音声合成システムについて

伊藤眞樹*1)・森澤孝光*2)

1. まえがき

我社が音声合成装置（VOICETECシリーズ）を世に発表してから早や4年の歳月が経った。この間，〝VOT-1024（アルボ・パック）〟，〝VOT-1501〟，〝VOT-1001〟，〝VOT-80J1〟，〝VOT-8001〟，〝VOT-1002（ハロー・トーク）〟，〝VOT-8002〟，そして〝VOT-8003〟と8機種が次々に発売され，FA，LA，OAまたは民生分野にも広く知られるようになってきた。

又，この間の半導体技術の進歩も凄じく，種々の音声合成用のLSIが各社から出され，その上メモリの集積度も増々大きくなって来たことにより音声合成装置もより身近な物になってきた。

そこで音声合成システムの基本と応用について述べてみたいと思う。

2. 音声合成の変遷

人間の声を何らかの機械的な装置によって，人工的に発生させる試みは，ずいぶん昔から行なわれてきた。我々が日常，耳にするような装置は，1877年に発明された蓄音器で代表されるが，それは今日磁気記録を利用したテープレコーダが主流になっている。これらの装置は，人間の発した声を，音の波形としてアナログ的に直接記録しておくものである。

これを再生する合成装置は，アナログ的に記録された音声波形を，ただ単に音響波形に変換する装置である。

ところが近年，半導体技術が急速に進歩し，1チップに数千～数十万素子を集積した大規模集積回路（Large Scale Intergtion：LSI）が，容易に実現されるようになると，以前には大型コンピュータを駆使して行なっていた音声のディジタル信号処理が，わずか数チップから1チップのLSIで構成したシステムで可能となってきた。

このようなディジタル処理による音声合成技術は，音声波形の合成の仕方及び，音声として発せられる言葉の組み立て方によって，各種の方法が研究され，その用途に応じて，すでに，いくつかの方法が実用化され，各社からそのチップが発表されている。ここでは音声合成技術について代表的なものを取り上げて述べることにする。

2.1 音声合成システムとは[1)]

音声合成システムを理解するために，人間の発声器官の働きをまず理解する必要がある。そこで，ここではまず人間の発声器官の働きを概略的に説明する。

人間の発声器官は**図1**に示されるように，肺から押し出された空気によって，声帯波を発生する声帯と喉頭，咽頭，口腔，鼻腔などから構成される声道，および音声を放射する役割をもっている口と鼻に大きく分類される。

有声音の声帯波は，肺からの空気による声帯の振動音で，多くの高調波成分を含んだ鋸歯状の波形である。この波形の周期を「ピッチ」と呼んでいる。仮りに，その音だけを純粋に耳で聞いたとすると，ブザーの音のように味もそっけもないものとなる。そのような単調な音が音声として聞こえるためには，声帯から口までの声道が重要な役割を演じている。

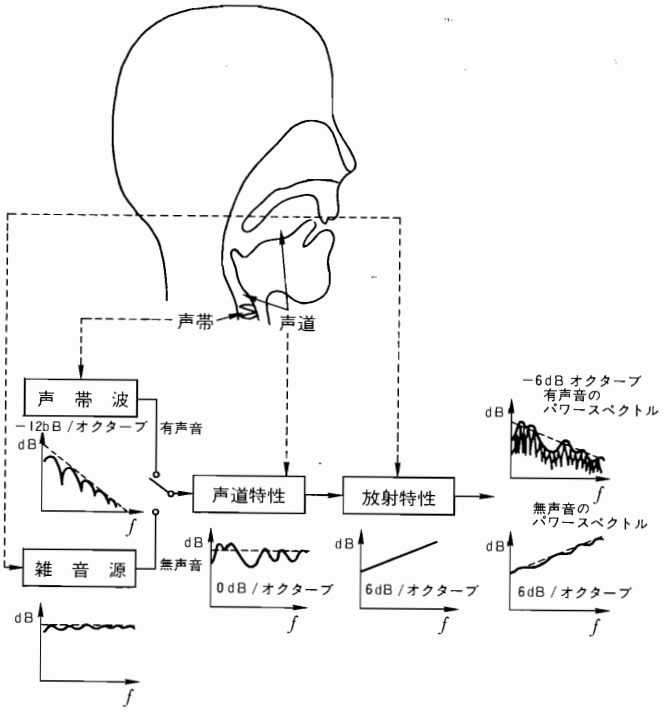


図1. 発声器官と音声合成モデル

*1) 新分野開発部　開発リーダー
*2) 新分野開発部　主任

声道は，舌とか喉頭の形の変化により，断面形状が時間とともに変化する導波管として考えられる。また，声帯波は，声道内部でその断面形状に応じて異なる共振を起こし，性別や年齢などの話者の個人的な特徴や〝a〟〝i〟といった音韻の違いが付加される。

