

声と音の技術戦略

森戸 誠

人は直立歩行によって変化した口腔と舌を使って声を使い分け、言葉を作り出した。音を利用して音楽を作り出した。今や声や音は人が生きていく過程の中で無くてはならない存在になっている。その使い方を分けると、図1に示すように論理的側面からの使い方と、感性的側面からの使い方がある。論理的側面とは声による情報の伝達や音によるアラーム、安全性の確保のような使い方である。一方、感性的側面とは、音楽を聞いて安らいだり、音楽を作って楽しむ、ライブを聞いてエキサイトするなど、感情、心理面を左右する使い方である。

産業革命以来の工業化社会、高度成長社会において、当社を含め多くの企業は、論理的な側面での音の技術開発に取り組み、生活のいたるところで利用されるに至った。社会が移り変わり、個の社会になったとき、電車内での音漏れ迷惑のように人の趣向がからみ、また、音が精神的なストレスの要因になるなど、ある種の公害のように感じられるようになってきた。これから我々がやらなければならないことは、論理/機能ばかりを追求するのではなく、感性的領域へ軸足を移し、人にやさしく、ゆとり、安らぎを持った声と音の追求ではないかと考える。

これから、当社における声と音の技術戦略を述べるが、

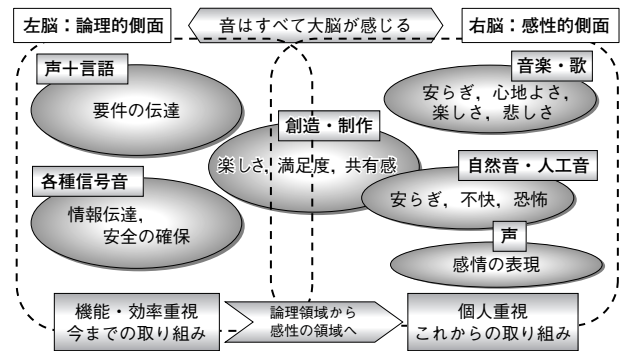


図1 音と声の使われ方

その前に、社会の移り変わりをコンピュータと通信の2つの観点から図2を用いて考える。

コンピュータの流れでは1990年代に入りパーソナルコンピュータが業務処理の中核を占めるようになった。それに伴い、文書作成、表計算など各種のソフトウェアが開発され、ビジネス、個人を問わずに使われてきた。2000年代では情報機器がますます小型化し、身につけて使われるようになる。身の周りにコンピュータが遍在するユビキタス社会の到来である。個人ツールとしてのウェ

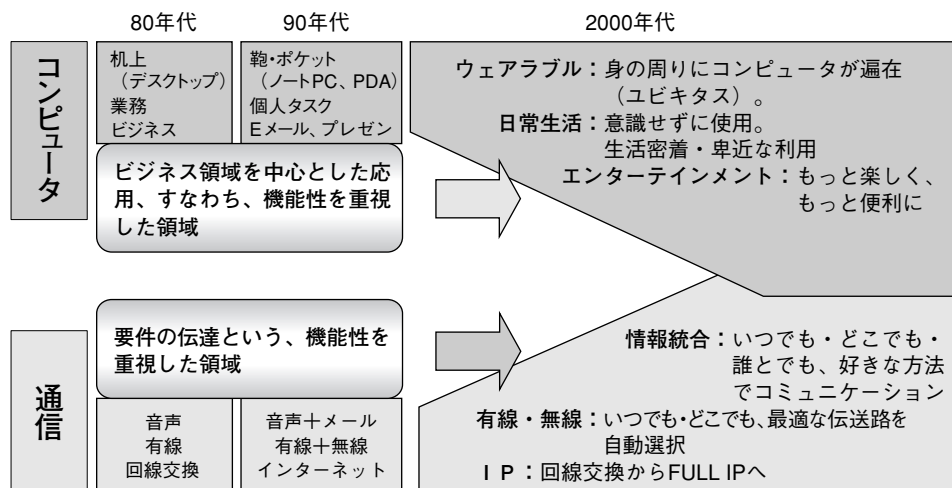


図2 コンピュータと通信の変遷

イトが高く、コンピュータを意識せずに生活の中で使う時代と言われている。音の領域で言えば音声認識や音声合成がヒューマンインタフェースの手段として、あたりまえのように使われる。また、超小型化が進み、可搬性に富んだ機器が開発され、エンターテインメント性に富む領域への使われ方もますます加速されるであろう。

