

回顧：「音声分析合成の研究」

ML—PARCOR—LSR—LSP

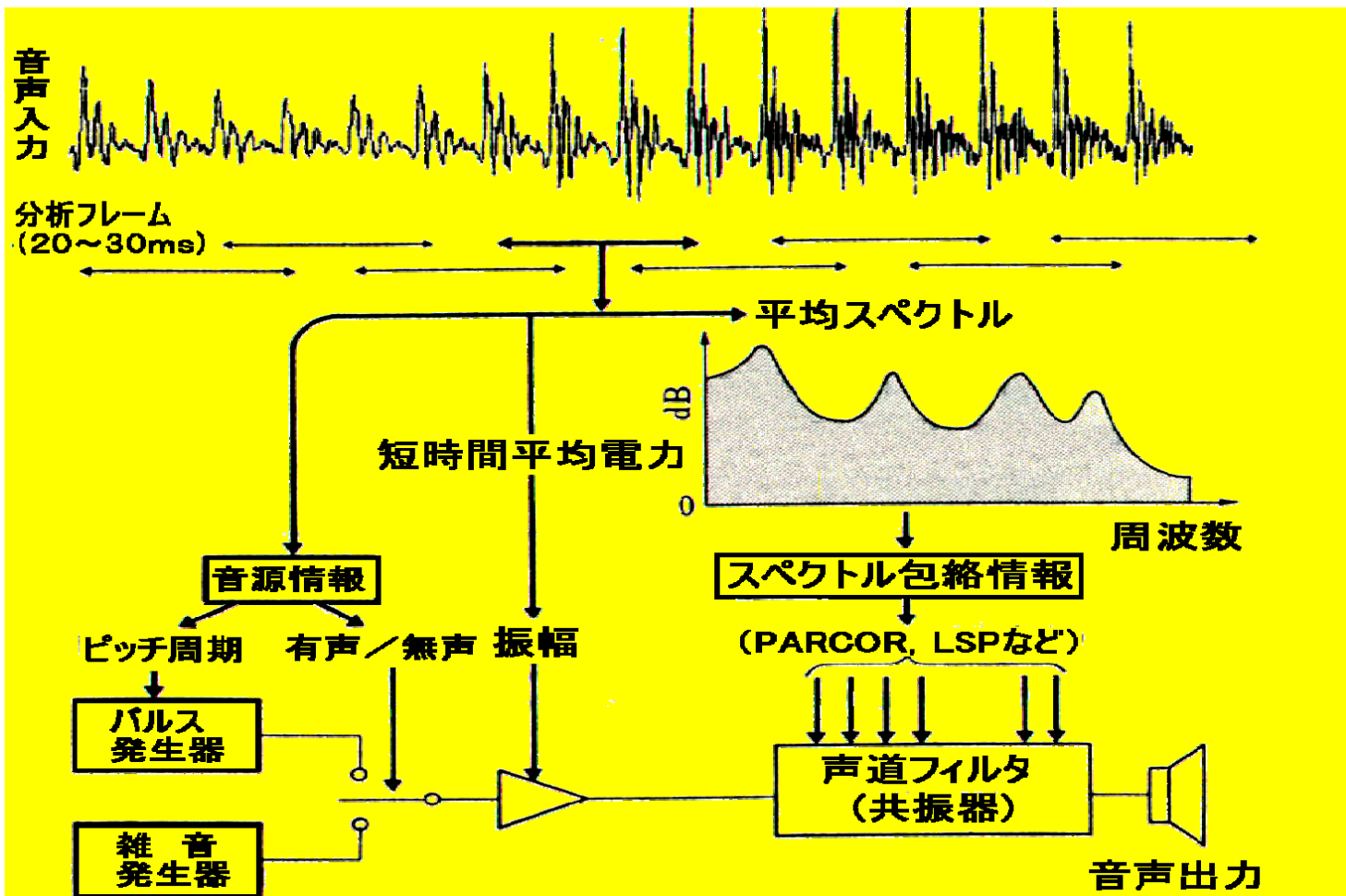
板倉文忠
(名城大学理工学部)

1. はじめに

音声分析合成系とは、その歴史

- 音声信号→音声のパラメタを分析抽出
低ビットレートで符号化伝送
再びもとの音声を合成する技術
- 音声帯域圧縮伝送, 音声合成, 音声認識の
前処理など、音声処理の基本技術になっている.
- PARCOR, LSP方式の誕生の経緯
出発点 最尤スペクトル推定法による
音声分析合成方式(最尤方式と略す)

低ビットレート音声符号化の原理



音声生成の理論 - 古典的・決定論的モデル -

Chiba, T. and Kajiyama, M. : **The Vowel - Its Nature and Structure**

(1941)

母音の発声理論とその構造： 物理学的手法

Fant, G. : **Acoustic Theory of Speech Production**

(1960)

音声発声の音響学的理論： 電気回路理論・通信理論的手法

初期の音声分析合成系

<http://www.cs.indiana.edu/rhythmsp/ASA/highlights.html>

VODER , BTL	the FIRST EVER electronic speech synthesis, demonstrated at the New York World's Fair	1939	
The Pattern Playback	designed by Franklin Cooper,	1951.	
OVE II G.Fant/J.Holmes	Comparison of synthetic and a natural sentences	1961	
'PAT' formant synthesizer	Copying a natural sentence Walter Lawrence	1962	
Parrarel Formant Syn. Dennis Klatt	Comparison of synthesis and a natural sentence	1986b	

合成音による歌

(BTL, Gerstman and Mathews, 1961.)

<http://emosamples.syntheticspeech.de/emoSamples/dazy.mp3>



合成音による歌とピアノ伴奏,
"Bicycle Built for Two" This
song was reprised by the
'decorticated' Hal in **'2001: A
Space Odyssey'**.

Bicycle Built For Two

*Daisy, Daisy,
Give me your answer do!
I'm half crazy,
All for the love of you!
It won't be a stylish marriage,
I can't afford a carriage
But you'll look sweet upon the seat
Of a bicycle made for two.*

2. 最尤スペクトル推定法

- 1965年3月: 名大大学院修士課程を修了

統計的パターン認識や時系列解析の勉強

- 同年4月に博士課程に進学、指導教授の宇田川先生が急逝

専門を音声・音響分野に変更

- 音声認識を統計的パターン認識の立場から追及

- 福村晃夫先生の指導(特別輪講)

Fant, G.: "Acoustic Theory of Speech Production"

- 1966年4月に, NTTの通研実習生として音声の研究開始

通研に基礎研究部が新たに設置

斎藤収三第四研究室長のもとに, 音声聴覚・画像・映像に関する新進気鋭の研究者が集まり, 活気に満ちていた.

<はじめの挑戦と失敗>

- 初仕事: 自分の声をサウンドスペクトログラムで分析
- きれいなフォルマント構造を示すスペクトルパターンが表れない!?

その時に音声は想像以上に複雑な信号で多くのランダムさを含んでいるに違いないと考えた. →

- 確率過程の時系列解析の手法による音声分析
- 時間的に連続な p 次の全極型定常確率過程としてモデル化
- そのパラメタの違いを最適判別する問題を考察
- この問題は母集団確率過程の十分統計量を求める事に帰着
- 連続波形モデル: 十分統計量が波形の $1 \sim p (=10)$ 階までの微係数の電力で与えられる

実際に 電子回路を製作して母音のスペクトルを分類する実験を始めた. 当然の事ながら, 波形の高階(10階!)の微係数は, わずかな雑音によっても大きく乱され, 見事に失敗に終わった.

< 転回 >

通研4研に音声研究用計算機(FACOM 270-20)が導入

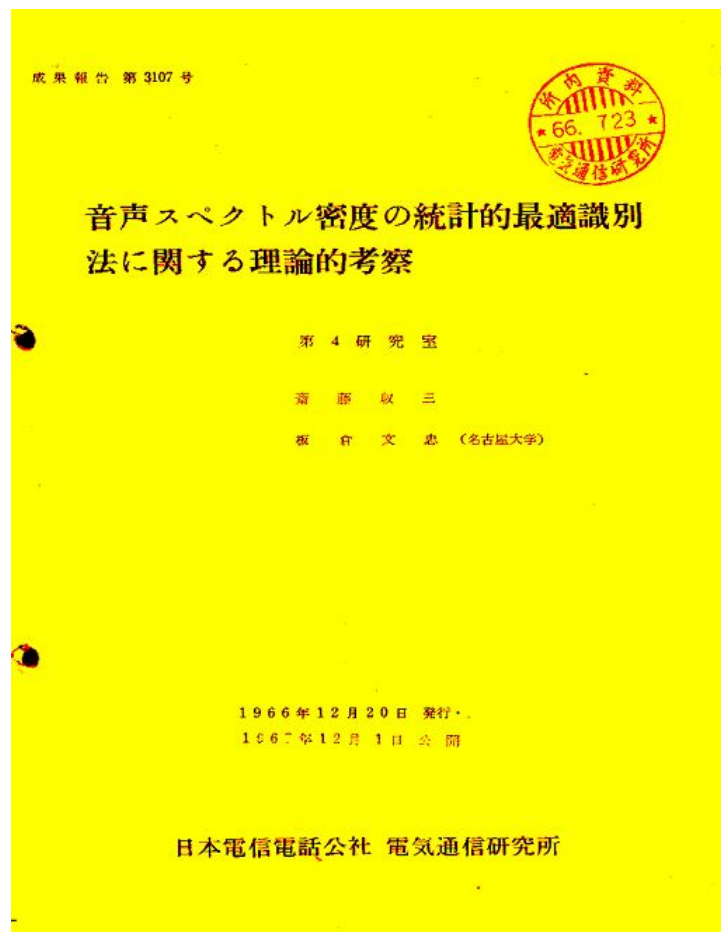
→ デジタル信号処理を意識

■ 時間的に離散的な p 次の全極型定常時系列モデル

モデルの最適判別に必要な十分統計量: 初めの p 個の自己相関係数値である. この研究の副産物として, 全極型モデルの最尤推定法と音声スペクトルの合成法が発見され, ML音声分析合成方式という概念が生まれた.(図2) これは, 後に線形予測符号化(LPC)と呼ばれる事になったものと数学的に等価である.

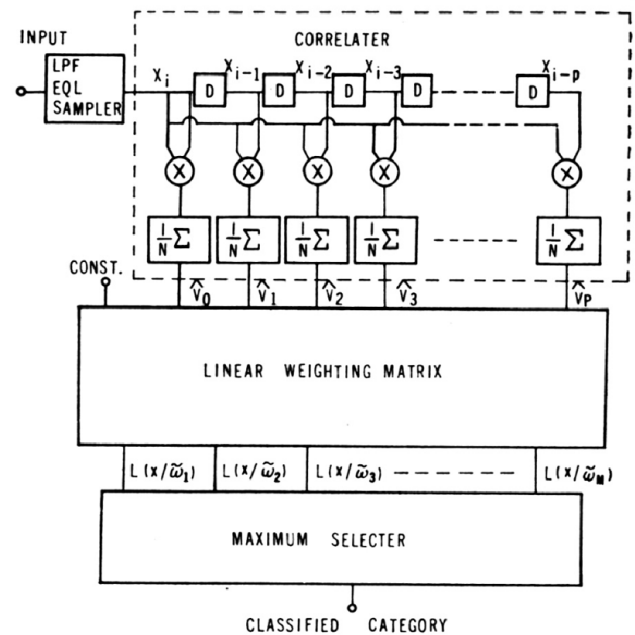
この理論によれば, 音声合成に必要なパラメタが, 入力波形の自己相関を係数とする p 元連立方程式(正規方程式)の求解により推定できる. またフォルマント周波数の推定は, こうして求めたパラメタを係数とする p 次代数方程式の求根に帰着される.

理論のまとめ



内容

$$L(\mathbf{x}) \doteq -\frac{N}{2} \left[\log 2\pi \sigma_e^2 + \frac{1}{\sigma_e^2} (A_0 \hat{v}_0 + A_1 \hat{v}_1 + \dots + A_p \hat{v}_p) \right] \quad (3)$$



D: unit delay with delay time ΔT .

