

ネットワークを支えるマイコン用ミドルウェア

—テレビ電話システムにおける例—

SuperH RISC Microcomputer Middleware for Network Systems

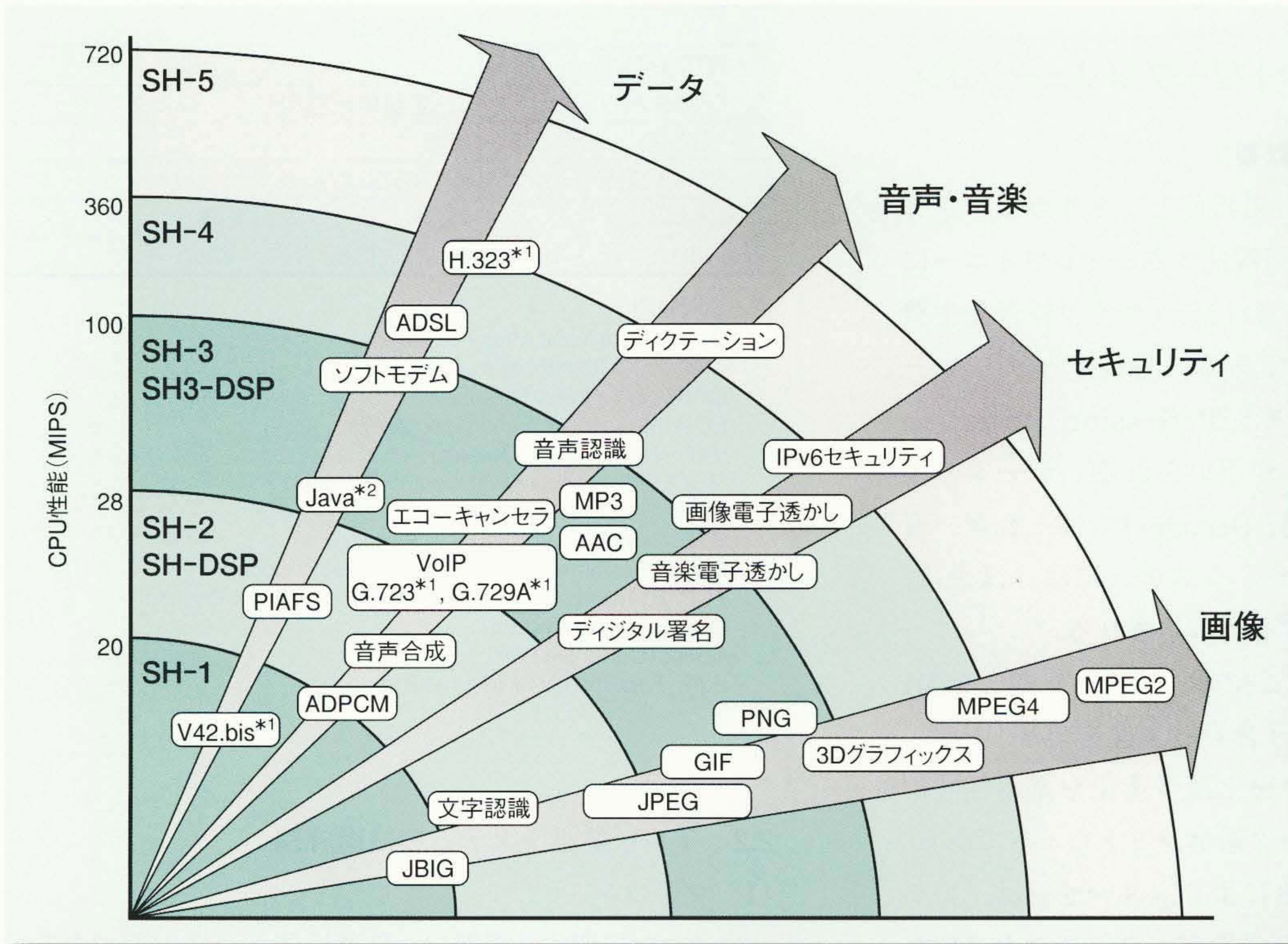
大場信弥 Shin'ya Ōba

富永浩安 Hiroyasu Tominaga

新村 篤 Atsushi Niimura

小島 昇 Noboru Kojima

久保英二 Eiji Kubo



注：略語説明ほか

CPU (Central Processing Unit)
MIPS (Million Instructions per Second)

ADSL (Asynchronous Digital Subscriber Line)

PIAFS (PHS Internet Access Forum Standard)

VoIP (Voice over Internet Protocol)

ADPCM (Adaptive Differential Pulse Code Modulation)

