

論文 / 著書情報
Article / Book Information

論題(和文)	ISDN時代のヒューマンインタフェース
Title(English)	
著者(和文)	古井 貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	電子情報通信学会東京支部講演会「ヒューマンインタフェースの動向」, , , pp. 1-10
Citation(English)	, , , pp. 1-10
発行日 / Pub. date	1991, 2
URL	http://search.ieice.org/
権利情報 / Copyright	本著作物の著作権は電子情報通信学会に帰属します。 Copyright (c) 1991 Institute of Electronics, Information and Communication Engineers.

ISDN時代のヒューマンインタフェース

古井 貞照

NTTヒューマンインタフェース研究所

1. はじめに

電気通信の究極のねらいは、距離と時間を超越して、相手との意志疎通を円滑に行えるようにすることにある。近年のISDN(デジタル総合網)の実用化によって、情報の伝達機能に関しては、ほぼ要求が満たされる状況にあるが、真に距離と時間を超越した一体感、臨場感、リアリティのある通信の実現という点では、まだ解決しなければならない課題が多い(1)。これらの課題は一言でいえば、ISDNの機能とコンピュータの能力をフルに活用した、高度かつユーザフレンドリなヒューマンインタフェースの実現と言うことができる。人間と機械との係わりにおいて、人間が機械に歩み寄ったこれまでの形態ではなく、機械が人間に歩み寄った人間主体のヒューマンインタフェースの実現とすることもできる(2)。ここではこのような観点から、現在行われつつある種々の新しい研究開発の状況を紹介する。

2. VI&Pサービス

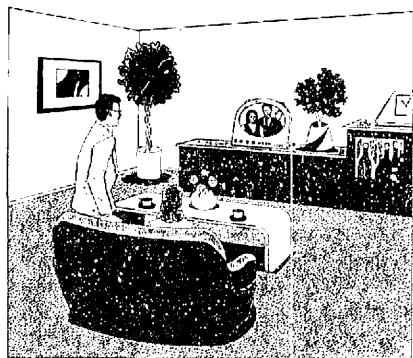
NTTでは、21世紀における社会・経済・文化などの変容や技術革新の動向を踏まえ、21世紀に向けた新しいサービスビジョン「VI&Pサービス: Visual Intelligent and Personal Communications Service」を掲げて、研究開発を進めている(3)。VI&Pサービスの中では、「V」のビジュアルテレ

フォン、「I」のテキストメール、「P」のポケットテレフォンの3つを基本通信サービスに位置付けている。これらの基本通信サービスは1:1の対話型通信で、これまでの電話同様、全国にあまねく公平に提供するものである。VI&Pではさらに、これらの基本通信サービスを発展させた高度通信サービスを考えている。高度通信サービスは、料金や内容で通信業者間で競争するものであり、ユーザは自己のニーズに従って選択的にサービスを楽しむことができる。

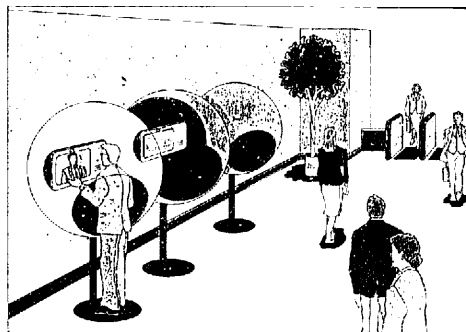
高度通信サービスまでを含めた研究開発の目標は次の通りである(表1に主な高度通信サービスを、図1に種々の例を示す)。まず、映像通信サービスとしては、自然かつ鮮明な映像を表示できる高精細映像通信、等身大の迫力ある大画面映像通信、好みで画面を自由に選択できる多画面映像通信などを目標にしている。また、インテリジェント通信サービスとしては、高度情報処理機能をネットワークに付加することによって、電子秘書や翻訳通信、さらにユーザ自らがサービス内容の一部を自由に変更できる各種オーガメイトサービスなどの実現を目指している。さらに、パーソナル通信サービスとしては、各人が自分の番号を持ち、どこにいても追跡して接続してくれる追跡通信、本人のみが受けられる親展通信、いたずらなどを防止できる選択受

表1 主なVI&P高度通信サービス

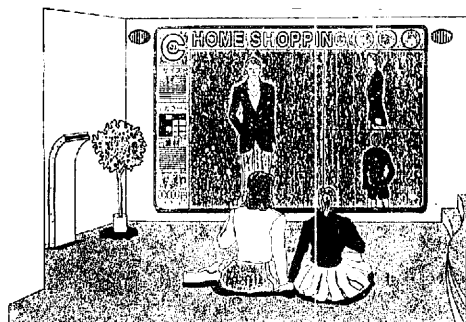
映像通信サービス	高精細・大画面・多画面映像	映像ショッピング、遠隔看視・医療検診、映像データベース(電子美術館、ファッション情報など)、高臨場感映像伝送(等身大映像伝送など)、遠隔環境監視、映像掲示板、多地点間映像会議
	立体映像	立体映像会議、高臨場感立体映像伝送、立体映像データベース、立体映像中継
インテリジェントサービス	異言語間翻訳	テキスト翻訳通信、翻訳電話
	知的支援	音声ダイヤル、時間指定通信、ハンディキャップ通信(文字、音声、手話、点字などの認識変換)、電子秘書(代行応答、スケジュール管理など)、検索代行(あいまい検索、自然語検索)、掲示板、戦略情報システム
	柔軟なネットワーク	カスタマイズドサービス、可変プライベートネットワーク
パーソナルサービス	受発信者の利便向上	発信者名表示、個人名呼び出し、自動指名通信、選択受信、個人番号
	プライバシー保護	親展通信
	柔軟な課金	選択課金、個人料金明細記録
	ICカード利用	電子金券、料金決済



