

2 状況共有コミュニケーション技術の実現に向けて

2-1 身体性に基づくコミュニケーションメカニズムの研究

2-1 *Research on Communication Mechanism of Embodied Interaction*

矢野博之 藤原義久 前川 聡 小嶋秀樹 善本 淳

Hiroyuki YANO, Yoshi FUJIWARA, Satoshi MAEKAWA, Hideki KOZIMA, and Jun YOSHIMOTO

要旨

我々は、けいはんな情報通信融合研究センターで、身体性に基づくコミュニケーションメカニズムの研究を開始する。この研究は生理・認知・行動の三つの分野で包括的に推進される。生理分野では、独立成分分析を用いることで筋電位信号からそれぞれの指の伸展・屈曲という動きの抽出を行う。この指運動の識別を、音楽、発声などの様々な計算機インタフェースへ応用する。認知分野では、前言語的コミュニケーションの認知メカニズムの研究を行う。この研究では、言語を獲得するまでの子供のコミュニケーションの発達と、その発達のヒューマノイドロボットへの実現を並行して行う。行動分野では、人間のインタラクションにおける非言語情報の分析を中心に研究を行う。表情や頭部の動きから、会話の活性度や意思疎通の度合いを計量化することを目指す。我々はこれらの研究を推進することで、コミュニケーションシステムが社会的な存在になること、さらに、私たちの生活の質を改善するものになることを目指す。

We launch an investigation of communication mechanism of embodied interaction in Kehanna Human Info-Communication Research Center. We conduct our investigation in the following three fields, physiology, cognition, and behavior. In the physiological field, we extract the motion of each finger, extension and flexion, from electromyogram data by independent component analysis. We apply separation of fingers' motion to computer interface. In the cognitive field, we investigate the cognitive mechanism of the preverbal communication. We focus on both the development of childhood communication and its implementation on the humanoid robot. In the behavioral field, we study nonverbal information in human interaction. We measure the activity of communication and a degree of mutual understanding of intention using the motion of our face and head. We pursue our research so that the communication system becomes the social being and the quality of life.

[キーワード]

身体性, 筋電位, 前言語的コミュニケーション, 共同注意, 非言語情報

Embodiment, Electromyogram, Preverbal communication, Joint attention, Nonverbal Information

1 はじめに

我々は関西支所で、「人と対話のできるシステムを作る」というテーマを掲げコミュニケーションの認知メカニズムを軸とした研究を続けてきた。そのテーマに沿ってこれまでに我々が構築した対話システムは、計算機ディスプレイを通してテキストや音声ベースで対話をするものであった。電子メールシステム利用支援システム^[1]や Webdaileague 対話システム^[2]はまさに計算機ディスプレイ上のウィンドウ内に現れる対話システムとの対話である。また、対話システムを高度化するための研究である協調作業時における対話の分析^[3]、意味と文脈の計量化^[4]、言語の意味の安定性の分析^[5]だけでなく、心の理論に基づく対話理解^[6]も、その時点ではテキストを中心としたコミュニケーション研究であった。

我々は計算機上に対話システムを開発する一方で、システムに対して真の意味で人と同じように対話しているという感覚を持つことはできなかった。すなわち、我々はシステムに対して Dennett の言う設計的構え^[7]を取ってしまう。たしかに、特定の状況に限定された質疑応答形式の対話用にプログラムされた表現内での発話であるということ、心を計算機システム上に導入できていないことなども原因として考えられる。しかし、より大きな原因は小嶋が赤ちゃんロボット Infanoid 3^[8]で相手の顔を注視することを実現し、そして Infanoid 4^[9]^[10]で上半身を実現できた頃から我々に明確に捕らえられるようになった。すなわち、これまでの対話システムには身体性が欠如していたのである。

私たちと真の意味でコミュニケーションができるシステムを構築するためには、システムが身体性を持つことが必要である。私たちが他者とコミュニケーションを行うときには、常に何かの身体動作を伴っており、ことばとして相手に情報を伝える以外に様々な情報が身体動作により相手に伝えられる。対面性を持たない対話、すなわち電話やチャットをしている時でさえも、相手の動作を想像していることがある。我々は相手の動作を自然に自分の動作に置き換えることで、相手の感情や意図の理解をことば以外でも補足することができる。したがって、コミュ

ニケーションシステムに身体性を持たせることで、岡田の言うインタラクションの奥行き^[11]をつけることが可能になる。

以上のことから、我々はけいはんな情報通信融合研究センター（以下、けいはんなセンター）において、身体性コミュニケーションメカニズムの生理・認知・行動の三分野に関してそれぞれ以下の項目を中心として研究を行う。

- ①表面筋電位による指運動の識別及び応用
- ②前言語的コミュニケーションの認知メカニズム
- ③インタラクションにおける非言語情報

以下、**2**では表面筋電位による指運動の識別及び応用について述べる。この研究では、前腕の表面筋電位信号から指運動を分離・推定することを実現し、これをコンピュータインタフェースに応用する。特に、上肢切断者への適用では、被験者のほぼ意図どおりに動く仮想の手をディスプレイ上に表示することで、被験者はあたかも自分の指がそこにあるかのような感覚を持ったという身体性について興味深い報告を得ている。

3では前言語的コミュニケーションの認知メカニズムについて述べる。この研究では、幼児の前言語期からのコミュニケーションの発達に注目して、共同注意、間接経験、模倣という認知モデルの構築と、赤ちゃんロボット Infanoid 上にそれらを認知モジュールとして実現するという二つの相補的な研究を並行して進めている。けいはんなセンターでは発達心理学研究とロボティクス研究の連携をより強めることで、人間と同じように物理的・社会的環境とのインタラクションを目指す。