MP3 (MPEG1 (Moving Picture Expert Group) Audio Layer 3)

AAC (Adaptive Audio Coding)

JBIG (Joint Bi-Level Image Expert Group)

JPEG (Joint Photographic Expert Group)

GIF (Graphics Interchange Format)

PNG (Portable Network Graphics)

*1 国際標準化方式

*2 JavaおよびすべてのJava関連の商標およびロゴは、米国およびその他の国における米国Sun Microsystems, Inc.の商標または登録商標である。

ミドルウェアの製品展開

日立製作所は、SuperH RISCマイコン用アプリケーションソフトウェアであるミドルウェアにおいて、現在、図に掲げるものをサポートしている。

日立製作所のSuperHシリーズRISCマイコンのうち、デジタル信号処理機能を強化した“SH-DSP”用に、テレビ電話システムを支えるミドルウェアを開発した。関連するミドルウェアは、プロトコル、音声圧縮、エコーキャンセラ、画像圧縮、サウンド(音楽)など多岐にわたる。

このミドルウェアは、高速処理、低消費電力、豊富な品ぞろえ、各種OSへの対応などを特徴としている。ソリューションエンジンの開発環境による評価では、音声圧縮の最大の課題である遅延時間が150 ms以下と良好な特性を得た。

このソフトウェアとシステムにより、いつでも、どこでも、さまざまな人やデータにアクセスすることができるようになる。

1 はじめに

近年、ISDN (Integrated Services Digital Network) やxDSL (x Digital Subscriber Line) および光ファイバの普及により、家庭への通信回線のブロードバンド(広帯域)化が急速に進展してきている。また、携帯電話のiモード利用者が3,000万人を超え、各種サービスが充実しつつある。これに伴い、画像や音声、サウンド(音楽)などさまざまなフォーマットのデータを、インターネットや

LANなどのネットワークを介してやり取りできるようにするシステムソリューションが求められている。

このニーズにこたえるシステムの代表例として、テレビ電話やそれに用いられるプロトコル“H.323”がよく知られている。しかし、このシステムでも、ハードウェアやソフトウェア、インタフェースの各技術、システムインテグレート技術、評価技術、低価格化技術など、課題はまだ多い。一方、最近のCPU (Central Processing Unit) 性能の進歩が目覚ましく、従来、複数のCPUが必要で

あったが、CPU1個でH.323を扱えるようになってきた。
ここでは、ネットワークシステムの代表としてテレビ電話を取り上げ、これを支えるソフトウェア技術、とりわけ、パッケージ化された応用ソフトウェアであるマイコン用ミドルウェアについて述べる。

2 マルチメディアネットワークソリューション

2.1 テレビ電話システムの概要¹⁾

テレビ会議などで用いられる機能とミドルウェアをまとめたものを表1に示す。電話機能であるテレフォニーに用いられるミドルウェアは、DTMF(ダイヤル音発生機能)、CID(発ID通知)などである。各種データフォーマットの調整には、H.323やSIP(Session Initiation Protocol)などのプロトコルが用いられる。データの圧縮・伸張を行うCodec(Coder-Decoder)では、音声、音楽、画像の各メディアを扱う。システムには、ほかに、ファクシミリ用ソフトウェアも組み込まれる。
プロトコルの一つであるH.323のソフトウェア構成を図1に示す。同図中、左上は扱う音声と画像を処理(圧縮・伸張)するミドルウェア、中央上は多重化や電話発呼用の、右はテレビ会議用のデータ通信ソフトウェアである。
このシステムを用いることにより、ユーザーは、いつでも、どこでも、さまざまな人やデータにアクセスできることになる。これは昔からの人類の夢であり、その一つが現実になえられることによるユーザーメリットは大きい。これを支えるための課題について以下に述べる。

