Image Signal Processing
Professor A N Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology Madras
Lecture No 16
Research Challenges

Whatever, we have done till up, before we shift gears and go through a real aperture camera I wanted to kind of open your eyes to some of these things what are called research challenges. So if you think that in the camera motion, okay (()) (0:27) done stuff, so it looks like some issues when there is a depth change some issues here and that but then beyond that probably resolved.

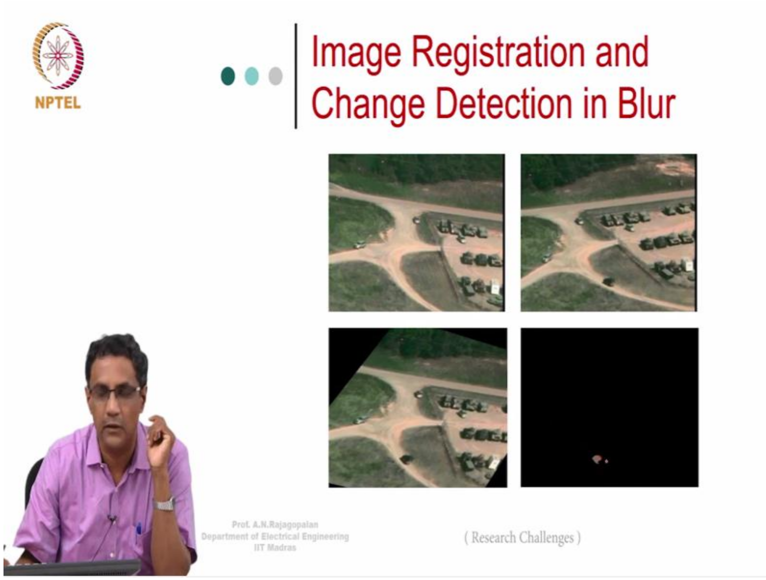
Is this all done? If you ask that question, the answer is no, it is wide open. It is wide open for those of you who are interested the unfortunately the world has come to a stage now that it whenever I see anybody on the, so it is like whenever we walk on the corridor, you just find everybody, you are an important man, if you have a cell phone, you are not an important man if you do not have a cell phone.

You will see that every other, every other student or every other person is so immersed I do not know WhatsApp such like that and so I felt that maybe out of this 40 odd people at least maybe one or two people who are not so probably that on to cell phones and maybe they also think that there is a world outside of the cell phone. So I mean, I am saying it is a fact when I see around that is what I see.

So, the idea is that if you, so what I thought I will just walk through some of these research challenges, this is not for everybody, this is just that this open your mind to things that are out there that are actually interesting that are many of them are still unsolved. And for those of you who are kind of a research oriented, you should also know that, there is one part which is like, like easy to follow and which is done stuff.

And then the other part where you sort of where if you want to escalate the bar, if you want to raise the bar and say okay now, if you raise the bar what else can happen. Then this, this will kind of give you a window, window for that. So prior tom so all this, so I will tell you a few of these things, what is that?

(Refer Slide Time: 02:05)



And so, so the first one that I just wanted to point out to you and these are all from, from my lab only, all these images are our own so it is okay so we can share with all of you. So here is an aerial image. And, and if you really notice, so what you want to do is you want to, you want to align the two images. So these aerials we can sort of assume that probably the world is flat, even though we know that world is really not flat there is of course variations and also there are cars and so on, trucks and so on.

And these are real images, this is not something that I have synthesized or something these are actually real. And it is called, it is called a virat, vrat is just a particular dataset. So it is available for everybody. So, so you see these two images now there is one on the left and there is one on the right. And, and you see that, you see that there is a change that has happened between the two, what changes do you see that when you see with your eyesight, what do you kind of, what do you kind of say notice?

You notice that that there is a vehicle right there is a new vehicle that has come in, which is which is this truck right? Some truck has come in here. Anything else that you notice? tThere is some kind of some kind of motion the camera, cameras undergone so it is not the same viewpoint. It is not like you are looking at it exactly from the same point. What else?

Student: Car is moving around.

Professor: Sorry cars is? Car is moving around. Yeah, agreed, so sorry. That car in the car and I already indicated, that is a new object in the scene. Then we are saying that there is some camera motion between the two images and what else? So what she is saying is the, the first one is slightly blurred, again, I mean sometimes if you see here, it will be much more it is easy to kind of supersede, otherwise it is going to blow it up on, on this thing unless you have a projector that can show it all.

You may not be able to appreciate but for those of you with, with satellites, it will kind of notice that notice that there is some blur in the in the first one. So the whole point is right. So, if you if you try to compare these two images, so if you go back and look at the sift thing, which I kind of talked about yesterday that you can still use, okay fortunately, some features work still work even, even if there is some blur, the sense that what it will mean is and, and yesterday by the way, somebody asked me later a question outside the class, I know as to how it is, what is this sift kind of, do you give it point?