(X): multiplier

M: number of categories (phonemes)

ω_i : parameter of spectral density of the i -th category

Fig. 1 Statistically optimum classifier of speech spectra, when all parameters are completely known.

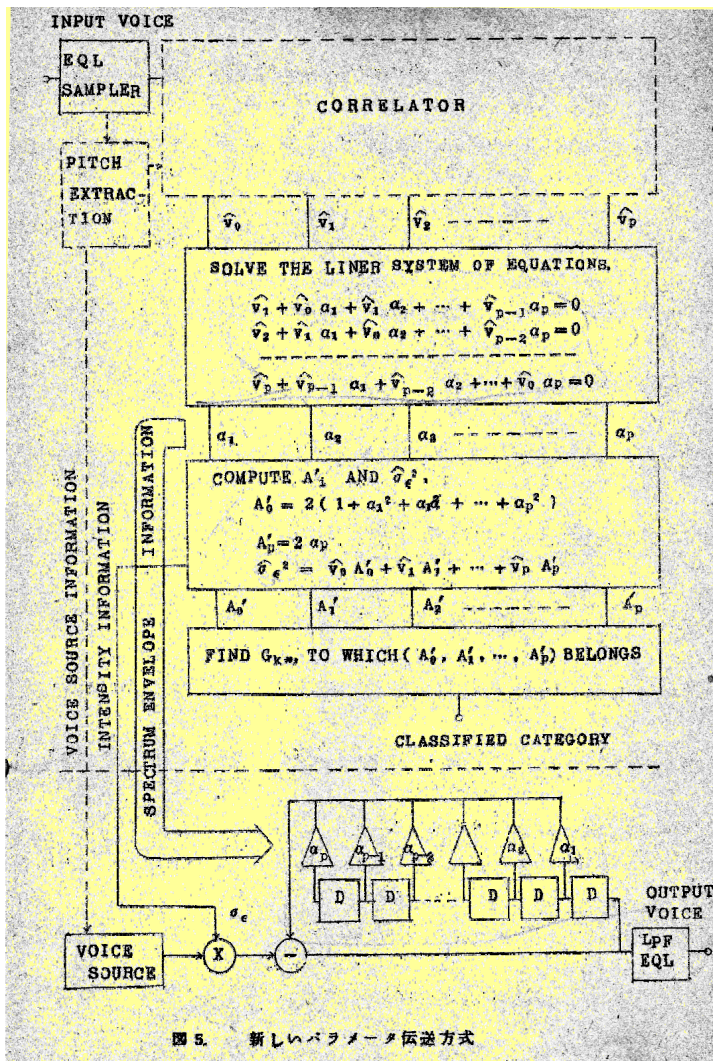


図 5. 新しいパラメータ伝送方式

3. 音声分析合成方式に挑戦

斎藤収三室長に相談:「ピッチ問題」を指摘される。

→最尤法:スペクトル包絡を逆フィルタの形で抽出

基本周期情報が抽出 シミュレーション

ケプストラム法に比べて遜色のない方法

(これを変形相関法と名付けた。)

最尤音声分析合成方式 + 変形相関法 = ML方式 完成 1967年夏

分析合成音は予想以上に自然・明瞭性

■ 1968年8月 第6回国際音響学会議(6th ICA)で発表

同じセッションで,音声の適応線形予測符号化(APC)の発表

(Atal, Schroeder)

→線形予測音声符号化の先駆的な論文

図3 MLスペクトル推定の例 (母音/a/,子音/s/の波形とFFTスペクトルとMLスペクル)

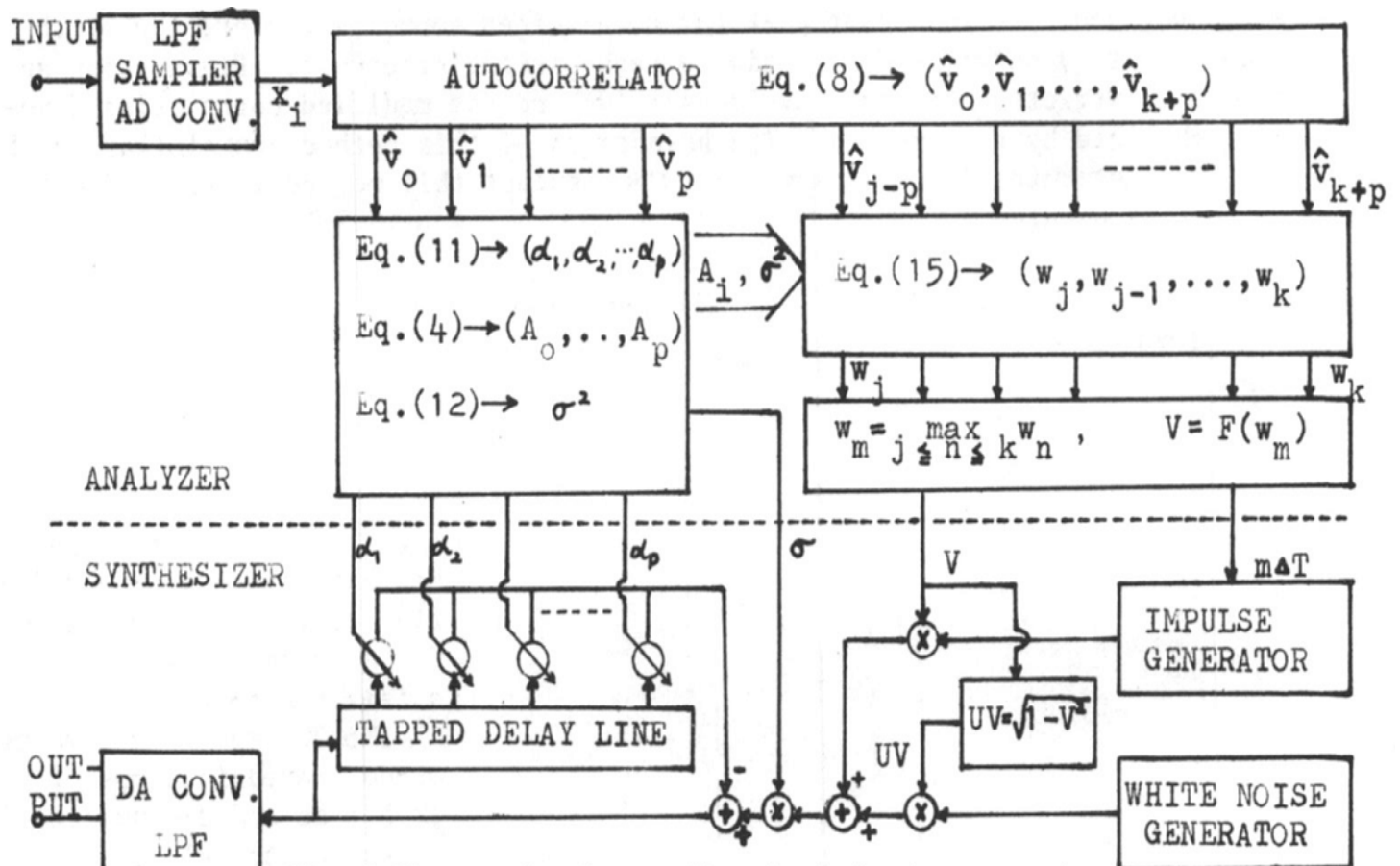
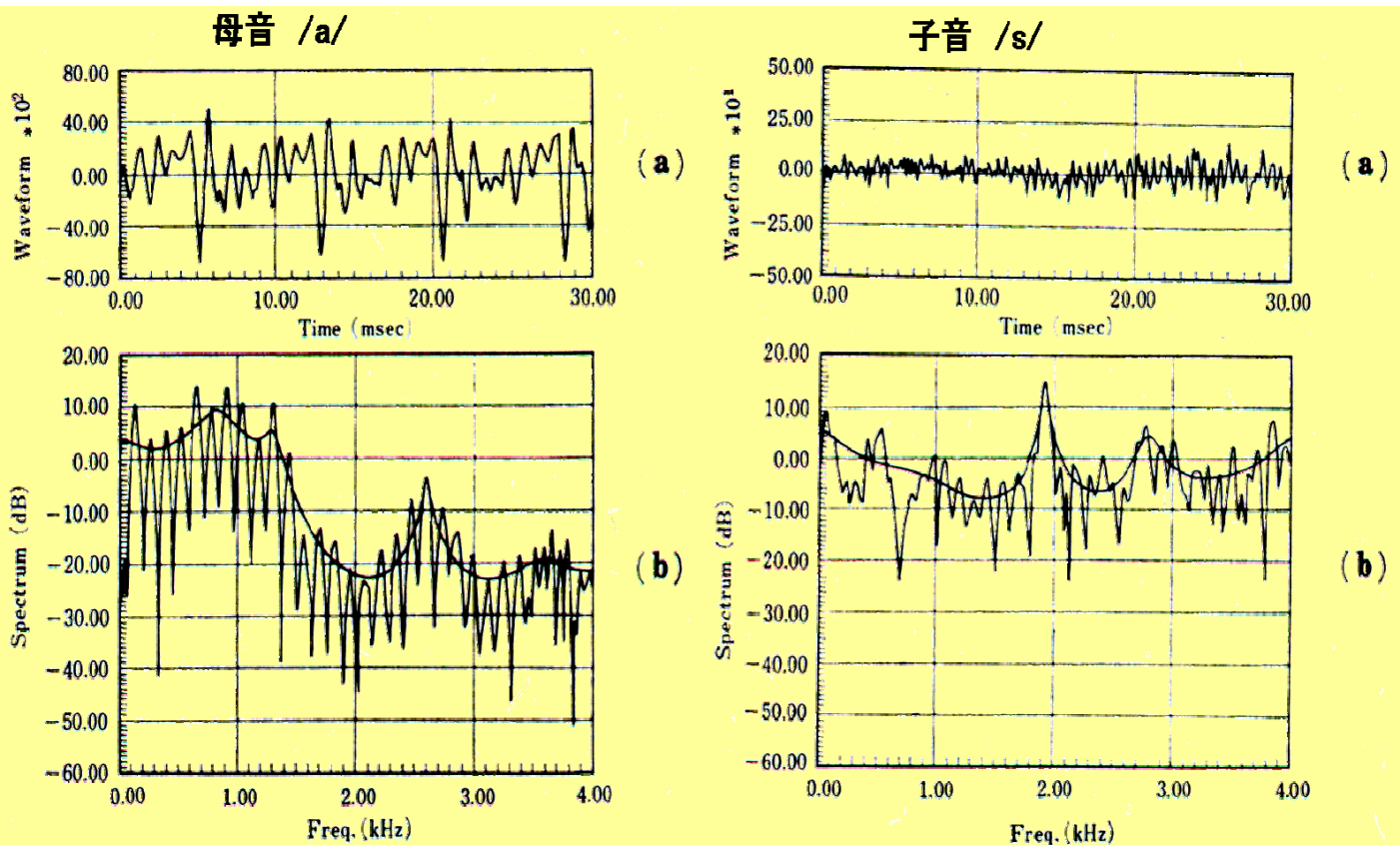


Fig.2 The analysis synthesis telephony system based on the maximum likelihood method

