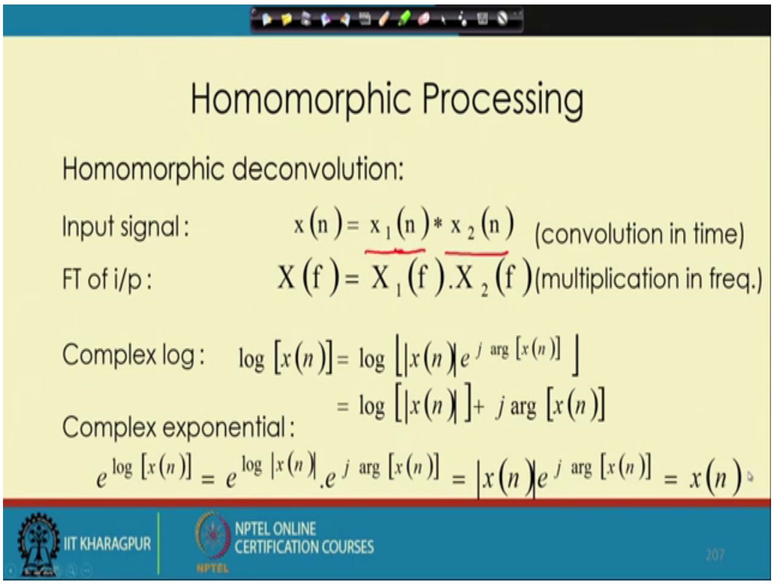
Biomedical Signal Processing
Prof. Sudipta Mukhopadhyay
Department of Electrical and Electronics Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture – 26
Homomorphic Processing (Contd.)

Now, we will look at that another application of homomorphic processing that is homomorphic deconvolution. Sometimes the situation is even more complex. The two signals which we are calling that they are interfering with each other; one is of signal of interest; another that is what we name as noise; they are actually getting convolved with each other. Now convolution means it makes it even more difficult to separate them up, but we can make use of the homomorphic filtering to take care of them or in other word to separate them to get rid of the interference.

(Refer Slide Time: 01:10)



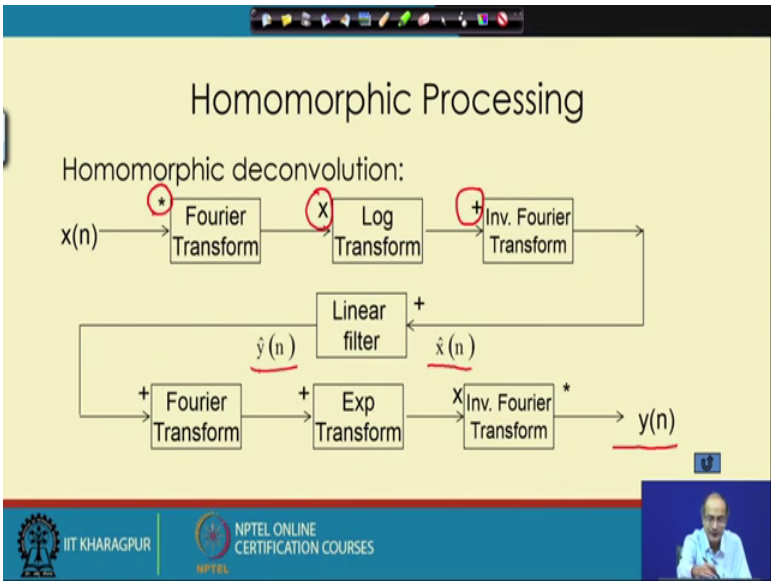
So, here let us look at that example of homomorphic deconvolution, in this case, say the input signal of the object signal xn, it is actually consists of the 2 parts again x 1 and x 2, but they are actually convolving with each other. So, we have actually that one of them we can say that the signal and other is a filter impulse response there the 2 signals they are getting convolved with each other that is the way they are we are getting them and one of them x 1 or x 2 is of interest depending on that what we are looking at, but unfortunately we get it in them in such a complex form and we cannot because they are not additively combined we cannot use the direct linear filters, they are not multiplicatively combined we cannot use the previous technique. So, we need to look for a different way to take care of them that.

