

われわれの身の回りには家電製品、携帯電話、カーナビゲーション(以下、カーナビと略記)などの車載機器にもマイクロコンピュータが組み込まれ、コンピュータを特に意識することなく使う時代が到来しつつある。この結果、個人の特性や場所に依存した、きめの細かいサービスがますます重要になってきた。音声認識/合成、画像処理、自然言語処理などのヒューマンインタフェース要素技術の統合により、ハンズフリーで(手を使わずに)即座な操作を可能とする対話技術を開発し、安心して使える情報サービス/機器を提供していく。

Many microcomputers are embedded into home appliances, automotive peripherals, and mobile phones, and users unconsciously use these computers. It is becoming increasingly important to provide detailed services related to users' profiles, locations, and situations. The integration of various human interface technologies such as speech recognition/synthesis, image processing, and natural language processing enables hands-free and real-time interaction, thereby realizing secure and comfortable services and appliances.

オフィスから家庭、移動中に拠点を移す対話技術

ヒューマンインタフェースは、人間と機械、情報システム、ひいては人どうしを仲介するものであり、音声や画像などのメディアの入出力、変換、認識、意図理解により一貫した使いやすさを提供するものである。まさに、対話技術こそヒューマンインタフェースの真髄である。

ヒューマンインタフェース関連技術の発展の歴史を振り返ると(表1)、工場などの生産機械や航空機など、いわゆる現場とそこに働くユーザーが対象であった。それが、コンピュータの登場により、CUI(Character User Interface)によりコマンドを与えると使えるようになり、生産部門だけでなく、設計/技術部門へと対象が広がってきた。更に、ワードプロセッサやパソコン(PC)が間接部門であるオフィスを中心に発展してきた。コンピュータの専門知識がないオフィスワーカーが自由にPCを使いこなせるようになったのは、画面に表示されるアイコン/メニューをマウスで選択

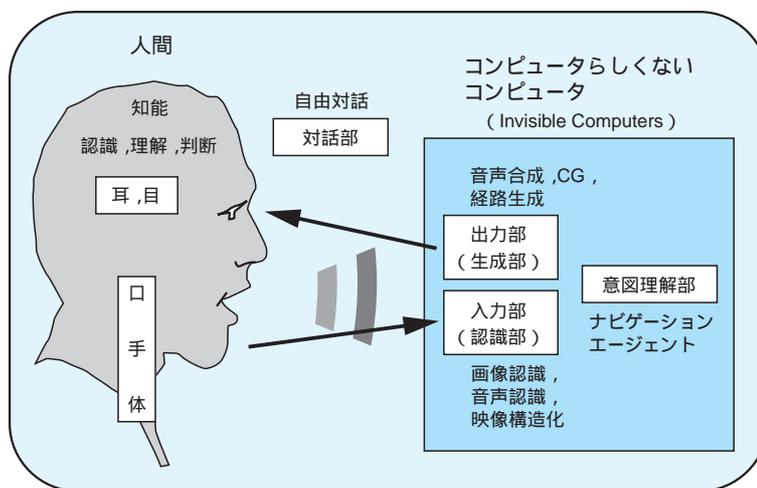


図1. 対話を支える要素技術 様々な要素技術により、対話インタフェースが成立する。
Fundamental technologies for dialogue-based human interface

するだけというインタフェース GUI (Graphical User Interface) の登場によるところが大きい。くしくも GUI の発明と時を同じくして、1981年にインターネットが開発され、現在のデジタルネットワーク社会が出現した。

このデジタルネットワーク社会は、94年に標準化された MPEG 2 (Moving Picture Experts Group 2) が

DVD プレーヤとして一般家庭に入り、その中心ユーザーはオフィスから家庭へと移行してきた。更に、99年の i モードサービス^(注1)開始で携帯電話機からもインターネットにアクセスできるようになり、移動しながらの乗換え案内や道案内サービスも現実のものとなりつつある(囲み記事参照)。