表1 テレビ会議用機能とミドルウェア
テレビ会議システムには、ほとんどのソフトウェア要素技術が含まれている。ここでは、カメラ関連のソフトウェアを省いている。

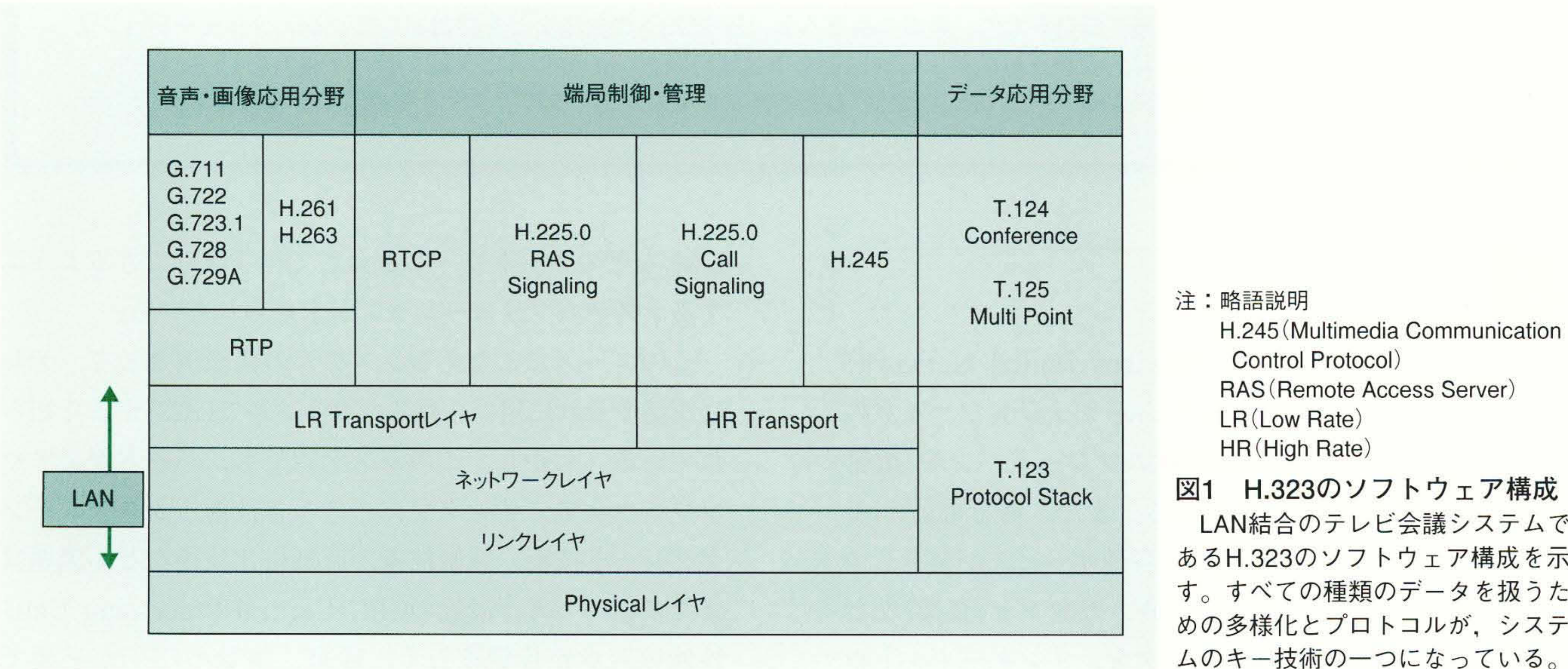
機 能	ミドルウェアほか
テレフォニー	DTMF, CID, CP, EC, VAD
発呼制御プロトコル	H.323 MGCP/SIP
RTP/RTCP	—
ジッタバッファ	日立製作所独自
TCP/IP	OS
デバイスドライバ	システムボード
Codec	音声, サウンド, 画像
FoIP	T.37, T38, V17, V34

