

国際シンポジウム第1日目 講演5

人とコンピュータを結ぶ科学と技術

東倉 洋一*

1. はじめに

コンピュータをはじめとする機械の存在は、日常生活においても当たり前の存在になった。この傾向は、21世紀には、さらに顕著なものとなろう。コンピュータの日常化の大きな要因の一つに、その劇的な価格低下と性能向上がある。この2つの要因の同時進行によって、コストパフォーマンスは数年で10倍を越えてしまう。

これだけ身近になったコンピュータであるが、使い易さに関しては、ほとんど変化が見られない。機能の複雑化によって、むしろ使いにくくなったとさえいえる。コンピュータや情報機器が私たち人間にとって、よりよきパートナーとなるためには、人とコンピュータの自然なコミュニケーションを可能にするインタフェース、人とコンピュータを結ぶ科学と技術の進展が不可欠である。

2. 人とコンピュータのコミュニケーション

人とコンピュータのコミュニケーションの理想の形態の一つは、コンピュータができる限り人に近づくことであろう。人と人のコミュニケーションの基本は、双方向 (bi-directional) で複数感覚を駆使する (multi-modal) にある。しかし、日常のコミュニケーションにおいて、種々のコミュニケーションの壁が存在し、この基本が全て満足されることは稀である。異言語間や異文化間では、言語や文化の壁によって、円滑なコミュニケーションは難しい。電話などの情報通信手段を介した場合には、使用できる感覚が技術的に制限されてしまう。人とコンピュータのコミュニケーションの現状は (図1)、この双方向、複数感覚の基本要素のうち限られた部分が実現されているに過ぎない。これらの基本要素を技術的に実現し、コミュニケーションの壁を取り去ることが人とコンピュータのコミュニケーションの理想となる。

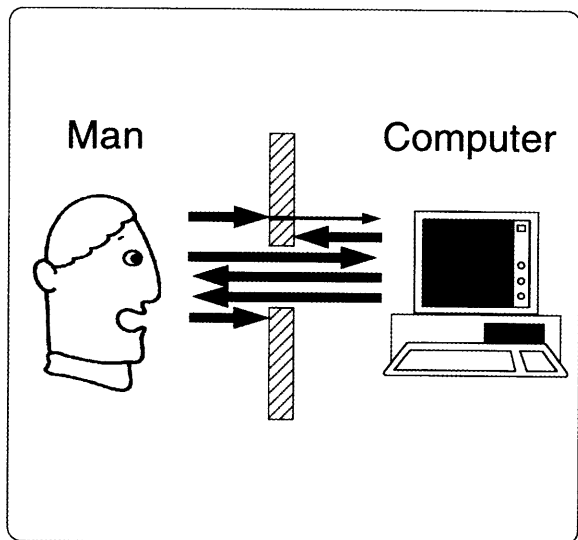


図1

3. 人間情報処理研究のアプローチ

人とコンピュータの理想的なコミュニケーションの基本要素の実現には、人のコミュニケーションの基本となる「人」そのものの研究、特に、人間、いいかえれば感覚や脳の情報処理メカニズムの解明が重要となる。本稿では、これらの重要性を以下の視点から述べる。

- (a) 人(脳)を知り、人を創る
- (b) 脳で見て、脳で聴く
- (c) 脳に学ぶ、脳でまねる

これらの視点で研究を進める上で、前提として、人間とその情報処理を以下のようにとらえる。

人間の情報処理において脳が最も重要な役割を果たすことは言うまでもない。しかし、人間の総体は、脳と心と体の三位一体として構成されている(図2)。そして、これらがインタラクティブでバランスのとれた関係を保ち、外界に対して優れた適応性を持つ複雑系(adaptive & complex)を形成している。また、脳の働きに関して極めて特徴的な性質として(図3)、その能動性がある。見たり聴いて理解するためには、視覚や聴覚によって外界の情報を取り入れ、これを脳に送る必要がある。しかし、脳は、これらの情報だけから対象を理解する訳ではない。むしろ、重要なのは、経験や学習によって予め脳に蓄積された知識である。外界からの情報がきっかけとなって、脳は膨大な知識を駆使した能動的な働きを自ら開始する。場合によっては、極めて限られた情報から大胆な予測を行うことによって、対象の正確な認識を可能とする。