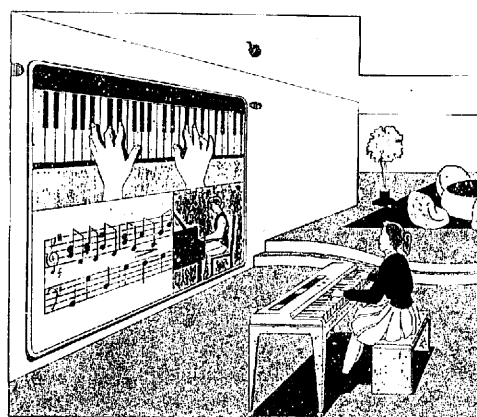
ビジュアルテレホン



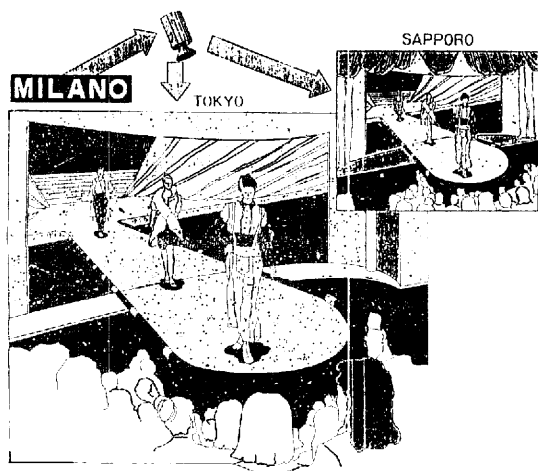
街頭型ビジュアルテレホン



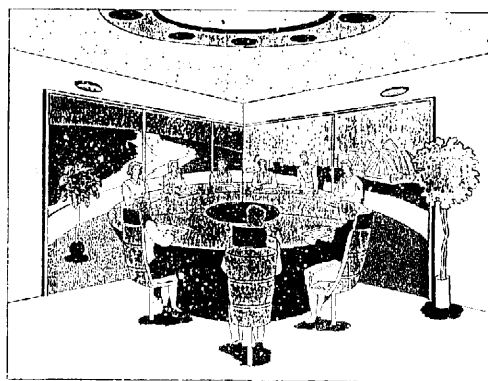
テレショッピング



テレレッスン



立体ライブシアター



・立体的映像会議

図1 VI&Pサービスの例

信、ICカードの記憶・情報処理機能を利用した自動料金決裁、などの実現をねらっている。

これらのシステムは、ただ単に距離や時間を克服して、情報交換を効率的に行うというのではなく、臨場感の伝達も行える通信システムであり、これにより、相手との意志疎通がより正確に、またより容易に行えるようになることを目指している。さらにコンピュータとの対話によって、データベース検索や、文書作成、ゲーム、教育、買物など、日常生活まで含めた種々の利便性を向上させることも目指している。これらのサービスは、単にビジネスのためだけでなく、快適な生活環境を創り出したり、豊かな余暇や興味を演出するのにも役立つと考えられる(4)。

3. マルチメディアによるヒューマンインタフェース

人と人のコミュニケーションにおいて伝達される情報の内、最も重要な情報は言語的情報であるが、相手の性別、年齢、体格、動作や視線の動き、声の特徴、表情、感情など、非言語的情報も情報伝達において重要な役割を果たしている。人と人のコミュニケーションが複雑になると、人は言葉以外に図や表を用いたり、実物を指し示しながらコミュニケーションする。これらの種々の情報が、歪や損失を受けることなく忠実に相手に伝達されることが、臨場感の伝達、すなわちあたかもその場に居合わせているような感じで情報交換ができるために不可欠である。これまでの電報、電話、データ、ファクシミリ等の通信において伝達される情報は、言語的情報がほとんどであったが、これからの通信は、映像と音を中心とする多彩なメディアを用いて行われるようになるであろう(マルチメディアあるいはマルチモーダルインタフェース)。種々のメディアの特徴を十分考慮した上で、メディアが効率的に使われ、またシステム全体としてメディアのバランスがとられているインタフェースが、優れたインタフェースであると言えよう。マルチメディアを自由に扱えるようになることで、人間の思考は幅広く支援されるようになるものと期待される。

表2に、通信会議における映像および音のインタフェースの発展段階を示す(6)。現在の通信会議システムは集中型であるが、今後投影型へ、さらに立体型へと発展することによって、臨場感のある通信会議が可能となる。

これら通信機能に加えて、人とコンピュータが、人と人がコミュニケーションするのと同じ様にコミュニケーションできるインタフェースが実現できれば、人は特別な訓練や学習を必要としないでコンピュータを使いこなすことができるようになるであろう。現在、コンピュータへの情報入力はキーボード、タブレット、マウスなどでなされており、コンピュータから人への情報伝達は、主にディスプレイ、プリンタ等を介しての言語情報によってなされていて、音声や画像情報はほとんど用いられていない。人は音声からその言葉を理解することや、図形からその意味や内容を理解することをいとも簡単にできるが、これらは現在のコンピュータには極めて不得意な機能である。人とコン

表2 通信会議の進化のイメージ

	イメージ	形態	概念
集中形		装置 列席者	1 次元的 利用者は 装置を意識する
投影形		虚像 ステージ	2 次元的 ステージ上に 過隔での状況を 再現
立体形		虚像	3 次元的 離れた人間が 一堂に集まった 状況をつくる (仮想会議空間)