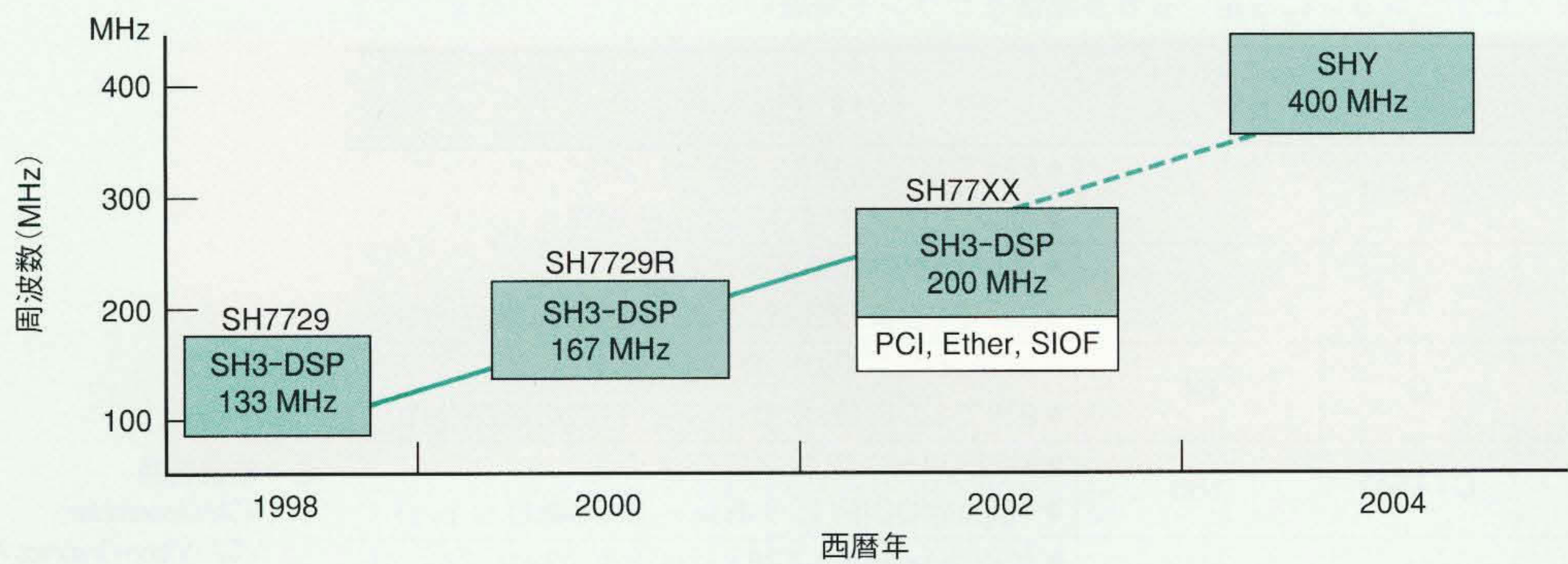
注：略語説明
DTMF (Dual-Tone Multi-Frequency)
CID (Caller Identification Tone)
CP (Call Progress Tone)
EC (Line/Acoustic Echo Canceller)
VAD (Voice Activity Detector)
MGCP (Multimedia Gate Control Protocol)
SIP (Session Initiation Protocol)
RTP (Realtime Transmission Protocol)
RTCP (Realtime Transmission Control Protocol)
TCP/IP (Transmission Control Protocol/Internet Protocol)
OS (Operating System)
Codec (Coder-Deoder)
FoIP (Facsimile over Internet Protocol)

2.2 テレビ電話システムの技術課題

(1) マイコン

マイコンに関する課題は、処理性能と消費電力である。
日立製作所のSH(SuperH RISC)-DSPの展開を図2に示す)。SH-DSPマイコンは、16ビット固定長の命令体系を持っている。これは、低消費電力、高速処理性能、高コード効率などを特徴としているSHマイコンに、DSP





注：略語説明
PCI (Peripheral Component Interconnection)
SIOF (Serial I/O with FIFO)

図2 SH-DSPの展開計画
メディアを処理するマイコンの代表であるSH3-DSPの高周波数計画を示す。周波数を上げることで、ソフトウェアを変更することなしに性能向上が図れる。

(デジタル信号処理)回路を内蔵したものである。1998年には133 MHzであった駆動周波数が、2002年までには通信用機能を集積化しつつ、200 MHzまで向上する。この高周波数化はソフトウェアを変えずにシステムの性能向上を図れることを意味しており、ユーザーの製品設計には重要な点である。これにより、従来処理できなかったテレビ電話などを1チップで構成できることになる。

(2) ミドルウェア

ミドルウェアの課題は、(1) 高速処理、(2) 低メモリサイズ、(3) 使い勝手のよいインタフェース、(4) 各種OSへの対応などである。

日立製作所のミドルウェアの特徴は、まず、アセンブラコードを用いてSH-DSP用にチューニングしてあるので、高速処理ができる点にある。したがって、処理時間が短縮でき、マイコンの低消費電力化が容易となる。また、iTRON、Windows CE^{※1)}、VxWorks^{※2)}、Linux^{※3)}など各種OSに対応している。豊富な品ぞろえを誇り、ソリューションエンジンによる充実した開発環境を整えている(表2参照)。ほかに、H.323までのトータルシステムサポートなども備えている。

ミドルウェアの機能のうち、特に重要な音声圧縮とエコーキャンセラについて以下に述べる。

表2 SH-DSP用音声圧縮ミドルウェア

今後の進展が期待されているVoIP (Voice over Internet Protocol) 向けミドルウェアのラインアップを示す。現在、ほとんどのミドルウェアがそろえられている。

ミドルウェア	内 容	ROM(kバイト)	MIPS
G.711	音声圧縮伸張A/μ-law	1	1
G.721	音声圧縮伸張ADPCM	3	11
G.723.1	音声圧縮伸張ACELP/MP-MLQ	57	27
G.726	音声圧縮伸張ADPCM	4	18
G.728	音声圧縮伸張LD-CELP	—	45
G.729A	音声圧縮伸張LC/CS-ACELP	36	15
G.165	ラインエコーキャンセラ	5	15
G.167	音響エコーキャンセラ	25	20