4ではインタラクションにおける非言語情報について述べる。この研究では、対面コミュニケーションにおける対話音声と頸部より上部の運動を計測することで、会話の活性度や意思疎通の度合いを計量化する。対面コミュニケーションの収録には、ハーフミラーを使うことで遠隔地対話でも視線の一致を可能にした、高いモダリティ情報を伝達可能なシステムを用いた。この研究を推進することで、ロボットのように身体性を持ったシステムで、感情のような内部状態の適切な表現への応用を図る。

これらの研究は、コミュニケーションの生

理・認知・行動の三つの階層での身体性に基づくコミュニケーションの研究である。我々は、この三つの階層で研究を進めていくことで、身体性を持ち社会の中で我々と対等な存在として真にコミュニケーションのできるシステムの構築を目指す。

2 表面筋電位による指運動の識別及びその応用^[12]

私達は普段何気なく指や手を曲げたり伸ばしたりしているが、これは運動神経繊維を伝わる電気パルスによって、腕の筋肉にある多数の筋繊維(運動単位とよばれる)を収縮させるコントロールを行っているためである。では、この筋電位信号(Electromyogram, EMG)を直接計測して、健常者であれば各指を動かす程度の、障害者であれば自分が意図的に動かせる範囲の、運動のパターンを認識させることで、様々なインタフェースに使うことはできないだろうか。

このような筋電位を用いたインタフェースの特徴は次のようなものである。

- ・ データグローブが指の位置を計測するのに対して、指の力の大きさの程度を計測し利用することができる。
- ・ 筋繊維に走る電気パルスを計測するので、ある程度高い反応性が得られる。
- ・ 上肢切断者に対しても、残存前腕筋から「指」運動に対応する筋電位パターンが得られ、それを利用できる場合がある。
- ・ 超多チャンネル化により電気パルスとその位相情報を解析し、運動単位の推定を非侵入に行える可能性がある。

以上の特徴から義手制御はもちろん、四肢障害者のための各種作業・コンピュータインタフェース等のために、筋電計測により意図する運動を補完する応用範囲は広い(様々な手法や応用として例えば^[13] ^[14] ^[15] ^[16] ^[17])。健常者にとっても生体信号を用いた何らかの操作(manipulation)のためのインタフェースやデバイスには様々な応用が想像できる。

我々は非侵入計測として、前腕の皮膚表面からの表面筋電位信号(surface EMG)を多チャンネル(16チャンネル)計測し、指や手首の運動を

分離・推定する研究を行っている。本稿では、各指の個別の運動に対して計測と信号分離を行った結果^[12]とそれを応用したシステムを紹介し、今後の研究計画について述べる。

2.1 各指運動パターンの分離

指の運動を支配しているのは手掌内及び前腕の伸筋、屈筋である。前腕部のみで計測した16チャンネルの表面筋電位から、各指の伸展と屈曲・手の握る離す等の10種類以上の運動を分離した。電極は、骨の直上は避けるなどの若干の配慮以外は、適当に配置した。

まず被験者は、リラックスした姿勢のまま腕を弛緩させて筋の緊張がなくなったことを確認した後、第1指(母指)から第5指(小指)の伸展及び屈曲と手の握る離す運動を、「日常的に行う無理のない」動作にて3回ずつ順番に行う。生じた表面筋電位を、遮断周波数250Hzのローパスフィルターを通し、1kHz サンプリングにて12ビットA/D変換を行って記録した(図1)。

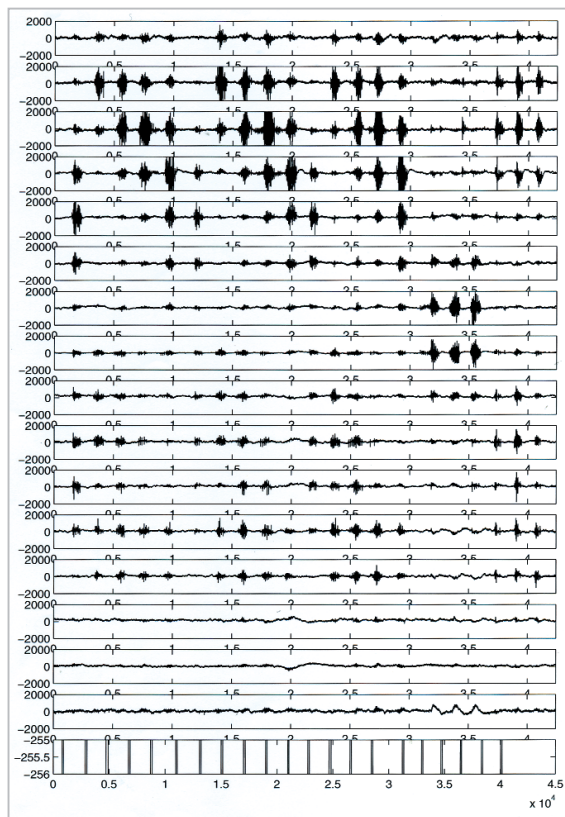


図1 各指伸展運動による筋電データの全体

各チャンネルごと縦軸に振幅と横軸に標本数(時間)を、最下段には各指の運動開始時刻のマークを示した。

我々の目的のためにはチャンネル間で信号がどのように分布しているかの特徴が抽出できれば十分であるので、筋電図の処理方法としてよく用いられている(例えば[18])フィルタと整流を組み合わせたノイズ除去とデータ平滑化を行った(遮断周波数20HzのHPF・全波整流・遮断周波数2.5HzのLPF)。そのデータに対して次に述べる独立成分分析を適用しパターン識別を行った。

