



世界に誇れる研究開発成果

高圧縮音声符号化の必須技術： 線スペクトル対 (LSP)

もりや たけひろ

守谷 健弘

NTTコミュニケーション科学基礎研究所

IEEEマイルストーンに認定された線スペクトル対 (LSP) について、その特徴と音声符号化の分野での世界への普及状況について紹介します。LSPは1975年に板倉文忠氏によって発明された音声の周波数スペクトル、すなわち声道（口の形）を表現する方法です。1980年にはLSPによる合成チップがつけられ、1990年代以降にはLSPは世界中の音声符号化の標準方式の主要な要素技術として組み込まれ、現在もほぼ世界中の携帯電話やIP電話で使われ続けています。

LSPとは

LSPは線形予測モデルの予測係数 $a[i]$ を等価変換したパラメータです。音声処理で主流となっている線形予測は自己回帰型あるいは全極型と呼ばれる形式で、音声サンプルの過去 p (例えば16) 時点前までのサンプルに係数 $a[i]$ をかけて現時点のサンプルを予測するものです。すなわち予測誤差の値 $\hat{x}[n]$ は現在の入力値 $x[n]$ と予測された $-\sum$ の項の値の差として(1)のように表現できます。

$$\hat{x}[n] = x[n] + \sum_{i=1}^p a[i]x[n-i] \quad \dots(1)$$

$a[i]$ は予測誤差の平均エネルギーが最小となるように決定します。これは z の多項式で

$$A(z) = 1 + \sum_{i=1}^p a[i]z^{-i} \quad \dots(2)$$

とも表現でき、この逆の $1/A(z)$ が合成フィルタの伝達式となります。この予測係数で定められるスペクトルの概形は音声の特徴、特に口の形に相当する周波数特性を効率良く表現できるため、音声符号化に限らず、合成、認識など幅広く使われている基本技術で

す。NTTの板倉文忠氏とAT&TのB. Atal氏、M. Schroeder氏が1966年の同時期に独立に線形予測モデルを音声分析に取り入れたとされています⁽¹⁾。

音声符号化では予測係数に相当する情報をできるだけ少ない情報で忠実に伝達するため、板倉氏は予測係数の符号化に予測係数と等価であり安定性や量子化特性に優れるPARCOR (Partial AutoCorrelation: 偏自己相関) 係数^{*1}を1972年に考案し、さらに優れたLSPの考案に至りました^{(2)~(4)}。 p 個のLSPは式(3)、(4)で定義される和と差の多項式 $F_1(z)$ 、