ピュータとのインタフェースと、人と人とのコミュニケーションには、現在非常に大きなギャップがある。

また、ミスを許容するインタフェース、解らなくなった時にいつでも助けが求められるインタフェース等も、最低限の機能として備えていることが望ましい。学習機能を有し、個人個人の知識や思考過程を吸収したり、こと細かく全てを指示しなくてもある程度のことは予測してくれたりする、個人に特化した自動誤り修正機能を持ったインタフェースもまた今後望まれる機能である。まず大局を表現することが得意な図形や画像等による表示(アイコン等を含む)を行い、細部については言語情報などで説明するインタフェースも、人の思考にマッチしたインタフェースと言える。

4. 新しいヒューマンインタフェースの実現技術

4.1 映像メディア

(1) 立体表示

現在の映像通信技術においてまだ実用化されていない最も重要な課題は、立体表示である。すでに可能性が実証されている技術でも、表示面の画面サイズ(視野の拡大と高精細化)については、まだ満足できるものはない。人間は、両眼に写る画像の見え方の差(両眼視差)、見る位置を変えた時の見え方の差(運動視差)、物の大小等種々の要因によって立体的な感覚を得ており、しかもそれらの要因が複雑に絡み合っている。このうち、心理的要因に基づく立体視(物の大小、物の重なり等による幾何学的立体視や、陰影、鮮明度等による光学的立体視など)については、現在のテレビジョンでも実現可能である。しかし、真に立体感のある映像通信を実現するためには、心理的要因に加えて、生理的要因(焦点調節、両眼輻輳、両眼視差、運動視差等)をも満

たす技術が必要である。

表3に、現在実現または実現可能性のある立体テレビ方式を示す(6)。メガネを必要とする方式については、濃度差方式を除いては実用段階にあるが、テレビ電話、テレビ会議等、通信への応用においてはメガネの使用は望ましくない。ホログラフィーは、写真技術の進歩により、高精細、フルカラーの立体像実現に成功しているが、まだ実時間で動画再生ができないため、通信への適用は難しい。

メガネが不要で両眼視差を用いる優れた方法に、レンチキュラ方式があり、現状技術では、通信への適用可能性が最も高い方式であると言える。図2に、レンチキュラ方式による立体テレビ電話実験装置の一例を、図3にその原理を示す(1)。本装置では、レンチキュラシートを、液晶ディスプレイ表示面にシートの1ピッチに左右眼用の2画素が正確に配置されるように張り合わせることによって、フルカラー動画の立体視を実現している。ただし一般に両眼視差を用いる方式では、焦点調節と両眼輻輳が一致しないことによる疲労感を生ずる、視点の移動で3次元形状が歪むなど、いくつかの問題がある。

(2)高精細・広視野映像

高精細・広視野映像の効果としては、立体感(奥行き感)や広がり感が生ずる、映像空間と観察者の空間が融合することによって臨場感や迫力が増す、視覚情報の増大によって通常人間が情報を受容している状態に近付き自然感が増す、観察位置や姿勢の制約が少なくなり解放感が増す、などの効果が報告されている(7)。これらはいずれも、臨場感通信において重要な効果であると言える。

(3)視線の一致

会話や会議において、目の動きや視線は重要であるので、通信会議においても、実際の会議同様にお互いの視線が違和感なく一致することが望まれる。単純にディスプレイの上にテレビカメラを載せると、視線はディスプレイに向かっているため、カメラの下を見ている形になり、相手を見ているつもりでも視線がずれてしまう。テレビ電話においては、ハーフミラーを用いた方法により、視線の一致

が可能なシステムが開発されているが、大画面には適さない。また、複数の人が参加する通信会議における参加者各人の視線の一致を可能にすることはかなり難しい問題である。参加者が複数人である場合には、複数の視点・視線が存在することになるので、参加者相互の視線の一致を実現す

表3 立体テレビ方式

メガネ	立体視	視点	方式名
要	主に 両眼視差	単一	立体鏡方式
			アナグリフ(2色)方式
			偏向方式
			時分割方式
			濃度差方式
不要	空間像	複数	バラックスバリア方式
			レンチキュラ方式
			バリフォーカルミラー方式
			ホログラフィー方式
			固体内発光方式