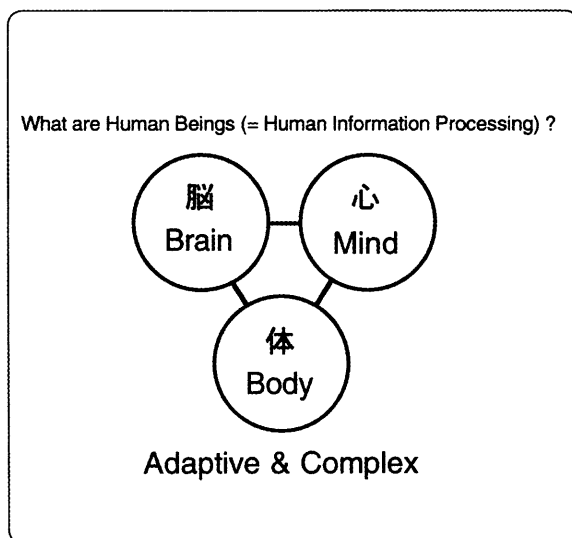


図2

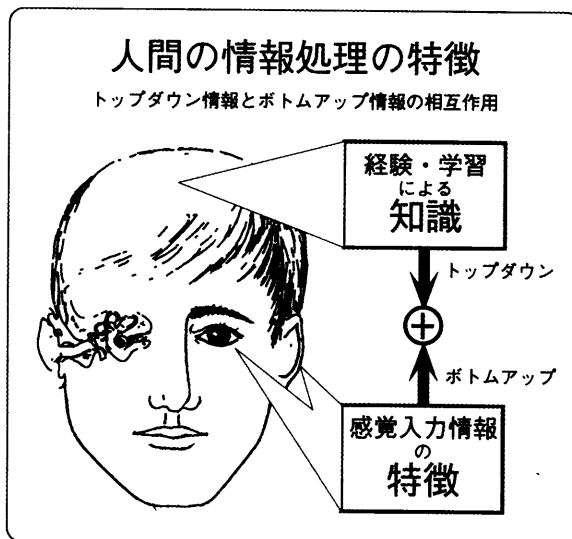


図3

3.1 人(脳)を知り、人を創る

人間の情報処理メカニズムの解明、すなわち「人(脳)を知る」ためには、どのような研究の視点が必要だろうか。最も重要な視点として、本稿では、以下の三つを議論する。

- (a) 複数感覚情報処理 (マルチモーダル処理)
- (b) 生成と知覚の相互作用
- (c) 知覚の能動性

人間の情報処理の総体は極めて複雑かつ膨大である。したがって、このような対象の解明には、「人を知る」研究だけでなく、これを「人を創る」研究と組み合わせた方法論が有効と考える(図4)。すなわち、人間情報処理の部分的な側面に関して、「人を知る」研究によって得られた成果と、これに基づいた大胆な仮説によって、「人を創る」ことを実現する。コンピュータやロボットなどを使ったシミュレーションによって創られた人間の機能は、直ちに完全な場合はない。ある機能が実際の機能に近い場合もあれば、懸け離れて異なる場合もある。しかし、この創られた人から学ぶものは、結果としての性能ではなく、「人を知る」研究において、次に何をどのように攻めるかという研究のポイントである。このような「人を知る」「人を創る」研究の循環によって、人間の情報処理の総体が段階的に解明されるものとする。

聴覚や視覚など種類の異なった複数の感覚が同時併行して働く複数感覚情報処理(マルチモーダル処理)は、人間の情報処理の本質である。これは、外界の情報を取り入れ、これを認識しようとする場合、できるだけ多くの感覚を総動員し、正確な認識結果を得ようとするためである。具体例を上げよう。物をつかむ動作(図5)では、まず、つかむ対象である物を見る視覚と手を伸ばしてつかむ運動の二つの感覚が協調的に働くことを必要とする。自分でいろいろな形の物をつかもうとするときの手の形に注目してみよう。手が物に近づくにつれて、無意識のうちに手の形がつかもうとする物の形にフィットした形になることがわかる。これは、脳において、視覚から得た物の形の情報からこれをつかむために最適な手の形の情報が作られ、この情報が運動神経を動かすことによって、実際の手の形が作られたためである。