注：略語説明
ROM (Read-Only Memory)
MIPS (Million Instructions per Second)
ADPCM (Adaptive Differential Pulse Code Modulation)
ACELP (Algebraic Code-Excited Linear Prediction)
MP-MLQ (Multi Puse-Multi Level Quantization)
LD-CELP (Low Delay Code Excited Linear Prediction)
LC (Low Complexity)
CS (Conjugate Structure)

3 システムを支えるミドルウェア

3.1 音声圧縮・伸張ミドルウェア

SH-DSP用に開発、チューニングした音声圧縮・伸張のミドルウェアを表3に示す。音声圧縮・伸張の基本アルゴリズムとしては、波形符号化とハイブリッド符号化が知られている。前者はG.711[PCM (Pulse Code Modulation)]やG.726[AD (Adaptive Differential) PCM]であり、前のフレームとの差分を符号化して伝送する方式である。G.723.1 (MP-MLQ)、G.728[LD-CELP (Low Delay Code Excited Linear Prediction)]、G.729A (CS-ACELP)などは後者に属し、音源をパルス

※1) Windowsは、米国およびその他の国における米国Microsoft Corp.の登録商標である。
※2) VxWorksは、ウインドリバーシステムズの登録商標である。
※3) Linuxは、Linus Torvaldsの米国およびその他の国における登録商標あるいは商標である。

表3 SH-DSP向け画像，サウンド関係ミドルウェアの概要
テレビ会議のほか、携帯電話やナビゲーションにも用いられる画像とサウンド関連のミドルウェアの概要を示す。

ミドルウェア	コードサイズ (kバイト)	言語	ワークエリア (kバイト)	備 考
JPEG/Codec	14	C+ASM	4	●2秒で1コマ〔100×50ドット(24ビット)〕 ●サンプリング 4：2：2，圧縮率： $\frac{1}{10}$
PNGデコーダ	110	C	64	●VGA〔640×480ドット(24ビット)〕 ●2秒で1コマ(表示サイズ，色数の最適化中)
GIFデコーダ	5	C	64	●VGA〔640×480ドット(24ビット)〕 ●2秒で1コマ(表示サイズ，色数の最適化中)
MPEG4 Viewer	90	C+ASM	180	●MPEG4シンプルフロファイル ●10fps@QCIF(ビットレート：64kビット/s)
MP3デコーダ	27	C+ASM	16.3	●ビットレート：128 kビット/s、32サブバンド ●サンプリング：44.1 kHz，ステレオ
AACデコーダ	—	C	—	●DSP命令を使用した最適化中

注：略語説明
ASM (Assembler)
VGA (Video Graphics Array)
fps (Frame per Second)
QCIF (Quarter-Common Intermediate Format)

の振幅や位置情報と雑音モデルで構成している。後者のほうが圧縮率は高い。

音声圧縮・伸張では、音声処理のために生じる遅延時間に注意する必要がある。パソコンをベースとしたテレビ会議システムに用いられているG.723.1では、37.5 ms (フレーム30 ms+先読み7.5 ms)の遅延時間が生じる。G.729Aは、G.723.1に比べてマイコンの負荷が軽く、音質特性がよいので、最近注目されている。この場合の遅延時間は15 ms (フレーム10 ms+先読み5 ms)になる。