In this convolution in time actually that what we can do to take care of it we can think of taking them in the frequency domain. In the frequency domain, we know that signals which are in the time domain related by convolution; they become actually product of each other now using that fact that multiplication in the frequency domain. In fact, now we are comfortable, we know because that that if the 2 constituents they are actually combined in a multiplicative way we already know that how to separate them out. So, only thing the operation becomes a little longer.

So, here faster than to separate these multiplicative components what we need to do we need to take complex log and out of the complex quantity x n if we have taken the log then we could actually separate the that real and the imaginary part of it that the real would be the log of the absolute value or the mod and the imaginary part would be the complex angle and then we need to take also the complex exponential in this that total game because we need to come back to the original domain.

So, we need to cancel the effect of complex log and if we take actually the complex exponential of the complex log of the quantity we can get back the same output and now we will look at the overall system.

(Refer Slide Time: 04:18)



First we take a Fourier transform of x n that is our the time domain signal and that the operation here before the transform it was convolution the operator is given in there in the left hand side after the Fourier transform, the multiplication changes to sorry the convolution changes to multiplication the constituents that now combined with the multiplicative operator.

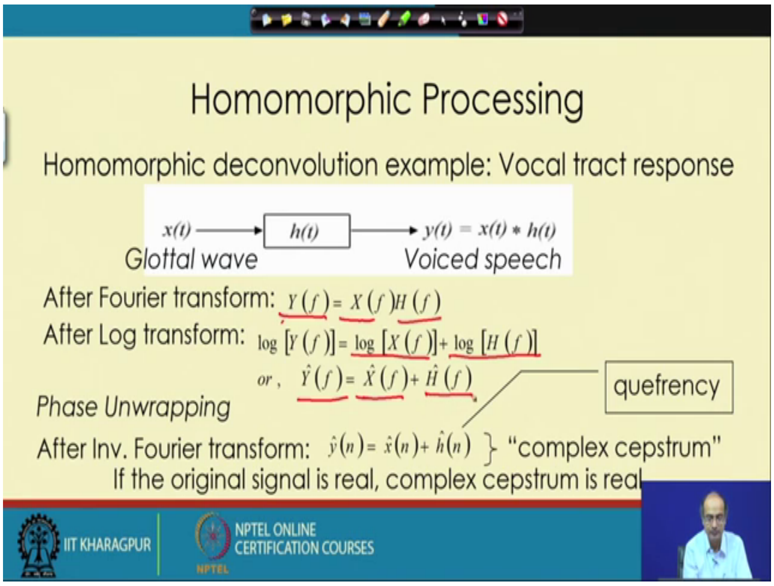
So, after taking the Fourier transform, we get rid of the convolution operator and we get the 2 constituents or multiple constituents they would be actually their multiplicative components now who remove the multiplicative operator we take the logarithmic transform or complex logarithm and that gives us the form that where the components would be combined with the addition operate now after getting that it seem that we are thus they are, but we need to do one more operation that initial data was in the time domain and just because that our log operator works on the that works better or that we need to get it before that taking the logarithm that the constituents should be multiplicatively combined, we have to take the Fourier transform.

So, we have gone already in the frequency domain. So, to do the linear transform in the time domain we need to come back to the time domain again. So, we need to take an inverse Fourier transform and then we get something called except n and these can be passed through a linear filter and we can get y and provided these components, they are well separated that they can be separated in that way and after that again we need to get rid of the effect of the logarithm what we have taken and as a logarithm has been taken after the Fourier transform.