$F_2(z)$ の根として定義されるものです。

$$F_1(z) = A(z) + z^{-(p+1)}A(z^{-1}) \quad \dots(3)$$

$$F_2(z) = A(z) - z^{-(p+1)}A(z^{-1}) \quad \dots(4)$$

根は z 平面状の単位円上にならび、その角度 (周波数) をLSP周波数 (LSFとも呼ばれる) として量子化などに使います。図1は音声の対数スペクトル

*1 PARCOR係数：線形予測係数と等価なパラメータで予測係数そのものより安定判別が容易で量子化特性が良いという特徴があります。

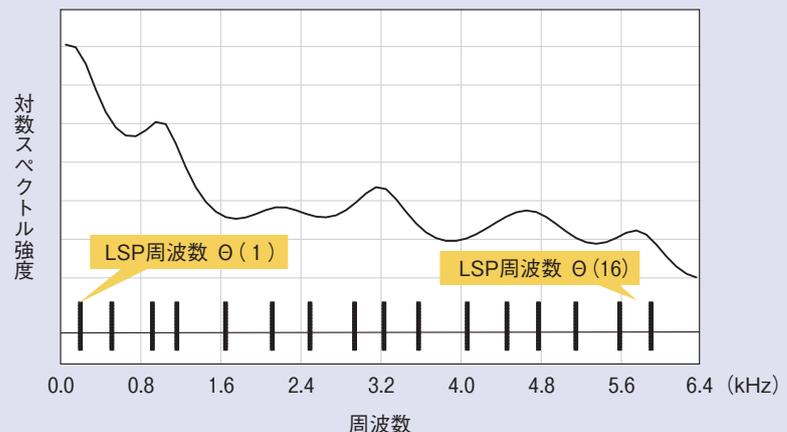


図1 対数スペクトル包絡と対応する16個のLSP周波数

包絡と対応する16個のLSP周波数, $\theta[1], \theta[2], \dots, \theta[16]$ を示したものです。LSP周波数の大きさは周波数軸上で低次から順に大きくなるように並ぶ性質があり, その順序が保たれることでフィルタが安定であることが保証されます。またLSPの変動がスペクトル包絡の変動に及ぼす影響がほぼ均一であり, LSPを粗い量子化してもスペクトルに対する影響が少ない(量子化特性が良い)という特徴があります。また2つのパラメータA, Bの中間値 $(A+B)/2$ に対応するスペクトルはA, Bのスペクトルをスペクトルの領域で平均したものと近い(補間特性が良い)という性質があります。さらに, このような性質はLSPの10個のパラメータを通常の波形の値のベクトルとして扱うことができ, さまざまな量子化, 補間の工夫が可能になりました。安定性や補間特性はテキストからの音声合成に有利な性質で, さらに量子化特性が良いことは符号化にとってもっとも有利な特徴です。

LSPの世界への普及の歩み

管村昇氏, 嵯峨山茂樹氏, 小林勉氏, (故) 東倉洋一氏らによって基本性能や理論の体系化, ハードウェア化など実用への準備が進められました⁽⁵⁾。1980年にはLSPによる音声合成LSIチップが試作され, それまで400枚もの回路ボードが必要であった実時間合成器が1チップでできたので音声応答装置などに応用されました。ちなみにこのチップの演算能力はまだ0.1MOPS (Mega Operations Per Sec: 秒当たり100万演算) 相当でしたが, 1990年代以降の携帯電話用処理LSIチップの演算能力はその100倍にも達するのでLSI技術の進展に敬意を表したいと思います。

LSPと音声の基本周期, 音量だけをパラメータとする分析合成符号化は, 4 kbit/s以下の20分の1程度の極

めて高効率の圧縮符号化を可能としましたが, 背景雑音による劣化や出力音声の不自然性などの問題が残りに, 公衆通信には採用されませんでした。1980年代には「分析合成符号化」と「波形符号化」を組み合わせたCELP (Code Excited Linear Prediction: 符号励振線形予測)^{*2}に代表される「ハイブリッド音声符号化」の研究が進められ, その中でもLSPの量子化など応用技術が進展しました。その一方で, 光ファイバや光ディスクの情報容量が急速に増大し, 1970年代末にアナログ方式でサービスが開始された移動体通信のデジタル化も不透明でした。ところが1990年代初頭に音声符号化の進展, 信号処理チップの高性能化と省電力化, 電池の高性能化, 電波のデジタル変調技術などの進展や秘話の必要性を背景に, 世界の移動体通信は欧州と北米を皮切りに一気にデジタル化に舵がきられました。これを契機に図2のようにLSPのさまざまな量子化法, LSPと組み合わせたさまざまな符号化の研究が進展し, さまざまな標準化に採用されるに至りました。マイルストーンの授与式に来日された現IEEE会長のRoberto de Marca博士ご自身も1990年代初めにLSPの効率的量子化につ

いて研究論文⁽⁶⁾を発表されていたことは感慨深いものがあります。

LSPが組み込まれた各種標準化符号化方式を表に示します。1991年には米国連邦政府音声符号化標準規格にLSPが組み込まれ, 日本, 欧州でも第2世代後半〔日本では1993年RCR (現ARIB) によるPDCハーフレート〕からの携帯電話用音声符号化標準規格に組み込まれました。1996年にはLSPが組み込まれた2種の音声符号化方式がITU-T勧告 (G.723. 1とG.729) となり, 特にG.729はゲートウェイやアダプタなどの装置に搭載され, 低ビットのIP電話 (日本では050) 用途で世界に広く普及しています。1999年には, 現在の世界で広く利用されている第3世代携帯電話向けの音声符号化方式である3GPP (3rd Generation Partnership Projects: 第3世代共同プロジェクト)^{*3}と3GPP2^{*4}の双方の音声符号化標準規格にLSPが組み込まれました。またLSPは音声だけ