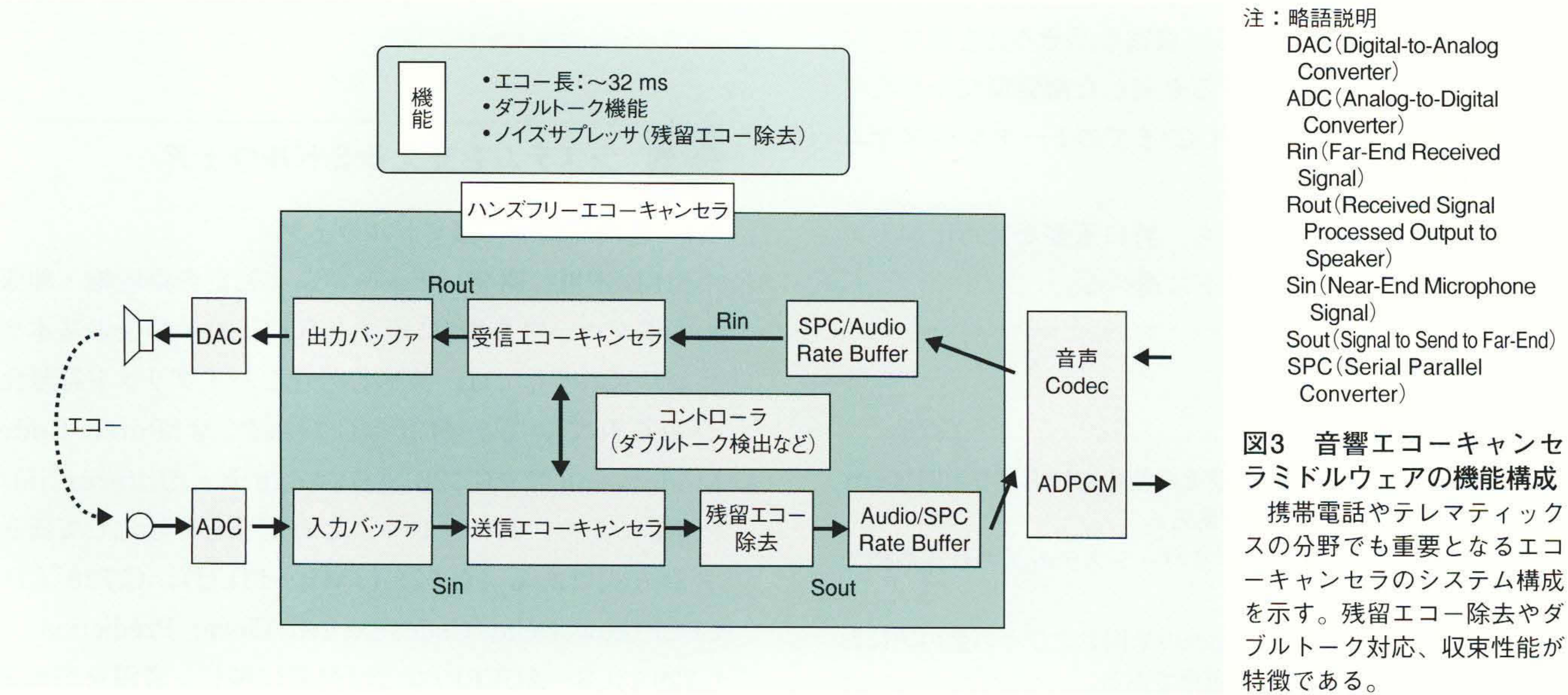
3.2 エコーキャンセラ

エコーキャンセラには、電話回線でのエコーを抑圧するライン エコー キャンセラ (G.165) と、電話端末のスピーカからマイクロホンへの漏れ込みを消去する音響エコ

ーキャンセラ (G.167) がある。最近では、ハンズフリースystemに対応するため、音響エコーキャンセラが注目されている。

音響エコーキャンセラのシステム構成を図3に示す。音響エコーキャンセラは、受信信号を逐一遅延させ、ゲインを調整し、送信信号と差分をとることで、エコーによる雑音が最小になるように動作する。このとき、収束時間、安定性、タップ数や係数の設定などが重要である。最終的には製品に近い形で評価し、ミドルウェアのパラメータを決めることになる。

音響エコーキャンセラのミドルウェアの主な仕様を表4に示す。日立製作所のエコー キャンセラ ミドルウェアには、送受者が同時に発話した場合のダブルトーク対策



注：略語説明
DAC (Digital-to-Analog Converter)
ADC (Analog-to-Digital Converter)
Rin (Far-End Received Signal)
Rout (Received Signal Processed Output to Speaker)
Sin (Near-End Microphone Signal)
Sout (Signal to Send to Far-End)
SPC (Serial Parallel Converter)

図3 音響エコーキャンセラミドルウェアの機能構成
携帯電話やテレマティックスの分野でも重要となるエコーキャンセラのシステム構成を示す。残留エコー除去やダブルトーク対応、収束性能が特徴である。

表4 エコー キャンセラ ミドルウェアの仕様
エコー キャンセラ ミドルウェアの仕様例を示す。自動車や電話で使える仕様にしてある。CPU負荷はタップ数(遅延時間)に依存する。

仕 様	ITU G.167
CPU	SH2-DSP, SH3-DSP, SH-4
サンプルレート	8 kHz 帯域：300～3,400 Hz
タップ数	256, 512
ブロック長	64 サンプル
コードサイズ(ROM)	12 kバイト
ワークエリア(RAM)	16 kバイト
MIPS	15～20 MIPS(タップ数に依存)
使用内蔵メモリ	4.0 kバイト：X-RAM 4.5 kバイト：Y-RAM