We need to follow the same set of operations, we need to first take the Fourier transform, then take the exponential transform that complex exponential to get rid of the effect of the logarithm transform and because we have taken the Fourier transform to nullify the effect of it we need to take the inverse Fourier transform and again we see that operators keeps on changing we started with that addition initially before the Fourier transform after the Fourier transform because it is a linear operator we had that addition operation now after taking the complex exponential, it becomes multiplication and that multiplication after the inverse Fourier transform it becomes convolution.

So, we get back to the original domain here and we get the output as y n. So, that is the way we can actually have that the deconvolution using the homomorphic processing and we can actually separate 2 components which are actually combined through that convolution operator now lets us take an example where actually such kind of thing is important let us take.

(Refer Slide Time: 08:32)



That we have some speech signal now when we utter the vowels that is called the voiced part of the speech that the signal comes from the glottis that is called the glottal pulse. It gives a repetitive kind of wave or sound which actually gets convolved with the transfer function of our vocal tract that found are that glottis to our lips that part that there is a cavity and if you think of a pipe in between these 2 points that the shape of the type of pipe is not uniform ok.

The cross section changes everywhere and we also modulate that at a different way to make different kind of sounds and because of the change in cross section that it gets convolved with that actually the transform function of the vocal tract and the ultimately if voiced sound comes and we know that what are the common voice sounds that we make or other vowels we utter in English they are a e i o u.

So, in each of these cases we have the voice sound that comes out of our that mouth now if we want to separate the 2 things one is the glottal pulse and another is the vocal tract that they are actually convolved with each other and we need to did the help of this homomorphic deconvolution. Now before that first try to understand that why you would be interested in such kind of activity there could be many applications which we will need only one component of them.

You see that when we are uttering a sound, there are 2 parts one part is specific to the person who is uttering that if we are interested in the that biometric identification; that means, we want to identify that person with the help of the voice that what that person is uttering; that means, whether he is uttering a or e or u or some other word that is not important we want to know that what is the actually person is there and we would try to get that from the vocal tract because the just like our retina or thumb print that the configuration of the vocal tract or that cross section is also used for each person ok.

So, from there, we would try to identify that person and there that the glottal pulse which is determining that what we are what kind of sound we are making that is actually becoming irrelevant on the other hand there could be some other application where we are trying to have a voice activated device in that case those part which are actually personalized that actually is causing problem for the detection or taking that command.

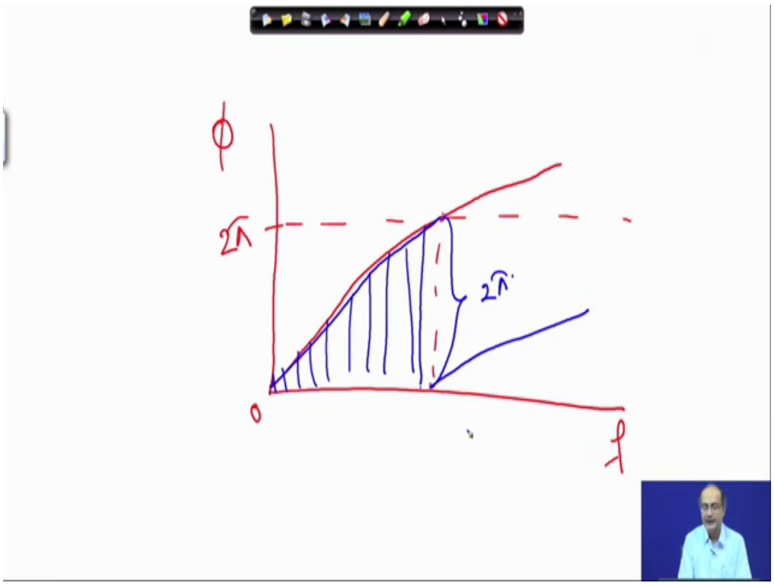
In fact, that is the person which is the sound part of the sound which is specific to a particular person it is actually is unwanted component we are more interested that what is the command given the command could be on, it could be off for a device or we may say play or we can say next we may say previous to change the channels.