近年、高次統計量を利用して、信号を統計的独立な成分へと分離する独立成分分析という手法が盛んに研究されている[19][20]。これによれば、生体信号のような非ガウス分布を持つ、幾つかの統計的に独立な信号の線形和として観測された多チャンネルの信号から、各独立成分に信号分離を行うことが可能である。統計学の言葉で言えば、独立な成分の分布を知らないまま混合行列のパラメータを推定する、セミパラメトリック推定である。

上述のように計測した学習データは、各指運動を個別に計測しているため、信号は実は混合していない。その意味では必ずしも独立成分分析を用いる必然性はない。また、複数の指の運動によって生じる表面筋電位は、一般的には線形加算で表現することはできないため、独立成分分析の結果、各指の組合せ運動が抽出できるわけではない。しかし、時間構造を無視した学習データ信号の確率密度分布は、独立成分分析の仮定をほぼ満たしており、組合せは期待できないものの、個々の指運動を簡便に分離するという目的には十分使用可能である。また、独立成分分析を用いることによって、いつ指を動かしているかということを知りたくなく、各指運動の分離が可能であることも使用上非常に有益であることが分かった。

独立成分分析の結果を(運動の順番でソートした)各成分ごとにプロットしたものが図2である。これによれば、第1から第5成分で各指の伸展運動を、第6成分で握る動作を、ほぼ完全に分離していることが分かる。手掌全体での開く動作は、主に二つの成分に分離しているが、これは、3回の開く動作が、必ずしも同じ運動にならなかったことによると思われる。第8成分以下については、若干の指運動との相関は見られるものの、ほとんどがノイズと言って良く、各運動に対す

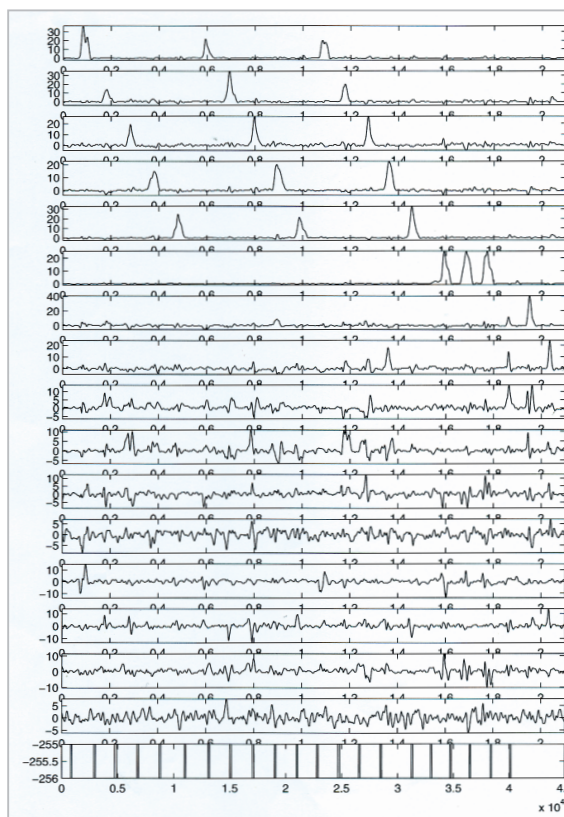


図2 各独立成分

最下段に各指の運動開始時刻のマークを示す。

るゆらぎあるいは外部ノイズ等の成分を抽出していると考えられる。さらに混合行列の推定結果から、各成分の各チャンネルへの空間的分布が生理学的知見とも整合的であることが分かった。

2.2 応用事例

次に独立成分分析によって抽出した指運動の応用例について述べる(参考のために[21]に動画などを掲げた)。

指運動の分離、識別ができるのであれば、それを単にマッピングすることによって、様々な応用例が可能となる。我々は以下のものに応用を行った。

- ・各指に音符を割り振った仮想ピアノの演奏
- ・子音・母音・実行の組合せを用いた発声インタフェース
- ・バーチャルハンド

前者二つは指運動によるオンオフ信号を利用したものであるが、後者では分離した信号強度に比例させて仮想の手を動かすことを試みた。

ここでは、全5指の屈曲及び伸展と、握る動作の全11種類の運動の分離を行っている。指同士の干渉が若干あるものの、おおむね運動を再現できることを確認した。

2.3 上肢切断者に対する適用

以上のシステムを実際に事故で前腕を失った上肢切断者に対して適用した例について報告する。事故で切断後3年が経過した切断者に対して図3のように電極を装着して、「指」を動かすように指示した。すると、各指の屈曲運動を識別することに成功し、被験者のほぼ意図どおりに動く仮想の手を提示することができた(図3)。この時点で被験者は、把持運動が可能な筋電義手を半年あまり装着しており、無いはずの手があるように感じる幻肢も持っていた。それは筋電義手の機能と同等の2～5指が固まったグローブのような幻肢であり、指が分離しているという感覚を持っているわけではなかった。しかし、このシステムによる仮想の手の提示により、そこに指が動く自分の手があるというリアルな感覚を感じることができたそうである。

このような結果は被験者のすべてに対して同様に得られているわけではなく、個人差がある。現在、上肢切断者に対する実験結果の収集に努めている。しかしながら、上肢切断者が前腕に残存する筋を随意的に動かし、異なる「指」運動に対する筋電位パターンを生成でき、かつ認識できるという事実は、非常に興味深い。この研究は、兵庫県立福祉のまちづくり工学研究所(神戸)の協力を得ている。