No you do not give any points. Saying that you show this image, you show this image and then it is job is to find out all such interesting points, whatever it things are interesting. It will it will automatically pick we do not go around and tell look at that region or this region and so on. And so, it will just pick for example, it, it might just pick that you know, it may just look at those car, the edge of the car or something and it may think okay that there is something interesting there, which it can which can then also relate here.

So, so, when I say correspondence, what it will do is it will tell that there is an interesting feature here, and that feature occurs there in the other image inside some other location and similarly, something here and something there may be something here it might pick maybe, maybe there is something out there, it might pick there. Pick at the other sometimes it can also go wrong, that is a reason why I said that we have to use some kind of an outlier handling otherwise we could we could go wrong.

Now, so in this case, what I am saying is even if you use SIFT, SIFT features, and suppose they are able to align the two, right, once you get your SIFT, that means you can align the two images. But then after you align them, if you slap them, one on another, they will be reasonably aligned now, because of the fact that you have finally you have already found what is that homography, that map's the two.

But if you try to subtract the intensity is thinking that that will tell me that that will should probably show up only this object it would not because of the fact that, that that the first guy has some has some kind of a blur in it. Now why it might have happened is because there was something which was just moving. You had an aircraft which is moving and it kind of and it kind of captured an image.

So, so when you actually move, you integrate scene points. If you are still right, you can imagine that there is only one scene point that is hitting a pixel as long as nobody moves, there is only that photon only coming from that scene point only. But suppose it start to move then it means that some of them so that it means that the same pixel also get from some other fellow from some other neighboring fellow, so all these will get integrated and then you start seeing the seeing blurring effects.

So, so similar, so, if you just slap one on top of the other and then say that I can subtract now and I will try to and then I can get the changes, so you can change the we detected, then you will go wrong and then it also indicate changes in many places, simply because intensities do not match. So, so which is the reason what is important here is that one of the key challenges is that when there is blurring, especially for aerial images, it is a problem.

Whenever you are moving, you are capturing and so on. You cannot simply assume that these images, images will be sharp and so on. And therefore, you can do it in different ways. The goal is not to really not talk to you about what is how you solve this and so on. It just wanted to show you that there is so much more to all of this, then what we actually do in the class, what we do in the class, is simply to lay the foundation just to know what lies out there, what simple things can you do.

But those of you who want to go for higher studies or those of you who are already in higher studies, who are already PhD students, who are MS students here or the dual degree or the B-tech guys who are going for higher studies. So, for you people, it is important to know that there is much more to it than, than what can I say meets the eye in the class.

So, which is why I thought I will actually show you these things, these are these are fairly interesting. So, here of course, you can see that this dark region has come because we have done some alignment, and then what you need to do is one of the ways is that you should actually compensate the intensities now.

Either you can, either you can make the other guy sharp, this in you can sharpen this and then match or you can actually blur this guy up to the same level as, as an unmatched what is called re-blurring another is called it is called de-blurring. So, de-blurring means you remove the blur so that they, so that both are sharp and you match them or you reblur which means that one guy which is already sharp you kind of blur on top of that.

You apply some blur so that it looks similar to the other one. Again what should be the amount of blur, is the blur a constant all over the image, what are the blur is changing all those will still read you can be I mean, I can talk on this one problem for an hour okay so, so each one is very involved but I just wanted you to know that there is something like that

So, if you do that kind of illumination compensation and then if you do a detection then your change then you of course rightly, rightly you find that there is only one object there that is actually the change between the two the rest of it is all artificially induced because of some blur because of whatever else, camera motion and all. So ideally, you should be only in detecting that object.

(Refer Slide Time: 08:40)

Next is this. What do you think this is? This is inside our campus. I hope you recognize your own campus. Now this is called, this is a pan of photography pan. Pan photography is a no name for this. Okay this is one of my PhD students work pan photography. If you go and check what is pan photography? What it will, what is says is that let us say that we are inside our campus we have we have very nice examples inside, inside our own IIT Madras.

So let us say you see a deer that is fleeting and then you take your camera, and what would you like to do? Ideally, you would like to move the camera in sync with the deer so that in all your images, the human deer looks, looks focused. I mean, as long as imagine, imagine that that if you can move in a manner that you are completely in sync, in sync with the deer as it moves, you cannot see any blur because you are actually moving at the same rate as a deer.

So all these relative points are all kind of locked so, so you cannot see any blur because there is no relative motion between the camera and the deer. Now if you are an expert photographer, is there anybody here who is, who is an avid photographer? It is okay you can raise your hand. Anybody, anybody that has heard about pan photography?

So what does actually means is that so if you are an expert photographer you can actually, you can pan in a manner that this object of interest stays locked as it moves. So, ideally what ideally what do you think? Suppose I suppose I showed you a picture, suppose I showed you one shot of that, what would you see? What would you show to a person or how would that image look like?