また、音声の聞き取りにおける聴覚と視覚の相互作用も特徴的である。読唇術のように、唇の動きを積極的に利用することが

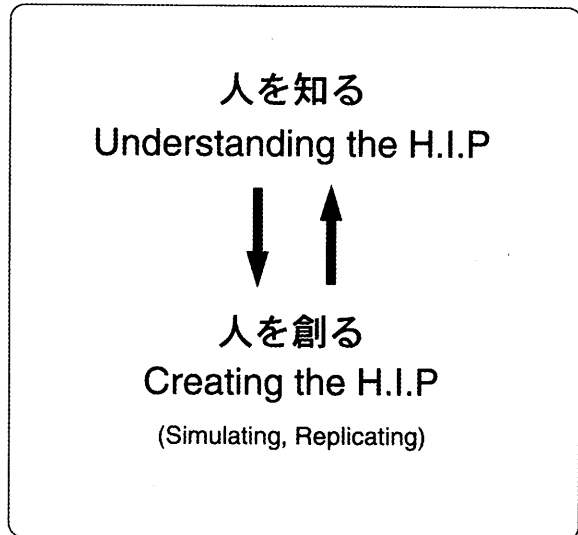


図4

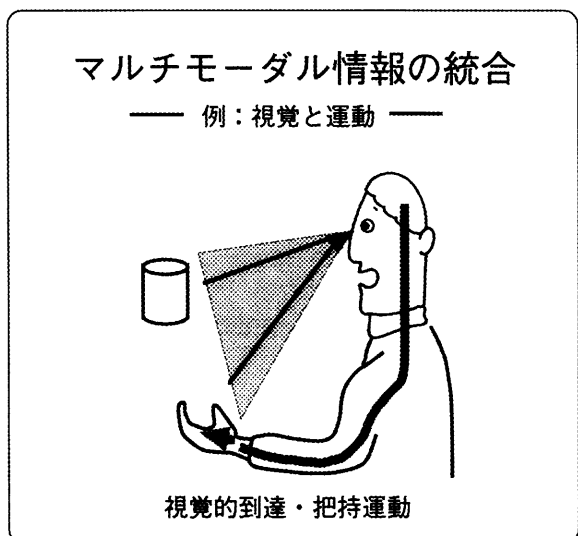


図5

できることはよく知られている。しかし、通常、ほとんど意識的には使っていないと思われる唇の動きが、音声の聞き取りに相当の影響を与えていることを示す実験結果がある。雑音の中で音声を聞き取る場合を想定しよう。雑音の影響で30%程度しか聞き取れない場合にも、唇の動きを見ながら聞き取ると70%を越えてしまう。このように唇による視覚情報は、音声という聴覚情報と融合し、これを補う働きを持つ。さらに興味深いのは、この視覚情報と聴覚情報の融合を逆手にとった実験がある。この実験を初めて行った心理学者の名を取ってマガーク (McGurk) 効果と呼ぶ (図6)。この実験では、まず、「ga」を発音する顔をビデオで録画する。そして、ビデオの音声だけを「ba」の音に入れ換える。これを再生すると、「ba」でもなく「ga」、2つの音の中間的な性質を持つ「da」の音に聞こえることが多い。すなわち、音声の聞き取りにおいて、聴覚と視覚という異なった二種類の感覚から送られた情報が脳で融合したことになる。

ここでは上げた具体例からも分かるように、マルチモーダル処理は、人間の情報処理の本質であり、この視点を強く意識した研究の進め方が重要となる。

マルチモーダル処理と同様に重要なのが、生成と知覚の相互作用である。生成とは、人間 (脳) が情報を創り、外界に送り出すことであり、情報の出力である。知覚とは、情報の入力であり、外界の情報を感覚を通して脳に取り入れ処理することである。生成と知覚の具体例は、話すことと聞くことである。生成と知覚には密接な関係があり、人間の情報処理は、生成知覚連関システムと考えた方がよい。