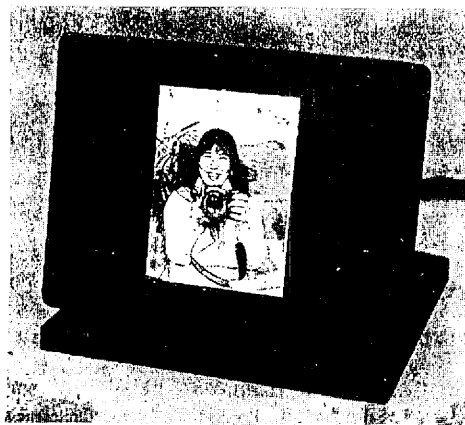


図2 レンチキュラ方式による立体テレビ電話実験装置

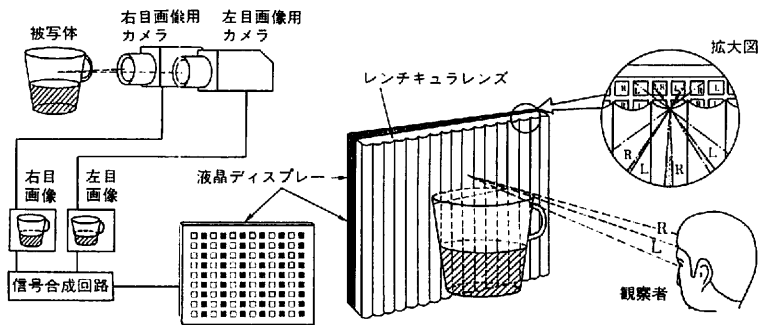


図3 レンチキュラ方式による立体表示の原理

るためには、各参加者に対して他の全ての参加者から見た画像を生成して伝送し、表示しなければならない。

4.2 音メディア

(1) 音像定位ステレオ通信方式

音の情報すなわち聴覚情報も、臨場感を表出する上で極めて重要である。臨場感のある通信会議を行うための一つの要素として、発言者の位置に応じた音像を会議相手に伝えることが必要である。従来のステレオ再生では、図4において、左右のスピーカから等しい位置Bの受聴者には、再生音に応じた音像が左右スピーカの間の広い範囲に定位するが、その位置からはずれた、例えばスピーカRの直前の位置Aの受聴者には、音像がスピーカRに引き付けられたように定位する。これは、左右のスピーカの出力が中央の受聴者に適合するように調節されていると、受聴者AにはスピーカRからの信号が音量も大きく、また早く到達するため、音像がスピーカRに引き寄せられる(先行音効果)からである。受聴位置がAの場合にもBと同様の音像を得るためには、左右スピーカからの音圧差および到達時間差を解消することが必要である。

図4に示す新しいステレオ再生法は、このような考え方に基いて、広い聴取領域において会議相手の発言位置に対応した音像が良好に定位するようにしたものである(8)。この方法では、左右のステレオ信号を、遅延器、位相器を用いて各々3種類の信号とし、それらを計6個のスピーカを用いて放射することにより、両端を含む広い範囲に位置する受聴者に対して、正しい定位をもたらす信号だけが到達するように工夫されている。

(2) 音場制御技術

人間は2つの耳を使って、音の到来方向を検知したり、特定の会話を聞き分ける能力(カクテルパーティー効果)を持っている。これには、室内の壁などでの反射音も重要な役割を果たしており、臨場感通信の実現、すなわちあたかもその空間に存在して聞いているのと同じ知覚が得られるためには、同じ反射音の構造を持つ音響環境を共有する必要がある。このためには、部屋の音響特性を正確に再現できる通信技術が必要である。これは従来の2チャンネルステレオによる2次元の音場再生から3次元への拡張であり、従来のステレオと区別して、音場創成と呼ばれる。またこれを実現するための技術を音場制御技術と呼び、デジタル信号処理技術と密接に結び付いている(6)。

音場創成技術は、図5に示すように、ある特定の音場を環境まるごと再現する「室内音場創成」と、2つの耳元で音場を再現する「両耳音場創成」に分けることができる。多人数で音場を共有する場合は前者の方が有効であり、次のような方法が具体化されつつある。すなわち図6に示すように、再生音場としては反射音をできるだけ少なくしたデッドな部屋を選び、模擬したい音場の反射音構造を測定して、それに基づいて、反射音を表現するための仮想音源(壁

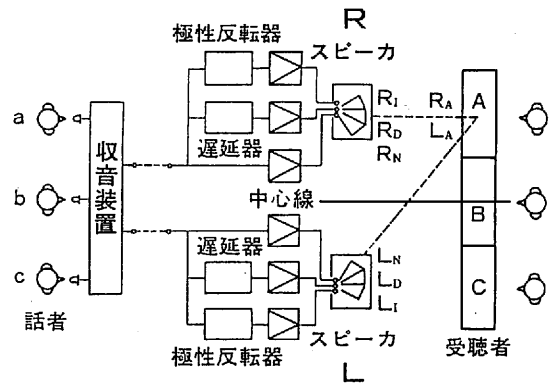


図4 音像定位ステレオ通信方式のブロック図

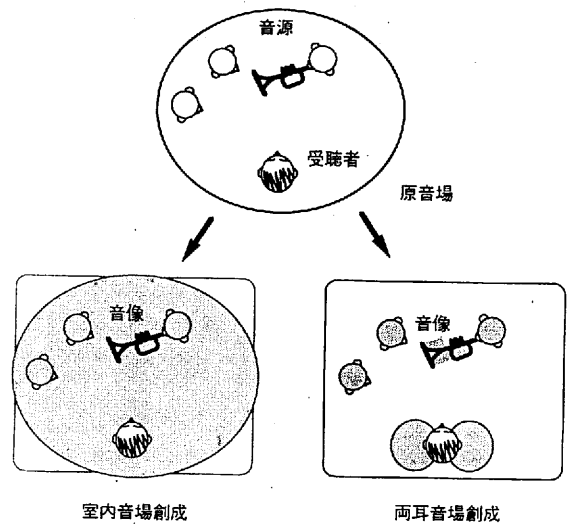


図5 音場創成の考え方

面を取り除いた自由音場内に存在を仮定した音源)を配置する方法である。ただし、この方法は映画鑑賞のようなAVへの応用には十分であるが、音声の明瞭性が重要でかつ定位感が重視される通信会議への応用の場合には注意が必要である。反射音構造の精密性が失われると、かえって明瞭性を損なう恐れがあるからである。

両耳音場創成はパーソナルな状況に適しており、その代表例はヘッドフォンで実現するバイノーラルステレオである(図7(a))。これは、実際の耳元に伝わる音をそのまま保存するか、あるいは合成し、それをヘッドフォン再生系で受聴する方式である。この考え方の延長として、スピーカで実現するトランスオーラルステレオがある(図7(b))。これは、耳元で実音場を再現するステレオ系のことで、耳元で目的の信号が得られるよう加工された信号を、空間的に離れた複数のスピーカで再生する方法である。トランスオー