これまでの GUI は卓上での操作が

(注1) (株)エヌ・ティ・ティ・ドコモグループが99年2月から始めた携帯電話を使った情報通信サービス。

表1 ヒューマンインタフェース関連技術の変遷
Advancement of human interface-related technologies

年号	入出力関連事象	世代	ユーザー		
1866	タイプライタの発明(C.L. Sholes & C. Glodden)	第1世代： 媒体インタフェース	工場の生産部門 (専門家)		
76	電話機の発明(A. G. Bell)				
1915	邦文タイプライタの開発(杉本京太)				
26	ブラウン管にテレビ映像を表示(高柳健次郎)				
44	リレー式コンピュータ MARK1 完成(ハーバード大)				
46	電子計算機 ENIAC 完成(ペンシルバニア大)				
55	ひらがなタイプライタの開発(川上晃)				
60	ワードプロセッサの開発(T. Nelson)				
63	対話型図形処理システム Sketchpad の発表(I. Sutherland)				
64	データタブレットの開発(M. R. Davis & T. O. Ellis)			第2世代：CUI	工場の設計 / 技術部門 (専門家)
68	パーソナルコンピュータ構想の発表(A.Kay)				
73	ワークステーションALTOの開発(Xerox 社)				
78	日本語ワープロ JW-10 の開発(東芝)				
81	オフィスコンピュータ Star の開発(Xerox 社) Internet 開設				
83	ファミリーコンピュータの発売(任天堂(株))				
84	Macintosh の発売(Apple Computer 社)				
85	英日機械翻訳システムの開発(東芝)				
89	日本初の音声ダイヤル電話機の発売(東芝)				
91	MPEG 1 標準化	第3世代：GUI	オフィス (オフィスワーカー)		
94	MPEG 2 標準化 家庭用ビデオゲーム機の発売((株)セガ・エンタープライゼス(株)ソニー・コンピュータエンタテインメント)				
95	フラッシュメモリによる小型メモ리카ードの開発(東芝) 英日 / 日英翻訳ソフトウェア(ASTRANSAC) Web 版の開発(東芝) 東芝音声システムの製品化(東芝)				
96	DVD 発売(ソニー(株) , 東芝 , 松下電器産業(株) , ほか)				
97	インターネットによる情報サービス“ 駅前探検クラブ ”の開設(東芝) IEEE Wearable Computer 国際会議				
98	MPEG-4 標準化 MPEG-4 ビューフ(MobileMotion)の発表(東芝) 日英・英日翻訳ソフトウェア“ The 翻訳オフィス ”の発売(東芝)				
99	iモードサービスの開始((株)エヌ・ティ・ティ・ドコモ) Microsoft®Windows®(注2) CE 用テキスト音声合成ミドルウェアの製品化(東芝)				
2000	家庭用高性能 3D ビデオゲーム機の発売((株)ソニー・コンピュータエンタテインメント) SDメモ리카ード の発売				
				第4世代： Invisible Interface PUI	家庭 (エンドユーザー)

東芝, 松下電器産業(株) , SanDisk 社が共同で開発したフラッシュメモリを使った小型記録メディアの一つ。

前提であった。しかし, 家庭や移動中での使用が中心となると, いつでもどこでも使えるインタフェース(Ubiquitous Interface), 更に, ユーザーの状況に合わせて情報提供を行えるインタフェース(PUI : Perceptual User Interface)が求められてきている。

ミドルウェア化するヒューマンインタフェース技術

人間は, 目や耳から入ってきた視覚 / 聴覚情報を頭脳で認識, 理解, 判断する。更には, 手で触って得られ

た触覚情報を用いる。その結果, 口から情報を話す。単にことばとして話すだけでなく, 身振り / 手振りを使って, 表現力を高める(図 1)。

これに対し, コンピュータ側(現状の PC だけでなく, 情報家電に組み込まれた Invisible Computers も含む)では, ユーザーからの音声や手振り / 身振りを認識し⁽¹⁾, 構造化するための音声認識, 画像認識, 映像構造化技術がある。この入力部(認識部)は, 人間で言えば, 目や耳に当たる。これらの認識された情報を理解し, インターネット上にある別の情報を参

照して, 最新の関連情報などを提供するエージェントやナビゲーション技術がある。この意図理解部が人間での知能に当たる。認識し, 意図理解した結果に基づき, 生成された情報を音声やテキスト, 図あるいは三次元情報として提示するための音声合成, CG(Computer Graphics), 経路生成などの生成技術がある。この出力部は, 人間の口や手, 体に相当する。従来のコンピュータには手足がなかったが, ペットロボットのようにアクチュエータを持つ新しいインタフェースも登場してきている⁽¹⁾。これらの認識部, 意図理解部, 生成部を統合的に制御し, 対話を成り立たせるのが対話部