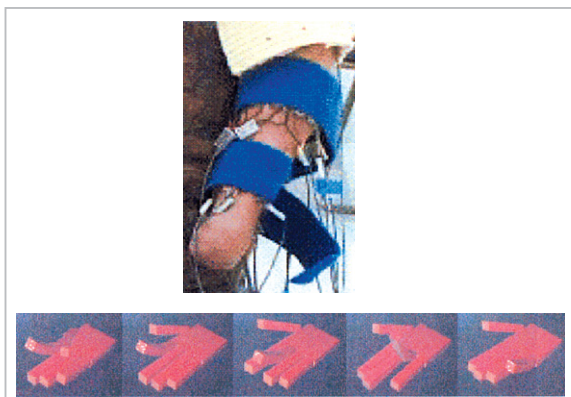


図3 上肢切断者における指運動の認識

上:電極装着の様子。下:仮想の手による指運動表示。

2.4 これからの課題

今後の展開として以下のような研究課題をあげる。

2.4.1 上肢切断者に対する適用

現在、障害者の方がどのくらいの指運動パターンを意図的に使うことができるのかを調べるため、兵庫県立福祉のまちづくり工学研究所と共同研究を行っている。特に上肢切断者の随意的な「指」運動が取り出せることは、被験者に画像や音楽によるフィードバックを視覚・聴覚にて与えることが可能であることを意味する。脳科学における体性感覚の再構成に関する最近の研究[22]によれば、これは幻肢痛の軽減や除去といった目的に役立つ可能性がある。また、上肢切断者のQOLの面での応用も考えていきたい。単なる仮想の手の提示にとどまらず、計算機アクセス支援、音楽演奏等の自己表現手段の提供等が可能であると思われる。

2.4.2 超多チャンネル計測による運動単位の推定

現在の指運動の分離は個別の指を排他的に識別するものであり、その組合せ運動については今後の課題である。一方、現在用いているような信号の強度だけではなく、微細な電気パルスとその位相情報を解析できるような多重電極(64チャンネル)を用いた計測を行い、一つの運動神経繊維に支配されている運動単位(複数の筋繊維)の特定[23]により、いわばCTスキャンの技術などの基礎研究も行う。予備的な解析を、神戸大学情報知能工学科小谷研究室と共同で行っている。

3 前言語的コミュニケーションの認知メカニズム

近い将来、介護ロボットや家事ロボットが家庭に入ってくるだろう。ロボットのようなシステムと人間とのコミュニケーションを円滑にし、システムを社会的存在(social being)として人間の社会的活動にコミットさせるには、どのような条件が必要だろうか。本研究では、乳幼児の前言語的コミュニケーション(指さし・ジェスチャなどの身体的コミュニケーション)がどのように発達するのかを参考として、人にやさしいシ

ステムの設計原理を探索する。

3.1 Infanoid — 研究プラットフォーム

コミュニケーションの発達モデルを論じる前に、著者が現在開発している赤ちゃんロボット Infanoid[9][10]を紹介したい。Infanoidは、人間のコミュニケーション能力の発達を工学的にモデル化するための研究プラットフォームである。図4に示すように、テーブルに載せられた上半身だけの人間型ロボット(ヒューマノイド)であり、3歳児とほぼ同じサイズ(テーブル面からの高さは48cm)の「身体」に23の自由度(動作軸)を持つ。

Infanoidの頭部には二つの眼球があり、それぞれに広角(水平画角は約120deg)と望遠(同約25deg)の二つのビデオカメラを持ち、上下左右に眼球運動(サッカード及び対象追跡)が可能である。ビデオカメラから取り込まれた画像は、超並列画像処理システムによって処理され、人間の顔やその構成要素の追跡、色・奥行き・形による対象抽出などが行われる。また、Infanoidは音声処理装置を持ち、イントネーションを伴った喃語及び分節語を発声することができ、発声される音素に合わせて上下の唇を適切な形に動かすことができる。

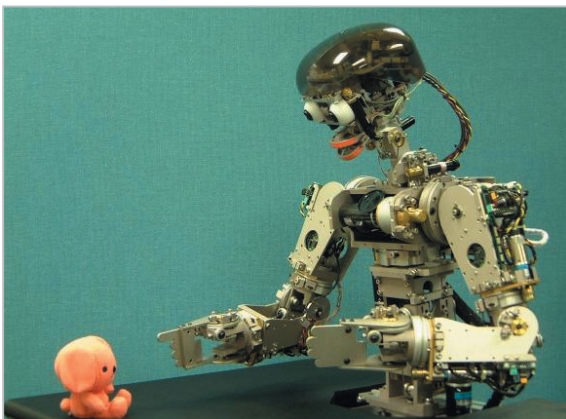


図4 赤ちゃんロボット Infanoid

現在のところ、Infanoidは①人間の顔及び顔方向の追跡、②特定のオモチャの追跡、③人間の顔とオモチャへの視線交替、④喃語の発声などが可能であり、⑤視線方向の追跡、⑥部分的な音声模倣(発声できる音素[列]に制約を加えた反復)などを準備中である。

3.2 共同注意から間接経験へ

新生児でも、泣き・ほほえみ・人への注目などによって、他者の行動をある程度コントロールすることができる。ただ泣くことによって養育者から食べ物を手に入れたり抱っこしてもらうことから始まり、指さしなどによって欲しいものを選択的に手に入れるようになる。やがて言葉を使うようになり、より複雑な社会的インタラクションに自らを巻き込んでいく。

最近の発達心理学研究[24][25][26]では、このようなコミュニケーション発達に共同注意(joint attention)が重要な役割を果たしていることが明らかになりつつある。共同注意とは、他者(養育者)が注意を向けている対象に自分も注意を向ける活動である。共同注意を実現するには、図5及び図6に示すように、①養育者の注意方向を視線・顔の向き・指さしなどから読み取り、②養育者が実際に注意を向けている対象を同定する(切り出す)ことが必要となる。

