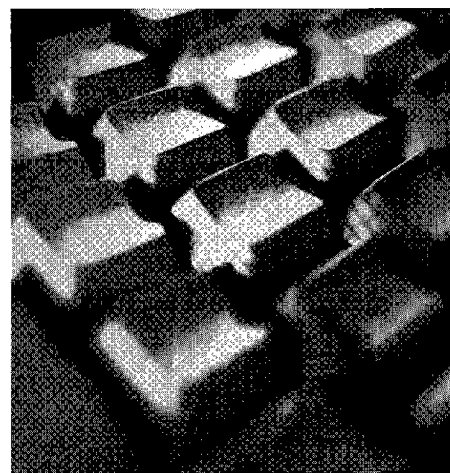


マルチメディア ホームコンピューティングの未来



—第3回 マルチメディアホームコンピュータ の考え方—

釜江 尚彦

(財) イメージ情報科学研究所
kamae@tokyo.image-lab.or.jp

【ゲームコンピュータの現状】

前々回（本年7月号）、パソコンは家庭の情報化の主役足り得るかで、パソコンの現状をある程度述べた¹⁾。ここではゲームコンピュータの現状を眺めてみよう。

ゲームコンピュータはどうか。ゲームコンピュータはCPUの他にマルチメディア処理専用のDSPを持っている。このDSPは積和を非常に高速に行う。たとえばセガのドリームキャストに用いられているSH4は図-1に示すように128ビットのレジスタ

に入った4つの32ビット浮動小数点の数値が4つの浮動小数点の掛け算器（FMUL）で対応する2数値が掛け合わされ、その結果がFADDで積算されるようになっている。これを200MHzのクロックで実行し、1クロックで4つの掛け算、3つの足し算を実行する。したがって性能は1.4GFLOPSとなる。典型的な3Dグラフィックスのジオメトリ演算は4×4のマトリクスと4×1のマトリクスの掛け算であり、それが4クロックで実行できるようになっている²⁾。

一方プレイステーション2ではCPUとDSPを合わせたエモーションエンジンで6.25GFLOPSの性能を持っているといわれている。

3Dグラフィックスではジオメトリ処理のあと、映像として完成させるためレンダリングと呼ばれる処理があるが、これはジオメトリ処理に比べて扱う数値のダイナミックレンジが小さく、16ビットの整数演算で間に合うことが多いのでハードウェアとして独立させることが多い。むしろレンダリング処理ではメモリとの間のデータ転送量がネックとなることが多く、それを軽減させる工夫がなされている。音響処理については種々の音源や音響効果のような特殊な演算が多く、ハードとして独立させることも多い。入出力についてはCDやDVD、ゲーム用のインタフェース機器などの入出力処理を必要とする。

このようにゲームコンピュータはパソコンに比べてCPUを小さくし、マルチメディア演算をDSPなど別の

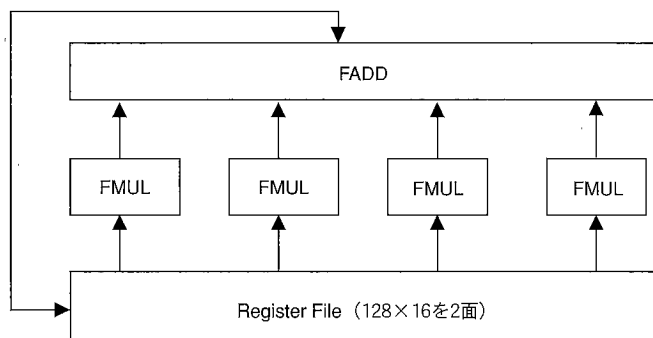


図-1 SH4のDSP演算

| 機種 | 用途 | DSP |
|--------|------------------------------|------------------------------------|
| スタンダード | 電子機器、カラオケなどリアルタイム演奏の残響付加はしない | 32ビット浮動3.5GFLOPS 16ビット整数10GOPS |
| 生演奏用 | 生演奏に残響付加を行う（2スピーカ） | 32ビット浮動23.5GFLOPS 16ビット整数10GOPS |