(注2) Microsoft, Windows は, 米国 Microsoft Corporation の米国及びその他の国における登録商標。

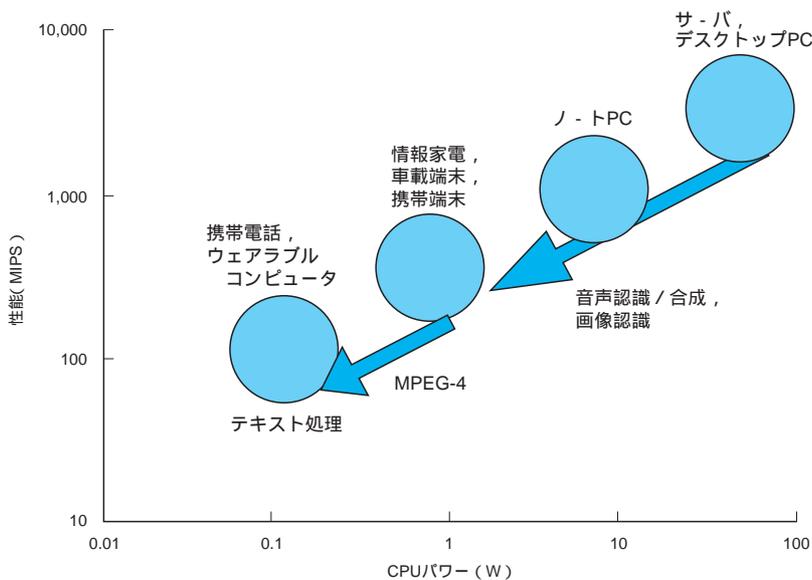


図2. ヒューマンインタフェース技術のミドルウェア化 ヒューマンインタフェースの機器への組み込みが進む。

Progress of middleware-oriented human interface technologies

である。対話部は、人間で言えば技術に当たるが、現在では、カーナビに向けて自由音声対話技術が実用化さ

れようとしている。

これら、認識/生成/意図理解/対話のための音声認識/合成、画像

認識などの対話の要素技術は、表1の第1世代や第2世代では、大型コンピュータや専用ハードウェアの上で開発されてきた。それが、近年のCPUの処理能力向上により、第3世代のGUIでは、デスクトップPCだけでなく、ノートブックPC上で稼働することが可能になってきている。それが更に第4世代で、家庭などで安価に使えるようになり、コンピュータを意識することなく、テレビ(TV)や電子レンジなどの家電製品やカーナビなどの車載製品、PDA(Personal Digital Assistants)などの携帯端末を操作できるように、少ないCPUパワーでの稼働が求められている。更には、家庭の中にとどまらず、移動中にもハンズフリーで、携帯電話やウェアラブルコンピュータなどを操作し、いつでも(anytime)、どこでも(anywhere)、だれでも(anyone)、使えるようになることが求められている⁽²⁾。0.1W程

情報端末化する携帯電話

今日、携帯電話は音声を伝送する単なる電話機というより、インターネット上の様々なサービスが利用できる総合的なモバイル情報端末へと変貌(へんぼう)を遂げつつある。この携帯電話とインターネットを結びつける仕組みが「iモード」と「WAP(Wireless Application Protocol)」である。

PCと違い、携帯電話では表示部が小さく(標準的なもので96×72画素)、通信速度も限られる(PDC(Personal Digital Cellular)方式で9,600bps)。

また、データ処理能力も高くない。したがって、インターネットに効率的にアクセスするためには、携帯電話専用のプロトコルやコンテンツ記述言語が必要となる。この観点から、新しく仕様の標準化を進めているのがWAP、一方、既存のインターネット標準プロトコルや言語を携帯電話向けに直したのがiモードである。両者の違いを示すと下表のようになる。

なお、現在はコンテンツとしてテキストベースの情報が主に扱われているが、

Java^(注3)対応になれば、より高度なコンテンツの表示、更には、携帯電話自体の基本機能の変更やバージョンアップも可能となる。また、IMT-2000^(注4)が実用化され、次世代携帯電話のサービスがスタートすれば、携帯を利用した音楽や映像などのマルチメディア通信の可能性も広がってくる。

文献

「ここが知りたい」日経コミュニケーションズ 2000.2.21

“iモード”と“WAP”の違い

項目	iモード	WAP
コンテンツ記述言語	HTML4.0ベース、iモード用HTML	XMLベース、WML/HDML
プロトコル(アプリケーション層)	HTTPの拡張型	WSP, WTP/HDTP
ネットワーク	パケット交換	回線交換/パケット交換
メール	オフラインで参照可	オンラインでの参照だけ
コンテンツアクセス	ページ単位	デッキ(複数ページ相当)単位

XML : eXtensible Mark-up Language
WML : Wireless Mark-up Language
HDML : Handheld Device Mark-up Language
WSP : Wireless Session Protocol
WTP : Wireless Transport Protocol
HDTP : Handheld Device Transport Protocol

(注3) Javaは、米国SunMicrosystems社で開発されたプラットフォーム非依存のオブジェクト指向プログラミング言語で、商標。

(注4) ITU-R(国際電気通信連合無線通信部門)で検討されている、次世代携帯電話システムの国際標準規格の名称。

度の消費電力で100 MIPS(Million Instructions Per Second)以下のCPUパワーで稼働する対話ミドルウェアとしての製品化が要望されている(図2)