So, those kind of thing we want that any person may be uttering those sounds we should be able to capture that identify that and for that purpose, it is best that we if we can rub of those parts which are specific to a particular person and take that which is common for all and take it forward for that understanding that what is the comment. So, such applications are there that both the components are important, but in different cases.

So, now, let us take that how we proceed here in this particular case. So, first we take the Fourier transform that that and after the Fourier transform that signal what we get the voice speech of that y f that we get 2 components x g and H f and once we take the complex log that we get the 2 parts separated that log of x f and log of H f. So, we write them in compact form that y hat f is equal to x hat f plus H hat f.

now comes that fact of phase unwrapping now let us try to understand that that what happens that is as we told that there is a actually that if we look at the one side is the frequency.

(Refer Slide Time: 14:29)



And the phase block that we know that there is some kind of continuity of it, say there is a continuous change in the phase this is the phase say phi and say here it is zero and here it is 2 phi. Now after taking the log, what will happen from here that once it goes to 2 phi 2 phi means the same as the point 0, we will see, it will look like this that it will again build up like this. So, that trajectory will look like this coming like from here to here and then going here.

So, here is a jump of 2 phi which gives a discontinuity and in some cases actually the discontinuities in the phase is not that actually important for example, when we are actually listen to the sound that phase has little actually or less importance though it gives some kind of artificial actually sound that effect, but usually that if we change the phase of say if we actually we can do an experiment that we can have hammer a fork and get a frequency out of it now if we do it at different time we will see that it is giving the same sound, we are unable to get that what phase it is starting that could be a simple actually it takes to get confused that our ears are less sensitive to the phase.

So, such kind of applications you can forget about the phase, but there are other applications like if we take an image if you take the Fourier transform of it and you modify the phase you take the phase part of the Fourier transform another image and then you go back again to the that original domain by taking the inverse Fourier transform you will find that image has totally change. In fact, the phase has more actually important clues in the visual system. In fact, that image from where you are take the phase part of it these image will look like more like that rather than that which has donated the intensity component.

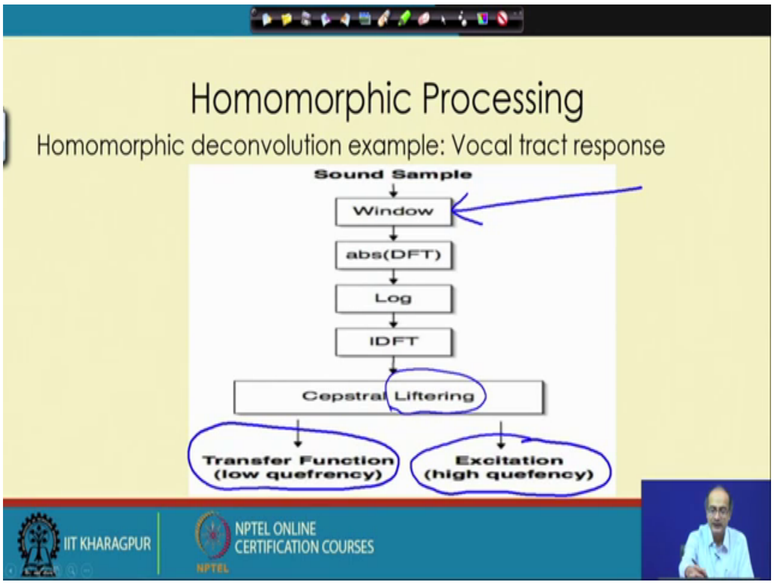
So, in such case, you cannot ignore the phase and you need to restore the phase the simple way that can be done that what we need to do we need to follow that for each frequency, how much is the change in the phase when the phase actually due to this complex transform there is a jump is occurring the value of the jump is above 2 pi right. So, we know it is a huge jump. So, when such a big jump occurs then we simply add 2 pi and we put it back in the real place.