ラルステレオは、膨大な計算処理が必要なため、まだ提案のみの段階であるが、今後有力な両耳音場の創成手段となろう。

音場制御技術の今後の新しい展開のためには、先行音効果など、人間の両耳聴覚特性の活用が必要である。

(3)高品質・広帯域サウンド

高精細・広視野映像の効果と同様に、音(サウンド)の高品質・広帯域化により、定位感や広がり感を含む臨場感の向上、さらに自然感や迫力の向上効果が得られる。今後は、デジタル符号化技術により、音声に関しては7kHz帯域、音楽を含むサウンドに関しては20kHz帯域の高品質伝送再生が普通になるであろう。

4.3 ハイパーメディア

従来リニアまたはトリー状の構造で記述されていた情報を、関連する情報間にリンクを張ってネットワーク構造にしたものがハイパーテキストであり、リンクをたどることによって、次々とリンクのノードにある関連情報を参照することができる。基本的には、グラフィックスを駆使した表示画面、アイコン、およびマウスをインタフェースとするGUI(グラフィカル・ユーザー・インタフェース)をベースにしながら、これに情報をネットワーク化して整理する手段を付け加えている⁽⁹⁾。コマンド等の操作手続きではなく、取り扱いたい物・対象を中心に、操作方法やシステム全体の仕組みを考える、オブジェクト指向・ユーザー・インタフェースの一種である。

このハイパーテキストの扱う情報が文章(文字列)であったのに対して、図形、画像、音声などをノード情報として扱うようになったのがハイパーメディアである。これにより異なるメディア間に直接リンクを張ることができ、画像や図形に対してキーワードを付与しなくても、直感的に次々に情報を検索(参照)する直接操作(ダイレクト・マニピュレーション)が可能となる。

アップル社が提案したコンピュータの未来像の一つであるナレヅナビゲータ(ボイスナビゲータ)は、タッチパネルと音声認識、応答によるインタフェースを持ち、小型ビデオカメラと電話機能を有しており、ハイパーメディアの一例と言える。ハイパーメディアによる大規模なデータベースの構築と情報検索に関しては、まだまだ難しい課題が多いが、今後大いに期待される技術であり、文書管理、文書作成、電子出版、オンラインマニュアル、CAIなどに、今後広く用いられるようになると思われる。将来は、オブジェクトに知性、すなわち自律的行動能力や学習能力を持たせた、知的オブジェクトへの発展が期待される。

4.4 認識・理解・合成技術

(1)基本技術

人と人のコミュニケーションと同様に、メディアを自由に、また選択しながらコンピュータとコミュニケーション

できるようにするためには、次のような技術の確立が必要である⁽²⁾。

- (a) 人がコミュニケーションする際に用いる音声、文字、図形、画像等のメディアで表現された意味、概念、意図、感性等を、コンピュータが理解できる言語(コード情報)に変換する各種の認識・理解技術、さらには人の身振りや視線の動きを検出、計測する技術や、それが何を意味するかを理解する技術(ヒューマン・イメージリーダ、図8参照)。

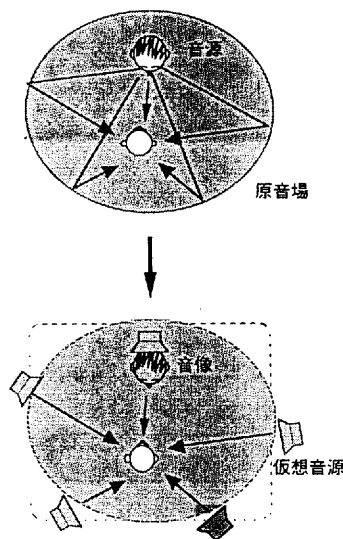


図6 室内音場創成の方法

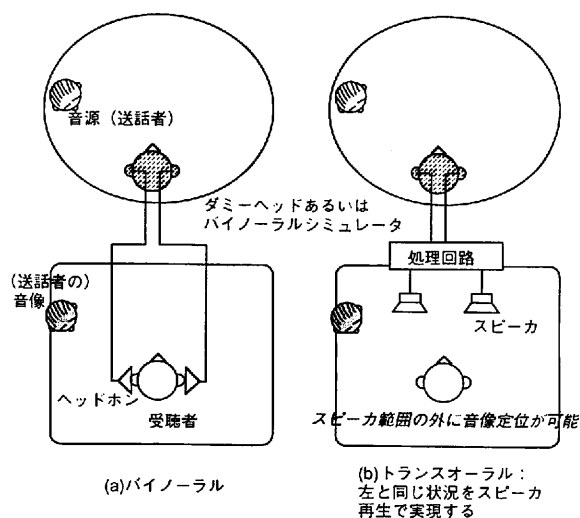


図7 バイノーラルとトランスオーラルステレオによる両耳音場創成

- (b) コンピュータが人に伝えたい意味、概念、意図、感性等(コード情報)を、文字、図形、画像、音声等、様々なメディアに変換するメディア生成、合成技術。
- (c) 複数メディアを統合して認識・理解するメディア融合(メディアフュージョン)技術。

人に近いインタフェースを実現するためには、知識表現、推論、学習等に係わる人工知能の研究や、人の行動、思考過程のモデル化等を行う認知心理学の研究の発展も必須である。

(2)文字認識技術

文字認識技術を読み取り対象文字の種類という面から見ると、数字、英字、仮名(カタカナ、ひらがな)、漢字等に分類でき、後者は互いに似た文字が多くなるので、より高度な技術が要求される。また、文字変形の程度という面から見ると、印刷、制限手書き(ガイドラインに沿って書くなど)、常用手書き(JIS等の見本字形にならって書く)、自由手書き(文字記入枠内に書くものから枠のないものまで)の順に変形の多様性が増すので、高度な技術が必要になる(10)。

