-Lagrange equation. So, we will derive Euler-Lagrange equation. Of course, you get it, I will not go through every step because some of the steps are just calculating derivatives. So, I ask you to do that. Once you know what the Euler-Lagrange equation are, you know what f is and you know there are two, there is again, like you saw here, that the curve propagation equations for this, bear with me.

(Refer Slide Time: 14:25)







So, if we actually apply. So, remember, just to be clear I think I am using this f too much. So, I am going to just switch notation a bit, because if you look at the, I am sorry, I have to go back and forth, but you have to when you are trying to do the derivation. So, when you actually formulate the Euler Lagrange equation, this is what you will get. So now, how are you going to solve this problem? So again, gradient of the external $C(s)$ is equal to, I am just using this term as it is, gradient with respect to with respect to x and y, actually, that is what you have.

So, and these terms will only with respect to s and delta s and the second derivative with respect to. In fact, second and fourth derivative, delta s and delta s square. Ideally, this alpha should be inside. Because if it is a function of s, you can take that out, but assuming it is constant in this case, so what is it? Now we have to understand this, so the way to solve this, because you are solving there are no analytical solution. So, to solve it using numerical approximations.

So, I urge you to go and actually derive this, it is not, you just have to plug in the correct function expressions in the Euler-Lagrange equation that we saw in the previous video. And then you can get to this very quickly, it is not too hard to do. So, how do we discretize this. So, the discretization scheme is what I found is that if this is the curve, the blue points are the discretization.

So, s of i minus there, you take, let us say capital n points were discretization is capital n points and you have to choose it, so that they are equal distance of a typical that is how it is sample, this length and this length are the same, even though in my picture, they are not drawn the scale, ideally, you choose those points, “n” points on the curve for a discretization so that they are equally spaced and of course, the first and last points coincide. That is your boundary condition, the first time the last points coincide.

And the way you calculate the derivatives with respect to s any parameterization is parameterization, is that you actually, they use your discretization itself. So, for instance, the first derivative with respect to if you can zoom in a bit, allow me to zoom in bit, it might get tilted. So, the first derivative is calculated by a finite difference approximation. So, the delta c or delta s at si is nothing but C of si, which means the C of si is nothing but the xi, yi at that point, minus C of si minus 1 divided by h, or h is the distance between these two points that you have discretized. So that is the first derivative.

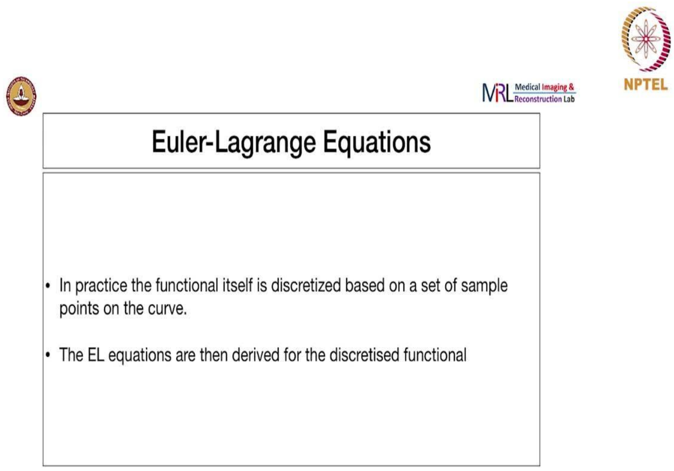
Similarly, you can write down for a secondary derivative also so this again if you are using some numeric analysis, these are standard formulas. Similarly, for the second derivative of the curve at s site, we have this formula. Again, I have made a mistake here. It is not r it is C. It is again I was referring to multiple textbooks and I was trying to use C because it seems to be more commonly used than r so somehow it slipped in there, sorry about that. So, h is again the distance between successive points.

Now, these are for the first and second derivative. What you do is, you take this and you plug it back in this equation. So now we are going to calculate delta or delta s of delta C or delta s, now you have delta C or delta s that formula here, that formula is there. Sorry, there is something I have done which I need to take care of. So, now, we have delta C or delta s the expression for this.