So, that is called phase unwrapping and after that we follow the rest of the operation the main intention is to get back that phase in a proper way; that means, removing the discontinuity and that way we restore that the phase and after the phrase unwrapping we can do the inverse Fourier transform and we get back the signal that which actually is given as y hat that is sum of x hat and H hat and here; however, we need to keep in mind that n we cannot say same as the original that n.

That here we take a term called that it is called Quefrency, actually we are changing the word frequency jumbling that word and getting this new term and that the new name is given for that the new signal that is called complex Cepstrum again you get that Cepstrum part is coming from jumbling the word spectrum.

So, this complex Cepstrum, why we are telling in that way because it has gone through some transforms and it is not same as the original the domain that the time domain because we have though that that Fourier transform inverse Fourier transform is taken in between some log was also taken which makes it different and these complex Cepstrum, it has some good actually properties in the original signal is real which would be the case for the biomedical signal the complex Cepstrum, it is also real that is one of the property that makes the life simple in many cases for example, that that we will see that later bit that that how that can help us.

(Refer Slide Time: 20:23)



So, here we take the sound sample and the first thing is to window that sample the reason is that first we need to ensure the signal is stationary the signal in real life is not stationary. So, we take the other definition of stationarity. We check for the week send stationarity; that means, only the first 2 moments, they would be constant over the time. In fact, that condition is also violated in practice especially for the speech signal.

So, we go for again a more loose definition that is we tell that the signal is quasi stationary quotations. So, quasi stationary means that it is usually stationary and the change is limited. So, if we take a small window within that there is no change. So, we will select such a window of time and limit the signal within the window. So, that the assumption of stationarity is maintained.

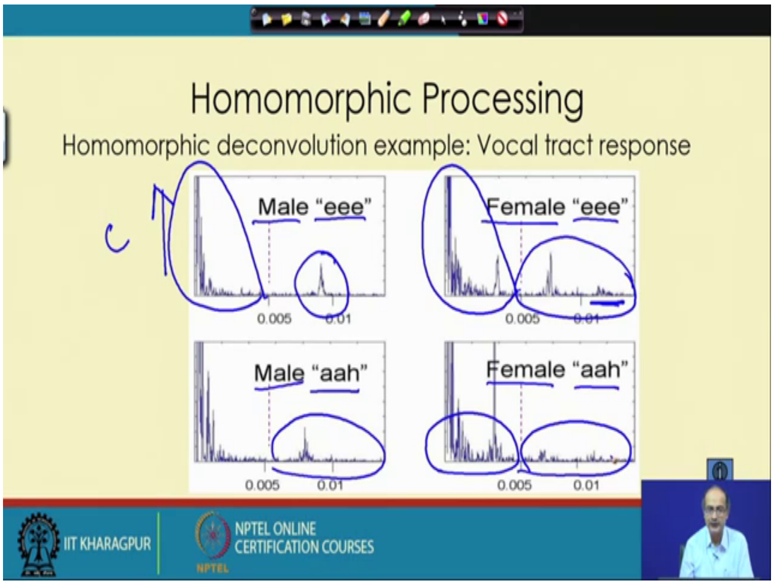
So, first job is to limit the signal within the window next we go for Fourier transform and we have taken the absolute value of it and then we take the logarithm and then it passed through the inverse Fourier transform and for a efficient implementation of the Fourier transform and inverse Fourier transform we have used the DFT and IDFT and then that we get the Cepstrum.

Now the filtering we have a different term for that we call it is liftering, again the term filtering is jumbled here sorry, liftering and the first thing to note here that we have not done the phase unwrapping and we have not told it is the complex that Cepstrum that the reason is that original signal was real. So, the complex Cepstrum also it would be real and because it is real that we do not bother about the phase or the phase unwrapping that we are not interested in that. So, we have just removed that and then go for the filtering in the Cepstral domain what is called the Cepstral liftering and we are separating the 2 parts one side is the transfer function which is a low quefrency part.

