

L P C

linear predictive coding

# Was ist LPC?

Prinzip:

- Ein Klang (vorzugsweise Sprache) kann durch LPC analysiert und resynthetisiert werden
- Anhand der Kombination der Analysedaten mehrerer Inputsamples wird ein Weiteres ausgehend von einer Linearität vorausgesagt

Klangbeispiele:

- Paul Lansky „Idle Chatter“ (1985) 
- Texas Instruments, Speak&Spell (1980) 

# Geschichte und Entwicklung

- 1966: erste Ideen einer automatischen Phonemanalyse mit Hilfe von Wahrscheinlichkeitsrechnung zur Sprachkodierung von S. Saito und F. Itakura
- 1969: M. G. Culler arbeitet an Echtzeit-Sprachkodierung
- 1969: B. S. Atal präsentiert ersten LPC-Coder auf dem Treffen der Acoustical Society of America
- 1971: Philco-Ford stellt erstes 16-bit-basiertes Echtzeit-LPC vor
- 1974: es entsteht ein erstes Zwei-Wege-LPC-System über das Arpanet (3500bit/s) zwischen Culler-Harrison und den Lincoln Laboratories
- 1980: LPC findet Verwendung in Texas Instruments' „Speak&Spell“, ein Spielzeug zur Spracherziehung

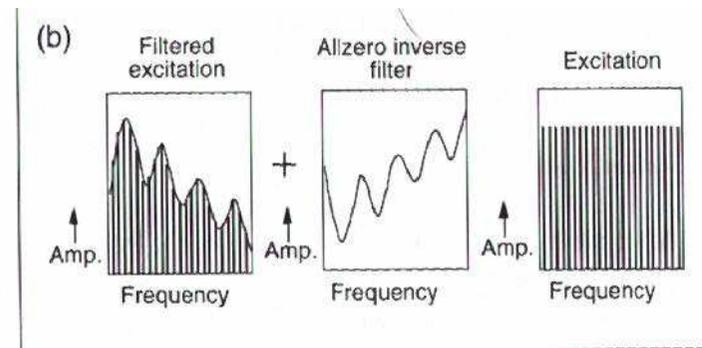
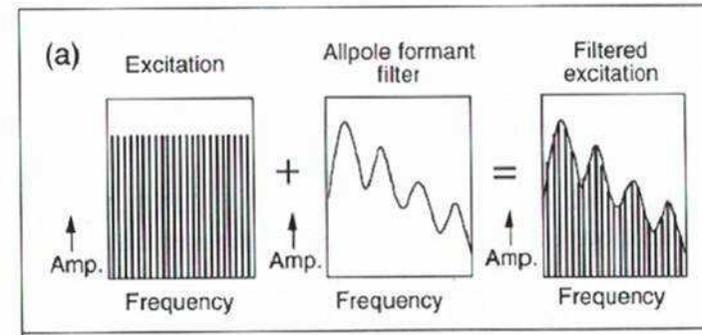
# Praktische Anwendungen

- Digitale Sprachübertragung
- Sprachkompression
- Spielzeug/Unterhaltung/“Spracherziehung“
- Spracherkennungssysteme
- Musikalische Anwendung