Now, we have to do delta, I have put 1 over delta s, but I was I should have put delta over delta s, I am sorry, there seems to be again a mistake here, we will correct that. So, delta over delta s of delta that is of delta C or delta s. So, now, all you have to do is in order to do this instead of C of si, you just put delta C over delta s, C of si minus 1 you put delta c or you can directly use this formula better still directly use delta over delta s.

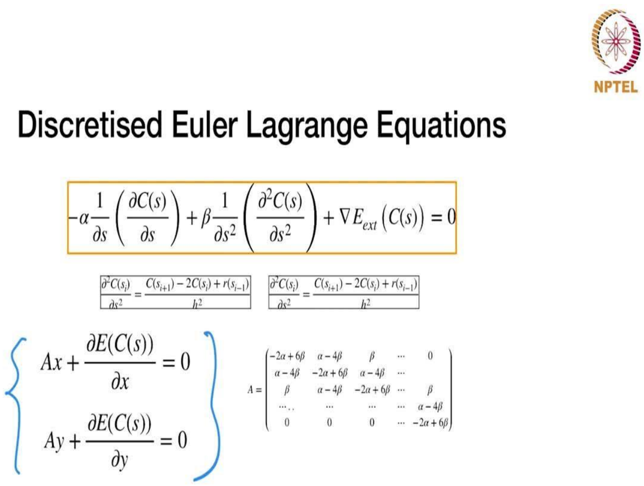
So, because alpha is not a functional base, you assume needs to be constant. We assume it to be constant when directly plugged this in this equation. For the second derivative again of the secondary way to for every instead of C of si plus 1, you have to substitute delta square C of delta s square at i plus 1 and delta square C of delta s square at S of i and delta square C, square root i minus 1, so again, I seem to change this. You put this in there, write down these expressions and you can simplify.

(Refer Slide Time: 19:50)



I have chosen to provide it we can look at it later time. In practice you can also discretize the functional itself so it is possible to do that straight away the original expression, we had the functional expression, we can discretize that and then take the derivative then based on the discretize function.

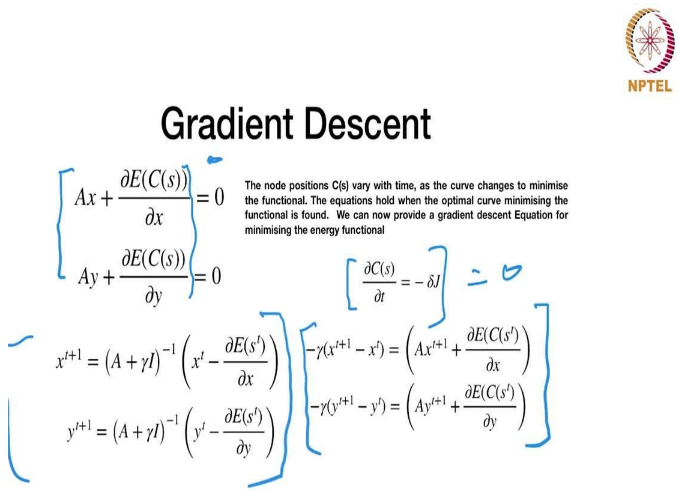
(Refer Slide Time: 20:10)



So, if you do that we end up with the following set of equations, you can actually derive them, it is very, not too hard to do, again, are you to do that. And if we find time, I will go through this in slightly more detail than what I am doing now. So, you will end up with this set of two equations. Again, the E is basically the image term. It is the gradient of the image evaluated at the curve, A is a Penta diagonal banded matrix, which basically carries the weights of the finite difference terms. And these are the two simplified or the final form for the Euler-Lagrange equation.

Now that we have this, what do we do? So, it is not done yet. So, this just gives you some relationship between or like I said, the tension between how you move the points because A, is the regularization matrix, you can think of it that way A is the matrix that regularizes your curve, delta u delta x is gives you the gradient information. So just like this tension between the two, so how would you propagate the curve that is what we are going to look at but it is not done there.

(Refer Slide Time: 21:19)



What we have to do is something called gradient descent. So, because what we have learned remember that at the optimal position of the curve, once a curve has latched on to the boundary of the object, then the curve no longer moves. So, which means that this delta s c over delta t equal to 0. But this delta c or delta t equal to 0 is, again, that is how we have defined the equilibrium position. So, which means that we are tracking the time curve over time, tracking the curve over time.