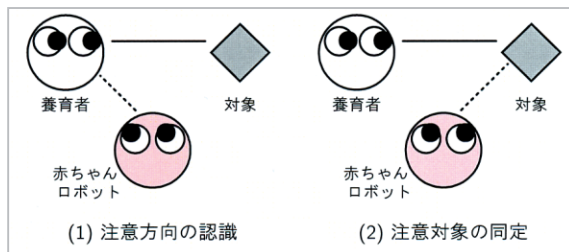


図5 共同注意のプロセス

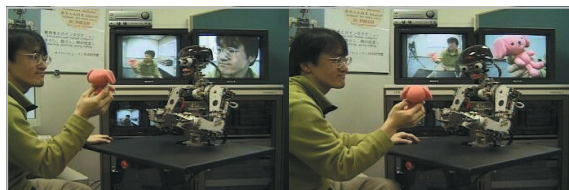


図6 Infanoidによる共同注意

共同注意の役割とは何だろうか。それは他者の行動を観察するためのデバイスであろう[8]。図7に示すように、養育者が対象と何らかのインタラクションをしているとする。共同注意によって、乳児若しくはロボットは、この養育者がどのような感覚入力*i*を対象から受け取っているのかを推定することができる。同じ対象に同時に注意を向けるのだから、両者の感覚入力*i*と*i'*はほぼ同じものになる。また、養育者の注意方向を読み取ろうとするとき、乳児は養育者がど

のような運動出力 o (表情・発声・動作など)を対象に対して働きかけているかを観察することになる。

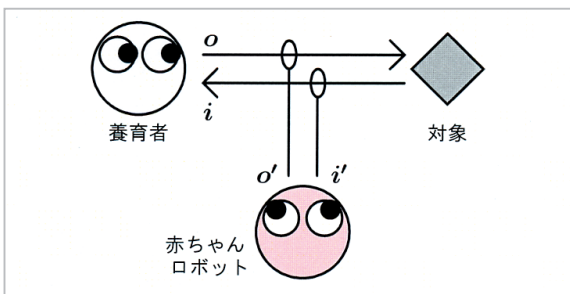


図7 他者の感覚入力と運動出力の観察

ここで注意が必要なのは、乳児が観察した他者の運動出力 o は、あくまで遠感的(主に視覚的)なイメージにすぎないことである。観察した他者行動を模倣可能にするには、この遠感的イメージ o' を乳児自らの筋肉運動イメージ μ (o')に変換するマッピングが必要となる。このマッピングは、新生児にみられる表情模倣や共鳴動作のような原初模倣[27][28](すなわち視覚=運動的な反射)を土台として、養育者との模倣あそびや自分の身体の一部を見ること(例えばハンドリガード)などをとおして構成されると考えられる。

他者の感覚入力 i を共同注意によって i' として取り込み、また他者の運動出力 o を遠感的イメージ o' として捉え、それを運動イメージ μ (o')として取り込む。これが他者行動 $\langle i, o \rangle$ を $\langle i', \mu(o') \rangle$ として疑似体験すること、すなわち間接経験(indirect experience)[8]である。間接経験した他者行動を再生すれば模倣(imitation)となる。

3.3 間接経験による社会的発達

間接経験や模倣は、他者がどのような状況にどのように対処するのかを学習するためのデータを(正例だけでなく負例も)提供する[8]。この学習によって、他者がある対象(例えば電話器)を前にしたとき、どのような行動をとるのか(例えばダイヤルする)を予測することができる。また、他者が何らかの動作(例えばバイバイと手を振ること)をしているとき、その動作がどのような対象(例えば視野外にいる人物)に向けられたものなのかを推定することができる。また、他者に何らかの行動をさせたいければ、それに対応

する感覚入力(又は対象)を提示すればよい。

間接経験によって、外因的に誘発される行動のパターンを学習することができる。しかしこれだけでは目の前にキャンディがあっても、誰かがキャンディを食べるのを模倣せずには、それを手に入れることができない。では、自発的にこのキャンディを食べ、その心地よい甘さを知るにはどうしたらよいだろうか。すなわち、自発的に環境(対象)に働きかけ、報酬を得るにはどうすればよいだろうか。おそらく、間接経験によって得た運動出力(及びその変形版)やランダム運動(オペラント反応)を実行し、その結果として返ってきた感覚情報を(例えば快・不快として)評価し、それを強化子として学習を進めるというストラテジーであろう。

3.4 社会的環境の探索

養育者も探索活動の対象(上の例ではキャンディ)になり得る。図8のように、乳児(若しくはロボット)が既存の運動レパトリの変形版やランダム運動を提示すると、それを受けた養育者はその社会的意味を解釈し、それを乳児に返してやる。社会的に適切な提示であれば報酬(注目・物の受渡しなど)が得られ、社会的に不適切あるいは無意味な提示であれば罰(無視・怒り声など)を受けるだろう。

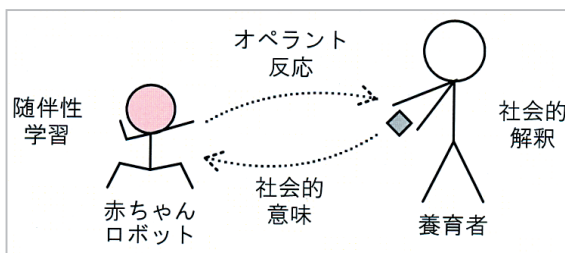


図8 養育者をととした社会的探索

さらに、この探索活動によって「意図」というものが発現する[10]。例えば、乳児が何の意図もなしに喃語(例えば「バブー」)を発声したとする。それを聞いた養育者は、たまたま乳児の手の届かない所にあったおもちゃを取り、乳児に渡してやる。結果として乳児は願ってもない報酬を得ることになる。この行動—発声を原因としおもちゃを結果とする—は強化される。何度も試行を繰り返す中で、ある発声はあるおもち