有声音に対して，無声音とは，声帯振動によらない音のことで，声道中に発生する乱流による音（雑音）が音源となり，声道による共振特性が付加されたものである。

声道を通ったこれらの音は，口あるいは鼻から音声として空中に放射されることになり，我々の耳に到達する。つまり音声は，声帯から発生された声帯波や雑音を音源として，それらが声道を介し，放射されると有声音や無声音となるわけである。

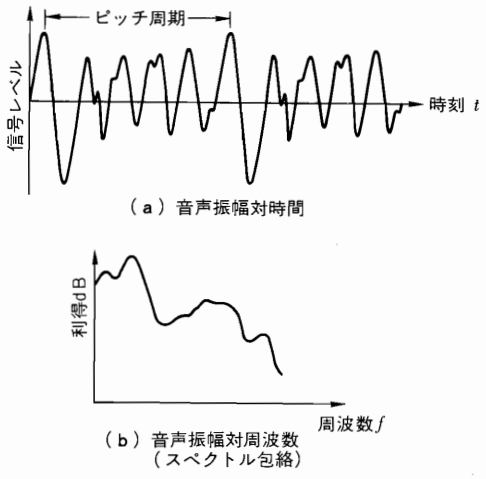
口から発せられた音声波形を時間に対する１次元信号として観測する方法には２つがある。すなわち一つは，**図2(a)** のように音声信号を時間のパラメータとして，直接音声波形の振幅値を表示する方法と，**図2(b)**のように，音声信号に含まれている周波数成分の振幅特性を周波数のパラメータとして表示する方法である。後者の表示方法を「パワースペクトル特性」または「スペクトル包絡特性」と呼び，その中には微細構造がある。

この音声信号のパワースペクトル特性には，生の音の波形に存在する位相に関する情報は含まれていないが，どのようなものでもその音を聞いたとき，正常な音として聞くことができる。それは，人間が同じ形をしたパワースペクトル特性の生の音で，かつ，波形の位相の異なる２つの音を聞いた場合，位相の区別ができないためである。

従って，音声合成を行なう場合，位相に関する情報は不要（冗長）になることが分る。このように，パワースペクトル特性は音声を観測する場合，それだけで意味があり，生の音声波形よりも人間の音声の区別（認識）と結びついた観測方法といえる。ここで，もう一度**図1**を見ると，有声音，無声音のそれぞれに対して，声帯波特性，声道特性，放射特性を周波数対利得図の形で示していることが分る。

有声音は，音源特性，声道特性，放射特性の和で表わされる。即ち，約－12dB/オクターブの音源特性（周波数が１オクターブ上がると，利得が12dB落ちる）と0dB/オクターブの声道特性（利得は一定で部分的に強弱がある），及び約6dB/オクターブの放射特性をもち，結果として口，鼻から約－6dB/オクターブの右下がりの特性をもった音声として発せられることになる。

一方無音声は，音源が0dB/オクターブの特性を示すことから，約6dB/オクターブの右上がりの特性で発せられる。例えば，典型的な場合，周波数特性の傾斜をみることによって，有声音と無声音が区別されることになる。



（a）音声振幅対時間

（b）音声振幅対周波数（スペクトル包絡）

図2. 音声波形を時間のパラメータとして表現した特性(a)と周波数のパラメータとして表現した特性(b)

このように，人間の音声の特徴を各特性として的確に把握することによって，合成音声を出力するシステムが種々考えられるわけである。

音声合成システムとは，一般的に何らかの情報を言葉（言語化）に直し，その言葉を人工的に装置によって，人間の発する音声に似せて発声するシステムと定義付けられる。

もっとも古くから実用化されている方法として，人間の音声の波形データをそのまま単語あるいは文節単位で記憶し（これを「分析」という），再生時（これを「合成」という）に発声したい文章あるいは言葉になるように，適当な順序で単語あるいは文節を並び換え編集し，時系列的に音声波形を再現してゆく「録音編集方式」があった。