So, we can also if you think about it, if you want to evolve the curve over time then we can actually do this so-called gradient descent. So, if till we find the optimal position, we have to evolve the current how do we evolve the curve, we evolve the curve in the direction of increasing gradient, which is what the minus delta J tells you. That is how you propagate the curve. So, delta c, over delta t is minus delta J is just the gradient descent equation.

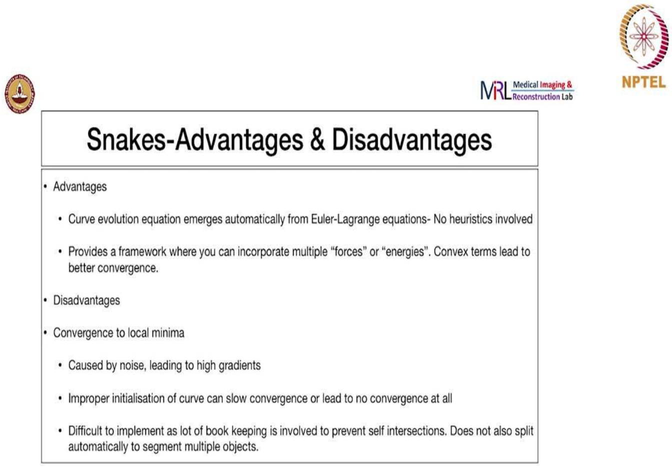
So, you are propagating the curve in the direction of the, opposite to the direction of increasing gradient, which is basically the gradient of the function that is what we have, because we are trying to estimate the curve itself. So, we have delta c over delta t is minus delta J. And minus delta J is what we estimated here. This is minus delta J, these terms, so then we can write this delta c over delta t as time stepping. So, this is again, you do a finite difference approximation.

Like I said, $C(s)$ is nothing but pairs of vertices x, y. So, then we can get this time step equation right here, which we can simplify to get the equation in this final form. So here is a matrix that can be there is again, a small confusion here as to why I put x^{t+1} here, but I put t there. So, this is called a semi-implicit formulation. Because one way to look at it is, you propagate the curve, you move the curve based on a gradient equation at the current time step.

And then, you move it and once you move it at the once after you move it, then you regularize it, that is what we are doing. So, it is semi-implicit, it turns out it has much more desirable numerical convergence properties. So, this is again techniques borrowed out of numeric analysis. So, this is actually a very complicated method to actually understand and code.

So, you start with calculus of variations, you form the current functional where there is a tension between how smooth your curve has to be or regularized your curve has to be and how badly you want it to latch on to the gradients in the image, the curve to latch on to the gradients in the image. And then, from there the Euler Lagrange equations come out, but they are non-trivial to solve. So, you solve them by using the semi-implicit numerical analysis technique. By and of course, you have to discretize these things properly.

(Refer Slide Time: 24:22)



So, what are the advantages of this Snakes? So, the curve revolution equations emerges automatically, you did not do anything other than intuitively decide this. You intuitively decided I want a smooth curve, and I want it to be attracted to gradients. That is the condition you imposed. And the rules for moving the curve around. The mathematics for moving the curve around came out automatically.