ヤと結びつき、別の発声は別のオモチャと結びつく。やがて乳児は、欲しいオモチャを選択的に手に入れる行動を獲得する。つまり、原因から結果をつくりだす行動が、手段によって目的を達成する行動として機能するようになる。意図的行動とは、欲求(オモチャが欲しい)つまり目的に対応する手段(喃語の発声)の行使である。

間接経験(模倣)によって行動パターンの「タネ」を取り込み、ときに全くランダムな行動を「タネ」として、養育者からのフィードバックに応じて社会的に意味のある(報酬をたくさん得られる)行動をシェイピングしていく^[29]。可能な行動からなる空間の中で、「タネ」にランダムなゆらぎを与えていき、その各ステップで養育者をおとしてその行動を評価し、社会的に有意な行動クラスを探索することが考えられる。もっとも、この行動クラスは明確な輪郭をもつ二値的なものではなく、むしろファジー集合的な性質を持つかもしれない。

このような社会的行動の探索によって、社会の中で共有されているコミュニケーション=プロトコル(ジェスチャ・言語など)を獲得することができる^[30]。あるレベルのコミュニケーション能力を獲得できれば、それを利用してより複雑な探索活動が可能となる。このメカニズムが人間のコミュニケーション能力を特異的に高度なものにしたのだろう。

3.5 これからの課題

本研究は、コミュニケーションの発達モデルを構築する発達心理学的な研究と、それを研究プラットフォームである赤ちゃんロボット Infanoid として実証するロボティクス研究という、二つの相補的な研究項目からなる。これまでのロボティクス研究では、学習はあっても発達という視点はなかった。また、発達心理学だけでは、赤ちゃんを観察することはできても、操作可能(特に侵襲的に操作可能)な実験はできない。これらの点で、本研究はこの二つの研究分野に新しい視点と実験パラダイムを提供できるだろう。

今後の研究は、発達心理学及びロボット工学との連携をより深め、現状での両者のギャップを縮めていく方向をとるべきだろう。具体的に

は、健常児及び発達障害児のコミュニケーション能力^[26]^[31]をより詳細に調べ、特にコミュニケーション発達的前提条件となる生得的な能力は何か、そしてどのような養育環境がその発達に不可欠なのかを特定することが必要である。また、現状の Infanoid は、まだ生身の赤ちゃんが持つ感覚運動モダリティ^[32]すべてを具現化したものではない。より機能的に等価な身体性を持たせ、人間と同じように物理的・社会的環境とインタラクションできるようにしたい。

4 インタラクションにおける非言語情報

現実世界での我々のコミュニケーションは、顔を見合わせた形での対話が主である。もちろん、互いに顔を見合わせないコミュニケーションも存在しているが、親密さを必要とするコミュニケーションや重要情報を伴ったコミュニケーションこそ、筆談よりも電話、電話よりも顔を見合わせた形での対話を多くの人は希望している。これは会話の内容そのものである言語情報よりも、音声のレベルの強弱、周波数の高低、話す速さなどの会話に含まれる言語に付随するパラ言語情報、そして視線を含む表情やうなずき、指さし等の身体的変化という範囲まで広げた非言語情報が、相手に対する意思伝達に重要な役割を果たしていることを示唆している^[33]。

4.1 頭部運動から会話の場の状況を知る

非言語情報の研究の一つとして、身体の上半身、特に頸部より上部の頭部運動と会話、意思疎通の効率等を調査することで、高いモダリティーを工学的に実装するための下準備となる研究を行う。工学的に実装するには2段階の流れがあり、1段階目は認知、2段階目は情報発信である。

まずは我々研究者が、高いモダリティーに含まれる情報の「認知」過程を知る必要がある。そのために一対の被験者が実際に対話を行っているときの言語情報と、視線を含めた頸部より上部の頭部運動の記録を行った。従来、一対となった被験者の対話の身体的情報を記録するために、被験者の斜め正面からの動画を記録する

方法が行われていた(図9)。この従来法では斜め正面からの情報であるが故に、実際に対向している被験者の観ている情報とは異なった情報を記録していることになる。対向している被験者の観ている情報に近似させるための変換をするためには、斜め由来のズレの補正をしなければならず、そこには誤差があることも考慮しなくてはならない。その場合、視線一致に関する正確な情報取得は困難である。

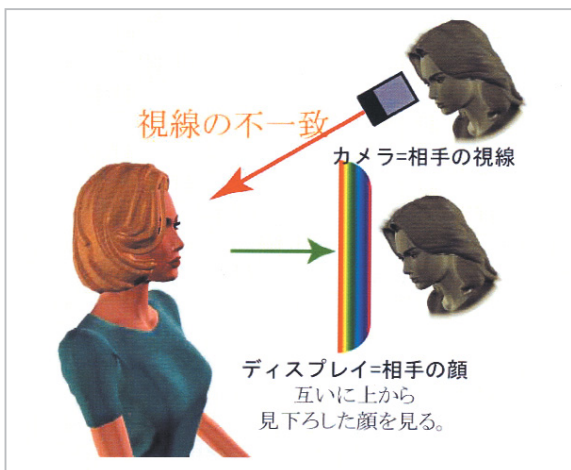


図9 従来法での視線の不一致

従来法での問題点を回避するためにハーフミラーを使い、一対の被験者を真正面から撮る実験を行った(図10)。また、情報を補うために、ほぼ真横からの動画の情報も得た。記録した実際の対話は、あらかじめ与えられたタスクを解くタイプの対話であり、通常の音声由来の言語情報、パラ言語情報のみならず、視線、表情(笑い、困惑)、頭部運動等の非言語情報を含め、良好に記録を行うことができた。