表-1 マルチメディアホームコンピュータの種別

プロセッサで演算させるコプロセッサ形式を採用しており、家庭におけるエンタテインメントにより適した構成となっている。しかしその欠点は主たるベンダのコンピュータに互換性がなく、家庭用のコンピュータプラットフォームになりきれていないことである。ゲームコンピュータのもう1つの欠点はエンタテインメントに適しているとはいえ、3Dグラフィクスに焦点を当てすぎており、音声/音響などのオーディオ、さらにはデジタルカメラを中心とした静止画像などを扱ったり、囲碁、将棋といったスタティックなエンタテインメントとのバランスを欠いていることである。さらに全体に若者に焦点を合わせすぎている。

【家庭用コンピュータプラットフォーム】

●条件

家庭に入るコンピュータ（ホームコンピュータ）の条件を整理してみよう。

a) プラットフォームとしての信頼感
メーカー間互換性、世代間互換性などの互換性の確保、ハードウェアやソフトウェアの世代交代の周期を家電並みにするなどの安定感の付与、ソフトウェアハウスやコンテンツプロバイダがプラットフォーム化への協力をしやすい環境の整備など。

b) 使い勝手のよさ

ログインに時間がかからず、しか

も統一した思想で設計されたヒューマンインフェース、誤操作に強いヒューマンインタフェース、置き場所に対する制約の小さい形状とサイズ、AV機器などとの接続の容易さ、ネットワークへの接続と立ち上げの容易さ。特に高齢者が大きなストレスなしに使えるインタフェースが大切である。

c) エンタテインメントの多彩さ

家庭での利用の相当部分がエンタテインメントになる。コンピュータゲームだけでなく幅広いエンタテインメントがサポートできることが大切である。

d) コストと消費電力

エンタテインメントに利用するとすると専有時間が長い。1つの家庭に複数台設置することができるコストにすること。さらに複数台設置できる大きさや消費電力も大切である。

e) 幅広い利用者層への配慮

人口の25%が高齢者になる時期がせまっている。利用者として高齢者も視野に入れたコンセプトが大切である。特に高齢者用に音声インタフェースやマルチモーダルインタフェースの採用が大切である。

●マルチメディアホームコンピュータプラットフォームコンセプト

パソコンとゲームコンピュータの現状を考えたときマルチメディアホームコンピュータのコンセプトは次

のようなものになるろう。

- マルチメディアの向いているゲームコンピュータに使われている技術を中心にしたコンピュータとする。
- マルチメディア機能は映像、3Dグラフィクス、音響、音声でバランスのとれたものとする。
- メーカー間互換性を確保し、ホームコンピュータプラットフォームとして安定性のあるものとする。
- しかし技術進歩には遅れないよう4~5年でバージョンアップを図る。過去との互換性は確保するが、そのため複雑性が増しすぎないように適当な段階で切るべきものは切る。
- 周辺機器、ソフトウェアメーカーの参入を促すため、インタフェースなどはできるだけオープンにする。
- プラグアンドプレイを原則とし、周辺機器、ソフトウェア、ネットワークなどのインストールは分かりやすいやり方を採用する。
- 他のアプライアンスとホームネットワークなどで相互接続するのを容易にし、協調動作がやりやすい構造とする。
- 環境を考慮しバージョンアップしても使える部分は使えるよう配慮する。さらにパワーセーピング技術をできるだけ採用する。