■ ネットワーク隣組との対話

ミドルウェア化された対話技術を使って対話する相手は、コンピュータに限らない。ペトロボットやアバター^(注5)と呼ばれる仮想的なキャラクタなども新しい対話対象である。更には、ネットワークを通じて、世界中のユーザーとチャット(おしゃべり)を交わすことも可能である。図3から通常の音声通話の伸びに比較し、ショートメッセージは1年間で10倍に伸びていることからわかるように、安い回線を使って相手に常にコミュニケーションを求めるユーザー層が増えている。現在は、9,600 bpsでの回線に対し、テキストだけの通信であるが、T(テラ)ビットオーダの大容量回線になったときには、MPEG-4レベルでいつでも自分の存在や環境情報を送り続ける“ながらユーザー”が増えることが

予想される。これらのユーザーが簡単に自己の個性を発現できる新しいオーサリング(マルチメディアコンテンツを作る)やAQU・Pal!(本誌p.13参照)のようなメーリングツールへの需要が高まることが予想される。

■ 場所 / 個人依存による地域性 / 個性の復活

インターネットは、世界中のどこにでも発信できる。しかし、その分、情報洪水などの弊害も生じている。情報洪水に対しては、情報フィルタリング技術がユーザーの欲しい情報だけをより分けることを可能にした。21世紀にはデジタル放送となり、100チャンネル時代が到来する。デジタルTVでは、100チャンネルのTV番組のほかに、インターネットの情報の閲覧も可能となり、オフィスだけでなく、家庭でも、個人のプロフィールに応じた情報フィルタリングが必須となる。

21世紀には、車載向け道路網の整備、家庭に向けた情報インフラの整備が一段落し、歩行者向けの情報インフラ、街角の情報インフラの整備に

重点がおかれるようになる。これとあいまって、乗換え案内や道案内がPCだけでなく、携帯電話でも行われるようになってきている。インターネットでは、世界中どこにいても同じ情報を得られることに価値があったが、乗換え案内などでは、地域性に重点がある。

■ より円滑なヒューマンインタフェースを目指して

人間とシステムとの円滑な対話を実現するためのヒューマンインタフェース技術とその動向について述べた。人や環境に関する新たな知見、新しいデバイス、革新的なアルゴリズムを活用し、“いつでも”、“どこでも”、“だれでも”使え、かつ、個人や場所に適応したシステムの開発、サービス創出に取り組んでいく所存である。

文 献

- (1) 土井美和子;住田一男.特集「実世界に近づくインタフェース技術」.情報処理学会誌. 41, 2, 2000, p.122-150.
- (2) 竹林洋一.ヒューマンインタフェースの観点から見た気の利いた情報システム.電子情報通信学会誌. 82, 4, 1999, p.310-318.

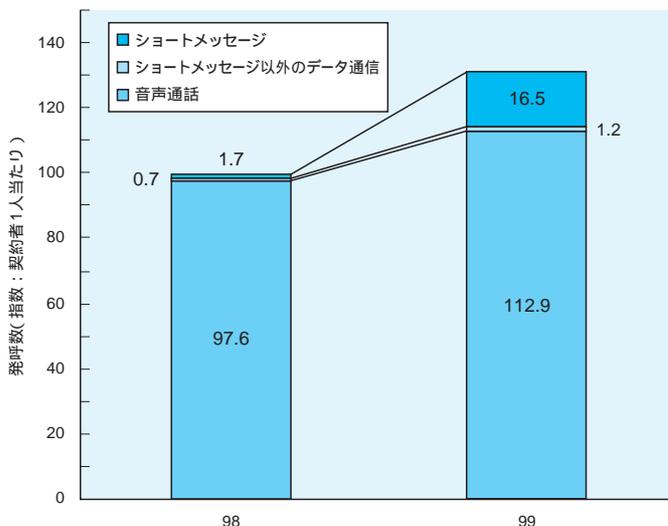


図3 . 通話の変化 手軽なショートメッセージを利用した通話が増大している。
Trends in net-based communication

(注5) Avatar(化身): 仮想のヒューマノイド(人間もどき)で、自分の代わりに3Dのチャットルームやネット上の仮想空間を歩き回らせる。



土井 美和子
DOI Miwako

研究開発センター ヒューマンインターフェースラボラトリー研究主幹。次世代ヒューマンインタフェースの研究・開発に従事。電子情報通信学会, 情報処理学会, HI学会, 計測自動制御学会, ACM 会員。

Human Interface Lab.



亀山 研一
KAMEYAMA Ken-ichi

研究開発センター ヒューマンインターフェースラボラトリー研究主務。次世代ヒューマンインタフェースの研究・開発に従事。情報処理学会, 映像情報メディア学会, HI学会, VR学会, ACM 会員
Human Interface Lab.