現在、この記録を画像処理装置で処理を行い、非言語由来の情報を大雑把に工学的に認知させることを行っている。話の内容そのものの情報を使わずに、工学的に「場の盛り上がり」や「共感」という曖昧な状態をリアルタイムに評価することを当面の目標としている。これは言語獲得が完了していない幼児や、未知の言語を話す地域に足を踏み入れたときの、言語内容以外から「場」の雰囲気を知る場合と状況は同じであると思われる。話の内容そのものの情報を使用しないために、言語の違いから生じる差異に影響されず、むしろ対話文化の差異を知る指標

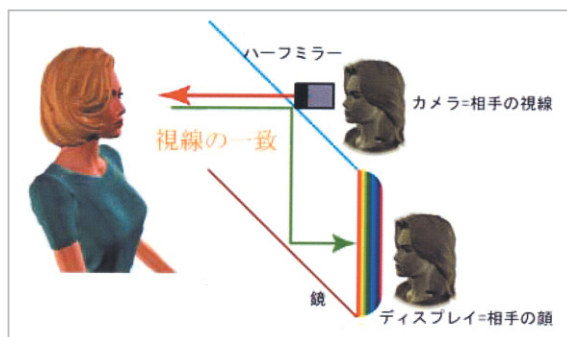


図10 ハーフミラーを使った視線の一致

となると思われる。

具体的には発話に付随する音量や声の高低等の音声情報とともに、情報由来の非言語的な動きの画像情報を総合的に評価する。また、笑い、共話、協調的首振り、アイコンタクト等の従来の言語観測では特殊な例としてやや置き去りにされがちであった「場」の感知も組み込みながら、工学的認知の一応の完成を目指している。

4.2 これからの課題

今後の計画として、例えば実際に身体性のある機械(ロボット)等にこの非言語情報認知モジュールを実装することで、非言語情報から話者の状態や感情を知る手だて(認知手段)が得られる。また、ロボット側から非言語情報を発信する場合のロボット内部のフィードバックに応用することで、場に与えるであろう影響をあらかじめ推測したり、リアルタイムに自分を含んだ場の現状を知ることにも可能になると思われる。これは将来、人とロボットの間で行われるであろう通常対話の受け答えの中で、ロボットが故意に話の場を盛り上げたり、反対に下げたりと、対話の場をコントロールすることが可能になることを意味している。さらに、現在普及している動き優先の動画の圧縮(例えばMPEG [34][35][36])とは異なった方法である場の盛り上がりを優先した動画の圧縮等に応用し、高いモダリティーを保ったまま双方向の通信を行う技術等への利用を考えている。

5 おわりに

けいはんなセンターでの我々の新しいプロジェクトへの取組として、身体性を持ったコミュ

ニケーションの研究の紹介を行った。これまでの研究に身体性を付与することで、コミュニケーションシステムを意図を持った存在、コミュニケーションにおいて我々と対等な社会的な存在として認めることができるようになるだろう。さらに、コミュニケーションシステムが身体性を持つことで、物理的・心理的な状況の把握が

容易になり、相手との状況を共有したインタラクションへの道が開かれる。

我々の研究が、コミュニケーションシステムの社会参加、その結果として我々の社会生活の質の改善の一助になるように、今後の研究を推進しなければならない。

参考文献

- 1 熊本, 伊藤, “支援対話の解析に基づく対話処理方式の提案”, 電子情報通信学会誌 DII, Vol.J77-D-II, No.8, pp.1492-1501, 1994.
- 2 熊本, 伊藤, “WWWを介した対話システムとユーザとの対話の分析—発話レベルの分析—”, 言語処理学会第4回年次大会, S2-6, Mar., 1998.
- 3 矢野, 伊藤, “創発的な対話に関するコーパスの構築”, 自然言語処理, Vol.6, No.4, pp.117-137, Jul., 1999.
- 4 小嶋, 伊藤, “文脈依存的に単語間の意味距離を計算する一手法”, 情報処理学会論文誌, Vol.38, No.3, pp.481-489, Mar., 1997.
- 5 善本, “マルチエージェントを用いた辞書安定性の分析”, 信学技報, TL2000-10, pp.25-32, 2000.
- 6 伊藤, “[心の理論]に基づく対話理解の提案”, ソフトウェア学会第13回大会, 1996.
- 7 D. Dennett, "Kinds of minds", Harper Collins Publisher, 1996 (土屋俊訳, “ころほどこにあるのか”, 草思社, 1997).
- 8 H. Kozima, "Attention-Sharing, Behavior-Sharing, and Acquisition of Language", International Workshop "Ecology of Language Acquisition (EcoLang-99)", The Netherlands, 1999.
- 9 H. Kozima and J. Zlatev, "An epigenetic approach to human-robot communication", International Workshop on Robot and Human Interactive Communication (ROMAN-2000), pp. 346-351, Osaka, 2000.
- 10 H. Kozima and H. Yano, "A Robotogenetic Model of Human Social Development", International Workshop on the Relationship between Cognitive and Linguistic Development, Bangkok, Thailand, 2001.
- 11 岡田, “完全無欠なコンピュータ HALは何を欠いていたのか”, 人工知能学会誌, Vol.16, No.1, pp.98-101, 2001.
- 12 藤原, 前川, “独立成分分析による筋電データからの各指運動の分離”, 信学技報, MBE99-7, pp.41-46, 1999.
- 13 A. Hiraiwa, K. Shimohara and Y. Tokunaga, "EMG pattern analysis and classification by neural network", IEEE Int. Conf. on Syst., Man and Cybern., pp.1113-1115, 1989.
- 14 辻, 市延, 伊藤, 長町, “エントロピーを用いた誤差逆伝搬型ニューラルネットワークによる EMGからの前腕動作の識別”, 計測自動制御学会論文集, Vol.29, No.10, pp.1213-1220, 1993.
- 15 K. A. Farry, I. D. Walker, and R. G. Baraniuk, "Myoelectric teleoperation of a complex robotic hand", IEEE Trans. on Robotics and Automation, Vol.12, No.5, pp.775-788, 1996.
- 16 梶谷, 星野, 西川, 横井, 梶原, 樋口, “障害者に適応可能な筋電操作型義手の開発”, 信学技報, PRMU98-85, pp.9-16, 1998.
- 17 西川, 横井, 嘉数, “筋電位を利用した動作認識デバイスの設計”, ロボティクス・メカトロニクス講演会予稿集, Jun., 1998.
- 18 A. L. Hof and Jw. Van den Berg, "EMG to force processing III : Estimation of model parameters for the human triceps muscle and assessment of the accuracy by means of a torque plate", J. Biomechanics,