この方式は，少なくとも単語単位で生の音声波形データを記憶するために，合成された音声の了解性，自然性が優れている。

しかし，記憶すべきデータ量は，他の方式に比べて大きくなるという欠点がある。また，この方式は記憶する方法によって，PCM,[注1] DPCM,[注2] APCM,[注3] ADPCM,[注4] DM,[注5] ADM,[注6] CVSD,[注7] CSM[注8]といった各種に分類されている。従ってパルス符号化方式，直接波形記憶方式あるいは，波形符号方式等と呼ばれている。

注1） PCM(Pulse Code Modulation：パルス符号変調方式）
注2） DPCM(Differential PCM：差分 PCM)
注3） APCM (Adaptive PCM：適応PCM）
注4） ADPCM （Adaptive Differential PCM：適応差分PCM）
注5） DM (Delta Modulation：デルタ変調方式）
注6） ADM (Adaptive DM：適応DM)
注7） CVSD (Continuously Variable Slope DM：連続変動公配ＤＭ）
注8） CSM (Composite Sinusoid Modeling：複合正弦波方式）

表1. 音声合成方式の分類

分類		入力情報	記憶内容	具体的方法	情報量（ビットレート）
録音編集方式	波形符号化方式（パルス符号化方式／直接波形記憶方式）	自然音声	波形パラメータ	PCM方式	64～200
				APCM（適応PCM）方式	～ 32
				DPCM（差分PCM）方式	10～32
				ΔM（デルタ変調）方式	10～32
				ADM（適応デルタ変調）方式	10～32
				CSM（複合正弦波）方式	1.2～2.4
	スペクトル符号化方式	自然音声	スペクトルパラメータ	PARCOR（偏自己相関係数）	1.2～9.6
				LPC（線形予測符号化）方式	～4.8
				LSP（線スペクトル対）方式	～1.2
				ボコーダ方式	0.6～4.8
規則合成方式		音韻記号または文字	記号系列とパラメータ生成規則	合成音声素片編集 スペクトルアナログ 構造アナログ	～0,1

これに対して，音成波形をそのまま記憶せずに，何らかの処理を施して記憶すべきデータ量を減少させて，より能率の良いパラメータを時系列として記憶しておき，音声の合成時にパラメータから音声波形を再現する方法がある。これを「スペクトル符号化方式」と呼んでいる。

この方式は，音声からパラメータへの変換方法によって，PARCOR[注9]，ケプストラム，LPC[注10]，LSP[注11]等といった各種の方式に分類できる。この方式の特徴は，波形符号化方式と同じように，自然音声を分析し，単語あるいは文節単位で，パラメータを記憶しておき，単語単位で編集することによって希望の言葉を発声するが，スペクトル上の性質を保存しながら，波形そのものを直接再現しないことである。

以上の2つの音声合成システムの他に，これらと異なる「規制合成方式」がある。この方式は，単語単位の自然音声を必要とせずに，文字列からのみ音声波形を合成しようとするものである。前二者の方式が自然音声を模擬するという性格が強いのに対して，この方式は，まさに機械が〝しゃべる〟という音声合成方式である。

従って，システムへの入力は，任意の通常の文章が考えられ，システムがその文章を読み上げるといった動作を行なうことになり，各単語を認識し，アクセント，イントネーション，リズムといった音声上必要な物理量をシステム内部で導き出し，音声合成を行なう。この方式は，1秒間の音声を合成するために必要とされる情報量が，他の方式に比べてきわめて少ない利点がある。しかし未解決の問題が多く，いまだに実用の段階に至っていないのが現状である。

以上，これらの方式をまとめると**表1**のようになる。

我社が採用している音声合成方式は，音声波形符号化方式中のADPCM方式又は，ADM方式である。次にこの音声波形符号化方式の基本となるPCM方式，それに，ADPCM，ADM方式について説明する。