種々の制約条件下では、文字認識は実用化の第一歩に達しているが、人間の認識能力に比べればまだまだ幼稚である。種々のフォントを自由に認識できる技術や、筆記制限を緩和して自由に書いた文字でも認識できる技術の開発が望まれている。このためには、文字や語の意味を解釈しながら認識する手法が必要であり、知識処理を取り入れた方法が研究されている。オンライン手書き文字認識は、漢字を入力する簡易な手段として、広汎な普及が期待されている。

(3)文字生成技術

コード情報(概念)が数字や記号を含む言語で記述されている場合は、単なる文字発生装置があれば文字生成は可能であり、技術的課題はほとんどない。すでに各種の文字発生ROMが安価に製造、販売されている。

(4)音声認識技術

音声は人間にとって最も自然な情報伝達手段であるので、音声の自動認識が可能であれば、コンピュータによる各種の予約、情報検索、問い合わせなどが電話などを用いて容易に行えるようになり、音声入力ワープロも可能となるなど、期待される効果は極めて大きい。音声認識技術は、認識対象が区切って発声された単語であるか、文章であるか、認識対象語彙数は何単語(10~数万語)であるか、発声者が固定されているか(特定話者)、固定されていないか(不特定話者)、などによって難しさが異なる。

音声は草書(行書)で書かれた文字列で例えることができ、単語や文章を発声したときの各音素は、時間的に伸縮する上に、いわゆる調音結合のために、前後の音素の影響を受けて大きく変形する。語彙数の小さい単語音声認識の

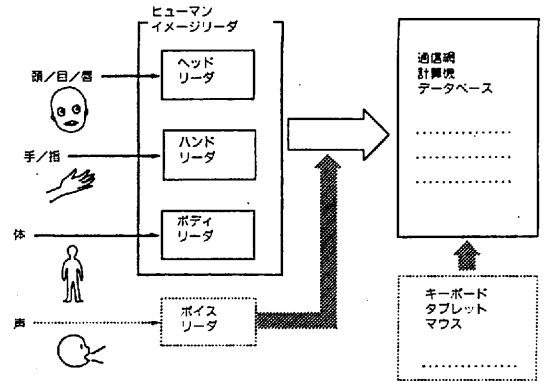


図8 ヒューマンイメージリーダ

場合は、単語音声をそのまま蓄えておくことによって、調音結合の問題を回避できるが、大語彙の単語音声認識や文章音声の認識の場合は、調音結合による変形を考慮した多種類の音素単位を用いることが多い。人間が文章を発声するときには、文法、意味、文脈、タスク(話題)に依存した一般常識、などの言語的制約を無意識のうちに考慮して発声し、聞く側もこれらの言語的知識に基づいて次の言葉を予測しながら聞いている。音声認識においても、これらの言語的知識を、音としての情報に適切に組み合わせて用いることが必要であり、このための研究が行われている。不特定話者の音声の認識のためには、声の変動を統計的モデルで表現する方法や、声の変動に自動的に適応できる方法の研究が行われている。

タスクをある程度限定すれば、高い性能で文章音声で認識できる装置に近い将来に実現されるであろう。ただし、人間どうしの会話音声は、書き言葉に比べて極めて大きなあいまい性を有しており、自然な会話音声でのコンピュータとの対話を実現するためには、まだ研究すべき課題が多い(11)。

(5)音声合成技術

コンピュータを用いた種々の情報案内システムにおいて、情報内容を音声の形で多くの人に伝える手段として、音声応答装置に対する需要は極めて大きい。電話による各種情報案内サービスが、その最も典型的な例である。例えば駅の案内のように、出力すべき語彙や文章の種類が限定されている場合には、アナウンサの声を録音したものをつなぎあわせて出力する録音編集方式、あるいはその音声を分析してパラメータの形で蓄えておくパラメータ編集方式の方法で対処できる。これらの方式に関しては、すでに1チップの音声合成器も作られ、自動販売機、電化製品、玩具等に組み込む安価なものから、高品質の音声を出力する装置まで、種々の製品が実用化されている。

任意の文章を音声に変換するテキスト合成装置(文・音声変

換装置)の高品質化のためには、書かれた文章を人間が朗読する能力(言語解析と発声機構)を模擬することが必要であり、これらに関する種々の研究が行われている。音声認識装置と組み合わせた音声認識・応答システムである、バンキングサービス用のANSERシステムでは、定型文はパラメータ編集方式で、会社や人の名前はテキスト合成方式で出力されている。将来、自動翻訳電話を実現するためには、任意の文章(会話)音声の認識と合成ができるようになることが必須である(11)。

(6)図形・画像認識

図形・画像の特徴抽出、構造分析を行い、図形・画像の内容把握を行うのが図形・画像認識である。画像の特徴抽出処理は、図形から濃淡画像、自然画像、3次元シーン(情景)にいたる各種の画像に対して、古くから研究がなされており、細線化、曲線の記述、領域形状の記述、2値図形化、輪郭線抽出、テクスチャ解析、領域分割等が研究されている。これら抽出された特徴を用いて、画像から対象物を検出し、色、形、テクスチャ等を用いて、対象物を分類・認識する。さらに画像認識から一歩進んで思考、推論を行い、画像の内容を解釈して記述することを画像理解と呼ぶ。

図形・画像の認識技術は、対象をかなり限定した場合に初めて可能になるものであり、あらかじめ認識対象を登録しておいて、それとのマッチングにより認識を行うものに限られている。範囲を広げると、認識が極端に困難になるのが現状である。今後の研究課題は、画像を解析して対象世界に関する知識に基づいて推論を行い、その画像が表す実世界の情景に関する記述を作成できるようにすることである(10)。