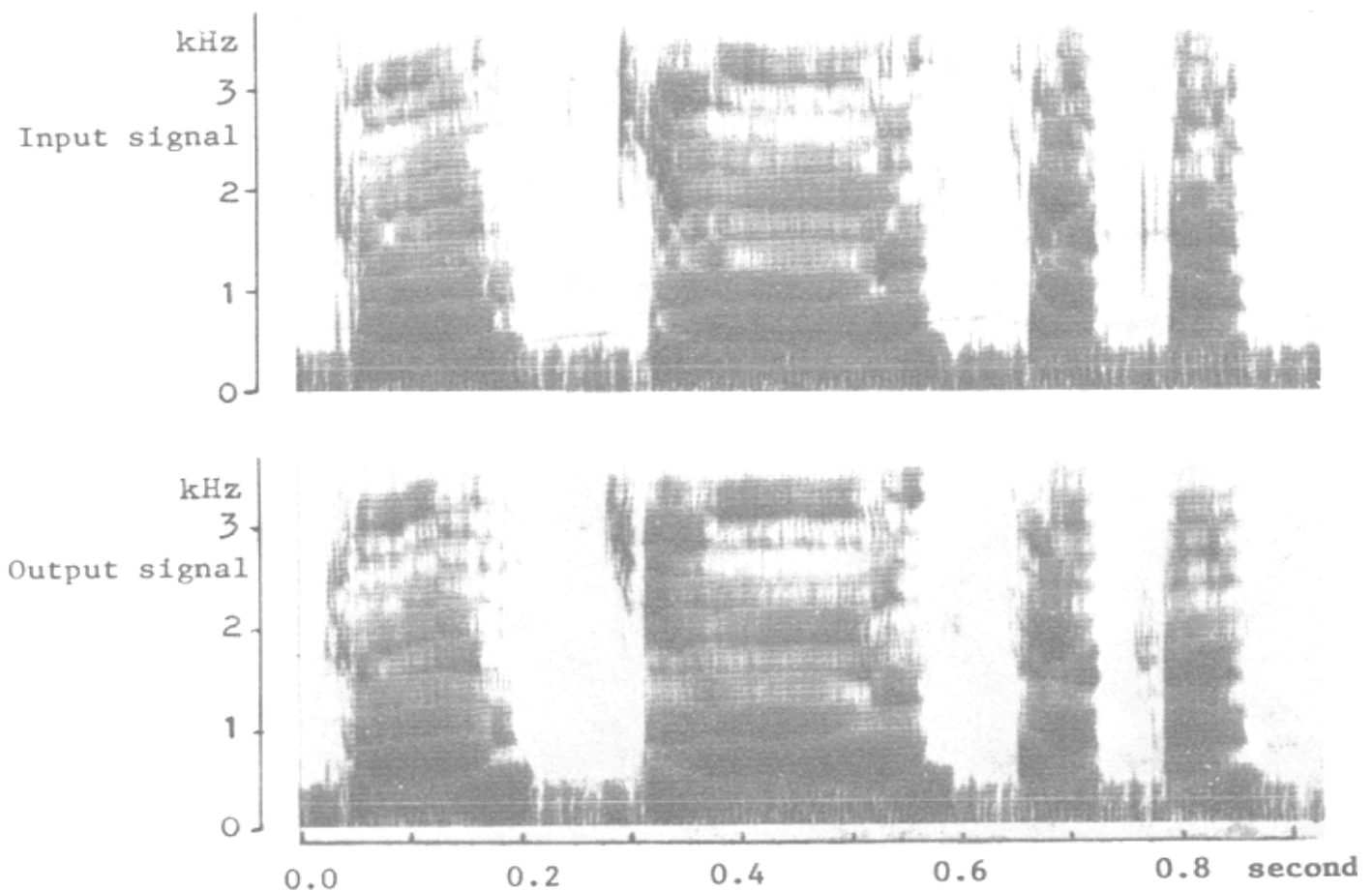


Fig.3 Sound spectrograms of input and output signals of the analysis-synthesis system. $\Delta T=0.125$ msec, $p=10$, $\theta=0.2$ and $K=20.0$

6th ICA, Tokyo, 1968 by Itakura and Saito (2)

音声信号の統計的解析

• 音声信号の自己回帰モデル

時間領域:

$$\sigma \cdot \varepsilon_t = x_t + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \cdots + \alpha_p x_{t-p}$$

周波数領域:

$$H(z) = \frac{\sigma}{1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \cdots + \alpha_p z^{-p}}$$

外部駆動信号

線形システム

音声信号

決定論的

確率論的

フーリエ解析

最大尤度基準による推定

自己回帰モデルの尤度関数

自己回帰モデルの尤度関数:

$\mathbf{X} = \{x_1, x_2, \dots, x_n\}$ の対数尤度関数は

$$L(\mathbf{X}, \Theta) = -\frac{N}{2} \left[\log(2\pi\sigma^2) + \frac{1}{\sigma^2} \sum_{i,j=0}^m \alpha_i \tilde{\phi}_{i-j} \alpha_j \right] \quad \text{for } N \gg 1$$

ただし $\Theta = \{\alpha_1, \alpha_2, \dots, \alpha_m, \sigma\}$ はパラメータ、

$\tilde{\phi}_{i-j}$ は短時間自己相関

$$\tilde{\phi}_{i-j} = \frac{1}{N} \sum_{n=1}^{N-|i-j|} x_n x_{n+|i-j|} \quad (|i-j| = 0, 1, \dots, m).$$

自己回帰モデルの最尤推定 (正規方程式, 線形予測分析)

入力信号 $\mathbf{X}$

自己相関の計算:

$$\tilde{\phi}_{i-j} = \frac{1}{N} \sum_{n=1}^{N-|i-j|} x_n x_{n+|i-j|} \quad (|i-j| = 0, 1, \dots, m)$$

自己回帰係数 (LPC) の計算:

$$\begin{bmatrix} \tilde{\phi}_0 & \tilde{\phi}_1 & \cdots & \tilde{\phi}_{m-1} \\ \tilde{\phi}_1 & \tilde{\phi}_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \tilde{\phi}_1 \\ \tilde{\phi}_{m-1} & \cdots & \tilde{\phi}_1 & \tilde{\phi}_0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = - \begin{bmatrix} \tilde{\phi}_1 \\ \tilde{\phi}_2 \\ \vdots \\ \tilde{\phi}_m \end{bmatrix}$$

$$\text{強度 (残差パワー): } \sigma^2 = \sum_{i=0}^m \alpha_i \tilde{\phi}_i$$

最尤推定パラメータ $\Theta = \{\alpha_1, \alpha_2, \dots, \alpha_m, \sigma\}$

周波数領域における解釈 合成による分析(A-b-S)との比較

尤度関数 $L(\mathbf{X}, \Theta)$ の最大化

板倉-斉藤距離の最小化

$$D_{IS}(S(\omega) : S_N(\omega)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [\exp(-d(\omega)) + d(\omega) - 1] d\omega$$

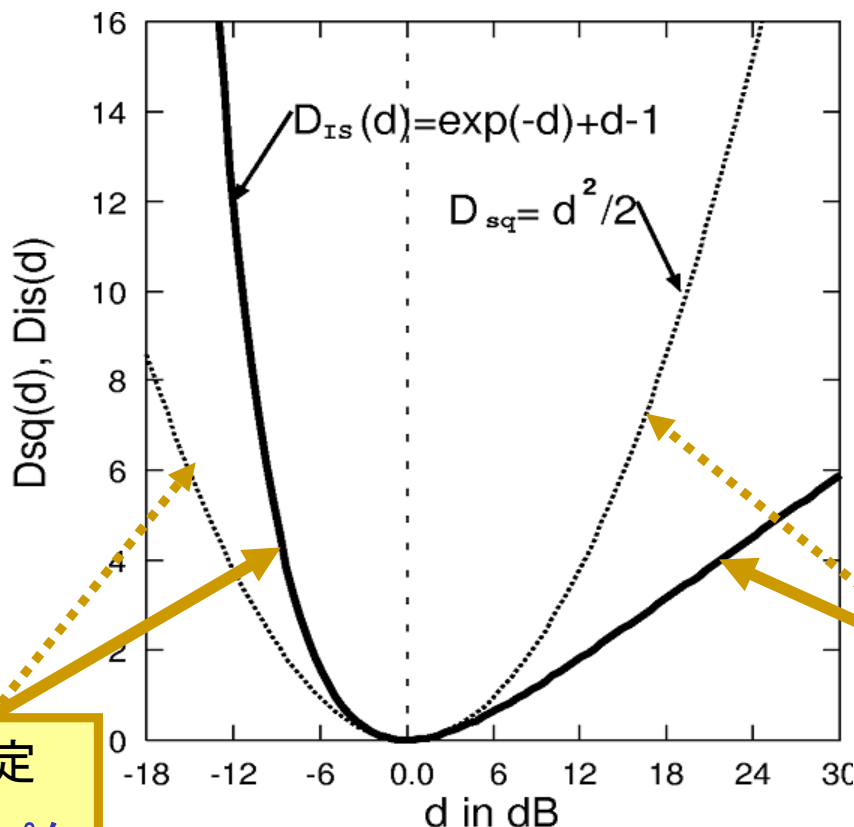
ただし $d(\omega) = \log[S(\omega)/S_N(\omega)]$,

$S(\omega)$ は AR スペクトル

$S_N(\omega)$ は入力信号 $\mathbf{X}$ の DFT

$$S_N(\omega) = \frac{1}{N} \left| \sum_{n=1}^N x_n \exp(-i\omega n) \right|^2$$

スペクトル距離尺度の対称性



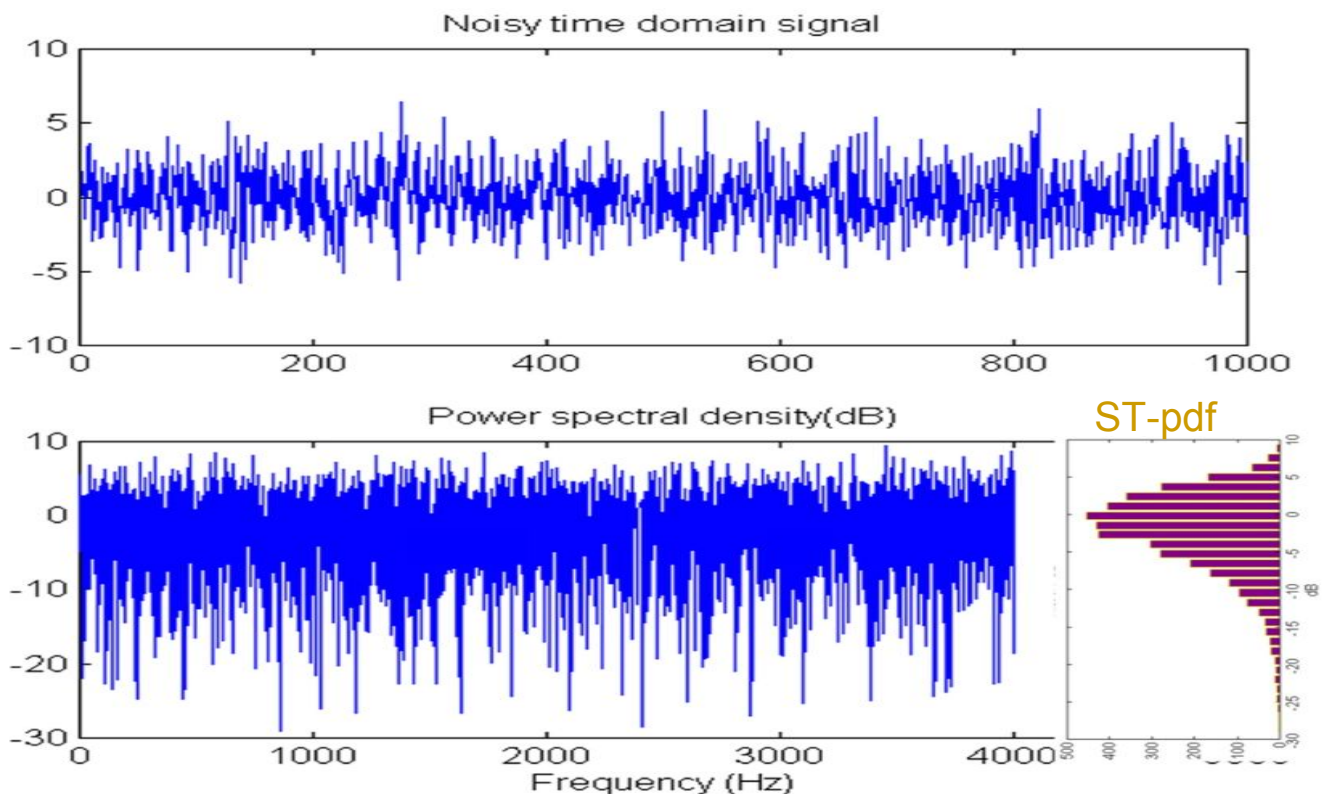
過小推定

存在するスペクトルの無視

過大推定

スペクトルの穴の埋め合わせ

短時間パワースペクトル(DFT) の統計的性質



4. ML方式からPARCOR方式への発展

- ML方式を使って, 低ビットレート符号化の詳細な実験
 - α パラメタ(LPC係数) $\rightarrow$ 10ビット以下に量子化すると, 時々著しい異音が発生する問題に遭遇
 - 原因: パラメタ量子化により合成用帰還形(IIR)フィルタが発振する
 - 1969年3月, Szegő著 *Orthogonal Polynomials*(AMS 1956)に遭遇.
 - 音声合成フィルタ伝達関数の分母の多項式:
 - 直交多項式により展開出来る事に気が付いた.
 - 直交多項式の漸化式によって, p 元の連立線形方程式がいわゆるLevinsonの再帰法によって解ける事
 - 直交多項式の漸化式 $\rightarrow$ 前向き予測誤差, 後ろ向き予測誤差を格子型フィルタによって構成できる事