2.2 PCM方式

音声をディジタルで取り扱う場合，電気回路系に入力されたアナログ信号である音声信号を，ディジタル信号に変換する必要がある。この変換を行なうのが，アナログディジタル変換器（A/Dコンバータ）である。

PCM方式は，A/Dコンバータの一種で，**図3**に示すように入力信号をPCM変換のビット数に対応するとびとびの値（ディジタル値）に変換するものである。この例では，10Vの入力信号を5ビットでPCM変換している。従って，10Vの入力レンジは，$2^5=32$に分割され，連続的に変化する入力信号は，32のきざみのステップに対応して，階段状に変化するディジタル値に変換される。この操作を「量子化」と呼んでいる。**図3**の例では，量

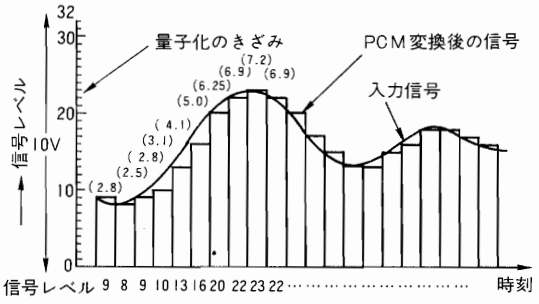


図3. PCM方式での入力信号の量子化

注9） PARCOR (Partial autocorrelation：偏自己相関係数方式)
注10） LPC (Liner Predictive Coding：線形予測符号化方式)
注11） LSP (Line Spectrum Pair：線スペクトル対方式)

子化の1ビットは約0.31Vに対応し，各信号レベルが同図（　）中の電圧値で表わされることになる。

このように，PCM方式は音声信号を適当な時間間隔tsで標本化して，適当なビット数で入力信号レベルを電圧値にそのまま変換するA/D変換方式である。標本化の周波数は，サンプリング周波数fsと呼ばれ，再生波形の帯域幅を決定し，A/D変換のビット数は，分解能を決定することになる。ここで，サンプリング周波数とビット数との積が，音声を1秒間に再生するのに必要とされるビット数となる。これを「ビットレート」と呼び，ビット/秒で表わす。

従って，サンプリング周波数を低くし，A/D変換のビット数を少なくすることによって，低ビット率で符号化できるはずである。

しかしながら，音声波形の標本化においては，音声波形の帯域幅と標本化周波数に関するShannon-染谷の標本化定理（Sampling Theorem）を満たすことが重要となる。即ち，音声帯域幅は，サンプリング周波数の1/2の周波数帯域までしか再生できないのである。このためサンプリング周波数は，必要とする音声周波数に対して2倍以上に設定する必要がある。

次式に標本化定理を示す。

$$g(t)=\sum_{n=-\infty}^{\infty} g\left(\frac{n}{2f_0}\right)\frac{\sin(2\pi f_0 t-n\pi)}{2\pi f_0 t-n\pi}$$

この式によれば，f_0(Hz)に帯域制限された音声波形g(t)は$1/2f_0$(Hz)ごとの標値$g(n/2f_0)$と帯域幅f_0を持つ理想低域ろ波器(LPF)のインパルス応答$\sin(2\pi f_0 t-n\pi)/(2\pi f_0 t-n\pi)$によって再現される。$f_0$(Hz)以上に残存成分があると$2f_0$(Hz)で標本化したときに$f_0$(Hz)で折り返されて歪が生ずる。これをエイリアシング（Aliasing: 折り返し）雑音という。これらの詳細については2.2のローパス・フィルタの必要性の項を参照されたい。

このように音声波形符号化方式の基本であるPCM方式に必要な情報量は最低64kb/s以上が必要とされている。

しかしながら，品質を同等に保って情報量をより少なくするための研究が続けられ，最近に至ってこれらの研究が大いに進展し，CCITT（国際電信電話諮問委員会）において32kb/s ADPCM方式の標準化が行なわれた。