Suppose somebody did that and he is an expert. Let us say this boy here an expert in photography. He does that, what would be the, what will, what will be the image look like there is a deer in front of Himalaya mess. So, you will see that the that the deer is focused, but what happens is the background, the background will be blurred by how much? By the same motion as that of the camera or the actual deer, whatever the deer emotion was because your camera motion is exactly in sync with the deer.

So that is how a pan photograph is supposed to look like. But then if you and I take a, take a camera today, and then if you try doing that, and you would not be able to do it. Not, it is not, it is not possible for every one of us to be able to be to be so exact. So, so what we do is, we simply take, but it can so happen that we get a moved that will be moved probably a little too slowly or moved and a little too fast, and so on.

And all that can happen. So the idea is that if you were to capture a bunch of rains like that, now what, what will typically happen is you can also have these two not being at the same depth, the sense that you can have the trees far behind, and there is a there is a deer, right to too much in front, in which case, where you have to worry about the fact that there is not, the world is no longer a single plane.

If you are fortunate, then maybe the deal is far away, and therefore the whole world looks planar. And then you move and there is a foreground object that is moving. So one one of the things that is actually done is even to get of locate where this moving object is, so what will typically happen is, you, you compute these SIFT features, across the frames, and then you match SIFT features across what is a moving object versus where the where you have a background.

And then you segregate the two and then you can kind of re, what you call realign this, this moving object such that it stays in the same place. See for example the second video this is playing. What do you see between these two, there is a kind of subtle difference. Do you notice that?

Student: The background looks stationary.

Professor: The background almost looks stationary. So, so it just looks like it looks like only this guy is moving and then the background is completely stationary. So what you get as input video is this and then eventually what you would like to show is this, where, where the this guy is completely in focus. And then the camera motion has been imbibed by, by the background. That is how if you have an experience if you are an expert photographer, it is what, what would have happened, this is how you would have caught that image.

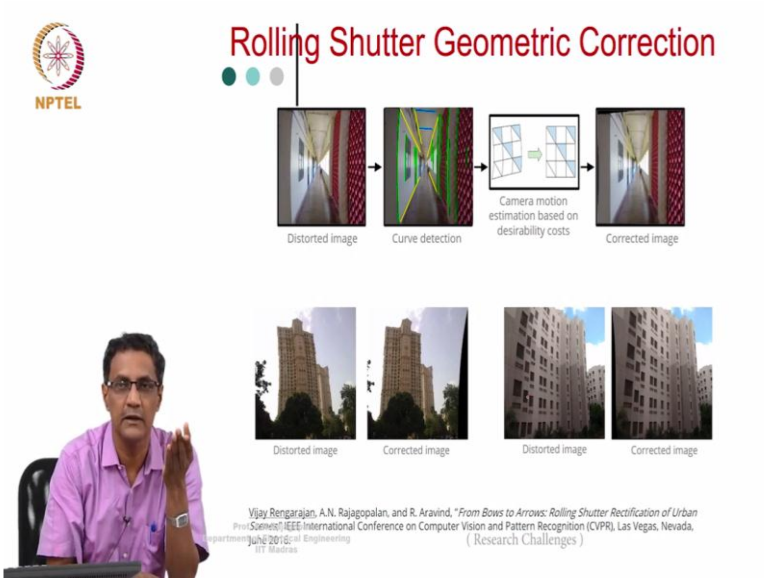
And of course the by and then you can complicate this problem. You can say what is the foreground is also blurred then what do you do and so on? Let us not get us into all that but the idea is that something using SIFT feature something based upon this hypothesis or homography idea that we are in, we are kind of realigning and so on.

I just wanted to show you as an example of something that you have learned something that you can apply, of course with some kind of background about, about this particular problem, but I am just tried to say that even with simple things that you have learned till now you can do things that are interesting.

(Refer Slide Time: 13:29)

Look at this, this is again, again a leopard or something.

(Refer Slide Time: 13:34)



Now by the way, so, this is a rolling shutter, I think I showed you when I maybe in the first class itself I showed you a rolling shutter example the rolling shutter is like is an extreme case, in the sense that and right now we are talking about one homography, right that matches two images. Now an extension of what you have already learned is when, let us say see right now what we are assuming is that all the sensors are kind a opening at the same time or when the camera gets, gets exposed, right, all the sensors, sensors get exposed.

And therefore if there is a camera motion that is global, so every, every sensor is experiencing, experiencing the same motion. Now what happens in a rolling shutter camera, which is what incidentally all of your cell phones have, is that you do not really expose all the sensors at the same time it is called a rolling shutter because it is like saying that if you had a shutter that you would that you would open slowly it is a rolling shutter No, it is like a car shed where you open the shutter.