PARCORの提案

■ 音声波形の偏自己相関係数: 2つの予測誤差の相関係数

(PARCOR; Partial Correlation)

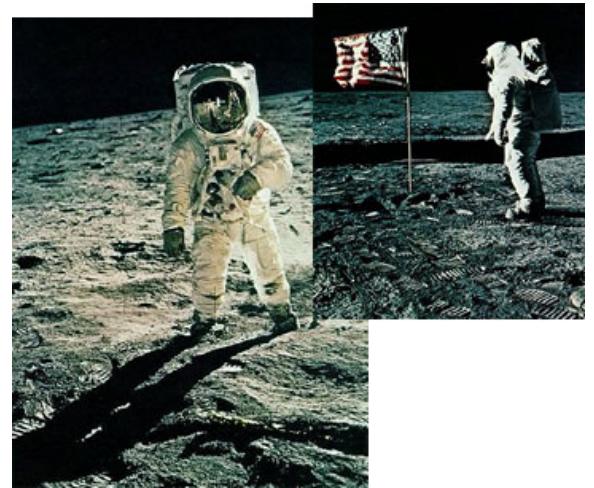
■ PARCOR係数の絶対値が全て1未満であれば, 分析フィルタの逆フィルタである合成用格子型フィルタは安定である事を理論的に示す事ができた.

1969年7月21日に開催された日本音響学会の音声研究会で発表

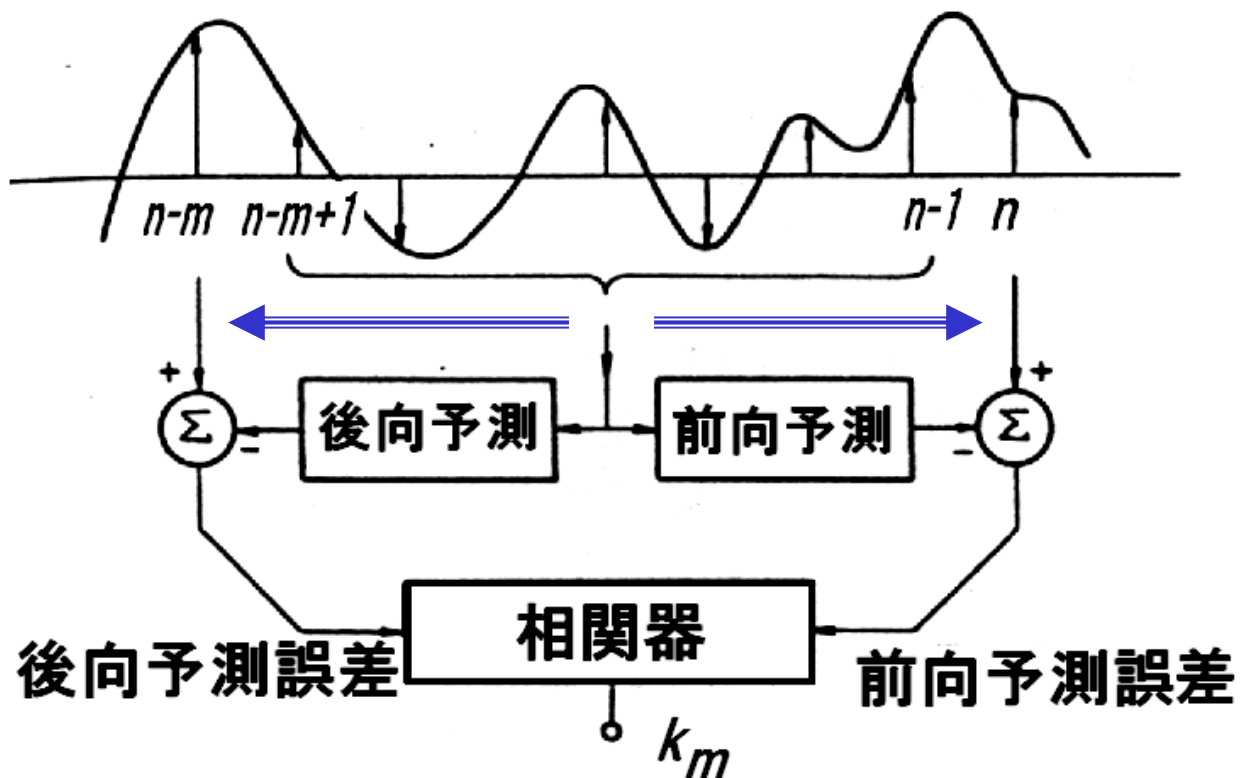
(A4版20ページの詳細な資料)

当日が月ロケット・アポロ11号の乗組員の人類初の月面着陸

研究会参加者はわずか数名、何の反響なし.



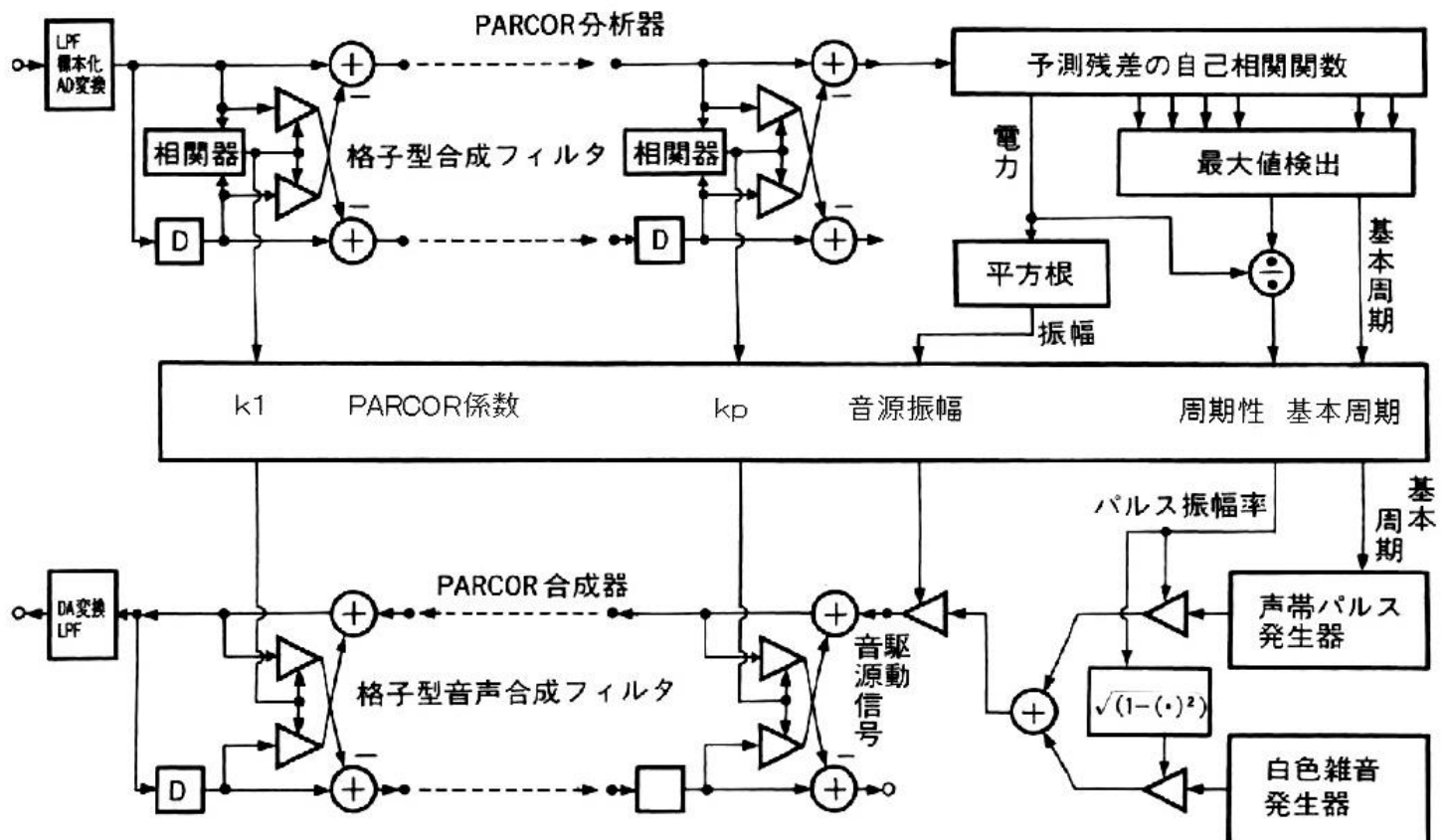
PARCOR係数の定義



偏自己相関関数の定義

格子形デジタルフィルタ

PARCOR音声分析合成系



海外発表

■第7回国際音響学会議

(7th ICA , Budapest 1971年, 出席不可)

格子型音声分析合成フィルタの理論

■音声通信と処理に関する会議

(1972年, Boston, 板倉初海外)

PARCOR係数の最適量子化の論文

国際的にも認知

PARCOR方式が誕生した1969年当時は, LSI技術は未熟であり, そのハードウェアによる実現は簡単ではなかった.

PARCOR音声合成器の試作

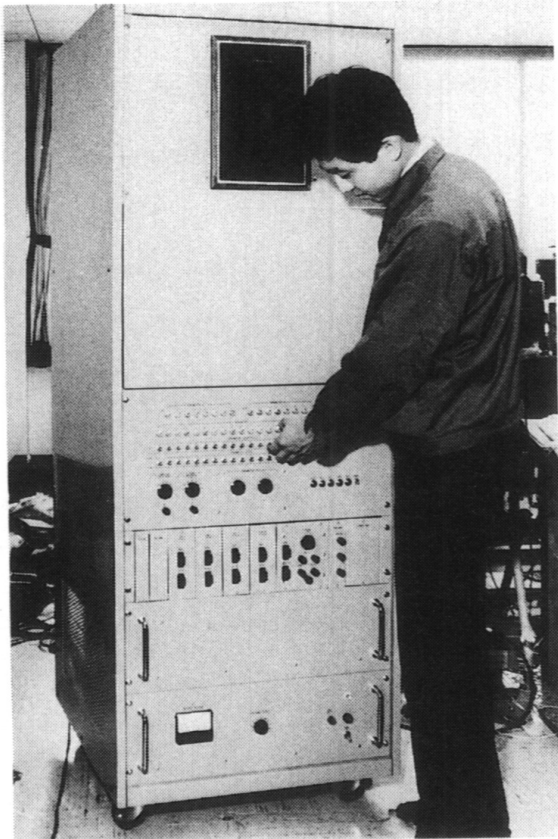


図2 PARCOR 音声合成器
(1970 年試作, NTT 提供)

汎用プリント基板約400枚

サイズ 70x80x180cm

5V・50Aの直流電源

(試作:NEC通信研究所)

この装置の実演に促され, 通研の基礎研究部と宅内研究部の協力により開発され, PARCOR方式は, パラメタ編集方式の音声応答装置として, 1971年に実用化され, その後, 各種の電話データサービスに供された。

PARCOR音声合成器を操作する
伊藤憲三君 (現岩手県立大教授)

7th ICA PARCOR Lattice Filter(1971)

Fig.(1)
Self-controlled time-varying DF to extract the spectral parameters $\{k_i\}$.