2.3 ADPCM方式[2)]

ADPCM方式は信号の隣接サンプルの差分(dn)を量子化，そして符号化することにより情報量を削減する手法である。

差分dnを量子化するときの量子化幅△n を適応的（Adaptive）に変化させることを特徴としている。（DPCM方式では量子化幅△nは固定）

すなわち，差分dnが大きいときは△nも大きく，dnが小さいときは△nも小さくなるように適応変化させる。

2.3.1. ADPCM分析

n番目のサンプル点における入力をXn，（n−1）番目のサンプル点における波形再生値を$\hat{X}_{n-1}$とすると両者の差分dnは，

$$dn=Xn-\hat{X}n-1 \qquad \text{（差分算出）}$$

となる。

これを現時点（n番目時点）での量子化幅△nにより符号化をする。

$$Ln=dn/\Delta n \quad \text{（符号化 Ln：ADPCMデータ）}$$

これを量子化し波形再生すると，

$$qn=\left(Ln+\tfrac{1}{2}\right)\Delta n \text{（注）} \qquad \text{（量子化）}$$

$$\hat{X}n=\hat{X}n-1+qn \qquad \text{（再生）}$$

となる。

次に（n+1）番目のデータのために量子幅を△n から△n+1に変更すると，

$$\Delta n+1=\Delta n\cdot M(Ln) \qquad \text{（量子化幅変更）}$$

〔ここでM(Ln)はLnの関数形〕となる。

このように量子化幅は過去データの累積により適応的に決定されることになる。

以上の演算プロセスを**図4**に示す。

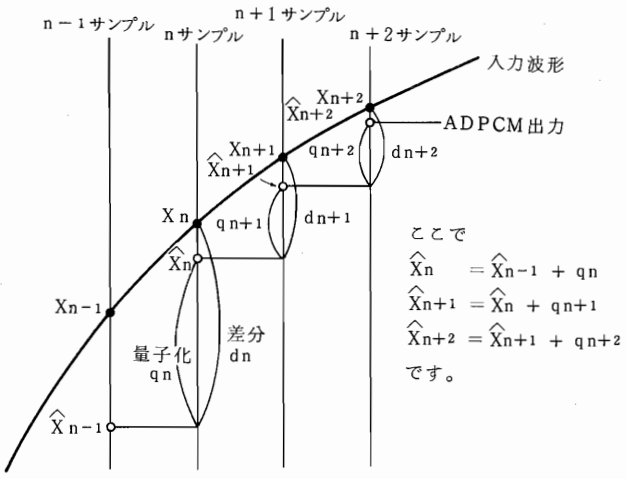


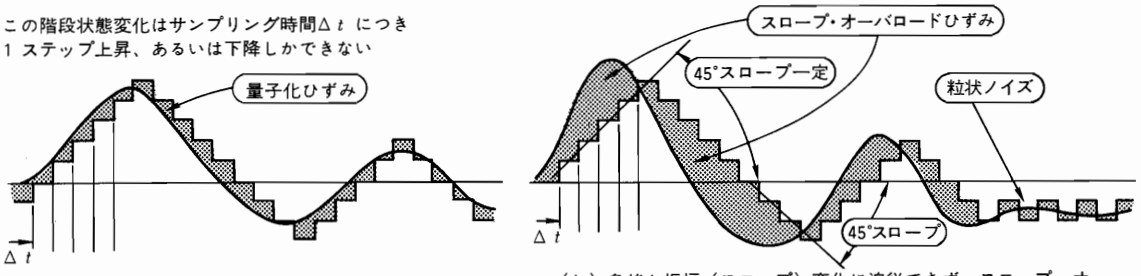
図4. ADPCM方式の演算プロセス

2.3.2. ADPCMデータ

ADPCMデータLnは極性ビットを含むMビットのデータとして表現される。

例えば，4ビット表現の場合は

$$Ln=\{B_3,\ B_2,\ B_1,\ B_0\}$$



（a）比較的ゆっくりした振幅変化のためDM方式での追従性が良い例

（b）急峻な振幅（スロープ）変化に追従できず，スロープ・オーバロード（SLOPE OVER LOAD）ひずみを生じている。また，緩慢な変化に対しても粒状（GRANULAR）ノイズが現れる。