(7)図形・画像生成

コード情報(概念)が、具体的に幾何学的コマンド(点、直線、円弧等の集合とそれらの色等を規定する属性情報)等で記述されている場合には、図形・画像生成技術は実用化レベルに達しているが、コード情報が少しでも高度な概念レベルになると、図形・画像生成は極端に困難になる。しかし最近では限定した対象について、モデルに基づいたCA(コンピュータアニメーション)が研究されるようになっており、音声に合わせた顔画像の生成等が試みられている。今後は、より複雑で質感のある図形を高速かつ安価に生成するためのハードウェア、ソフトウェア技術の開発、人工知能(知識処理)の導入などが重要になってくるであろう(10)。

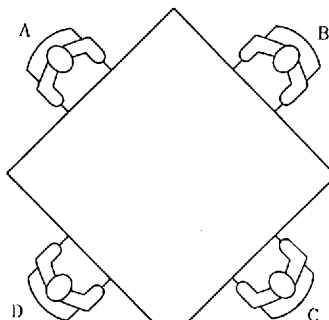
4.5 人工現実感とテレプレゼンテーション

人工現実感(アーティフィシャルリアリティ)あるいは仮想現実感(バーチャルリアリティ)は、CG(コンピュータグラフィックス)等の技術により、全く仮想の世界、あるいは仮想と実世界を融合した3次元世界を人間の周囲に作り、その世界があたかも実世界であるかのごとく人間がアクションをとれる環境を提供するものである。テレプレゼンテーション

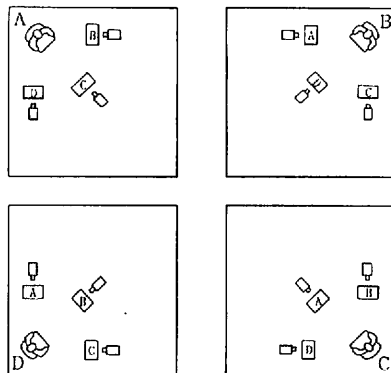
スは、人間が離れた場所にあるロボットを操作する際、あたかもそのロボットになりきったような感覚で、操作あるいは行動できる環境を提供するものである。このような技術を用いて、現実にはあり得ないような場をシミュレーションにより設定し、体験することにより、人間の創造性を飛躍的に助長することができるものと期待されている。これらの詳細に関しては、本講演会において館氏から詳しく説明がなされるので省略するが、具体的な場面として、次のような仮想空間での会議が研究されている。

4.6 仮想空間会議方式

大画面・高精細、立体表示による映像通信と、部屋の音響特性を含む音場の通信によって、通信者の存在する空間の臨場感は相手に伝わる。しかし、通信者は互いに異なる空間を見て、互いに異なる空間からの聴覚情報を聞いており、あたかも同一の場に居合わせて情報交換をしている感じとは異なっている。通信相手が複数箇所にわたっている場合には、さらに異なる空間が多くなり、臨場感はさらに希薄となる。このため、あたかも同一の空間を共用して情



実際の会議



仮想会議システムの構成

図9 BNR(Bell Northern Research)の多対地仮想会議システムの構成

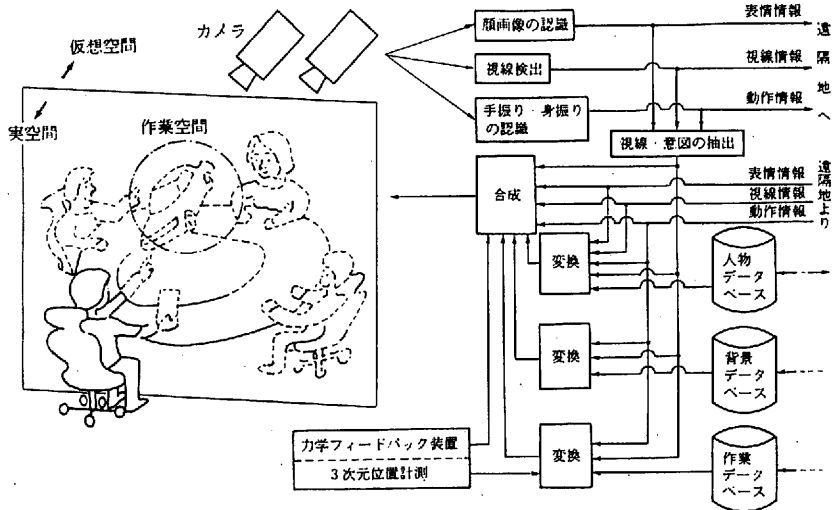


図10 臨場感通信会議のイメージ

報交換を行っている感覚を生み出すための研究が行われている。

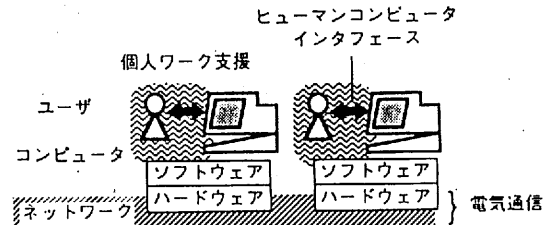
図9は、単に遠隔地間に情報バスを設定するという立場ではなく、遠く離れた人達を一箇所の仮想空間に集めて情報交換を行う、多対地仮想空間会議システムの構成を示したものである(12)。テーブルを囲んで4人で会議をしている環境を作り出すために、四角いテーブルの3面にそれぞれ1個のテレビモニタを設け、モニタの上に各々1個のテレビカメラを設置して、対地毎に異なる角度から撮影した映像が伝送されて来るようにしている。図10は、3次元CGで生成された仮想の会議室に、高度な画像処理技術を用いて、異なる場所にいる会議参加者を写し込む方式を示したものである(13)。