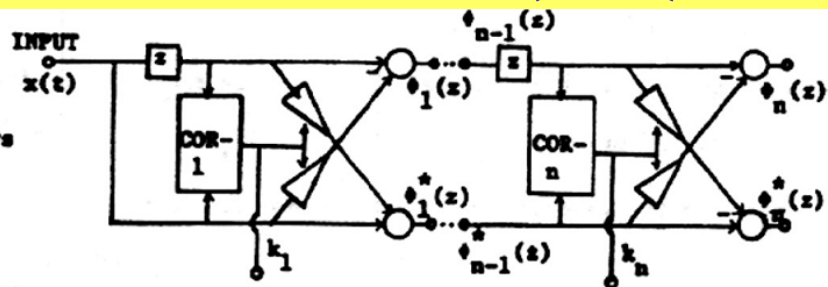


Fig.(2)
Ladder-form DF to synthesize speech spectrum from $\{k_i\}$.

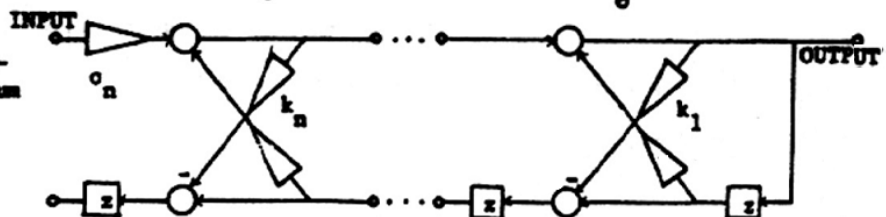


Fig.(3)
DF equivalent to Fig.(2), this configuration is the same as that of Kelly.

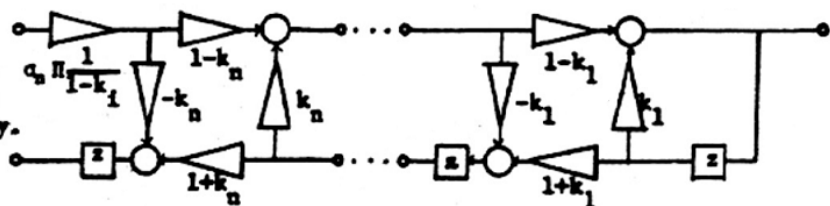
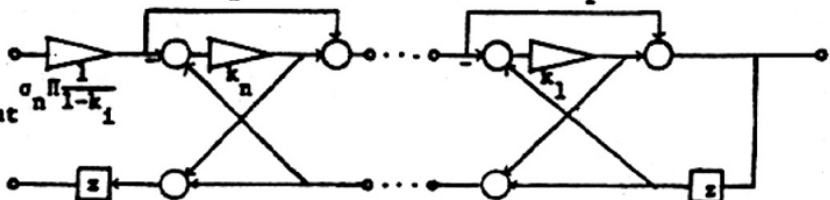


Fig.(4)
This DF is also equivalent to Fig.(2), but requires fewer multiplication.



Speak&Spell

1976年には、米国TI社からSpeak&Spellという商品名の形で、PARCOR合成の機能をもつ音声合成LSIが民生用に作られ、広く実用化されるようになった。



DOD標準方式

また、米国国防省は、1976年LPC10-eと呼ばれるFederal Standard 1015を制定し、軍用秘話通信方式を実用化した。この方式の原理はPARCOR方式そのものであり、音声を2400bpsで符号化する方式である。この方式を実行するために必要な計算速度は約7Mipsであり、比較的少ないが、実用的な騒音環境での音声品質は、MOS値で2.3程度であり、一般の商用通信用の音声符号化方式としては、不十分であったと言わざるを得ない。

その後、1996年になって、音源情報として混合駆動を用いたLPC(MELP)が新しい軍用通信の標準として採択された。この改良によって、計算量はLPC-10eの6倍くらいに増加するが、2400bpsにおけるMOS値は、3.2に向上した。

5. 最尤スペクトル推定法の音声認識への応用

PARCOR方式の基礎的検討が一段落した後、1973年8月から2年間、ベル研究所音響研究室の客員研究員として過ごすことになった。

ベル研は、当時は、世界の電子情報通信のメッカとも呼ばれていた。音声分野においても、Fletcher, Dudleyほか多数の聴覚・音声・音響分野のパイオニアがうまれた事は周知の通りである。

Flanagan氏から与えられたテーマは、"Digital Processing of Speech"ということで、ほとんど何の拘束もないようなテーマであった。ベル研での滞在期間は限られていたので、手っ取り早く最尤スペクトル推定法で求めたスペクトル距離尺度とDTW(Dynamic Time Warping)を統合した単語音声認識システムを試作した。

最小予測残差原理の音声認識への応用

(from IEEE Trans.ASSP-23,No.1, 1975)

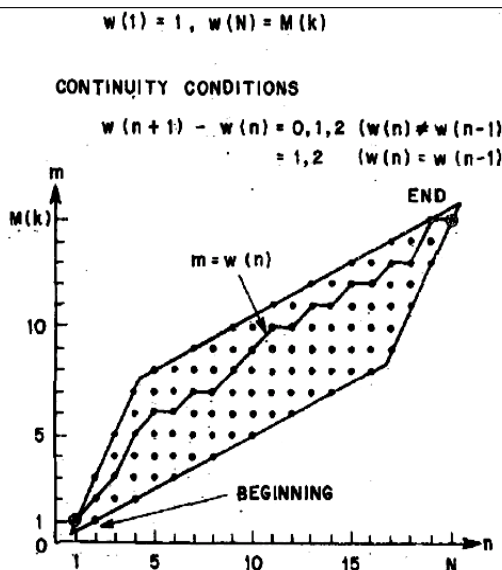
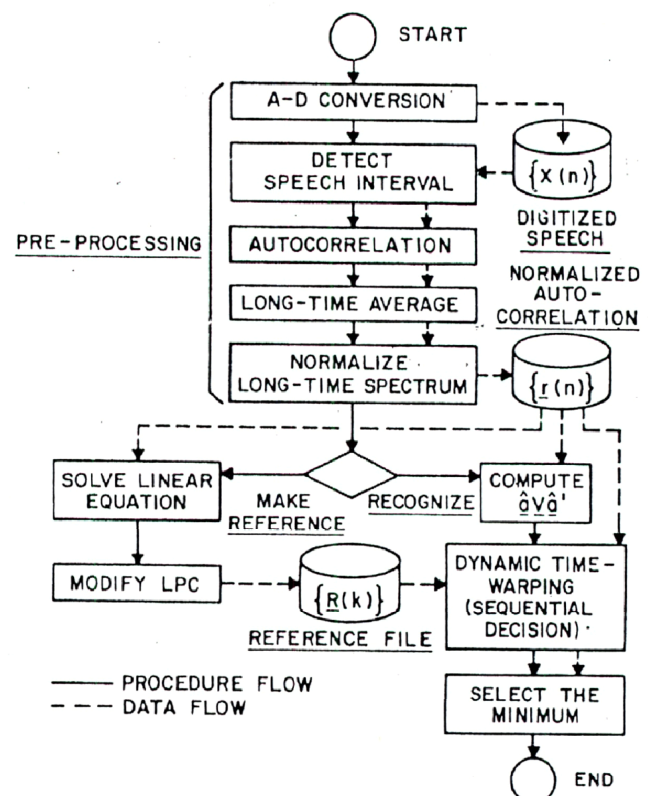


Fig. 1. An example of time-warping function. The parallelogram shows the possible domain of (n, m) coordinates.



3-modeシステム

ベル研での滞在の2年目には、Rosenberg氏と協力して、音声認識、話者認識、音声応答よりなる3-modeシステムを作成し、航空座席予約システムとして実験公開し、好評を博した。これらの研究開発が端緒になり、しばらく停滞していたベル研究所の音声認識の研究が息を吹き返し盛んになった



6. LSP方式の誕生

ベル研滞在中の1974年の暮れ頃、図書室に行って数学書を漁っていたところ、たまたま発行されたばかりの *Gesammelte Abhandlungen von Issai Schur* という本があり **Toeplitz行列のスペクトル分解**に関するものを見つけた。この理論をPARCOR合成フィルタと結びつけると、**全極型のフィルタの線スペクトル表現が可能である**ことがわかった。

声道断面積関数を p 個の等長音響管でモデル化し、**PARCOR係数**は、断面積の不連続点における波動の**反射係数**であると見なされる。唇端は体積流に対して完全開放のため**反射係数 $= -1$** と見なし、声門端は、**開放時には反射係数 $=1$ 、閉鎖時 -1** とモデル化する。音響管の両端の**反射係数を ± 1** にすると、**波動エネルギーが閉じこめられ**、声道共振系は**無損失系**になる。従って、**音響管の伝達関数は線スペクトル状**になる。

PARCORからLSP(1)

唇端からかぞえて $i = 1, 2, \dots, p$ 番目の反射係数を k_i とする. m 次の前向き予測演算子, 後向き予測演算子を $A_m(D), B_m(D)$ (D は単位遅延) とするとき, 次の漸化式が成立する: |

$$\left. \begin{aligned} A_m(D) &= A_{m-1}(D) - k_m B_m(D), \quad A_0(D) = 1 \\ B_m(D) &= D[B_m(D) - k_m A_{m-1}(D)], \quad B_0(D) = D \end{aligned} \right\} (1)$$

声門端反射係数 $k_{p+1} = \pm 1$ の時, 全極型伝達関数の線スペクトル周波数は, 次式の ω に関する方程式の根として与えられる.

$$\left. \begin{aligned} P(z^{-1}) &\equiv A_{p+1}(z^{-1}) \Big|_{k_{p+1}=1} = 0 \\ Q(z^{-1}) &\equiv A_{p+1}(z^{-1}) \Big|_{k_{p+1}=-1} = 0 \end{aligned} \right\} (2)$$

$$\text{ここで, } P(z^{-1}) = 0, \quad Q(z^{-1}) = 0 \quad (3)$$

の根が単位円周上にあることに注意すると,

PARCORからLSP(2)

$P(z^{-1}), Q(z^{-1})$ は, 次のように因数分解できる:

$$\left. \begin{aligned} P(z^{-1}) &= (1 - z^{-1}) \prod_{i=2,4,\dots,p} (1 - 2 \cos \omega_i z^{-1} + z^{-2}) \\ Q(z^{-1}) &= (1 + z^{-1}) \prod_{i=1,3,\dots,p-1} (1 - 2 \cos \omega_i z^{-1} + z^{-2}) \end{aligned} \right\} (4)$$

ここで p は偶数とした. このとき, LSP (線スペクトル対周波数) $\{\omega_i\}$ は, 次のように奇数番目と偶数番目のものが, 必ず交互に並ぶように順序づけることができる. すなわち

$$\{0 < \omega_1 < \omega_2 < \omega_3 < \omega_4 < \dots < \omega_{p-1} < \omega_p\}$$

