WAVEFORM CODING AND REPRESENTATION

Linear Predictive Coding

Linear predictive coding (**LPC**) is a tool used mostly in audio signal processing and speech processing for representing the spectral envelopment of a digital signal of speech in compressed form, using the information of a linear prediction model.

Linear prediction is a mathematical operation where future values of a discrete time signal are estimated as a linear function of previous samples. In digital signal processing, linear prediction is often called linear predictive coding (LPC) and can thus be viewed as a subset of filter theory.

Filter design is the process of designing a signal processing filter that satisfies a set of requirements, some of which are contradictory. The purpose is to find a realization of the filter that meets each of the requirements to a sufficient degree to make it useful.

The filter design process can be described as an optimization problem where each requirement contributes to an error function which should be minimized. Certain parts of the design process can be automated, but normally an experienced electrical engineer is needed to get a good result.

In system analysis linear prediction can be viewed as a part of mathematical modelling or optimization.

Optimization is the selection of a best element (with regard to some criteria) from some set of available alternatives.

In the simplest case, an optimization problem consists of maximizing or minimizing a real function by systematically choosing input values from within an allowed set and computing the value of the function. The generalization of optimization theory and techniques to other formulations comprises a large area of applied mathematics More generally, optimization includes finding "best available" values of some objective function given a defined domain or a set of constraints), including a variety of different types of objective functions and different types of domains.

LPC starts with the assumption that a speech signal is produced by a buzzer at the end of a tube (voiced sounds), with occasional added hissing and popping sounds. Although apparently crude, this model is actually a close approximation of the reality of speech production. The glottis the space between the vocal folds) produces the buzz, which is characterized by its intensity (loudness) and frequency (pitch). The vocal tract (the throat and mouth) forms the tube, which is characterized by its resonances, which give rise to formats, or enhanced frequency bands in the sound produced. Hisses and pops are generated by the action of the tongue, lips and throat during sibilants and plosives.

LPC analyses the speech signal by estimating the formants, removing their effects from the speech signal, and estimating the intensity and frequency of the remaining buzz. The process of removing the formants is called inverse filtering, and the remaining signal after the subtraction of the filtered modelled signal is called the residue.

The numbers which describe the intensity and frequency of the buzz, the formants, and the residue signal, can be stored or transmitted somewhere else. LPC synthesizes the speech signal by reversing the process: use the buzz parameters and the residue to create a source signal, use the formants to create a filter (which represents the tube), and run the source through the filter, resulting in speech.

Because speech signals vary with time, this process is done on short chunks of the speech signal, which are called frames; generally 30 to 50 frames per second give intelligible speech with good compression.

It is one of the most powerful speech analysis techniques, and one of the most useful methods for encoding good quality speech at a low bit rate and provides extremely accurate estimates of speech parameters.