ハードウェア

(1) コプロセッサ形式

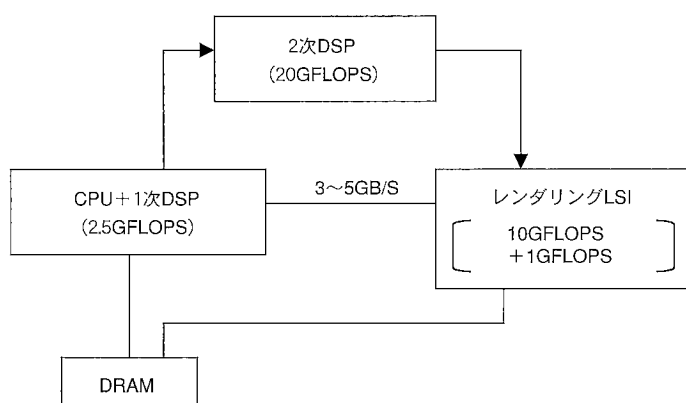


図-2 マルチメディアホームコンピュータのハードウェア構成

マルチメディアホームコンピュータに必要なハードウェア構成はコプロセッサ形式とする。DSPはグラフィックスのジオメトリ計算に必要な32ビットの浮動小数点用のものとグラフィックスのレンダリング用の16ビット整数演算用のものの2種が欲しい。音響計算についてはたたみ込み演算を時間軸で行うときには1スピーカーあたり15GFLOPSあれば十分であるが、非常に粗い推定なので精細につめることにより10GFLOPSまでは下げられよう(単純に残響時間を2秒程度に減じることで10GFLOPSとなる)。したがって2スピーカーで20GFLOPSの能力が要求されることになる。こうするとグラフィックスジオメトリとレンダリングに必要な1.5+9GFLOPSは同時に行わない限りクリアしていることになる。すでに述べたように遅延が許されるレコーディングされたオーディオではここまで処理能力を必要としない。こう考えるとハードウェア構成は大きく表-1のように2種類とすることが考えられる。

(2) バス

3Dグラフィックスのレンダリング処理には毎秒8Gバイトのデータ処理能力が必要であると推定した。一

方生演奏用のたたみ込み演算ではDSPの要求性能は高いが必要とするデータ転送は、残響データをローカルに持てば50KHzでサンプリングした音を2チャンネル分だけで済む。その代わり3秒分の残響データと音のデータを2チャンネル分、すなわち4バイト×150K×2×2=24Mバイトのバッファが必要である。8Gバイト/秒のバスは結構大変な技術である。インテルのAGPの最新バージョンが1Gバイト/秒、プレイステーション2が1.2Gバイト/秒であることからみても8Gバイト/秒の要求能力が最もきびしい。前回の表-2においてテクスチャマッピング用のデータ転送量が多いのでこれをローカルでバッファリングすることにすれば要求能力は約半分になる。しかもこれをデータ圧縮して転送することも考え、一方データ転送能力はプレイステーション2の4倍程度の5Gバイト/秒を最大とし、当面3~5Gバイト/秒とする。

(3) ハード構成

するとハードウェアは図-2のような構成になろう。ここに1次DSPとはCPUチップに内蔵するDSPで2.5GFLOPS程度のDSP内蔵はできよう。レンダリングLSIはテクスチャ

のローカルバッファを含め、グラフィックス用の16ビット整数演算DSPとMIDIのソフトウェアシンセサイザ用として32ビット浮動小数点用の1GFLOPS DSPを内蔵するLSIである。次に生演奏用に必要な2次DSPのチップは20GFLOPSのDSPと約3秒分の32ビット浮動小数点データを4面バッファリングするメモリが欲しい。20GFLOPSはプレイステーション2のエモーションエンジンの3倍にあたるが、ここではCPUが必要ないので実現可能であろう。こうして2チップないし3チップに主記憶DRAMを加えたものがマルチメディアホームコンピュータ用のチップの概略構成となろう。CPUチップとレンダリングチップを結ぶバスは3~5Gバイト/秒としたいが実現性の検討が残っている。

非常に先端的な利用者はこれでは満足しないだろう。これについてはマルチメディアホームコンピュータのハードウェア構成のレンダリングLSIを標準として最低限保証しなければならない部分以外にメーカー固有の部分としてより高度な機能を持たせることが考えられる。これもある種の認定を行い、ソフトウェアを含め分かりやすい分類とする。

●映像と音響の記述

マルチメディアホームコンピュータを一般家庭のエンタテイメント指向のホームコンピュータプラットフォームとして広く普及させるにはソフトウェア、映像(テレビ動画像や3Dグラフィックス)、音響/音声、さらには文書などの記述をプラットフォームハードウェアとは独立しておくことが大切である。こうすることによってメーカーの相異によるハードウェアやソフトウェアの若干の差異に対応できるだけでなく、ハードウェアやソフトウェアの技術進歩にも

対応しやすくなる。ソフトウェアについてはポータブルソフトで書き、ハードウェアはバーチャルマシンにしておく。

テレビ動画像および静止画像についてはMPEG-1, MPEG-2, MPEG-4, JPEGといった世界標準ができている。マルチメディアホームコンピュータではこれらを採用するにあたってその仕様のどのプロフィールにするかを決めればよい。音響信号についてはMPEG-1と付随した音響としてMP3, MPEG-2とMPEG-4に付随したAAC (Advanced Audio Coding), さらにはインターネットに普及しているReal Audio, 圧縮率が高いことで知られるNTT開発のTwinVQなどが知られている。一方楽譜の符号化を中心としたMIDIはその拡張形としてローランドのGSやヤマハのXGが存在する。マルチメディアホームコンピュータではこれらのどれを標準として採用するかを定めればよい。