この条件が満たされるとき, またそのときに限り,

$$1/A_p(z^{-1}) = 2 / [(P(z^{-1}) + Q(z^{-1}))] \quad (5)$$

で表されるフィルターは安定である。

LSPの特徴

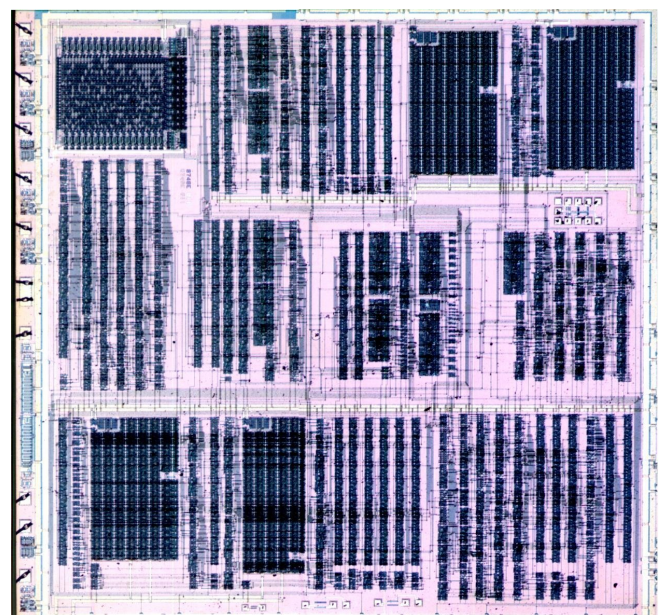
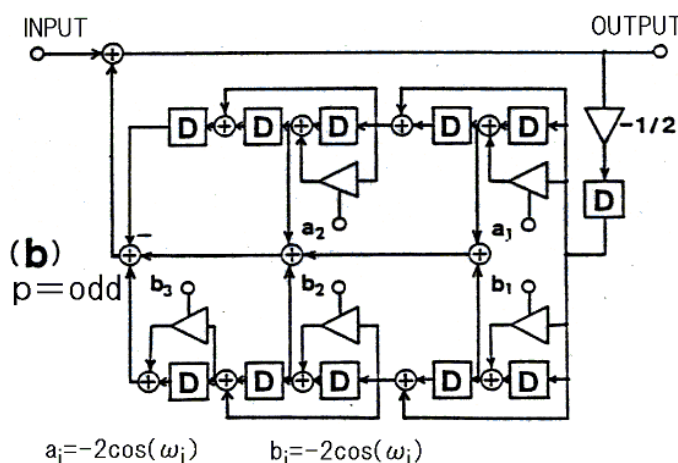
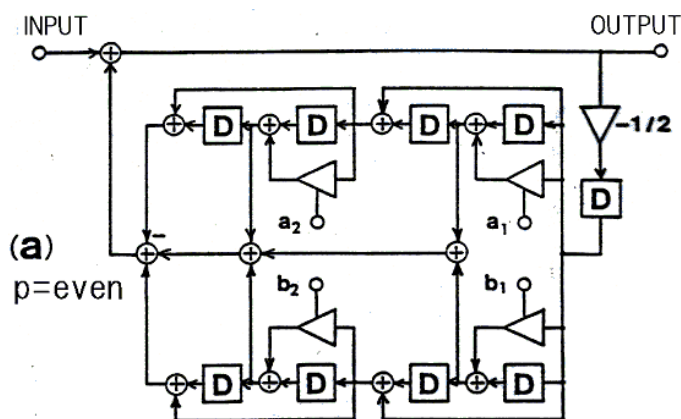
図6は、式(4), (5)により求まるLSP方式の音声合成フィルタのブロック線図である。この原理に基づいて、音声分析合成系を構成し、その性能を把握する努力が、1976年から1981年にかけて精力的に行われ、その成果に基づいて、LSP音声合成用LSIの開発も行われた。

LSPの長所:

- ① PARCOR係数に比べて、パラメタ量子化誤差の影響が少なく、
- ② パラメタを補間したときのスペクトル再現精度が高い

この特徴により、LSPパラメタは、現在も多くのデジタル携帯電話の線形予測を利用した音声符号化方式に利用されている。

LSP音声合成フィルタの構造



LSP Speech Synthesizer 1980

F. Itakura
S. Sagayama

各種LPCパラメータ間の相互関係

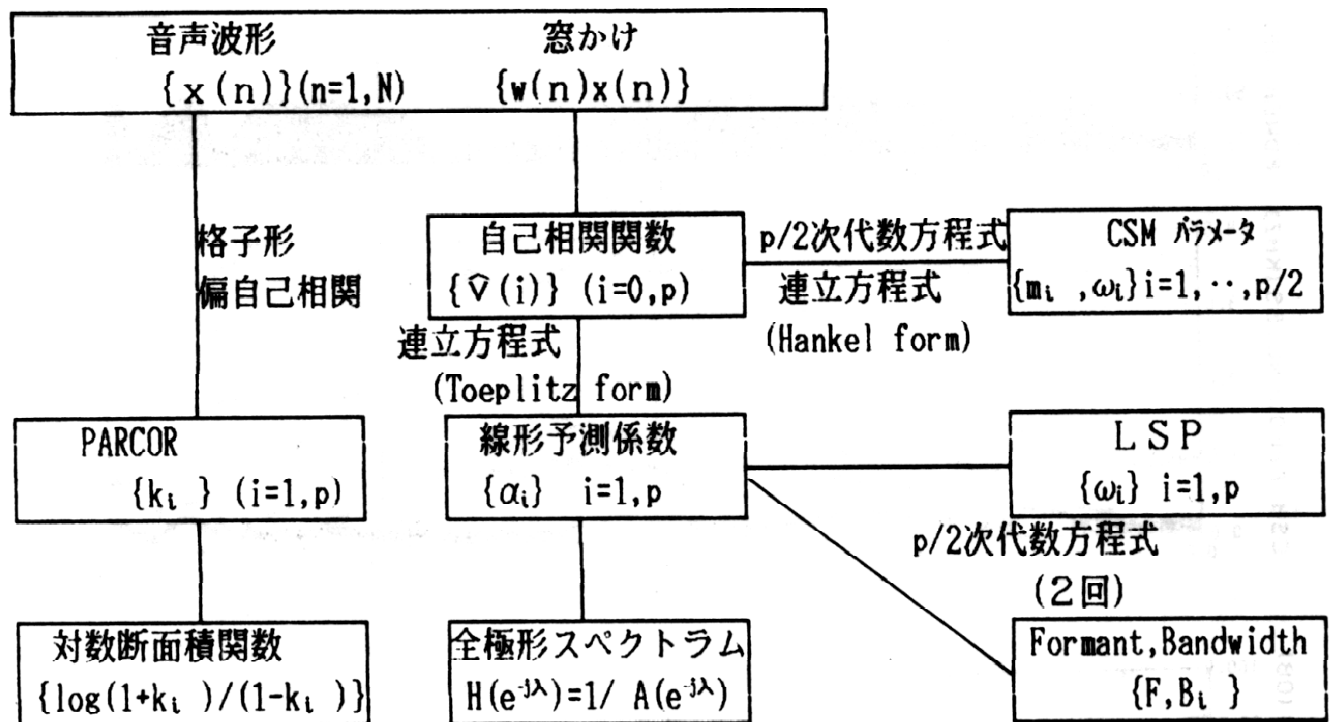


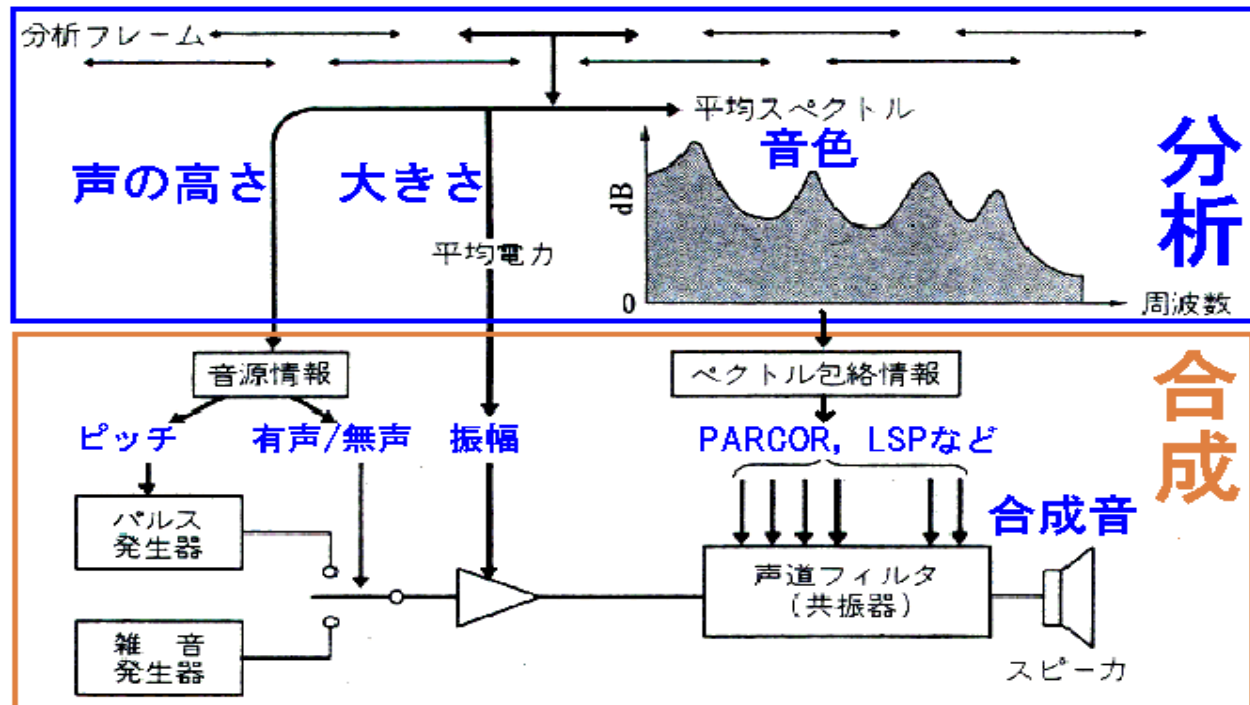
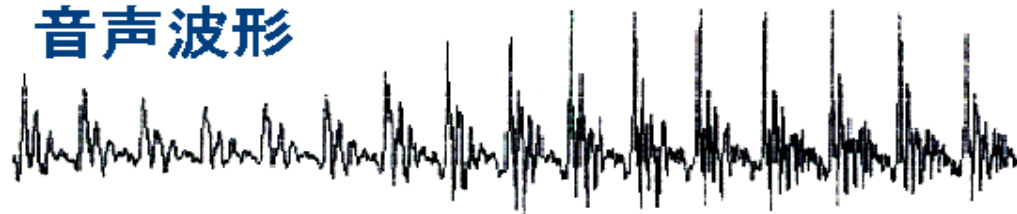
図10 各種LPCパラメータ間の相互関係