So what happens when you initially some things are exposed then some things and some things and so on it. So what will really happen is as the as the scene points keep coming in the, the, the let us say if you think of the first row as being exposed first, then after some time you expose the second row, after some time, expose the third row and so on. That what will happen is as long as there is no motion does not matter.

Then the problem is exactly identical to what you had. At the end of it, is there was no motion then of course you will get a nice image, but then what can happen is your camera could have actually moved during the time that these rows are being exposed. So what will happen is the first row the motion that it saw, imagine that you are opening it for a certain amount of time the exposure time is the same for all the rows.

But what will happen is when you get a delay it and then open this next row you will see that there is some motion that they both see, but there is some motion that this guy never saw which the first floor saw and there is some motion that this guy sees later that the first row never saw.

So, what will happen is when you do like this what can happen is a straight line which ideally should come as a straight line can now start to bend or can start to look skewed and all that it because it is all completely dependent upon how you move the camera it is all a function of what you do with the camera. So what will happen is so you actually end up, end up with, with images like this, this is our own ESB corridor.

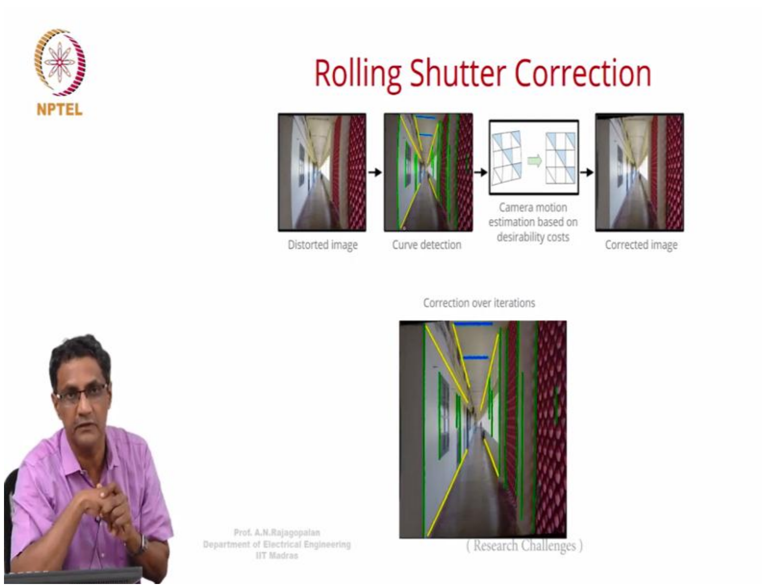
So you see that is you see this bend here. Now ideally, what you need to do is if you want to bring it all back correctly, then you then you are almost saying that you know there is a (homography), there is a complete, there is a continuous change in motion as you walk through the rows and you need to know what is that motion on a kind of row to row basis in order to be able to rework and be able to get back, get back what should be a clean image.

Do you see the point? I mean, on the one hand, you had this one homography is the simplest case. Now, ideally here you could have if you have n number of rows, you could have n number of n number of homographies to solve for that is not the way we normally do it, because people do some interpolation to make the problem simpler, they will pick some key rows for which they find the motion for find the homographies, and then the rest of it, they will simply interpolate across the across the rows because typically camera motion is not is not very arbitrary, it is not very abrupt.

So you can assume that no, there is some smoothness in the motion. Therefore, you can have interpolate the camera motion and that is how the holography that you need and then and then you can still align and then get back get back something like as you see this guy so this would be a corrected image.

And again how you do the correction whether, now there is a lot, lot to this, it is not so straightforward as I am telling but I see the other building they slant and fall off anytime. But that is because of the way you have captured and then if you straighten it up then I think that, there you get it. Similarly, this is I think one of your hostel blocks if I am not mistaken. But it is all this so this rolling shutter correction is again completely based on this homography idea and so on, that you have already learned but except that now you can escalate this, this challenge.

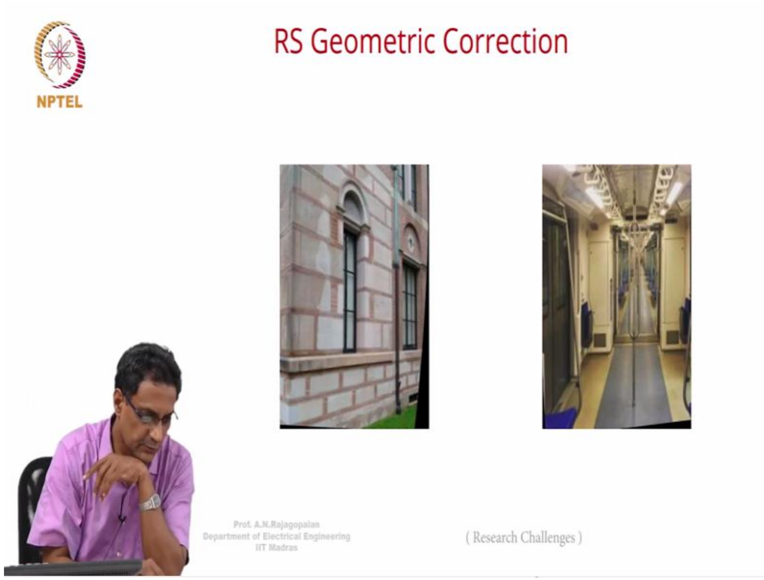
(Refer Slide Time: 17:37)