図5. DM方式の波形

となり，

B_3 （極性ビット）は差分dnの正負

B_2 （MSB） は変化分の4Δnの桁の有無

B_1 （2SB） は 〃 2Δnの 〃

B_0 （LSB） は 〃 Δnの 〃

をそれぞれ表わしている。

2.3.3 ADPCM再生

ADPCM再生はADPCM分析の一部として表現することができる。

n番目のADPCM入力データLnに対して，

$$q_n \fallingdotseq (L_n + \tfrac{1}{2})\Delta n \quad (注) \qquad (量子化)$$

$$= (1-2B_3)(4\Delta nB_2 + 2\Delta nB_1 + \Delta nB_0 + \tfrac{1}{2}\Delta n)$$

$$X_n = X_{n-1} + q_n \qquad (再生)$$

と再生される。

さらに（n+1）番目のADPCMデータのためにΔn+1を演算する。

$$\Delta n+1 = \Delta n \cdot M(L_n)$$

このことからADPCMコードは予め設定された量子化幅に対し，新規なPCM値を算出するためのデータであると同時に，次に設定されるべき量子化幅を演算するためのデータにもなっていることが分る。

また，

Xnを12ビット

Lnを 4ビット

とするとADPCMコード化により情報量が4/12に圧縮されたことになる。

2.4 ADM方式[3)]

ADM方式はDM方式の変形である。即ちDM方式とは音声信号のサンプリング1回をたった1ビットで符号化する方法である。（**図5(a)参照**）

サンプリングΔt =一定でも，量子化ステップ幅（QUANTIZATION STEP SIZE）が可変となっている。このステップ幅を決定するためには，2ビット以上のステップ幅信号が必要になる。

図6. ADM方式

しかしながら**図5(b)**においては急峻な振幅変化があるため，Δt=一定とした場合，ディジタル量の増減分を正か負（つまり1ステップ）にしただけでは斜線部のとり残しが生じてしまうため，正確なディジタル化ができなくなる。この追従性の悪さは，再生時のひずみとして現れるため，スロープ・オーバロードひずひといわれる。一方，あまりにも緩慢な振幅（スロープ）変化に対しても別のひずみが観測される。これを粒状ノイズ（granular noise）と呼んでいる。

これらDM方式の欠点をカバーすべく考案されたのがADM方式である。**図6**に示すようにサンプリングのΔtは一定にしたままで，振幅方向のステップ幅を原波形の変化の大きさに適応させて変化させようとするものである。

3. 音声合成システムの設計

音声合成システムは，音声信号をディジタル信号に変換してから，メモリに記憶し，必要なときに必要な音声を取り出すことを目的とするものである。

従って基本的な要素としては，テープレコーダとほぼ同一であるが，録音時間と音質を除外すれば，非常に優れた特長を備えている。

① ランダム録音／再生が出来る。

（注） 量子化$q_n=(L_n+\frac{1}{2})\Delta n$の$\frac{1}{2}$の項は極性変化分についてリニアに等分するための手法である。

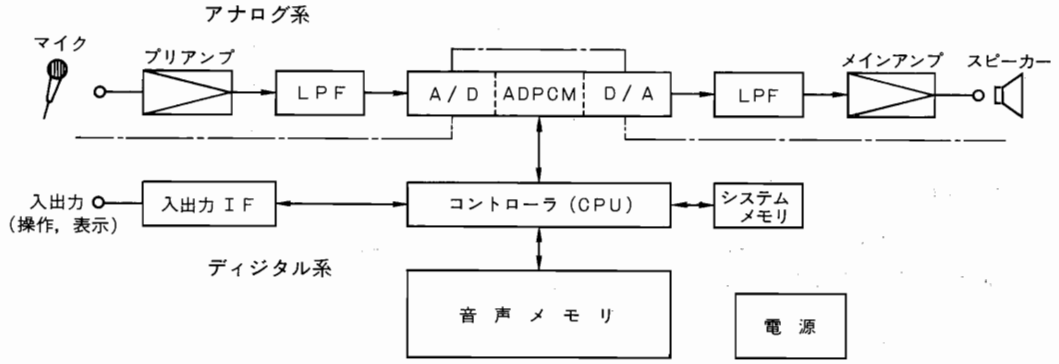


図7. 音声合成システム構成ブロック図

② 寿命が半永久的である。
③ 耐衝撃，耐振動性が良い。
④ 磁気の影響を受けない。
次にシステムの具体的例を紹介する。

3.1 音声合成システムの構成

システム構成は**図7**のようなブロック図から成っている。

大別するとアナログ系とディジタル系に分離することができる。

アナログ系は，

プリアンプ，メインアンプ，LPFと音声合成用LSIの一部であり，

ディジタル系は，

音声合成用LSIの一部，コントローラ、入出力のインターフェイスとメモリ系である。

これらの内で特に音質に直接影響のあるのは，サンプリング周波数と，A/D変換器の分解能(ビット数)，およびローパス・フィルタ(LPF) の特性である。

現在市販されている音声合成用LSIは，サンプリング周波数4K～16KHz，A/D変換器8～12ビット，フィルタは2～5次のOPアンプ又はスイッチト・キャパシタ・フィルタを採用したものがほとんどである。