MIDI, GSやXGを生かすためにホールに残響特性をたたみ込んでスピーカに出すには残響特性の記述を標準化する必要がある。標準化にあたって考慮すべきパラメータは音響のサンプリング周波数(典型的には44.1kHzのCDサンプリング周波数), 残響時間, チャンネル数(典型的には1, 2, 4), 量子化ビット数と浮動小数点の割振り, データ列とチャンネルの対応, 時間領域の識別子などがある。量子化ビット数については32ビットで, それを24ビット+8ビット(指数)という風に分割するのが妥当であろう。チャンネル数はホールの残響効果の迫力という意味では4チャンネルまで含めたい。残響時間は典型的には3秒であるが時間領域データは時間0から順にデータを配列して必要に応じて後の方をカットして使うことで対応できるが, 周波数領域ではフーリエ変換する関

係で残響時間は固定する必要がある。このような記述にはパラメータ記載の方法やデータ配列の方法を決めなければならないのでXML (extensible markup language) を採用し, 配列などをDTD (document type definition - ドキュメント型定義) として標準化しておく。

3Dグラフィクスについては特に映像品質が問題になるコンピュータゲームを考慮するとあまり高級な記述を採用すると記述効率が問題になる。しかし一方では利用者とのインタラクションや音響などと同期関係まで記述する必要がある。現時点では全体のフレームワークにXMLを採用し, 同期関係やインタラクションなどの記述にはDTDを標準化する。XMLは他の標準が貼り付けられるので音響はその記述法に従ったものを貼り付けられるようにする。最後に残るのは3Dグラフィクス, すなわち3Dアニメーションになるが動きの記述を含むコアの部分だけはきちんとした標準をXMLのフレームワーク内で定める作業が必要となる。

3Dアニメーションに記述の標準化にあたってはMPEG-4が参考になる。MPEG-4ではオブジェクトごとに符号化できる仕組みになっている。3Dアニメーションでもまずバックグラウンドとオブジェクトに分離する。バックグラウンドについては画面の動きがシステムティックであることが利用できよう。バックグラウンド全体をまず記述し, フレームごとにその見える範囲と視点を記述する方法などいろいろな方法があるが, ジオメトリ演算やレンダリングの際前後のフレームとの相関関係が分かりやすい記述法が望ましい。オブジェクトはDirect DrawやDirect3Dのsprayやフレームとカラーキーの考え方を利用し, オブジェクトの外接長方形をとり, その中でオブ

ジェクトを記述する。しかもオブジェクトをさらに部品に分割できるよう, フレームの階層構造を取り入れることになる。

動きの描写についてはテレビフレーム(毎秒30フレーム)またはテレビフィールド(毎秒60フレーム)ごとにポリゴンモデルを計算した結果を記述する方法と, ポリゴンモデルで動きの描写にキーとなる頂点をレファレンスポイントとして選別しておき, ポリゴン全体はたとえば10フレームごとに1度記述し, その間はレファレンスポイントの動きのみを記述する方法が考えられる。後者の方法ではレファレンスポイント以外のポリゴンの頂点はレファレンスポイントからの内そうでその位置を決める。しかもレファレンスポイントの動きの記述は動きの関数で記述する方法とフレームごとの(x, y, z)アドレスの差分で記述する方法が考えられる。

代表的な3Dアニメーションにおいてこれらの記述法で描写した映像の品質に, 記述に必要なデータ量とコンピュータで補間するときの処理量などを考慮してその記述を評価して標準化する必要がある。いずれにせよマルチメディアホームコンピュータプラットフォームの標準化とともにこの記述法の標準化は最重要事項でこれらによってマルチメディアホームコンピュータの成否が決まるといっても過言ではない。

なお標準だけで記述されたゲームは陳腐化し, しかも競争による品質向上を阻害するという批判がある。そのような先端的なゲームソフト記述のためにXMLの外にそういうメーカー固有のものを記述する余地を設け, それが認定されたメーカー固有のハードウェアと連動することによって対応することができよう。