Look at this case this is the same thing but then this is like, like some animation which will tell you how these like all these curves are getting straightened up. If you watch this figure, you will see how eventually so these curves, okay that are kind of read initially there. So for example, if you look at this guy, it is like curve no. So as you keep going, you see that this kind of straightening, straightening it up. So it is an optimization problem, this one is solved as an optimization problem where you try to maximize the number of vertical lines that you can expect to see.

And all these assume some prior and so on, it is like an indoor the same thing will not work I mean, if the if actually everything was curved in the real world, you would go totally wrong because you may think that everything should be straight. Whereas if everything is curved in the world, and you are then you are finished, so some assumptions are there. Again, so these are fairly difficult problems. So, therefore, one has to make some assumptions to be able to see this through.

(Refer Slide Time: 18:31)

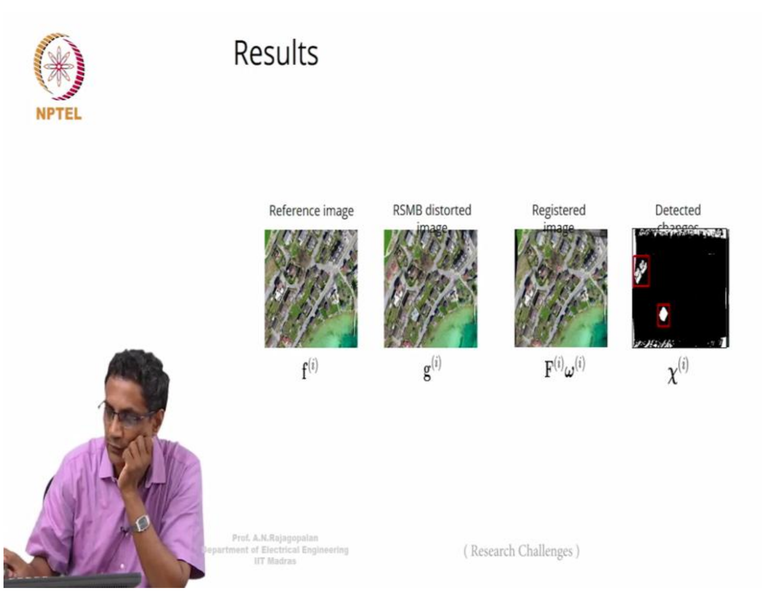


Yeah, this is again another example about a geometric (()) (18:36) this is inside a train, that is a building. So, so this animation will tell you how it kind of straightens up can you feel that? So how things are especially the pole if you see in the that metro train, the middle one, I think that that may be the one that that will give you a good idea and simulate the building I think you can probably look at look at this edge and you will see how it straightens up. So, this is again an extension of this homography idea, but extension to the case when you have a much more complicated situation.

(Refer Slide Time: 19:10)

This how does how would the poles and all will start to look like and so, you have this our rolling shutter effect will, will cause a pole to look like. And then if you actually do the correction, this is how it will look like, this should how, this is how it will become. Similarly, some building here, they are examples too.

(Refer Slide Time: 19:27)



And yeah, and I also wanted to show you so you the show the case now on top of this, if you want to escalate the challenge further, you get the rolling shutter plus blur. That is like the other extreme, it is not only that you have a rolling shutter, but then the exposure time is so long that you are starting to incur blur now.

Blur actually depends upon your exposure time also, the more, the longer your exposure time, then your motion is also going to be that much longer, and therefore you are likely to integrate points from like many, many scenes, you are likely to integrate many, many scene points at a pixel. So in this case look at this so, so here is here is an image that you captured when you were still and then when you are moving, the blur maybe you are not be able to appreciate it so much but there is, believe me RSMB that is why it is called Rolling Shutter Motion Blur.

And then, then when you align the two right now when you align, you have to account for the rolling shutter motion. So, what we normally do is we would kind of forward warp this guy according and so that so that, it looks like looks like the one that is already that is because they are doing the inverse warping is actually harder, if you go the forward way, forward warp this was to be able to make this look like that.

(Refer Slide Time: 20:41)

And then let us see. So, is this is your original reference and this is a distorted guy, then the original reference you try to, now you can, now we can see the loss in sharpness, Can you? Between the two, now, you take the original reference, and if you compute is homographies and all, you compute. I would not call them all homographies, typically you have to know, know what is what is the camera motion row wise and for which of course you know that and there is also correspondingly, correspondingly homography.