この内でローパス・フィルタだけは内蔵されたものは使用せずに独自のものを採用した。次にローパス・フィルタについて述べる。

3.2.1 ローパス・フィルタの必要性

PCM(ADPCMも同様) 装置には，必ずAD変換前とDA変換後にローパス・フィルタを使用しているが，ここでその必要性について説明する。

図8は，原信号を時間軸で波形，周波数軸でそのスペクトラム分布を示している。次に，**図9**は標本化された信号とそのスペクトラム分布を示している。標本化により追加されたスペクトラムが見られるがこれをエイリアス（折り返し）現象と言う。**図9**は原信号faと標本化周

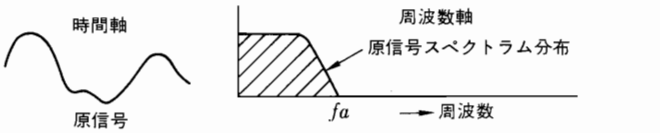
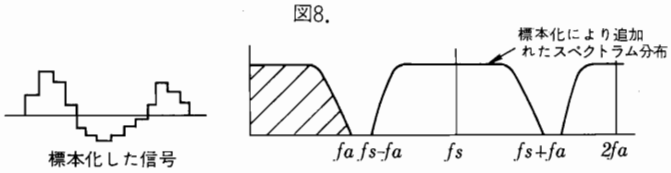
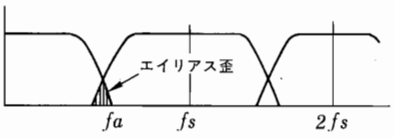


図8.



$fs > 2fa$

図9.



$fs < 2fa$

図10.

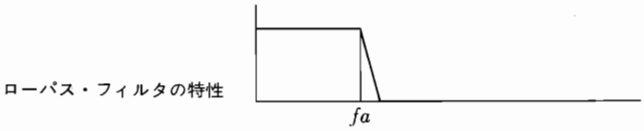


図11.

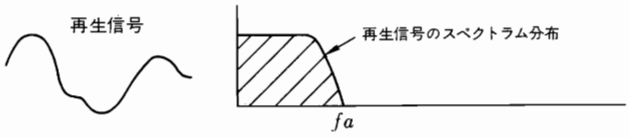


図12.

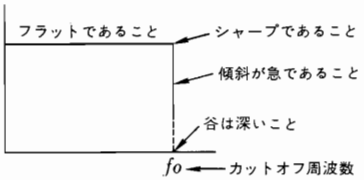


図13.

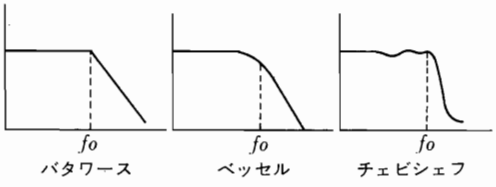


図14.

波数fsとの間をfs＞2faの条件としているが，もし標本化周波数を低くして fs＜2fa とすると**図10**のような原信号スペクトラムと折り返しスペクトラムの重なりが生じる。これをエイリアス（折り返し）歪と言い，この場合原信号に再生する事が困難になる。

音声合成装置の場合，標本化周波数を8KHzに選ぶ事が多いのであるが，この場合音声信号は，4KHz 以下としそれ以上の周波数は除去する必要がある。ここに、ローパス・フィルタを必要とする理由がある。人間の声のスペクトラム分布は，4KHz 以上は極めて少ない為に標本化周波数8KHz を選択したが，高忠実度再生，音楽等への使用を目的に標本化周波数を上げて商品化されている。