注：略語説明 ITU(国際電気通信連合)

や、背景雑音を抑圧するための雑音抑圧機能などを組み込んでいる。これにより、サービスの品質が高く維持できる。

3.3 動画像再生MPEG4

デジタル スチル カメラに用いられている静止画圧縮のJPEG(Joint Photographic Expert Group)のほか、最近ではMPEG4(Moving Picture Expert Group 4)による動画像が注目されている。DVD(Digital Versatile Disc)並みの画像はまだ無理としても、携帯電話向けの画面で動画を再生することが一つのマイコンで可能になってきたからである。SH-DSPでは、VGA(Video Graphics Array)サイズで每秒5～10フレーム、携帯電話などで用いられているQCIF(Quarter-Common Intermediate Format)サイズで每秒25フレームの性能が得られる。

3.4 サウンド(音楽)

MP3(MPEG1 Audio Layer 3)やAAC(Advanced Audio Coding)で圧縮した音楽を携帯端末で楽しむ人々が増えており、テレビ電話でもサウンドの効果は大きい。聴くだけのデコーダなら、約30 MHzのCPU負荷で音楽を楽しむことができる。エンコーダでは100 MHz程度の負荷になる。

4 システム性能評価結果

4.1 ソリューションエンジン

ソリューションエンジンの外観を図4に示す。これは、マザーボードとドータ(増設)ボードから構成しており、それらの組合せを変えることにより、各種マイコンと応用分野に対応することができる開発環境である。OSとし