*2 CELP: 数多くの種類の用意された波形を入力として線形予測合成フィルタで試しに音声を合成して, 原音にもっとも近くなる波形の番号を伝送する符号化の枠組み。1985年のベル研究所からの提案以降, 実用的改良が加えられ多くの低ビット音声符号化の基本となっています。

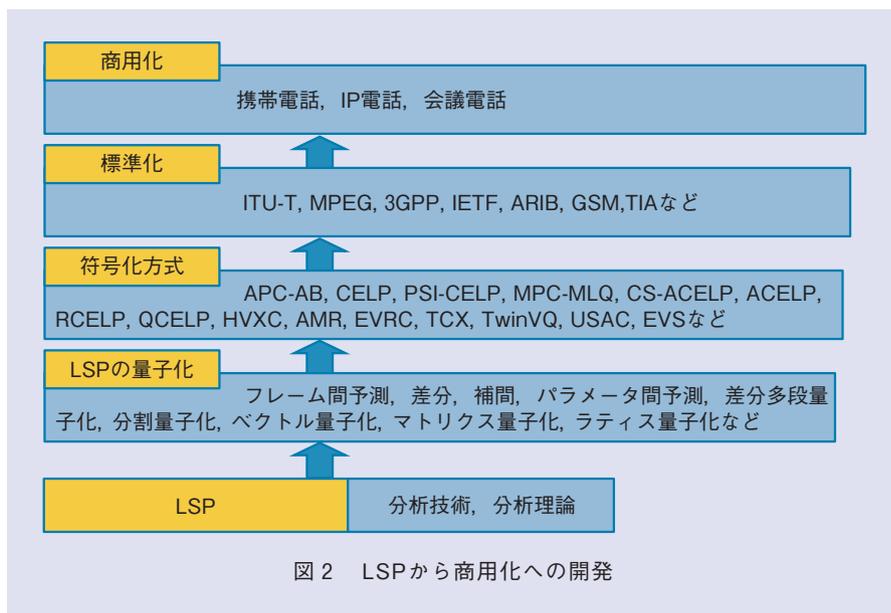


図2 LSPから商用化への開発

表 LSPが組み込まれている主な標準規格

標準化団体	方式名	情報量 (kbit/s)	主な用途	制定年
米国連邦政府	FS1016 CELP	4.8	秘密通信	1991
	FS1017 MELP	2.4	秘密通信	1995
日本RCR (現ARIB)	STD-T27 PSI-CELP	3.4	日本第2世代 ハーフレート	1993
米国TIA/EIA	IS-95 RCELP	2, 4, 8	米国第2世代 ハーフレート	1995
欧州GSM	GSM-EFR	12.2	欧州第2世代 改良フルレート	1997
ITU-T	G.723.1 MLP-MLQ/ACELP	5.3/6.3	TV電話・IP電話	1996
	G.729 CS-ACELP	8	IP電話 携帯電話 (PDC)	1996
3GPP	AMR	12.2	第3世代携帯	1999
3GPP2	EVRC	9.6	第3世代携帯	1999
ISO/IEC MPEG-4	14496-3:2009 CELP/HVXC/TwinVQ	2~16	音声音響符号化	1999
ISO/IEC MPEG-D	23003-3:2012 USAC	8~256	音声音響符号化	2010
3GPP	AMR-WB	8~23	VoLTE	2001
	AMR-WB+	6~48	音声音響符号化	2004
	EVS	5.9~96	VoLTE	2014

に限らず、あらゆる信号の周波数特性の包絡を表現することができます。これは音響信号の周波数領域での符号化TwinVQ⁽⁷⁾で有用性が実証され、ISO/IEC MPEG-4で部分的に採用され、またMPEG-D USAC (Unified Speech and Audio Coding) にも組み込まれています。