# Das Prinzip im Detail

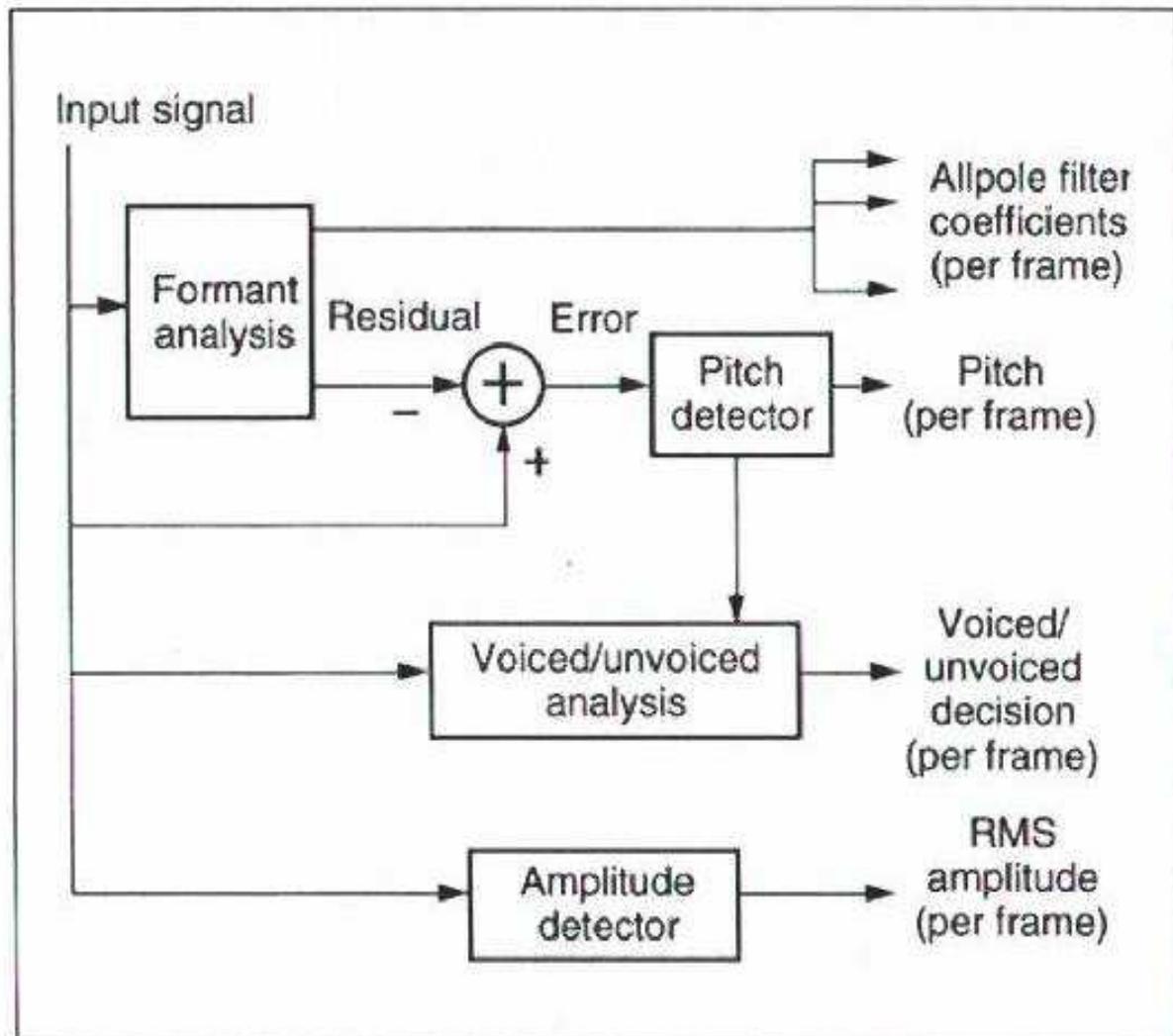
## 1. Analyse

- Inputsignale werden betrachtet als eine Summe aus ihrer Quellschwingung (excitation) und eines spezifischen Resonanzfilters (siehe (a))
- das Spektrum eines Inputsignals wird analysiert (Formant-, Fourier-, autoregressive Analyse)
- ein zum Spektrum inverser allzero filter reduziert das Inputsignal auf das residual, was nach Möglichkeit Quellschwingung sein sollte, tatsächlich aber die Summe aus derer plus durch Ungenauigkeiten hervorgerufene Störsignale ist (siehe (b))



## ... zur Analyse

- anhand des residuals, kann der Grundton bestimmt werden (Cepstrumanalyse)
- die Amplitude wird anhand des ungefilterten Inputs analysiert
- Entscheidung, ob Input ton- oder geräuschhaft (voiced/unvoiced) ist
  - Die Deutlichkeit einer Tonhöhe wird als „pitch estimation error“ in Form eines Wertes in der Regel zwischen 0 und 1 ausgegeben. Zu definierende Schwellwerte entscheiden über die Anteile von Geräusch und Ton.
  - die Amplitude des residuals deutet an, ob voiced oder unvoiced. Ein residual mit niedriger Amplitude lässt ein voiced Signal vermuten. Auch die Cepstrumanalyse gibt Aufschlüsse bzgl. voiced/unvoiced.