一方、通信の流れでは、回線交換による電話やファックス、モデムによって代表されるように、要件や情報の伝達など、機能性を重視した領域での情報伝送が行われてきた。また、携帯電話の出現により移動性が大きく向上し、いつでも・どこでも・誰とでも、好きな方法でコミュニケーションを楽しむことが可能となった。

2000年代の大きな変化は、インターネットの大衆化と無線ネットワークの進展であろう。この時代の大きな特徴はモバイルブロードバンド通信である。コミュニケーションの不変要素は、なんとと言っても『聞く・見る』であり、ブロードバンド通信は、音と映像を使ったリアルなコミュニケーションを実現する。そこで大切なのは、単に伝送するといった機能面からだけでなく、やさしさ、安らぎを与えるといった感性面からの満足度を高めることではないかと考える。

次に、声と音の技術について、当社が事業を営む通信、LSI、情報の3つの領域に関して、これからの技術開発の方向を述べる。

広帯域音声符号化，オーディオ符号化

インターネットの進展は、従来の電話の世界にも大きな影響を与えた。通話距離に応じた送信者課金である交換回線電話のビジネスモデルは、インターネット電話に

よって崩れつつある。この分野で、当社では、いい音にこだわりながらインターネット電話の基になるVoIP (Voice over Internet Protocol) 技術を開発してきた。

一方、インターネット電話は新たな可能性を生み出している。従来の回線交換電話では300Hz以下、3.4KHz以上の帯域の音声は伝送されていなかった。この制約は音声を多重伝送するための制約項目だが、こもった声質を生む原因になっていた。

インターネット電話の世界では音声信号はパケット化されて送られる。そのため、従来では送ることができなかった広帯域の音声を送ることができる。もちろん、送り側も受け側も同様の処理を施さなければその効果は発揮されないが、感情・感動が相手に伝わっていく、臨場感ある音声通信が可能となる。また、声だけではなく音楽も伝送可能である。さらにはテレビ電話のように、画像のデータといっしょに送れば、よりリアルな親近感でのコミュニケーションが実現される。

また、いい音の伝送は ネットワーク上での音声認識の利用を加速する。従来の電話帯域での音声ではなかなか認識性能が向上しなかったが、広帯域化によりマイク入力と同様な認識性能が期待できるからである。

今後の技術開発として、

- 広帯域音声符号化，広帯域エコーキャンセラ技術
 - オーディオ符号化技術
 - 音と画像の融合技術
- を開発していくが、さらには、感性面から、
- やさしさ，安らぎを作り出す音響処理，符号化技術の検討を行う (図3)。

