Linear predictive coding

From Wikipedia, the free encyclopedia

This article includes a list of references, but **its sources remain unclear** because it has **insufficient inline citations**. Please help to improve this article by introducing more precise citations. *(March 2010)*

Linear predictive coding (**LPC**) is a tool used mostly in audio signal processing and speech processing for representing the spectral envelope of a digital signal of speech in compressed form, using the information of a linear predictive model.[1] 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.

Contents [hide]

Overview [edit]

Main article: source–filter model of speech production

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 (sibilants and plosive 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 formants, 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 analyzes 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 modeled 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.

Early history of LPC [edit]

According to Robert M. Gray of Stanford University, the first ideas leading to LPC started in 1966 when S. Saito and F. Itakura of NTT described an approach to automatic phoneme discrimination that involved the first maximum likelihood approach to speech coding. In 1967, John Burg outlined the maximum entropy[*disambiguation needed*] approach. In 1969 Itakura and Saito introduced partial correlation, May Glen Culler proposed realtime speech encoding, and Bishnu S. Atal presented an LPC speech coder at the Annual Meeting of the Acoustical Society of America. In 1971 realtime LPC using 16-bit LPC hardware was demonstrated by Philco-Ford; four units were sold.

In 1972 Bob Kahn of ARPA, with Jim Forgie (Lincoln Laboratory, LL) and Dave Walden (BBN Technologies), started the first developments in packetized speech, which would eventually lead to Voice over IP technology. In 1973, according to Lincoln Laboratory informal history, the first realtime 2400 bit/s LPC was implemented by Ed Hofstetter. In 1974 the first realtime two-way LPC packet speech communication was accomplished over the ARPANET at 3500 bit/s between Culler-Harrison and Lincoln Laboratories. In 1976 the first LPC conference took place over the ARPANET using the Network Voice Protocol, between Culler-Harrison, ISI, SRI, and LL at 3500 bit/s. And finally in 1978, B. S. Atal and Vishwanath *et al.* of BBN developed the first variable-rate LPC algorithm.

LPC coefficient representations [edit]

LPC is frequently used for transmitting spectral envelope information, and as such it has to be tolerant of transmission errors. Transmission of the filter coefficients directly (see linear prediction for definition of coefficients) is undesirable, since they are very sensitive to errors. In other words, a very small error can distort the whole spectrum, or worse, a small error might make the prediction filter unstable.

There are more advanced representations such as log area ratios (LAR), line spectral pairs (LSP) decomposition and reflection coefficients. Of these, especially LSP decomposition has gained popularity, since it ensures stability of the predictor, and spectral errors are local for small coefficient deviations.

Applications [edit]

LPC is generally used for speech analysis and resynthesis. It is used as a form of voice compression by phone companies, for example in the GSM standard. It is also used for secure wireless, where voice must be digitized, encrypted and sent over a narrow voice channel; an early example of this is the US government's Navajo I.

LPC synthesis can be used to construct vocoders where musical instruments are used as excitation signal to the time-varying filter estimated from a singer's speech. This is somewhat popular in electronic music. Paul Lansky made the well-known computer music piece notjustmoreidlechatter using linear predictive coding.[1] A 10th-order LPC was used in the popular 1980s Speak & Spell educational toy.

LPC predictors are used in Shorten, MPEG-4 ALS, FLAC, SILK audio codec, and other lossless audio codecs.

LPC is receiving some attention as a tool for use in the tonal analysis of violins and other stringed musical instruments.[2]

See also [edit]

- Warped Linear Predictive Coding
- Akaike information criterion
- Audio compression
- Pitch estimation
- FS-1015
- FS-1016
- Generalized filtering
- Linear prediction
- Linear predictive analysis
- Code-excited linear prediction (CELP)

Notes [edit]

1. ^ Deng, Li; Douglas O'Shaughnessy (2003). *Speech processing: a dynamic and optimization-oriented approach.* Marcel Dekker. pp. 41–48. ISBN 0-8247-4040-8.
2. ^ Tai, Hwan-Ching; Chung, Dai-Ting (June 14, 2012). "Stradivari Violins Exhibit Formant Frequencies Resembling Vowels Produced by Females". *Savart Journal* **1** (2).

References [edit]

- Robert M. Gray, IEEE Signal Processing Society, Distinguished Lecturer Program

External links [edit]

- real-time LPC analysis/synthesis learning software
- HawkVoice open-source LPC software and API
- 30 years later Dr Richard Wiggins Talks Speak & Spell development

Further reading [edit]

- Alexander, O'Shaughnessy, D. (1998). *Linear predictive coding*. pp. 29–32.
- Lincoln Wallen, Alan Bundy (1984). "A Generilisation of the Glivenko-Canttelli Theorem". *Symbolic Computation* (Springer): 61. doi:10.1007/978-3-642-96868-6_123.
- Alexander, Amro El-Jaroudi (2003). "Linear Predictive Coding". *Encyclopedia of Telecommunications* (Wiley). doi:10.1002/0471219282.eot155.

v · t · e **Data compression methods** [hide]

Lossless	Entropy type	Unary · Arithmetic · Golomb · Huffman (Adaptive · Canonical · Modified) · Range · Shannon · Shannon–Fano · Shannon–Fano–Elias · Tunstall · Universal (Exp-Golomb · Fibonacci · Gamma · Levenshtein)
	Dictionary type	Byte pair encoding · DEFLATE · Lempel–Ziv (LZ77 / LZ78 (LZ1 / LZ2) · LZJB · LZMA · LZO · LZRW · LZS · LZSS · LZW · LZWL · LZX · LZ4 · Statistical)
	Other types	BWT · CTW · Delta · DMC · MTF · PAQ · PPM · RLE
Audio	Concepts	Bit rate (average (ABR) · constant (CBR) · variable (VBR)) · Companding · Convolution · Dynamic range · Latency · Nyquist–Shannon theorem · Sampling · Sound quality · Speech coding · Sub-band coding
	Codec parts	A-law · μ-law · ACELP · ADPCM · CELP · DPCM · Fourier transform · **LPC** (LAR · LSP) · MDCT · Psychoacoustic model · WLPC
Image	Concepts	Chroma subsampling · Coding tree unit · Color space · Compression artifact · Image resolution · Macroblock · Pixel · PSNR · Quantization · Standard test image
	Methods	Chain code · DCT · EZW · Fractal · KLT · LP · RLE · SPIHT · Wavelet
Video	Concepts	Bit rate (average (ABR) · constant (CBR) · variable (VBR)) · Display resolution · Frame · Frame rate · Frame types · Interlace · Video characteristics · Video quality
	Codec parts	Lapped transform · DCT · Deblocking filter · Motion compensation
Theory	Entropy · Kolmogorov complexity · Lossy · Quantization · Rate–distortion · Redundancy · Timeline of information theory	

Compression formats · Compression software (codecs)

Categories: Audio codecs | Lossy compression algorithms | Speech codecs | Digital signal processing