さて，標本化→量子化・符号化された信号が再生のため，DA コンバータで複号化された時は，**図9** と同様の波形と周波数スペクトルを有している。原信号にする為にはfa以上の成分を除去すれば良い訳で，前記同様のローパス・フィルタ，**図11**を通せば原信号が再生できる事になる。**図12**

以上が音声合成装置にローパス・フィルタを必要とする理由の説明であるが，いわゆる標本化定理「信号のスペクトラム分布の最高周波数の２倍以上の周波数で標本化を行なえば元の波形は完全に再現できる。」により，又標本化により生ずるエイリアス現象の除去の為にローパス・フィルタを必要とするのである。現実には，標本化周波数の0.4倍のしゃ断周波数特性を有するローパス・フィルタを使用するのが通常である。

3.2.2 ローパス・フィルタの設計

理想的なローパス・フィルタは，**図13**に示すようなものであるが現実には不可能である。

そこで，シャープな優れた特性を実現させる為，ＯＰアンプを利用したローパス・フィルタについて説明する。**図14**のような種類があるが，各々の特長を簡単に言い現わすと，まずバタワース型は通過帯域の振幅特性の平坦を重視したもので，ベッセル型は，位相特性（位相角を角周波数で微分した遅延特性）の良い波形の忠実伝送に優れたものである。又チェビシェフ型は，通過帯域の振幅特性を多少犠牲にし，帯域外のしゃ断特性を向上させるものである。

我々の音声合成装置では、６次のチェビシェフ型をリップルを極力小さく設計しており比較的バタワース型に近いチェビシェフ型ローパス・フィルタとなっている。

3.3 メモリ

音声信号をディジタル信号に変換された音声データは，符号化されて半導体メモリに記憶される。

音声合成用ＬSIの種類および扱い方により種々のメモリが使用可能となっている。

しかしながら基本的には，

再生のみならば ＲＯＭ

録音／再生ならば ＲＡＭ

ということになる。

又，ある種のＬSIにはメモリを内蔵しているタイプもすでに販売されている。

従って，用途，価格，スペース等を考慮して選定する必要がある。

ちなみに現在までに使用したメモリは，ＥＰＲＯＭ，ＳＲＡＭである。

3.4 音声合成用 ＬＳＩ

現在合成方式，ＲＯＭ外付・内付，マニアル制御・コンピュータ制御それに録音・再生等により，各メーカーから種々なＬＳＩが販売されている。ちなみに我社の製品をこのＬＳＩにより分類すると次のようになる。

ＡＤＰＣＭ方式（外付ＥＰＲＯＭ）

再生のみ ＶＯＴ－１０２４

ＶＯＴ－１５０１

ＶＯＴ－１００１

ＶＯＴ－８０Ｊ１

ＶＯＴ－８００２

ＡＤＰＣＭ方式（内蔵マスクＲＯＭ）

再生のみ ＶＯＴ－８００１

ＡＤＰＣＭ方式（外付ＳＲＡＭ）

録音／再生ＶＯＴ－１００２

ＡＤＭ方式（外付ＳＲＡＭ）

録音／再生ＶＯＴ－８００３

3.5 音声合成システムの今後

メモリの集積度増加により長時間の音声合成が可能になると共に，サンプリング周波数の向上で，より高音質化が今後の流れとなるであろう。また，それとは相反して１チップ化による周辺デバイスの省略化が増々進み，安価で少スペース化がより可能となるであろう。即ち，より高級化へ進むか，より軽薄短小化へと進むかのどちらかであろう。

これらの状況下で，常に市場ニーズを先取りした，ユーザーが納得する製品を開発していかなければならない。

4. おわりに

音声合成システム製品については，今までのカタログやマニアル等で既に紹介済みなので詳細については割愛した。

そのかわり，音声合成とは何かについて，音声発生の

しくみから順に，分り易く説明したつもりである。

これを機会に一人でも多く，音声合成に興味をもたれることを期待しておわりとする。

参考文献

1） 鈴木八十二編著，ディジタル音声合成器の設計，初版，産報出版，1982年7月，P. 9～P. 15

2） ＯＫＩ音声合成用ＬＳＩ技術資料，沖電気工業，1987年10月，P. 2～P. 4

3） 青木伸吾，音声録音再生ボードの製作，トランジスタ技術，ＣＱ出版，1987年9月，P. 488～P. 489