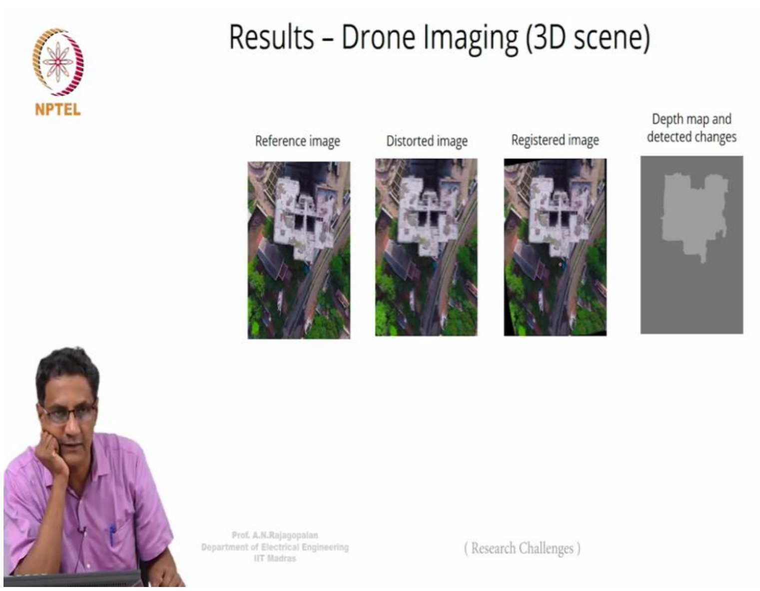
(Refer Slide Time: 21:10)

And then now, you see, the reference guy has been brought to the same blurring level as that of the distorted image and also in terms of the geometric this one there is a distortion, it will also match in terms of that.

(Refer Slide Time: 21:29)

And now, if you take a difference that you will find where the objects are where the actual changes have happened. Because you cannot subtract these two directly you will go totally wrong. These two if you subtract everything will show up as a change when ideally some of these changes and all, so for example it this road has become like that that is not a change nobody I mean physically went and changed the road there. So things like that, so which we can attribute to camera and not the not to the world. this is one example.

(Refer Slide Time: 21:59)



And then another thing this is a kind of a drone. So here I wanted to show I keep on saying about, about the parallax effect I keep telling so here I will show you the parallax effect so look at this building which is kind of relatively tall right compared to the ground.

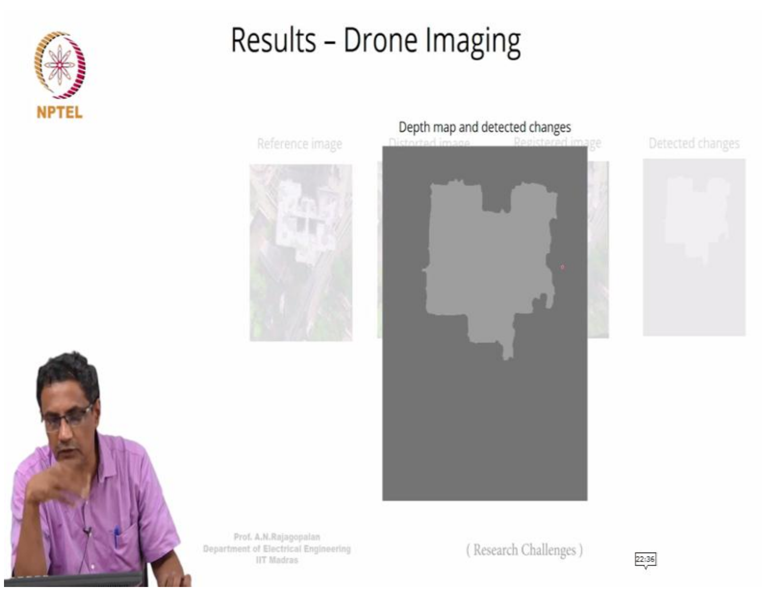
(Refer Slide Time: 22:14)

On this if you move okay this, this let us first image.

(Refer Slide Time: 22:18)