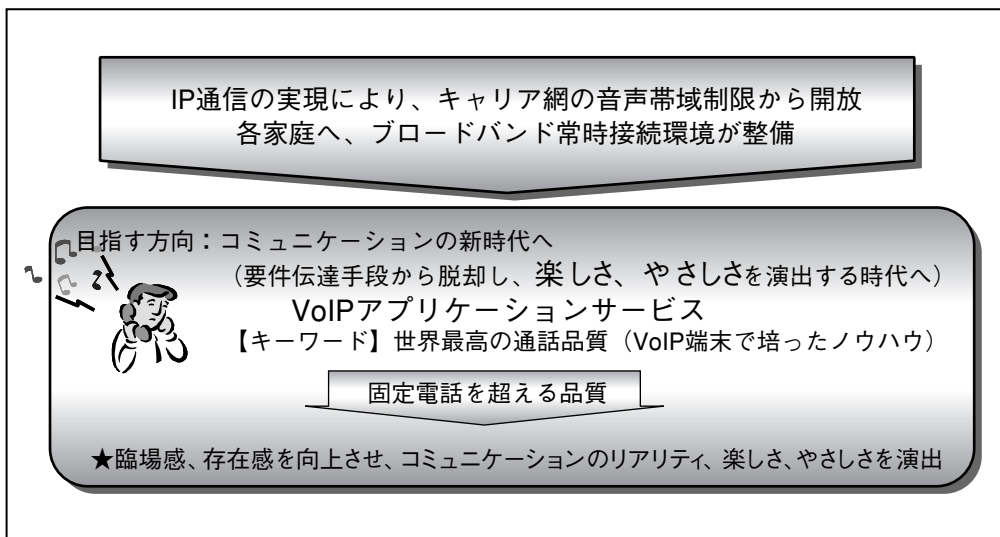


図3 広帯域音声符号化

マルチメディアLSI

音や音声に関するLSIの事業は音声LSIに始まる。この事業では音声出力方式にADPCM (Adaptive Differential Pulse Code Modulation) 方式を採用し、いい音にこだわった戦略をとった。ADPCM方式は波形符号化と呼ばれ、情報圧縮の点では他方式に劣るが波形そのものを符号化するため高品質の音声を出力することができた。メモリはどんどん安価で大容量化していく。そのため、音声の圧縮度より音質を重視した。また、ADPCM方式は音声以外（たとえば 動物の鳴き声、生活音など）の蓄積も可能なことから、情報機器以外のあらゆる生活用品に展開され、ビジネスの幅をさらに広げる結果になった。

次のステップでは、着メロ用の音源LSIへ事業を展開した。音源LSIの開発においても、いい音へのこだわりを継承させ、楽器音そのものを蓄積するPCM音源を採用し、リアルな楽音を実現した。

今、携帯電話の加入者数は全世界で12億加入に達し、個人ツールの位置付けを確立しつつある。単なる通信目的にとどまらず、インターネットと連携して多くのサービスが提供されている。近年ではデジタルカメラ機能が高機能化し、メガピクセル携帯電話と呼ばれるごとく100万画素の携帯電話も出現している。音楽再生機能を搭載し、単一なポータブル音楽再生機能も統合化されようとしている。通信環境におけるブロードバンド化がこ

の流れに拍車をかけているともいえる。

このように、通信以外の利用ニーズがますます高まっているため、携帯電話は通信の処理を行うLSIに加えて、マルチメディア処理などを行うLSIを搭載した2プロセッサ構成をとりはじめています。

当社では 音源機能を足がかりに楽器・声・音・画像を統合処理するマルチメディアLSIの領域に進んでいく。その中には、いい音にこだわった各種のマルチメディア処理、やさしさを与える音響処理を搭載しようと考えている（図4）。

声のソリューション素材

情報システム分野の事業はテキスト音声変換ソフトSMARTTALK^{TM*1)}に始まる。SMARTTALKTMも音質の良さにこだわり開発がなされ、メール読み上げなど多くのアプリケーションで利用されている。

一方、調査機関The Radicati Group, Inc.によれば音声を用いた音声サービスは2006年に1億5400万人の利用者、659百ドルの市場があると報告されている。当社のこれからの方向は、図5に示すように、音声サービス事業者に対して満足度の高い『声のソリューション素材』を提供することを目指している。現状のテキスト音声変換（TTS）ソフトでは声で情報を伝える論理的側面からの機能は実現されているが、利用したくなる、聞きたくなるといった感性的側面からの音質の追求が不十分である。よ