- Vol.14, pp.771-785, 1981.
- 19 A. J. Bell and T. J. Sejnowski, "An information maximization approach to blind separation and blind deconvolution", *Neural Computation*, Vol.7, pp.1129-1159, 1995.
- 20 S. Amari, A. Cichocki, and H. H. Yang, "A new learning algorithm for blind signal separation", *Advances in NIPS*, Vol.8, pp.757-763, MIT Press, 1996.
- 21 <http://www.crl.go.jp/jt/a134/maekawa/emg.html>,
<http://www.crl.go.jp/jt/a134/yoshi/myo/>
- 22 V. S. Ramachandran and S. Blakeslee, "Phantoms in the Brain", William Morrow, New York, 1998.
- 23 G. Rau, C. Disselhorst-Klug, and J. Silny, "Noninvasive approach to motor unit characterization", *J. Biomechanics*, Vol.30, pp.441-446, 1997.
- 24 M. Tomasello, "The Cultural Origins of Human Cognition", Harvard University Press, 1999.
- 25 G. Butterworth and N. Jarrett, "What minds have in common is space: spatial mechanisms serving joint visual attention in infancy", *British Journal of Developmental Psychology*, Vol.9, pp.55-72, 1991.
- 26 S. Baron-Cohen, "Mindblindness: An Essay on Autism and Theory of Mind", MIT Press, 1995.
- 27 A. Meltzoff and M. K. Moore, "Newborn infants imitate adult facial gestures", *Child Development*, Vol.54, pp.702-709, 1983.
- 28 A. Meltzoff and A. Gopnik, "The role of imitation in understanding persons and developing a theory of mind", in Simon Baron-Cohen, et al. (eds), *Understanding Other Minds*, Oxford University Press, 1993.
- 29 小嶋, "ロボットの社会的発達と「心の理論」の獲得", *情報研報*, ICS122-3, pp.13-18, 2000.
- 30 小嶋, "状況に埋め込まれた身体—ロボットによる言語獲得の条件", 発達心理学会シンポジウム「言語獲得の根幹を問い, 研究の視点を広げる— joint attention は言語獲得にどうかかわるのか—」, 2001.
- 31 U. Frith, "Autism: Explaining the Enigma", Blackwel, 1989.
- 32 R. A. Brooks, et al., "The Cog Project: building a humanoid robot", In C. L. Nehaniv (ed.), *Computation for Metaphors, Analogy and Agents*, Lecture Notes in Artificial Intelligence, Vol.1562, Springer-Verlag, 1998.
- 33 川浦康至, "メディアコミュニケーション", 現代のエスプリ, 至文堂, 1993.
- 34 渡辺, "MPEG/H.262", *テレビジョン学会誌*, Vol.18, No.1, pp.44-49, 1994.
- 35 R. Aravind, G. L. Cash, D. L. Duttweiler, H. M. Hang, B. G. Haskell and A. Puri, "Image and video coding standards", *AT&T Technical Journal*, pp.67-88, Jan./Feb., 1993.
- 36 D. L. Gall, "MPEG: A video compression standard for multimedia applications", *Comm. for the ACM*, Vol.34, No.4, pp.46-58, 1991.



やの ひろゆき
矢野博之

情報通信部門 けいはんな情報通信融合研究センター社会的インタラクシ
ョングループリーダー 工学博士
対話処理、対話の認知モデル



まえかわ りんたろ
前川 聡

情報通信部門 けいはんな情報通信融合研究センター社会的インタラクシ
ョングループ主任研究員 工学博士
進化及び学習における創発機構



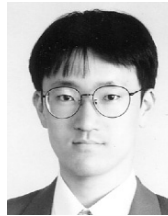
よしだ 淳
善本 淳

情報通信部門 けいはんな情報通信融合研究センター社会的インタラクシ
ョングループ研究員 博士(物質・生命
情報学)
個と社会、自己と他者、対面対話



ふじわら たかし
藤原義久

情報通信部門 けいはんな情報通信融合研究センター社会的インタラクシ
ョングループ主任研究員 理学博士
信号処理、社会的相互作用における創
発機構



おのづみ しげのり
小嶋秀樹

情報通信部門 けいはんな情報通信融合研究センター社会的インタラクシ
ョングループ主任研究員 工学博士
言語的及び前言語的コミュニケーションの認知科学