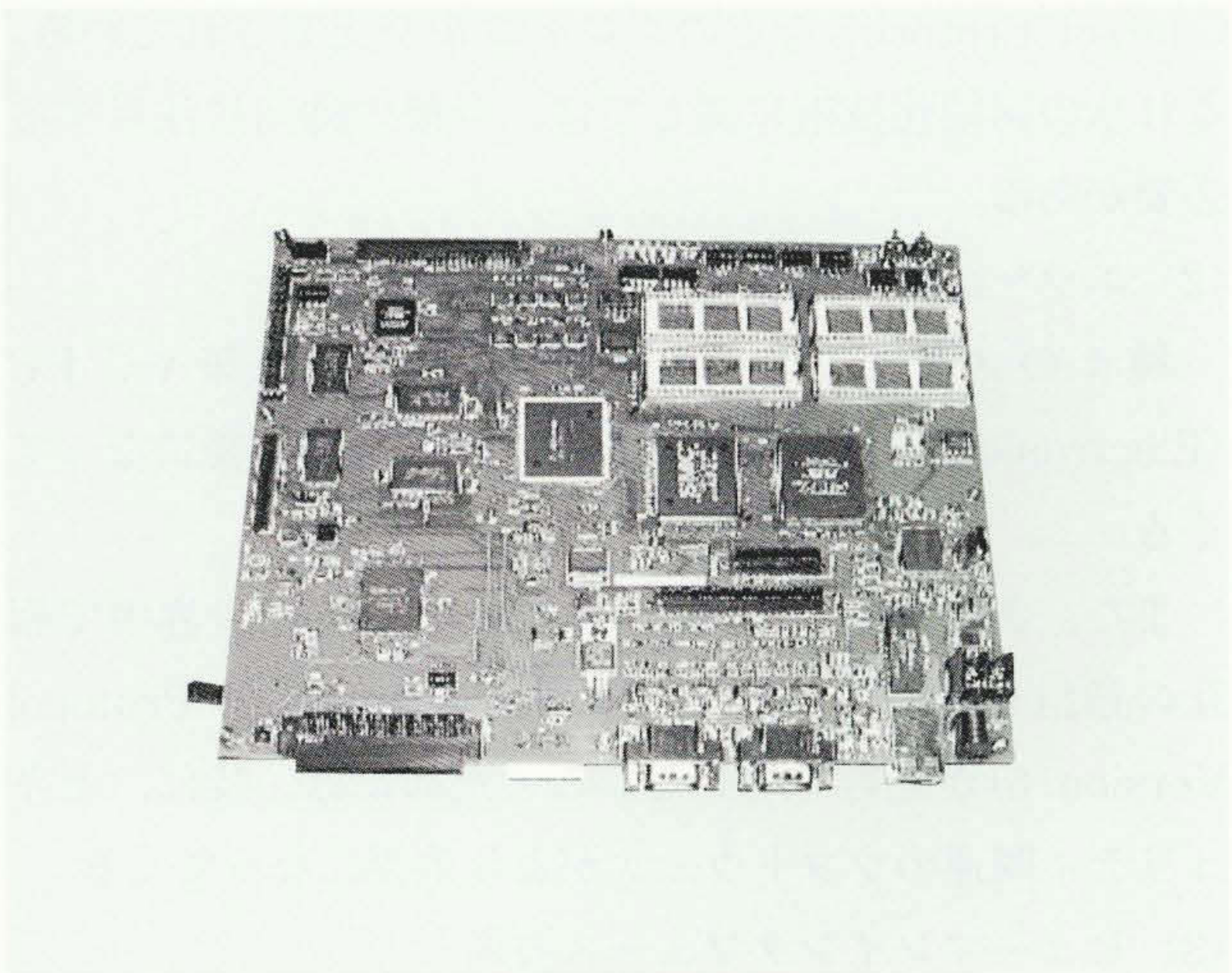


図4 ソリューションエンジンの外観
ユーザーのソフトウェア開発やデバッグになくてはならない開発環境のシステムボードの外観を示す。マザーボードを交換すればさまざまなSHマイコンに対応でき、また、ドータボードの交換により、適用分野を選ぶことができる。

表5 音声コーデックの測定結果

H.323で実際に評価した音声圧縮ミドルウェアの遅延時間の測定値を示す。遅延時間が大きいと音質が極端に劣化する。測定値はいずれも、許容値の200～250 ms以内となっている。

コーデック	遅延時間
G.729A	90 ms
G.723.1	110 ms
G.711(μ-law)	90 ms

てはiTRON, VxWorks, およびLinuxに対応している。

4.2 VoIPでの性能

ソリューションエンジンに搭載したSH3-DSPを用いて、音声圧縮・伸張ミドルウェアを模擬的にネットワークで伝送させた場合の遅延特性を表5に示す。ここでの遅延時間は、ネットワークの揺らぎ(ジッタ)がない場合の特性である。いずれの場合も、遅延時間は許容値の200～250 ms以内になっている。

5 おわりに

ここでは、テレビ電話を例に、ネットワークを支えるミドルウェアに使われるソフトウェアの課題と技術について述べた。

さらに豊かで優しく、アクセスが容易なネットワーク時代を創生するための今後の課題は以下のとおりである。

- (1) 画像、音声、サウンド(音楽)の同期化技術
H.323以外にも、SIPやMGCP(Multimedia Gate

Control Protocol)などのプロトコルが提案されている。これらの同期化技術に関しては、今後の動向に注目する必要がある。

(2) エラー訂正技術とセキュリティ技術

種々のメディアが欠落した場合の対応策や、EC (Electronic Commerce) などへの対応策が必要になってくる。

また、あらゆる端末にアドレスをつけて、世界中で相互の通信を容易にするために、IPv6 (Internet Protocol Version 6) が進行している。このためには、特に、セキュリティ関連のソフトウェアがより重要になってくる。

(3) ヒューマンインタフェース技術

複数の話者の中から特定の人物の声を強調して処理する方法や、話している人の位置を同定することにより、テレビ電話がさらに使いやすいものになることが求められている。

(4) 音声分析・合成技術

任意の声色で音声を合成したり、人物を特定する技術が重要なものになる。

参考文献

- 1) 北脇編：デジタル音声・オーディオ技術，電気通信協会，オーム社(1999. 12)
- 2) 藤原監修：画像&音声圧縮技術のすべて，CQ出版社(2000. 4)
- 3) 守谷：音声符号化，電子情報通信学会編，コロナ社(1998. 10)
- 4) 三木：MPEG4のすべて，工業調査会(1998. 9)

執筆者紹介



大場信弥

1971年日立製作所入社，半導体グループ 所属
現在，半導体グループでミドルウェアの製品化に従事
E-mail: ob-shinya @ sic. hitachi. co. jp



小島 昇

1977年日立製作所入社，半導体グループ カスタムビジネスユニット 所属
現在，マイコン用ネットワークのマーケティングに従事
E-mail: kojima-noboru @ sic. hitachi. co. jp



富永浩安

1984年日立製作所入社，半導体グループ モバイルネットワークビジネスユニット 所属
現在，通信システム用マイコンのマーケティングに従事
E-mail: tominaga-hiroyasu @ sic. hitachi. co. jp



久保英二

1977年日立製作所入社，半導体グループ ミドルソフト開発部 所属
現在，半導体グループでSHマイコン用通信ソフトウェアの製品化に従事
E-mail: kubo-eiji @ sic. hitachi. co. jp



新村 篤

1980年日立通信システム株式会社入社，IT事業センタ 所属
現在，インターネットプロトコル関連ソフトウェアの製品開発に従事
E-mail: atsushi-niimura @ hitachi-cs. co. jp