●高齢者／主婦用機能

現在の情報技術はせいぜい50代前半までの男性を主な利用者層と考えて設計されている。たとえば広まりつつある音声認識ソフトウェアの音声モデルもそのような利用者を想定して作られている。しかし高齢者にそのような必要はないのかという答えはノーである。たとえばYahooのゲームサイトを見るとウィークディの午前中に100人もが囲碁のプレイをしていることがある。オセロはもっと多い。最近はコントラクトブリッジのサイトも熟年主婦層に人気がある。これらは主として高齢者および子育てを終えた主婦によって占められていると考えられる。結構需要があると考えられるべきであろう。

音声認識をゲームに利用したい。コンピュータゲームを操作するのは大変だが、声でならできるといふ人は少なくない。ゲームに使うコマンドの数は多くないので認識率は非常に高くできよう。ついでにブリッジや囲碁、将棋、チェスもマウスを使わないで声でやれることだろう。

次に欲しいのは1人でできるエンタテイメントである。音楽を聴く、ビデオやテレビを見ることは今まで

からあった娯楽である。マルチメディアホームコンピュータはMIDI機能や残響付加の機能があるので電子楽器が楽しめる。従来の楽器に比べて上達が楽であり、しかもコンピュータは独習を援助することができる。1人で練習して上達するとそれをネットワークに通じて友人に聴かせるなど従来の楽器にはない楽しみ方ができる。人込みを1人で移動することが次第に億劫になる高齢者にとってネットワークを通じたプライベートコンサートはありがたい。次にゲームである。コンピュータゲームの大半がコンピュータと人の対戦であり、1人でも楽しめるようになっている。この種のゲームは一人住まいの高齢者または子育ての終わった主婦に向いている。我が国の伝統的なゲームである囲碁、将棋、麻雀がまず第一である。次に主婦に広がっているブリッジ、さらにはチェス。入門しやすいオセロなど。この種のソフトウェアを充実させ、もっと強くするだけではなく、きめ細かくレベル設定できるようにすることも大切である。

【ホームコンピュータプラットフォームの用途】

ホームコンピュータプラットフォームは何に使えるだろうか。エンタテイメントについてはすでに述べたのでそれ以外の用途を考察する。

(1) サイバー通信

電話、FAXやe-mailのように特定の相手との通信ではなく、インターネットのウェブ技術を使ってウェブサイトが提供するサイバースペース(cyber space)で自らを仮想の存在(アバター-avatar)として行動する形式の通信である。サイバースペースの他のアバタに話しかけたり、スペ

ース内にあるサイバーモール(cyber mall)で遊んだりショッピングを楽しむこともできる。囲碁、将棋、対戦型のコンピュータゲームなどの対戦相手もそこでみつけることができる。サイバー映画館の広告を見ながら見たい映画を選択することもできよう。

最初に接続するサイバースペースは3Dグラフィクスで映像化された画面として提供される。当然音響環境も室内は室内らしく、モールはモールらしい音がする。このような通信のためにはマルチメディアホームコンピュータプラットフォームのように標準的なマルチメディア端末が必要である。

(2) アニメ放送／データ放送

3Dグラフィクスの記述と音響の記述が標準化されるとそれらを利用した3Dアニメ放送が考えられる。テレビと異なりそれほどの周波数帯域を必要としないのでインターネットを利用した放送とすることが考えやすい。テレビ放送にある2Dアニメーションと異なり、映像の視点が自由に選択できよう。デジタル放送に伴って話題になっているデータ放送の一部にもこのような3Dアニメが取り入れられよう。

(3) 教育

すでにパソコン用の多くの学習プログラムがあるが、学習する側が受け身のものが多い。ATRの成果を下にパソコンで音響分析し、学習者の英語の発音を評価し採点するソフトがある³⁾。このように学習者の積極的な行動にそれをコンピュータが評価する形式の教育用のソフトが作りやすくなる。

ホームコンピュータプラットフォームのDSPの能力は音声分析には

威力を発揮する。音声認識／合成ソフトと自然言語処理技術を駆使すればコンピュータと英会話をすることもできよう。喋った結果を再生し、悪い個所を指摘するなどアクティブな教育ができる。3Dグラフィックスの機能は動物や植物をリアルに映像化することができるので生物の教育なども映像化して行うことができる。さらに植物を育てたり、動物を飼ったりするソフトウェアにより生物界の仕組みをアクティブに学習することもできよう。実際に植物を育てるには学習に時間がかかるがコンピュータではたとえば1日を10秒ぐらいに短縮するなどフレキシブルな学習が可能になる。