7. 音声情報圧縮への応用

- デジタル携帯電話
- IP電話

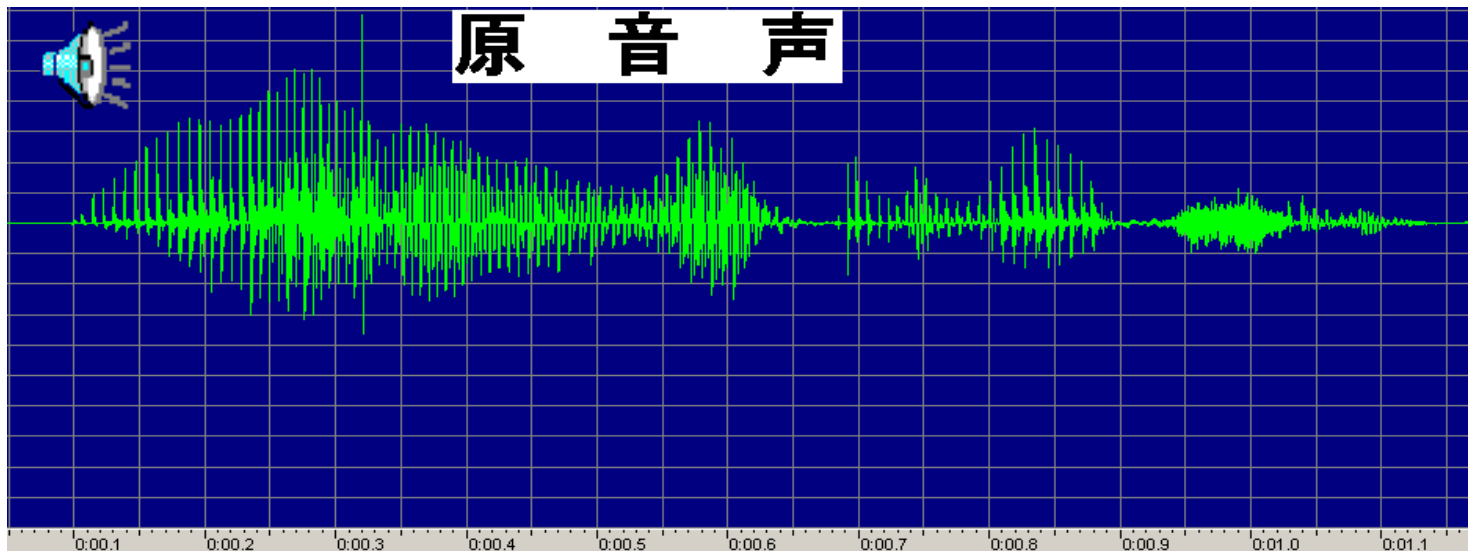
音声分析合成システム

音声波形

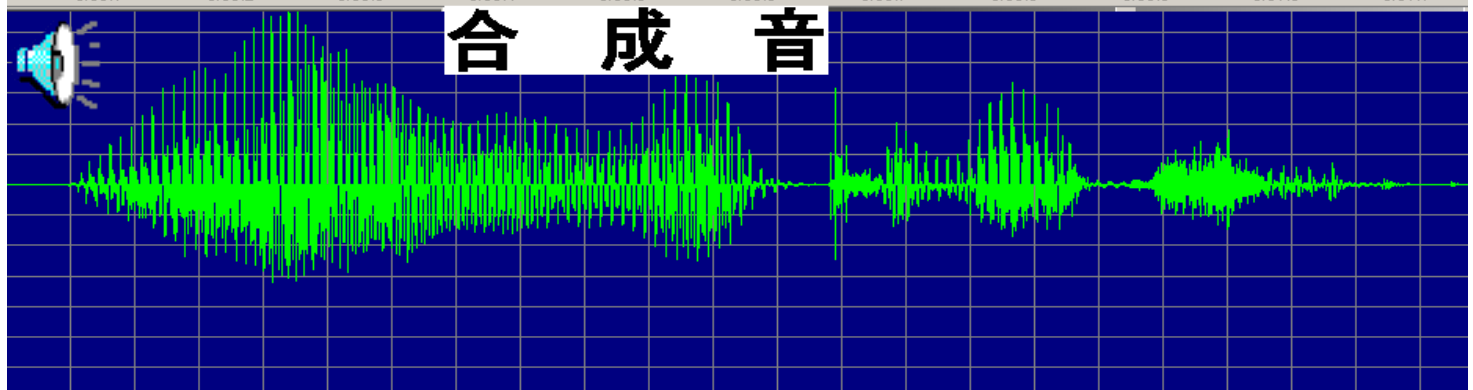


分析結果から合成した音声

原 音 声



合 成 音



携帯電話の主な音声符号化方式

地域	規格	符号化方式
日本	PDC	VSELP PSI-CELP CS-ACELP
北アメリカ	cdmaOne	Q-CELP EVRC
欧州・中近東・中近東 ・東南アジア	GSM	RPE-LTP VSELP

ITU-T標準方式

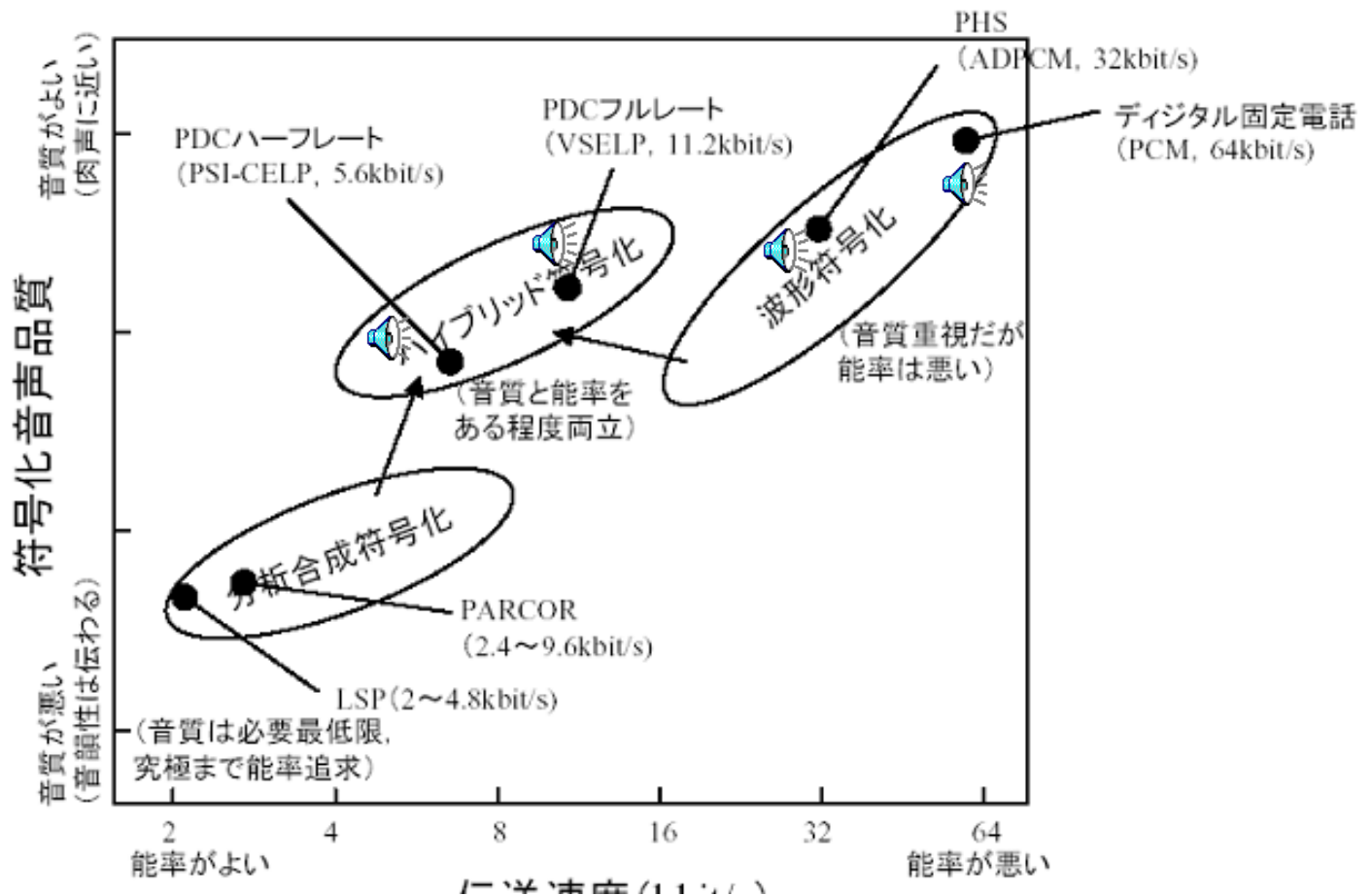
規格名	方式名	情報量 (kbps)	遅延 (ms)	処理量 (MOPS)	主な用途	デモ
G.711	Log-PCM	64	0.125	0.1	基幹伝送	
G.726	ADPCM	16,24,32,40	0.125	0.5	多重化, PHS	
G.727	ADPCM (embedded)	16,24,32,40	0.125	0.5	パケット伝送	
G.728	LD-CELP	16	0.625	20	多重化、TV電話	
G.729	CS- ACELP	8	15	15	多重化、次世代公衆陸上移動通信 (IMT-2000)	
G.723.1	ACELP/ MPC-MLQ	5.3 6.3	37.5	15	TV電話	

LSP使用

CS-ACELP: Conjugate Structure Algebraic CELP

MPC-MLQ: Multi-pulse Coding-Maximum Likelihood Quantizer

符号化方式と音声品質



8. むすび

音声情報処理の研究は、音声生理、知覚、音声学などの基礎的研究と関連しながら、最近のマイクロエレクトロニクスとソフトウェア技術をベースにして、過去40年の間に長足の進歩を遂げてきた。音声情報処理の目標は、人類の夢である人間とシステムの自然な対話(人間相互間と同様な)を実現することであろう。ここで紹介した音声分析合成技術は、「千里の道の一里塚」にすぎないが、数理的な基礎が強固であるため安心して応用でき、実用性も高いものと言えよう。

現在、精力的に研究開発が進められている音声対話システムは、巨大なソフトウェアシステムである。それ故、その構成要素の中に数理的に脆弱なものが多々見受けられる。こうした弱点を取り除き、理論的にも強化してこそ、信頼できる人間・機械間の対話が実現されると考える。

初期(1966-1978)の資料1

- { '66年度(12) } ('67.1.25) 斎藤収三(電通研), 板倉文忠(名大), ``音声スペクトル密度の統計的最適識別法に関する理論的考察'', 日本音響学会誌, vol. 23, No. 3, 4, pp. 181, July 1967.
- { '67年度(10) } ('67.9.29) 板倉文忠(名大), 斎藤収三(武蔵野通研), ``最尤法によるスペクトル包絡の推定値を用いた音声基本周期の抽出法'', 日本音響学会誌, vol. 24, No. 2, pp. 104, March 1968.
- { '69年度(3) } ('69.7.21) 斎藤収三・板倉文忠(武蔵野通研), ``偏自己相関係数による音声分析合成系'', 日本音響学会誌, vol. 25, No. 5, pp. 306, Sept 1969.
- { '70年度(8) } ('70.10.31) 板倉文忠・斎藤収三(武蔵野通研), ``PARCOR 形音声分析合成方式とその応用'', 日本音響学会誌, vol. 27, No. 2, pp. 114, Feb 1971.
- { '72年度(10) } ('73.2.16) 板倉文忠(武蔵野通研), ``時間領域における音声スペクトル情報の分析と合成'', 日本音響学会誌, vol. 29, No. 4, pp. 276--277, Apr 1973.
- {S74-07} 好田正紀・板倉文忠・斎藤収三(武蔵野通研), ``音声情報高速処理装置'', 音声研資, S74-07, pp. 1--8, June 1974.
- {S75-34} 板倉文忠(武蔵野通研), ``線形予測係数の線スペクトル表現'', 音声研資, S75-34, pp. 1--8, Dec. 1975.
- {S76-10} 板倉文忠・東倉洋一(武蔵野通研), ``擬ホルマント ボコーダ'', 音声研資, S76-10, pp. 47--54, 1975.
- {S76-11} 東倉洋一・板倉文忠・橋本新一郎(武蔵野通研), ``PARCOR方式におけるスペクトル平滑化の効果'', 音声研資, S76-11, pp. 55--62, 1976.
- {S77-07} 東倉洋一・板倉文忠(武蔵野通研), ``低ビットPARCORの品質向上'', 音声研資, S77-07, pp. 1--8, May 1977.
- {S77-36} 管村 昇・板倉文忠(武蔵野通研), ``可変フレームによる低ビット擬ホルマントボコーダ'', 音声研資, S77-36, pp. 1--8, Nov. 1977.
- {S78-02} 東倉洋一・板倉文忠(武蔵野通研), ``PARCOR分析合成系におけるスペクトル感度解析'', 音声研資, S78-02, pp. 9--16, Apr. 1978.
- {S78-13} 管村 昇・板倉文忠(武蔵野通研), ``パラメータの直線近似による音声情報圧縮'', 音声研資, S78-13, pp. 1--8, May 1978.