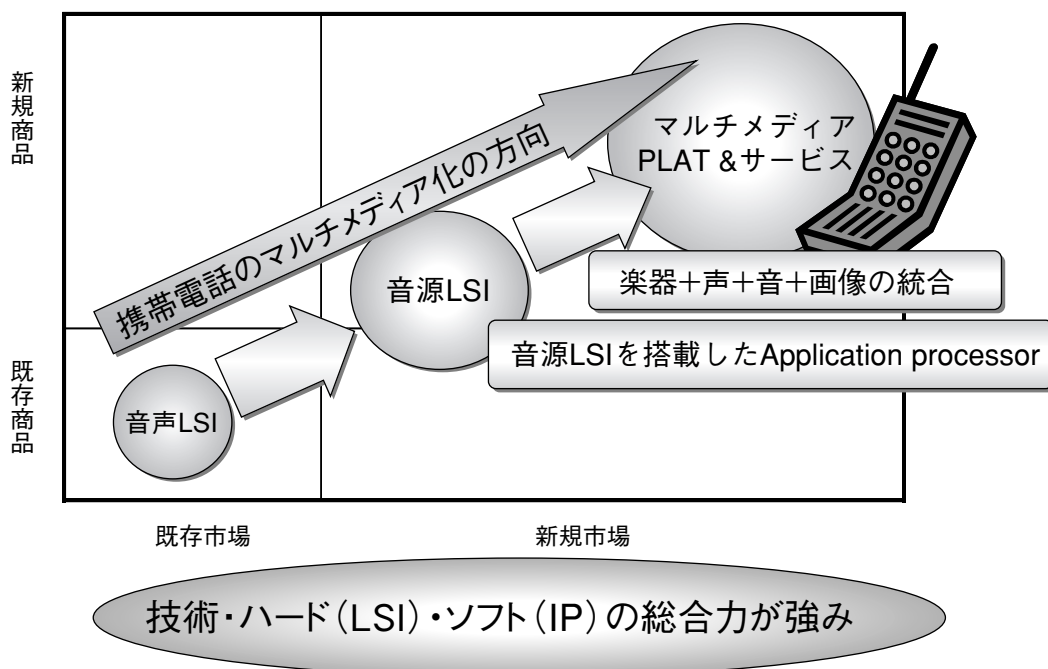


図4 マルチメディアLSI

*1) SMARTTALKは沖電気工業(株)の登録商標です。

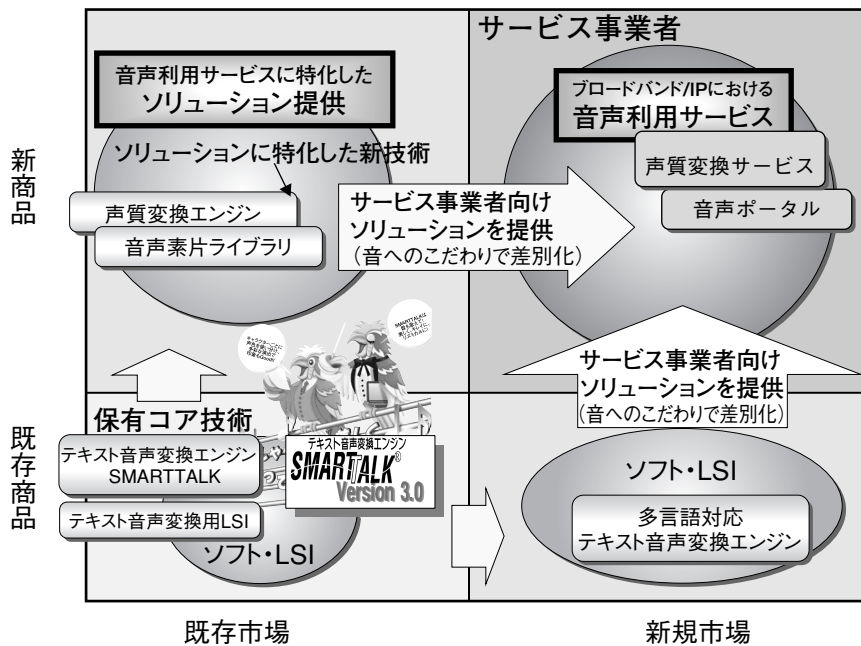


図5 音のソリューション素材

り多くの音声サービスへの適用を考えるならば、

- 多くの声質
- 声そのものに価値を与えるべく、特定のタレントやキャラクターの声質
- より多くの多様性を生むためにイントネーションの多様化
- 人間らしさを表現するために、泣き、笑い、怒り、楽しさの表現
- グローバル化を踏まえマルチリンガル対応を実現する必要があると考える。

音と声のホームページ

『いい音、いい声』にこだわりながら技術開発を進めていくが、開発した技術をいち早くみなさんのソリューションづくりに役立てていただくために、声と音の技術をわかりやすく解説した『声と音の技術』ページ(図6)を開設した。

当社のトップページ

<http://www.oki.com/jp/>

から容易にアクセスでき、代表的なVoIP技術、ADPCM技術、テキスト音声変換技術についてやさしく解説している。

声と音の技術に関し、従来から培ってきた『いい音、いい声』を、機能面からだけでなく、感性面・心理面からも追求し、『楽しさ、安らぎ、豊かさ』を実現する豊か



図6 音のホームページ

なコミュニケーションの担い手になる、というのが当社の技術開発の方向である。◆◆

●筆者紹介

森戸 誠 : Makoto Morito. 研究開発本部 ヒューマンインタフェースラボラトリ ラボラトリマネージャ