話すことと聞くことを例にとろう。私たちが話すとき、話し手であることは勿論であるが、自分の声の聞き手でもある。自分の声を聞くという聴覚フィードバック (図7) は非常に重要であり、この聴覚フィードバックによって、話すことに必要な種々の情報のコントロールが行われている。もし、この聴覚フィードバックを人工的に異常な状態に置いた場合には、正常な話し方が不可能になることが実験的に知られている。

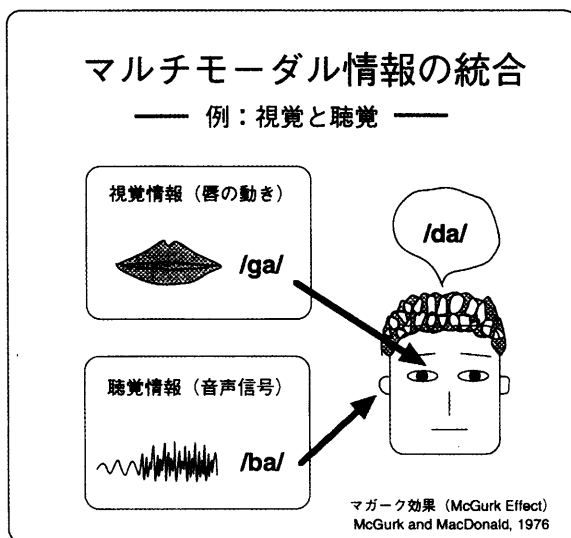


図6

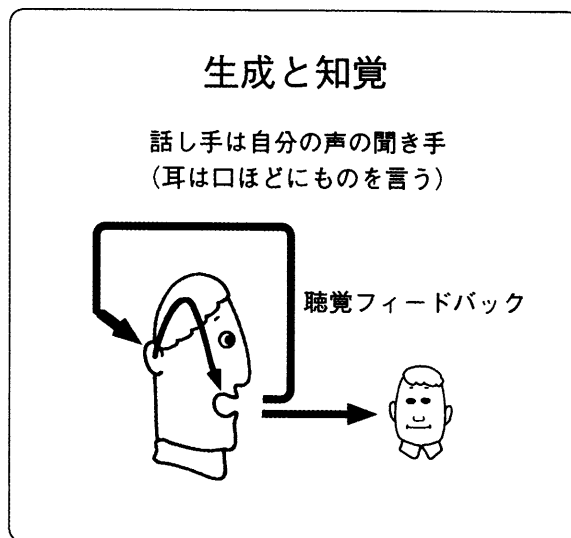


図7

実験では、自分の声が空气中を伝わって耳から入力される通常の聴覚フィードバックに代って、マイクロフォンからヘッドフォンを介して人工的にフィードバックする(図8)。このとき、途中に時間遅れを挿入すれば、通常のように自分の声がリアルタイムでは聞こえず、一定時間遅れて聞こえることになる。時間遅れを約200ミリ秒に設定した場合、どもるなど話し方に大きな乱れが起こる。話し方の乱れと時間遅れの程度には密接な関係があり、200ミリ秒より短くなる場合にも、長くなる場合にも、乱れの程度が減少する。このように、話すことと聞くことは密接に関係しており、この関係には時間的要素が重要な役割を持っていることがわかる。