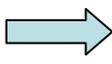
Phoneme	Frame	RMS2	RMS1	ERR	PITCH	DUR
S	197	813.27	1618.21	0.252	937.50	0.010
	198	1189.36	2090.14	0.323	937.50	0.010
	199	553.71	838.38	0.436	937.50	0.010
	200	742.59	1183.17	0.393	937.50	0.010
	201	1041.95	1918.33	0.295	123.95	0.010
	202	1449.16	2677.06	0.293	123.95	0.010
	203	1454.84	2920.50	0.248	937.50	0.010
	204	1430.03	2496.88	0.348	937.50	0.010
	205	1570.88	2981.21	0.277	142.84	0.010
	206	1443.27	2665.22	0.293	142.84	0.010
	207	1172.67	2150.50	0.297	150.00	0.010
	208	1200.73	2080.20	0.333	150.00	0.010
	209	1095.51	2055.25	0.284	116.26	0.010
	210	1260.36	2408.14	0.273	116.26	0.010
	211	1105.17	2293.05	0.232	937.50	0.010
212	809.10	1659.80	0.237	937.50	0.010	
213	428.20	784.93	0.297	250.00	0.010	
I	214	419.45	3886.15	0.011	250.00	0.010
	215	925.86	6366.20	0.021	208.32	0.010
	216	746.28	8046.81	0.008	208.32	0.010
	217	829.82	8277.42	0.010	192.29	0.010
	218	754.64	8049.50	0.008	192.29	0.010
	219	771.84	8001.70	0.009	197.35	0.010
	220	726.81	7955.17	0.008	202.69	0.010
	221	807.63	7835.20	0.010	202.69	0.010
	222	874.27	7732.59	0.012	205.42	0.010
	223	776.87	7491.86	0.010	205.42	0.010
	224	684.64	7317.04	0.008	205.42	0.010
	225	560.87	6297.36	0.007	102.03	0.010
	226	175.63	1842.81	0.009	102.03	0.010
	227	46.53	1329.09	0.001	197.85	0.010
T	228	38.25	793.00	0.002	197.85	0.010
	229	39.26	316.92	0.032	202.69	0.010

**Figure 5.38** A sequence of LPC frames as they might be displayed for editing purposes, after Dodge (1985). The Phoneme column is added for clarity in this figure. The RMS2 column indicates the residual amplitude, RMS1 is the original signal amplitude. ERR is an approximation to the ratio between the two and indicates an unvoiced signal if the ratio is high. PITCH is the estimated pitch in Hz, DUR is the frame duration in seconds.

## 2. Resynthese

- Amplitude und Tonhöhe werden übertragen auf eine beliebige neue Quellschwingung (jeweils eine für ton- und geräuschhafte)
- charakteristisches Spektrum entsteht durch Invertierung des allzero-Filters (→ allpole filter)

# Analyse-Frames

- Analyse wird durch frames gerastert (30-200 frames/s)  Datenkompression
- Pro frame werden Analyseparameter als Durchschnitt ermittelt
- Ab- und Reihenfolge der frames können verändert werden

# LPC mathematisch

- Ein  $n$ tes Inputsample kann aus der ausgewerteten Summe der  $p$  vorhergehenden Samples vorausgesagt werden
- $a_k$  sind die Filterkoeffizienten der vorangegangenen Samples
- Ein niedriger  $p$ -Wert kann ein zu glattes Spektrum bewirken, ein hoher kann unechte Formanten hervorrufen

$$\hat{s} = \sum_{k=1}^p a_k s[n - k]$$

- Das sog. Error-Signal  $e(n)$  im residual, lässt sich wie folgt ausdrücken:

$$e[n] = s[n] - \hat{s}[n] = s[n] - \sum_{k=1}^p a_k s[n-k]$$

- Über eine Z-Transformation erhalten wir:

$$E(z) = S(z) - \sum_{k=1}^p a_k S(z) z^{-k} = S(z) \left[ 1 - \sum_{k=1}^p a_k z^{-k} \right] = S(z) A(z)$$

- Wir sehen, dass das Error-Signal als ein Produkt aus Inputsignal  $S(z)$  und Transferfunktion  $A(z)$  ist. Die  $a_k$ -Koeffizienten entsprechen den zeros.
- Der resynthetisierte Output ist folglich der Quotient aus Error-Signal und Transferfunktion:

$$S(z) = \frac{E(z)}{A(z)}$$

# Fazit

- + Einfache Handhabung
- + Gute Resynthese von traditionellen Instrumente und Sprache
- + Vielseitigkeit durch Crosssynthese
- +/- gezielte Verfremdung / Replikat leicht als solches zu identifizieren
- Komplexe Filtermodelle können nicht simuliert werden

# Csound-Implementation

- **lpanal**
  - -s Samplingrate
  - -c channels
  - -b begin
  - -d duration
  - -p poles
  - -h hopsize
  - -P tiefste Frequenz
  - -Q höchste Frequenz
  
- **lpread**  
krmsr, krmso, kerr, kcps **lpread** ktmpnt, ifilecod
  
- **lpreson**  
aout **lpreson** asig

# Quellen:

- **Klapper, Magdalena**, Working with Csound's ADSYN, LPREAD, and LPRESOpcodes, in: The Csound Book, hg. von Richard Boulanger, Massachusetts 2000
- **Roads, Curtis**, the computer music tutorial, hg. von Curtis Roads, Massachusetts 1996
- [http://en.wikipedia.org/wiki/Linear\\_predictive\\_coding](http://en.wikipedia.org/wiki/Linear_predictive_coding), 07.07.2006, 18.20 Uhr
- <http://cnx.org/content/m12473/latest/>, 07.07.2006, 18.20 Uhr