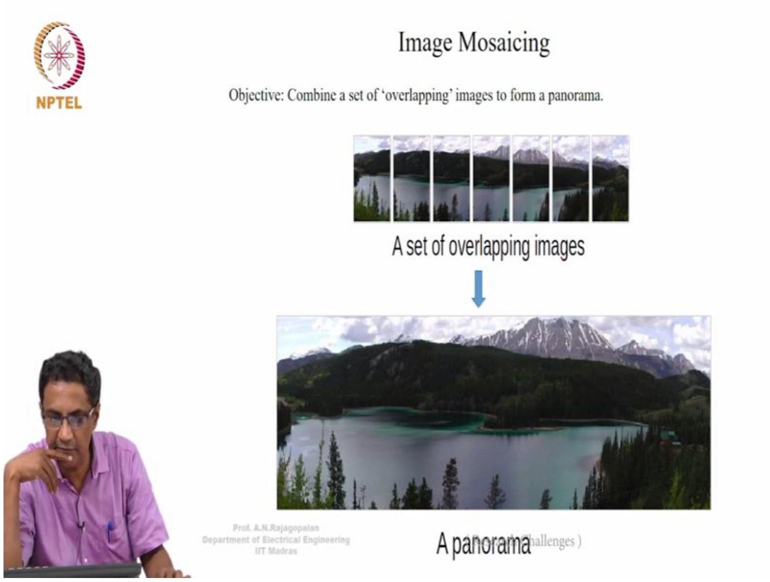
This is suppose second image which is actually a distorted so when you when you are actually align the first one that you will see that the building will move at a kind of a different pay, see can you a see the building moving much more than the ground. Are you able to see that? It is almost like somebody is taking the building and moving it around. So, so that is a parallax. That is parallax, okay so unless you translate it would not get that and when you translate pixels closer to the camera will tend to move more.

(Refer Slide Time: 22:51)



And then on this also you can do it so this for a 3D scene is something that I was talking to yesterday homography is you cannot apply if you have 3D. But then if we can segment is called a layered homography, in the sense that you have different, different planes. And then you kind of know, you kind of split the scene into multiple layers and then have a homography for each layer, and then you can do things. Again, this all this is all stuff to just be aware of not that we want to do all this and of course, but it is good to know that such things exist.

(Refer Slide Time: 23:24)



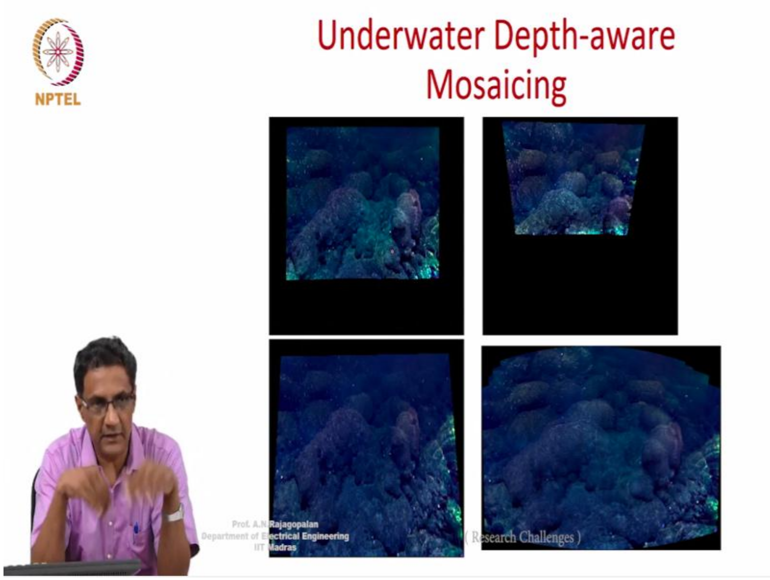
Then switching of course, this example I showed you earlier itself, so this is something like this is what is what you are going to do, except that we do not have something like this in Chennai. So you will have to live with whatever you can capture with your cell phone, go around, hunt for deer or whatever, but they will but they will move so okay.

(Refer Slide Time: 23:41)

Now here it is. So here it is another thing so rolling shutter rectification and if you had to stitch so if you see here right there is also a rolling shutter effect. See the see the sketch. See this one, (()) (23:56) getting so it is like saying you are actually moving with a camera now you have a rolling shutter camera which means that the motion as you are going through not all rows are experienced experiencing the same motion and therefore, if you take it with a regular camera this guy would have looked exactly the same in this picture also.

But now you see that some things are getting stretched and so on, that is all happening because of that rolling shutter effect. And therefore what you need to do is now you need also find out what, what motion happened row wise and then against it. So it is much more complex than what we saw, what is your lab assignment is much more simpler, you are not assuming anything like that. But really all they and anyway you can, any of you who is interested, you can go read these papers. These are all some CVPR, some CGLR journal papers and so on.

(Refer Slide Time: 24:46)