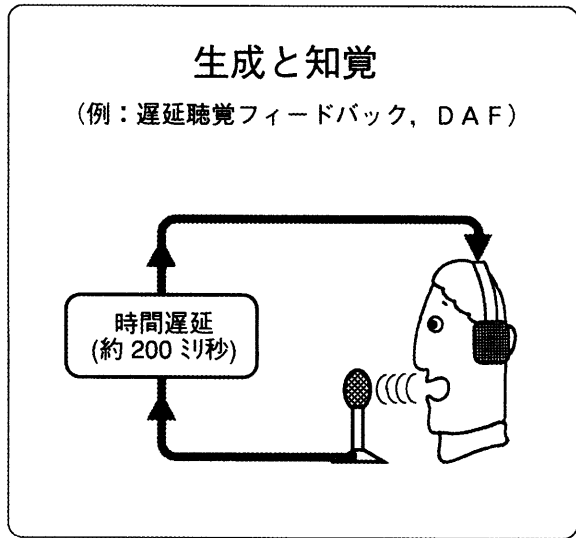


図8

聴覚や視覚などの感覚器は、通常、受容器と呼ばれるように、外界の情報を受け入れて脳に伝達する受動的な機能だけを持つものと考えられてきた。しかし、感覚器は自ら外界に働きかけ能動的に情報を取り入れる機能を持つ。これは感覚器が独立に機能する末梢器官ではなく、脳に直結した、いわゆる出先器官であり、脳と一体となって働くからである。聞き耳を立て、多くの音の中から自分の聞きたい音を選び出して聞くカクテルパーティ効果などの選択受聴や視線や視点を変えて見たい物に注目する行動は、感覚の能動性を表す具体例である。

3.2 脳で見、脳で聴く

前項で述べたように、感覚器官は脳の出先器官であり、情報処理の主体は脳にある。したがって、視覚や聴覚の働きも、目で見て耳で聴くというよりは、脳で見、脳で聴くという方が実態に合っている。

したがって、感覚によって得られる外界の情報が限られた場合にも、あるいは情報の一部が失われた場合にも、脳は、経験や学習によって蓄積された知識を用いることによって正確な認識を可能とする。立体視のメカニズムに具体例を見てみよう。

私たちを取り囲む3次元の世界で奥行きや立体を知覚するためには、左右の両眼視が重要な働きをする。しかし、この両眼視以外にも立体視の手がかりは多く、陰影や線の交わりなどがある。これらの手がかりを使った立体視のメカニズムでも脳が重要な働きを持つ。まず、3次元の世界は眼球の光学的な構造によって網膜に像を結ぶ。網膜は近似的に平面であるから、網膜像は2次元となり、外界の3次元構造は、情報の入口である眼球において2次元に変換されてしまう(図9)。この2次元の情報が脳に伝達され、脳の情報処理によって3次元情報である立体や奥行きを知覚するためには、2次元から3次元への変換が必要となる。この変換は、眼球における3次元から2次元への変換の逆変換にあたる。しかし、この逆変換は変換ほど単純ではなく、解が一意に定まらず、複数の解が求まる。これを数学的には不良設定問題と呼ばれ、どの解が真の解かを定め、解を一意に絞るための拘束条件が必要となる。

具体例として、陰影を手がかりとして、逆変換によって2次元情報から3次元情報を得るためには、光の方向が拘束条件となる。例えば、垂直面に凹凸があり、この凹凸が陰影で表現されているとしよう。この場合、下側に影があれば中心が凸、上側に影があれば中心が凹と知覚するのが、脳における立体視の働きである。これは、光は上方からという前提で、中心が凸であれば下側に影ができ、中心が凹であれば上側に影ができるという日常生活の経験と学習から学びとった知識が拘束条件となっている。

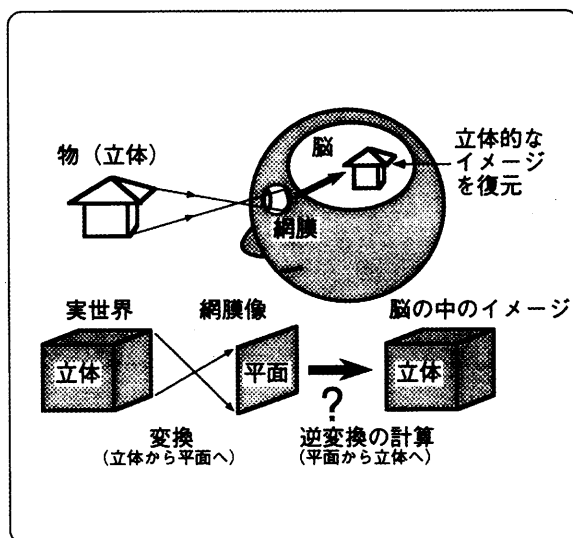


図9

3.3 脳に学ぶ、脳でまねる

人間の情報処理において学習は非常に重要な要素である。視覚と運動による協調的な学習を具体例にとろう。腕や身体の動きの学習では、手本となる動きを目で見るのが重要である。この学習は、日常的に私たちが行っている「見てまねる」ことによる学習であるため、「見まね」学習と呼ぶ（図10）。