【家庭でのソフトウェアの評価】

家庭にコンピュータを普及させる条件の1つはソフトウェア購入を容易にすることである。まずどのようなソフトを買うかに悩むことが多い。そのためソフトウェアの評価を分かりやすくしたい。

家庭の使うソフトウェアは1本5,000円から1万5,000円程度の価格帯のものが多い。一般家庭にとって、一万円前後の買い物はそれほど気楽なものではない。ソフトウェアはその性格上買ってしまえば返品することが難しい。したがって買う前にそれが納得のいくものかどうかを確認したい。ダウンロードして試行的に使ってみるといったソフトウェアもあるが、これは一部のものに限定される。そのソフトウェアの箱にはハードウェア条件は明確に書かれているようになったが、悩みは多い。囲碁、将棋、チェス、オセロなどのゲームはコンピュータが強すぎても弱すぎても面白くない。まず弱すぎるときは論外である。強すぎる場合は

何段階かにレベル設定できるようになって欲しい。たとえば将棋ソフトで6段階のレベル設定ができるものがある。それ以外に駒落ちでの対戦もできる。こういう「強さのレベル」を外箱に表示したい。その将棋のソフトではアマ3段程度の棋力があると外箱に書かれている。

ソフトウェアの品質の第2は音声認識ソフトや手書き文字認識ソフトの認識率である。たとえば音声認識ではその専門の技術者なら音声モデルと言語モデルがどういう風に作られたかが認識率を左右することを知っている。しかし一般の家庭人では購入する場合はどう判断するのだろうか。認識率だけではなく、誤認識の後の修正のインターフェースも気になるところである。多分一般利用者はそういうことを気にし始めると、時期尚早と判断し、買わないという結論になってしまうのではないか。

ソフトウェア購入は家族で食事に行くときのレストラン選び、温泉旅行に行くときの旅館選びと似ている。レストランの場合、ちゃんとしたフレンチ、イタリアン、和食だと1人5,000円から1万5,000円の価格帯であり、金額的にもソフトに近い。返品ができないことも類似している。レストランのサービスや味についてはインターネットの<http://www.jibaran.com>というサイトがある。これは「ジバラン～自腹覆面レストランサイト」と自称するだけあって、普通のサラリーマンとその家族が審査員となり、自腹を切ってレストランに食べに行ったときの評価を掲載している⁴⁾。ソフトウェアの評価にもこの“Jibaran software サイト”があれば様子は変わることであろう。

実は“owner's voice”というサイトがあり、PC本体や周辺機器とともにPCソフトの評価が寄せられてい

る⁵⁾。しかし一般家庭用ソフトという意味では物足りない。こういうものがもっと盛んになるのはまず利用者にとっていいことである。何を買うかの際、大いに参考になる。それだけでなく販売側にとっても有用である。まず利用者がある程度安心して買えるので利用が促進される。マーケットが大きく進展することが期待できる。第2に製造側にとっても大いに参考になり、次の製品にフィードバックする多くのヒントが得られることになる。

【むすび】

3回にわたって家庭の情報化についてその実態と考えられる方向を述べ、1つの候補であるマルチメディアホームコンピュータについて我々の検討結果を概説した。性能やアーキテクチャも大切であるが、何より大切なのは標準化である。標準化することによって家庭でのエンタテイメントだけではなく、通信や放送など幅広い応用が拓けてくる。本稿をきっかけに標準化に向けて関心が高まることを期待したい。

なおすでに述べたように、この3回の連載は情報処理振興事業協会(IPA)のプロジェクトの一環として行った「エンタテインメント指向マルチメディアホームコンピュータのフィービリティの調査」の報告書に基づいて書いた。これを可能にしてくださいIPAに感謝したい。

参考文献

- 1) 釜江尚彦、連載解説マルチメディアホームコンピュータの未来—第1回家庭の情報化の主役は何か—、情報処理、Vol.41, No.7, pp.826-830 (July 2000)。
- 2) 日立 SuperHTRISC engine SH-4ハードウェアマニュアルSH7750、(株)日立製作所電子統括営業本部 (平成10年4月第1.0版)。
- 3) 山田、足立: 英語スピーキング科学の上達法、講談社 (Aug. 1999)。
- 4) ジバラン～自腹覆面レストランサイト: <http://www.jibaran.com>
- 5) Owner's Voice: <http://www.ownersvoice.ne.jp> (平成12年6月29日受付)