Why low quefrency because the change in the that vocal tract, it is slow we cannot change it very quickly and thereby the transfer function, it is a slowly changing I would say that is that signal and the other part that is the that excitation that is coming from the glottis that is here a pulse of him actually that impulses are coming from the glottis or glottal waves are coming that is high frequency term or here, in these nomenclature, we should say that high quefrency part and the transfer function is low quefrency part.

So, by that the general concept of filtering by using the cepstral liftering we could separate them out and we can actually find out that what is the transfer function if we are interested in that if we are interested in the that to find out that what is the command there is given then we may go for the glottal pulse and here.

(Refer Slide Time: 25:11)



We see some example of it. To show the example that what we have done first we have taken that male speech and 2 vowels are uttered first is eee next is aah for the male speech again for the same 2 vowels eee and aah we have taken the female speech and we are going to the that cepstral domain in all these cases.

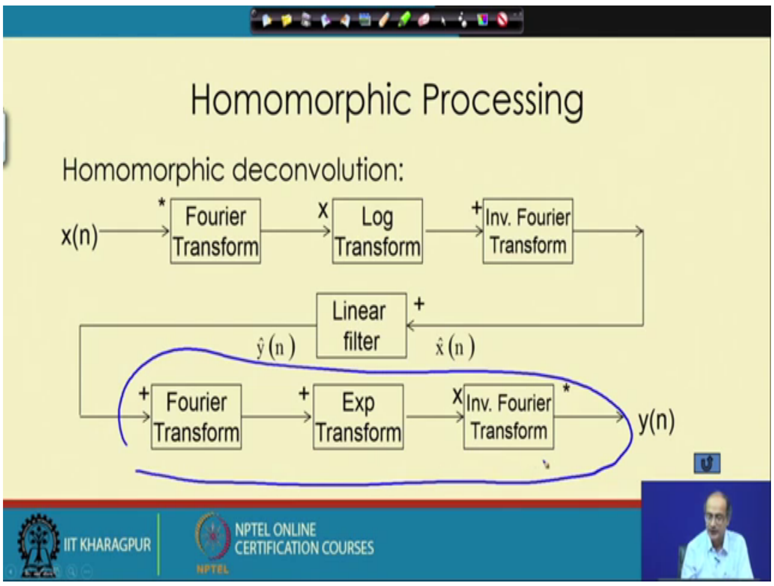
So, when we look at say for the eee for the male speaker we get actually this is the cepstral coefficient that y axis it is giving the that cepstral domain that that here we get actually that this part is the our; the low quefrency the part that is given as the vocal tract and this is the excitation that glottal pulse and for the female speaker we get this is the transfer function and this is the glottal pulse.

Now, when we compare these 2 for the male speaker and the that female speaker for the same utterance that the 2 the transfer function they have some similarity, but no difference is there in the excitation part of it and we see that more actually high quefrency components are there for the female speaker these components that what we are getting that at high quefrency that the parts are there and that give rise to the that comes to the fact that comparing the male speaker, the female voice is rich in high frequency component.

And same thing we go actually a notice for that the other utterance are here also if we look at the excitation part that for the male speaker and the female speaker we see that energy is more distributed to the high quefrency part for the female speaker and here by simple thresholding that we can actually separate these 2 parts out that is in the low quefrency part we are getting that this part we are getting the transfer function and the high quefrency part we are getting the excitation signal.

So, by Cepstrum liftering, we can get that and we can once more look at that the original.

(Refer Slide Time: 28:16)



Actually that homomorphic deconvolution that block diagram that can help us to remember that what are the steps we have to take and here one point just I would like to actually point it out that as we are doing the filtering that it becomes a simple actually a separation in this case that becomes a very simple operation and that that way, we could actually separate these 2 things out.

So, the that and again after that filtering then we need to take that Fourier transform then the exponential transform and Fourier transform again to get back the time domain signal whatever component maybe we were looking at maybe the vocal tract or the glottal pulse we can get that, but we need to actually apart from the rest of the part which we have not shown in the previous diagram. So, that is how we can complete the process and this is with this we actually complete the part of homomorphic processing.

Thank you.