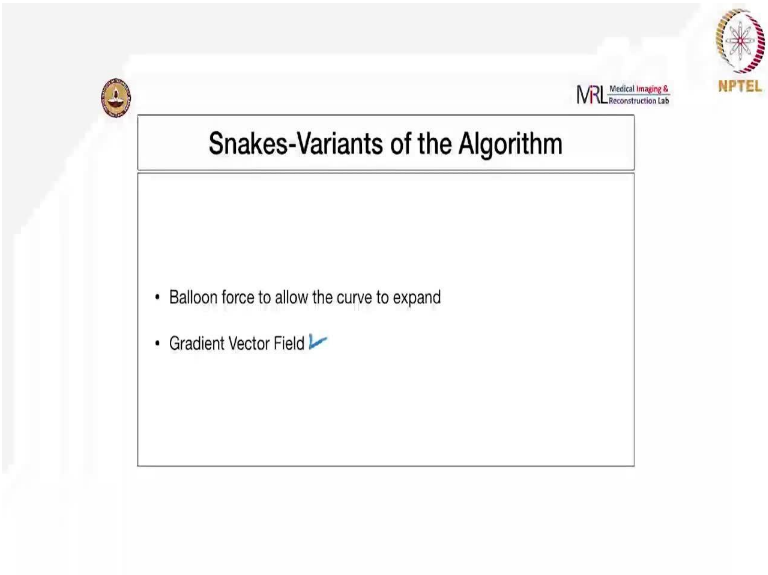
Of course, it is very complicated, but still it does come. And so, think about this, this actually provides a framework where you can incorporate multiple forces or other energies. And of course, we use convex terms, which basically something that has only a global minima. But again, we will not get into that. So, you add terms which are amenable to easy to differentiate and do not have too many local minimum.

So, you can like add more forces to make it move faster, et cetera, we will see a couple of them not in detail, as mentioned, they call balloon forces, and there is something called a gradient vector field. What is the disadvantage, it converges to local minima? Why would it converge to local minima? We will take any image to think about it, a lot of images will have noise and noises are typically high gradients.

So, which means that a curve can latch on to a noisy pixel or a group of noisy pixel, which might not correspond to the object of interest in the image. And improper initialization can cause slow convergence. So typically, this method is kind of sensitive to initialization, you pretty much have to start very near to the boundary that you want to segment, you might say, that is no fun why do you want to do that, it is better than drawing by hand, if you have a large number of images to annotate, it is best that you start varying, you initialize it, and automatically converges to a boundary.

Now, it is difficult to code because you will have to keep track of all the points, that is a double edged sword, because you do not have to keep track of it for an entire image, you are trying to segment the entire image, you have to keep track of only like a finite set of points, you do not have to take up every pixel, keep track of every pixel, how we are keeping track of every each of these points, because you are propagating them, it is hard because you do not want them to converge to one point, you do not want them intersecting all these are much more difficult programmatically to enforce.

(Refer Slide Time: 26:36)



So, there are variants of this, I will not go into the details, if I find time I will, there is something called a balloon force, remember, the first derivative term makes the curve shrink but sometimes you want them to blow to expand a bit based on certain image statistics. So, that is possible that has been done, that is called a balloon force, which makes the curve move in any direction, expand and compress, the gradient vector fill, also remember, sometimes image is pretty smooth, contains maybe one or two objects, but away from the objects, there is no strong gradient.

So, if initialize slightly further away from the boundaries of the object, your curve will never converge, so it will get, it will just keep going very slowly. Of course, if you continue to do time stepping for thousand of time steps, then maybe it will converge, but that is very slow and not practical. So, something called a gradient vector field, wherein you think of it as smoothing the gradients from the edges away from it you have a very strong edge, but then you can maybe smooth it a bit. So that the gradient values are kind of dispersed around the object a little bit farther than the boundary on either side.

So that lets the constraints convert smoothly, again, this is accomplished, this is non-trivial to accomplish, you have to solve one more functional, you have to derive one more functional, and you actually formulate the Euler Lagrange there, and then you get a gradient vector field, which you use for propagating. So, this is kind of two step, kind of think of it like a two-step problem. But again, this is very, very popular, and people have used it very effectively. Because for a long time, these methods were there, only one switch, which can be used without too many heuristics. Because for every image, you cannot make up heuristic every modality you cannot make up heuristic.

This comes from very simple sound principles. So thus, this concludes our overview of snakes. Like I said, some of the derivations I have skipped, I have showed you the general principle. So, you can plugging into the Euler-LaGrange equation actually take the derivatives, it is not too hard, if you have done plus two mathematically you should be able to do, the only thing the only tricks there would be there is a discretization, et cetera, you have to actually work out it again, that is also not too hard.

Some of these, like the only two formulas you need to know are the ones I showed you, and then you should be able to do. All right. So, we will end this class here. We look at further development on this, because like I said, the disadvantage is keeping track of the Snakes, hard problem. So how do we get over that? That is where we run into level sets, so we will talk about that in the next lecture.