仮想空間会議方式は、①あたかも同一の場所で会議をしているような感じが得られる。②立体視が容易に実現できる。③会議参加者相互の視線の一致が可能になる。④見たい方向の映像を容易に表示できる。⑤3次元映像のシミュレーションが同一空間で実現できる。⑥通信対地数に依存しない自然な会議ができる等、多くの特徴を有しており、臨場感通信のための多くの要素を含んでいる。しかし、任意背景からの人物像抽出、人物像の動き認識、3次元CG、仮想空間内操作、立体表示等、仮想空間会議実現には多くの困難な課題が残されている(4)。

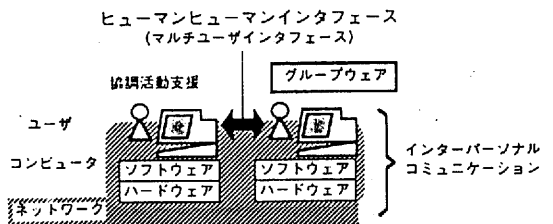
仮想空間における音場創成も今後の重要な研究課題であり、仮想音源の生成法、音源の移動の表現や頭部の動きに対する補正などの音場制御技術の発展が期待されている(5)。

4.7 協調活動支援システム「グループウェア」

組織における人間の活動は、基本的に複数の人や組織に



(a) パーソナルコンピュータ環境



(b) CSCW環境

図11 従来のパーソナルコンピュータ環境とこれからのCSCW環境

またがるグループワーク、すなわち協調活動である。このような人間の協調活動の、コンピュータによる支援に関する新しい学際的研究分野であるCSCW(computer-supported cooperative work)が、今後のISDN時代におけるヒューマンインタフェースの新しい視点として注目されている。図11に示すように、従来、人間とコンピュータとのインタフェースは、シングルユーザ対コンピュータという視点が

中心であったが、CSCWでは、コンピュータとネットワークにより媒介された人間どうしのダイナミックなインタラクションという視点に立っている。最近、高速LANやISDNの普及、電子メールシステムの発展と相まって、協調活動支援システム「グループウェア」の研究開発が活発化して来ている(14)。

CSCWの詳細に関しては諏訪氏のご講演があるので省略するが、上で述べた人工現実感、仮想空間会議の概念とも密接な関係を有しており、今後の展開が期待される分野である。

5. むすび

21世紀の情報通信の目標を一言で言えば、臨場感通信であるということができる。すなわち、人間にとって真に必要な臨場感を経済的に表出し、経済的に通信できる技術の開発や、臨場感の重要性を最大限に引き出すサービス(使用法)の検討が今後ますます重要になる。このためには、メディアの特徴を十分に生かした文字、図形、画像、音声、サウンド等のマルチメディアの使い分けや、さらにはメディアを意識しないで使えるメディア融合のなされたシステムが必須であるが、このようなシステムはまだ実在していない。メディアを意識することなく自由に使い、効率的でしかもユーザフレンドリーなマルチメディアインタフェースの実現を目指した、種々の革新的な挑戦が期待されている。

さらに、21世紀の通信においては、情報の伝達のみではなく、情報の共用と分配、さらに情報の認識、理解、推論を行う知識情報処理の役割が重要な鍵になるであろう。知的処理と電気通信の融合によるUICN(総合知的通信網)時代の幕開けである。このような新しい展開の究極のインタフェースは、人間の思考を支援するインタフェースであろう。

なお、インタフェース技術を考える上で忘れてはならない重要な視点に、人間にとって快適であるかどうかということがあがるが、これに関しては本講演会において、佐治氏から詳しい説明があるので省略した。

謝 辞 種々の資料を提供して頂いた、当所小林幸雄視覚情報研究部長に感謝する。

文 献

- (1) 秋山健二、小林幸雄、安田浩:21世紀の電気通信-臨場感通信-、信学誌、73、12(1990)
- (2) 小林幸雄:近未来のマルチメディア・マルチモーダルインタフェース、画像ラボ、1990.10、pp.19-22(1990)
- (3) 青山友紀:21世紀に向けた研究開発の展望、NTT技術ジャーナル、1990.10、pp.44-49(1990)
- (4) 小林幸雄:臨場感通信、テレビ誌、45、4(1991)(予定)
- (5) 小泉宣夫:音場制御(音場の創成)、テレビ誌、45、4(1991)(予定)
- (6) 元木紀雄:立体テレビジョン、テレビ誌、43、8(1989)
- (7) 畑田豊彦、坂田晴夫、日下秀夫:画面サイズによる方向感覚誘導効果、テレビ誌、33、5(1979)
- (8) 宮田裕之、青木茂明:聴覚の話題-音像定位-、信学誌、72、8、pp.860-864(1990)
- (9) 明日の情報社会を支える技術-ヒューマン・インタフェース、日経BP9誌合同特別編集版、pp.123-144(1990)
- (10) 小林幸雄:マルチメディア・インタフェース、1990年度人工知能学会全国大会、T6(1990)
- (11) 古井貞熙:デジタル音声処理、東海大学出版会(1985)
- (12) S.Sabri et al.: Video Conferencing System, Proc. IEEE, 73, 4(1985)
- (13) 竹村治雄、岸野文郎:人工現実感によるヒューマンインタフェース、テレビ誌、44、8(1990)
- (14) 石井治:協調活動支援とヒューマンインタフェース、テレビ誌、44、8、pp.967-972(1990)