見まね学習では、例えば腕の動きを目で見て、その運動情報を得る。この運動情報の解析によって、運動に重要な特徴点を抽出する。この特徴点に基づいて、同様な運動を自分の腕で再現するためには、腕の筋肉をどのように動かすかを決定するための運動指令を脳で創り出す必要がある。すなわち、特徴点から運動指令への変換である。これは、ある運動を創り出す場合に、通常行われる運動指令から特徴点への変換の逆変換に相当する。この逆変換を行うには、脳に逆モデルが必要となり、この逆モデルがあれば、特徴点から運動指令が計算できる。

逆モデルの計算によって得られた運動指令を筋肉に与えれば、腕が運動指令に従った動きを行う。通常、運動の学習が未熟な初期的な状態は、逆モデルが未熟であり、特徴点から運動指令を精度よく計算できない。このため、実現された運動も目標と異なる結果となる。この未熟な運動を目で見ることによって、特徴点の誤差を検出し、これに基づいて特徴点の抽出と逆モデルを修正する。修正後の逆モデルから計算される運動指令によって再び腕を動かす。これらの過程を繰り返すことによって、目標とする運動と実際の運動の誤差が減少し、最終的に学習が完了する。図10はロボットの

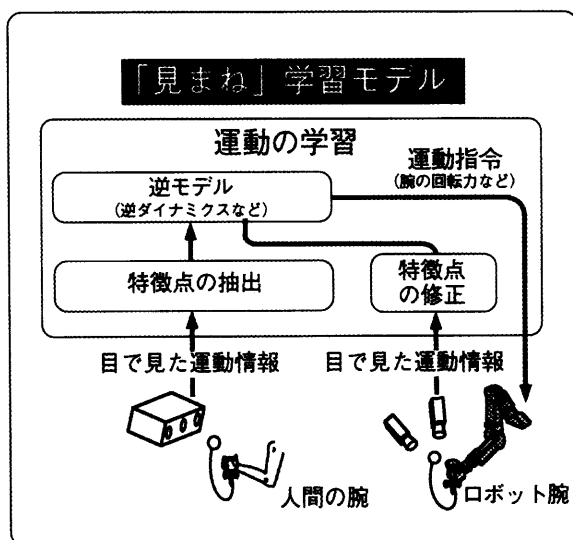


図10

腕に見まね学習を適用する場合を示している。

見まね学習は、人間の運動学習の本質の一つをモデル化したものと考えている。この学習をロボットに適用することは、運動の軌道を数値的に与えてロボットを制御する方法とは全く異なる。見まね学習では、視覚的に運動の情報を与えるだけで、ロボットは自己学習を行うことによって目標運動を達成することが可能となる。学習は柔軟的なだけでなく、効率的でもある。現在、見まね学習によって、ロボットがけん玉をはじめとする数種類の運動の学習が可能であることが実験的にも検証されている。

4. おわりに

人を知り、人を創ることによる人間情報処理の研究によって、人間情報処理の本質が段階的に解明されつつある。これらの研究成果に基づいた最先端技術によってシミュレーションされる人とコンピュータのコミュニケーションにおいて重要なのは、

- (a) Efficient (効率的)
- (b) Effortless (ユーザである人の努力なしに)
- (c) Empathy (感情移入)

の三つの要素、Eキューブと考えられる。これらの要素を満足させる人とコンピュータのコミュニケーションを実現するために解決すべき課題は多い。しかし、その基盤となる人とコンピュータを結ぶ科学と技術は、コンピュータ技術の更なる発展や脳科学研究を中心とする国際的かつ学際的で組織的な取り組みによって、今後も着実な進展が期待できる。