今後の音声通信に向けて

この夏からサービスを開始した

- *3 3GPP: 欧州のETSI (The European Telecommunications Standards Institute) と日本、韓国、中国、北米の標準化団体による第3世代移動体通信のための標準化規格制定共同プロジェクト。その後第4世代以降の規格制定活動も継続しています。
- *4 3GPP2: 北米のTIA (Telecommunications Industry Association) と日本、韓国、中国の標準化団体による第3世代移動体通信のための標準化規格制定共同プロジェクト。
- *5 VoLTE (Voice over LTE): LTEの移動体通信上で優先制御などを用いて音声通話を実現するシステム。

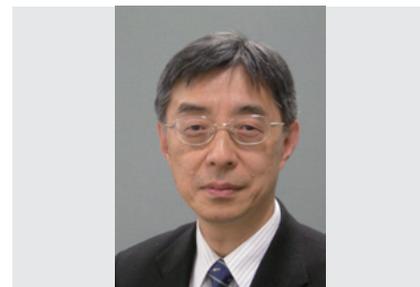
VoLTE (ボルテ)^{*5}には音声符号化として3GPP AMR-WBが使われて、16 kHzサンプルの音声 (8 kHz音声帯域: 中波ラジオ相当) に帯域が拡張されています。さらに次世代のVoLTEには現在標準化作業中のEVS (Enhanced Voice Service) が利用されることが想定され、32 kHzサンプルを入出力とし音楽も含めた品質改善が期待されています。VoLTE用のいずれの符号化方式でもLSPが使われることには変わりありません。将来はソフトウェアで実行可能な音声符号化や他のアプリなどとの統合などが図られると予想されますが、この場合でもLSPは引き続き必須の要素技術になります。このようにLSPは世界中のほぼすべての事業者の携帯電話機、基地局設備、IP電話機に搭載されており、さらに今後とも広く使われ続けるでしょう。

NTTコミュニケーション科学基礎

研究所では、世界の有力機関と連携して次世代VoLTE向けの音声符号化の標準化に取り組んでおり、近い将来に世界のお客さまにその高音質を喜んでいただけるよう努力したいと考えています。

参考文献

- (1) B. S. Atal: "The History of Linear Prediction," IEEE SIGNAL PROCESSING MAGAZINE, pp.154-157, March 2006.
- (2) F. Itakura: "Line Spectrum Representation of Linear Predictive Coefficients of Speech Signals," J. Acoust. Soc. Am., Vol.57, No.S1, 1975.
- (3) 板倉: "全極形デジタルフィルタ," 特許第1494819号.
- (4) 板倉: "音声信号に関する統計的・数理的手法—最ゆうスベクトル法からPARCORを経てLSPへ—," IEICE Fundamentals Review, Vol.3, No.3, 2010.
- (5) F. Itakura, T. Kobayashi and M. Honda: "A Hardware implementation of a new narrow to medium band speech coding," Proc. ICASSP '82, pp.1964-1967, Paris, France, May 1982.
- (6) J. R. B. de Marca: "An LSF Quantizer For the North-American Half-Rate Speech Coder," IEEE Trans. V.T., Vol.43, No.3, pp.413-419, August 1994.
- (7) N. Iwakami, T. Moriya, and S. Miki: "High quality Audio - Coding at less than 64 kbit/s by Using transform-domain weighted interleave vector quantization (TwinVQ)," Proc. ICASSP '95, pp.3095-3098, Detroit, U.S.A., May 1995.



守谷 健弘

LSPは世界の音声通信に広く貢献し、まさにIEEEマイルストーンにふさわしい技術です。NTT研究所は引き続き世界レベルの研究開発によって、お客さまに信頼とサービスを提供し続けていきたいと考えます。

◆問い合わせ先

NTTコミュニケーション科学基礎研究所
TEL 046-240-3141
FAX 046-240-3145
E-mail moriya.takehiro@lab.ntt.co.jp