初期(1979-1982)の資料2

- {S79-06} 嵯峨山茂樹・板倉文忠(武蔵野通研), ``複合正弦波モデルによる音声のスペクトル分析'', 音声研資, S79-06, pp. 41--48, May 1979.
- {S79-08} 管村 昇・板倉文忠(武蔵野通研), ``パターンマッチング符号化による音声情報圧縮'', 音声研資, S79-08, pp. 57--63, May 1979.
- {S79-39} 嵯峨山茂樹・板倉文忠(武蔵野通研), ``複合正弦波による音声合成'', 音声研資, S79-39, pp. 293--300, Oct. 1979.
- {S79-46} 板倉文忠・管村 昇(武蔵野通研), ``L S P 音声合成器の原理と構成'', 音声研資, S79-46, pp. 349--356, Nov. 1979.
- {S80-2} 菅田雅彰・板倉文忠(武蔵野通研), ``音声の適応ビット割当て予測符号化'', 音声研資, S80-2, pp. 9--16, Apr. 1980.
- {S80-11} 小林 勉・板倉文忠(武蔵野通研), ``L S P 音声分析合成装置の検討'', 音声研資, S80-11, pp. 87--92, May 1980.
- {S80-12} 嵯峨山茂樹・板倉文忠(武蔵野通研), 太刀川信一(長岡技科大), ``複合正弦波音声合成方式の検討と合成器の試作'', 音声研資, S80-12, pp. 93--100, May 1980.
- {S80-29} 管村 昇・嵯峨山茂樹・板倉文忠(武蔵野通研), ``L S P 音声合成の品質評価'', 音声研資, S80-29, pp. 229--233, July 1980.
- {S80-45} 管村 昇・板倉文忠(武蔵野通研), ``話者を限定した場合のパターンマッチングによる音声情報圧縮'', 音声研資, S80-45, pp. 355--361, Oct. 1980.
- {S80-46} 嵯峨山茂樹・管村 昇・板倉文忠・小池恒彦(武蔵野通研), ``音声合成器のL S I 化とその応用'', 音声研資, S80-46, pp. 363--370, Oct. 1980.
- {S80-54} 長渕裕実・小林 勉・板倉文忠(武蔵野通研), ``雑音環境下の音声分析合成における雑音抑圧処理の効果'', 音声研資, S80-54, pp. 425--432, Nov. 1980.
- {S81-39} 小林 勉・板倉文忠(武蔵野通研), ``汎用音声処理装置による音声分析合成系'', 音声研資, S81-39, pp. 307--312, Oct. 1981.
- {S82-14} 嵯峨山茂樹・板倉文忠(武蔵野通研), ``CSM, L S Pに関する理論と諸性質'', 音声研資, S82-14, pp. 105--112, May 1982.
- {S82-44} 菅田雅彰・小林 勉・板倉文忠(武蔵野通研), ``DSPによるAPC-ABコーデックの実時間処理'', 音声研資, S82-44, pp. 345--352, Oct. 1982.

付録

音声符号化と携帯電話

NTT技術史館資料による

<http://www.hct.ecl.ntt.co.jp/index.html>

PARCOR音声分析合成方式

PARCOR:Partial Auto-Correation

- 1969年に発明されたPARCORは、音声の特徴をスペクトル包絡と音源情報で表現する音声分析合成方式である。
- スペクトル包絡情報であるPARCORと、音源情報であるパルス列(有声音)とノイズ(無声音)との符号化によって、きわめて高い情報圧縮を達成し、音声の低ビット符号化の実用可能性を世界に先駆けて示した技術である。

PARCOR音声合成器

1970年（試作:NEC）

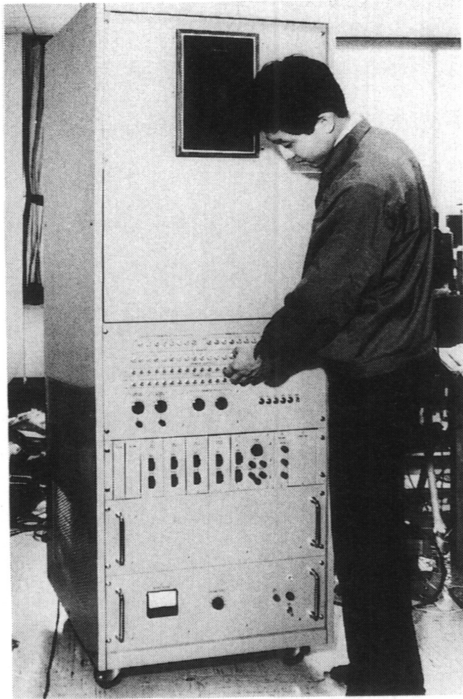
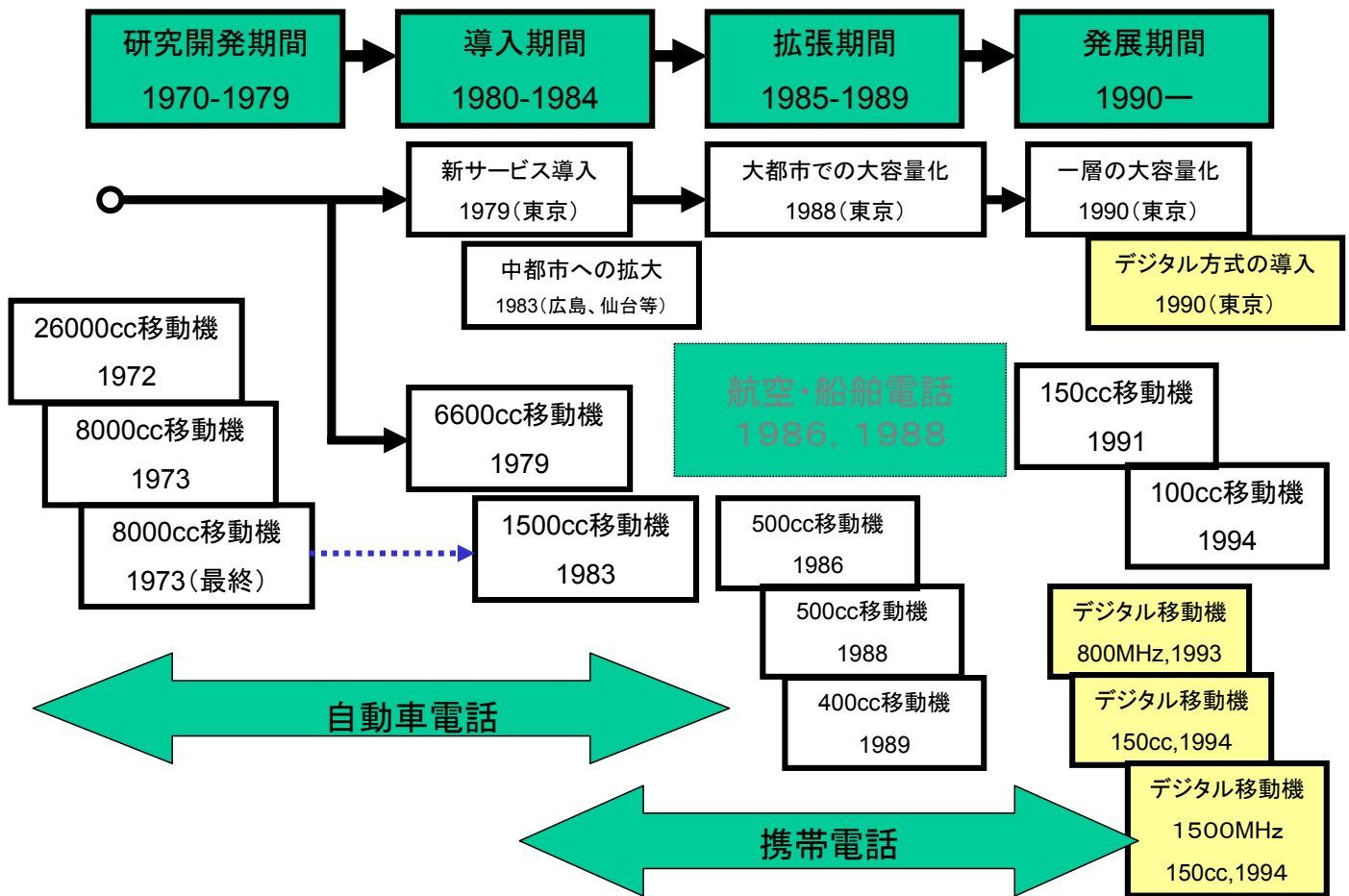


図2 PARCOR 音声合成器
(1970年試作, NTT 提供)

LSP音声分析合成法 LSP:Line Spectrum Pair

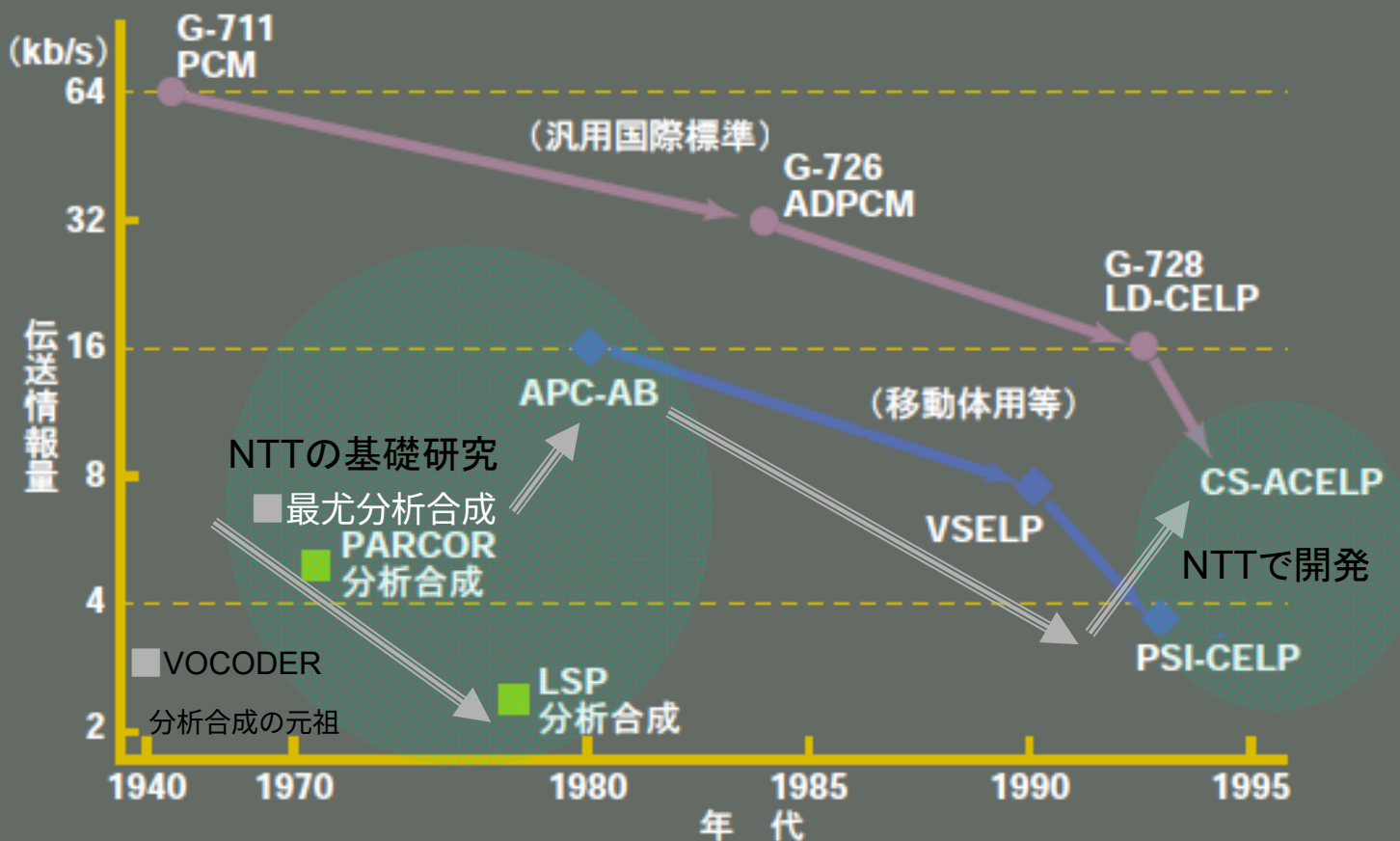
- 1978年、音声のLSP(線スペクトル対)と呼ぶ音声スペクトル包絡を表現する分析合成法を開発した。
- PARCOR分析合成系より40%も低い情報量で同質の合成ができる。
- また、このLSPパラメータの技術は携帯電話、インターネット電話などの低ビット音声符号化の必須の要素技術として世界中で使われている。

自動車・携帯電話の開発



音声符号化の研究の歩み

音声符号化における情報圧縮の進展



PSI-CELP方式 (携帯電話用音声符号化方式)

Pitch Synchronous Innovation-Code Excited Linear Prediction

➤ PSI-CEL方式は、移動体通信のデジタル化を実現する音声符号化(圧縮)技術である。1993年、電波システム開発センター(現電波産業会)が中心になって標準化された。

➤この方式はフルレート方式の品質を維持しつつ情報伝送量を半分にできるハーフレート方式である。

➤NTTはこの方式を考案し、ハードウェアを提案した結果、標準化となった。

PSI-CELP音声符号化方式の構成