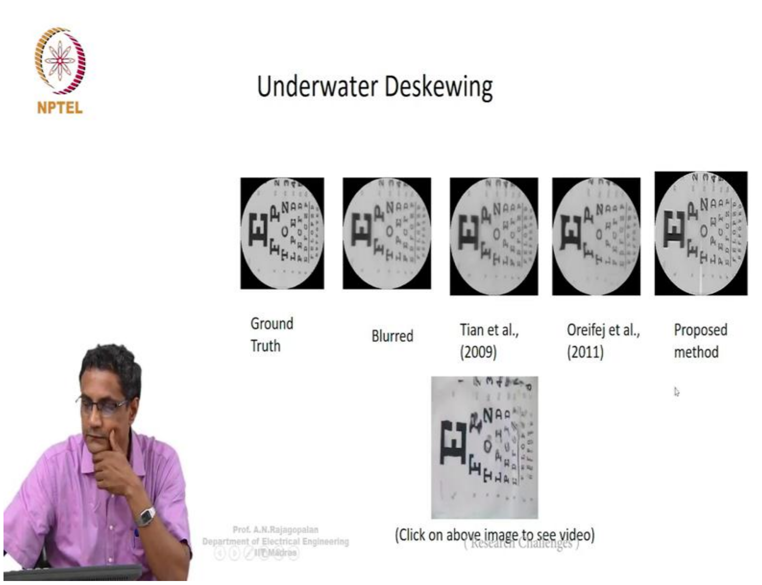
And then underwater, so this called the depth of aware stitching. And what does that mean a depth aware when we say because especially for sea, for underwater, it is more important, because for underwater one of the problems is you cannot be too far away from, from where you want a image. Typically, when you want a image of the floor of the ocean above it so people have remotely operated vehicles what are called ROBs.

So these ROBs is even in NIOT. Have you heard about NIOT, National Institute of Ocean Technology right here three kilometers from our campus, they have been ROBs that can goes way down and they can actually map the ocean bed. But then what will happen is because you have to take your own artificial source, let us say you are taking an optical source and because it has to be artificial that the sunlight does not come 20 meters you do not see anything.

So what happens is so, so if you stay too far away, to get by that I can avoid my what my homography issues then what level you would not see anything because the light is not reachable. So you got to be really close in order to be able to see the surface problem. When you are really close, that the ocean surface is not flat, it will have all these undulations, the rocks you see formations like that.

And therefore when you move and you want to map the bed, you have inevitably face this situation that you have to do a depth aware kind of a stitching. And all those things are again, these are all so it is should not to be restricted to surface scenes like outside surfaces by which I mean what you call (()) (26:20) surface there is also things that are happening under underwater and so on where again how these oceans are valid.

(Refer Slide Time: 26:26)



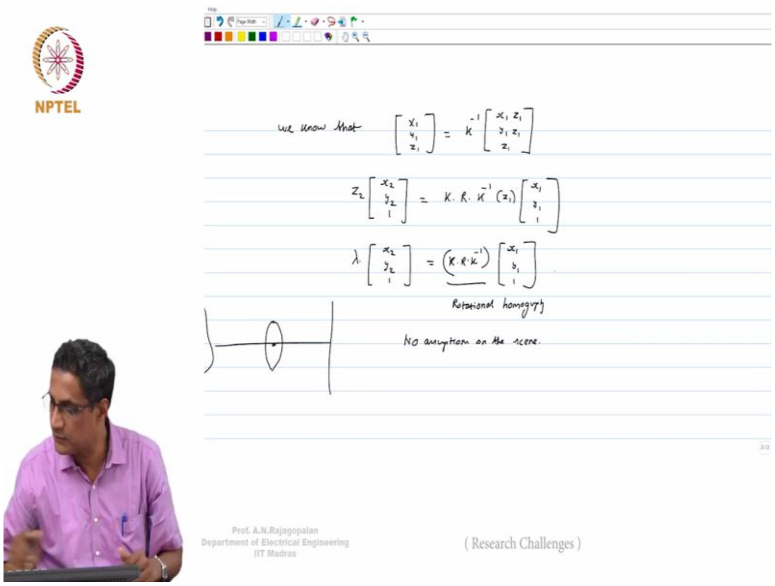
This again another example that I had given you that day. And, and I have shown you this image now if you see, this is like, I mean in this case also there is a motion but here it is here it is like. So it is like saying that if you had a wave, imagine that imagine that you had a water wave. And suppose, suppose I kept, I kept a flat thing like this a paper on which I have something written on it.

And suppose I keep it on the bed, imagine your swimming pool, imagine you are in the bottom of the swimming pool where I keep this and let assume that the sunlight is enough to fall on this. So, when you have still water you will see this as it is, but in the moment the water begins to flow, it will start introducing what is called a skewing effect. And because, because the rays will start to bend depending upon the slope of the water and then you have, you get what is called skewing effect.

Now, this is not exactly a homography but then it turns out that it can be simplified in some manner and you can actually arrive at a very, very elegant solution to remove that skew. So here it is almost like saying that a every pixel will move depending upon the one the slope of the water we saw at that time.

But, so, you see capture a bunch of frames, but you can actually eventually solve it as a bit of a de-skewing problem, but yeah this is how it look like so you see. So if you freeze it, you will see something like that is what you will see. now again, it is some skew. So, to stop the video in between, you will see each of those threads. There will be a few threads that are not able to stop in depth.

(Refer Slide Time: 28:06)



So all these are things like that that go far, far ahead of what we have learned but then this only for those that work in that area and so on. So for you it is enough to know what we have said already in the class, though let me now go back to, let me now go back to my get to loop what I call a real aperture camera to our next stop.