

大学の研究・動向

人間を支援するための行動センシングと協調のデザイン

電気工学専攻 情報メディア工学講座 複合メディア分野

教授 中村 裕一

助教 近藤 一晃

助教 小泉 敬寛

1. はじめに

身の回りの様々な場所で、家事や介護の補助をしたり、教育やトレーニングの補助をしたり、種々の目的で人間の行動を認識しながらそれを支援する情報メディアやロボットなどのシステムが急速に広がることが期待されています。このような分野の方向性を探る研究として、我々の研究室では、人間の動作・行動などの計測・認識・予測[3、5、6、8、10]、人間の行動の記録（ライフログ）や複数人の記録の統合と利用[1、7、9]、人間とシステムの協調モデル[2、4]等の研究に取り組んできました。本稿では、そのうち、筋電位計測と画像計測と生体計測との統合、人間にとって自然な機械との協調と役割分担のデザインについて紹介します。

2. 生体計測と画像計測の統合

2.1 筋電位計測の利用

人間の状況に応じた支援を丁度良いタイミングで提供するためには、その人の意図や生理的な状態など、様々な要因を考慮する必要があります。筋肉の活動を調べる筋電位計測はその一つの手段となります。

筋肉の収縮は筋繊維の興奮によって起こり、その際の膜電位の変化が皮膚表面に現れたものを表面筋電位と呼びます。その発生電位と発生する筋張力には大まかな比例関係が成立することが知られています。また、運動神経（ α ニューロン）から伝わった刺激により筋神経接合部付近に引き起こされた興奮が筋繊維に沿って伝わり始めてから実際に筋肉が収縮するまで最大 200msec 程度の時間があることから、運動が起こる前にその筋電位の変化を計測することができます。さらに、力が入っているが動いて

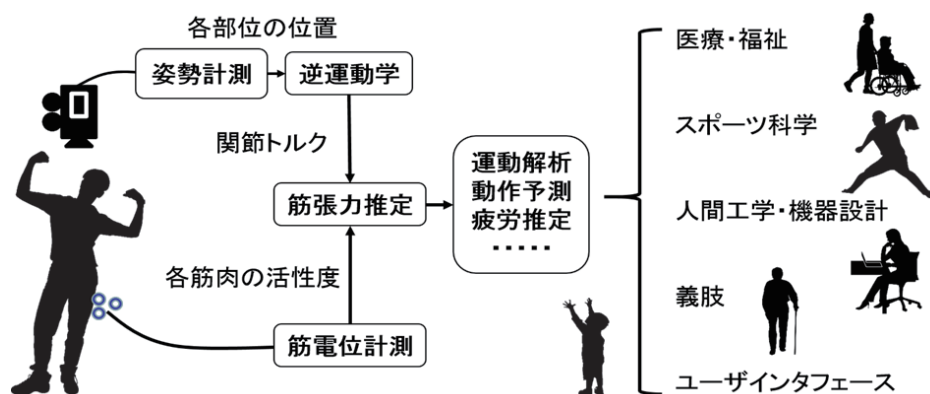


図 1：画像計測と筋電位計測の統合とアプリケーション

いない場合など、外部からはわかりにくい情報も得ることができます。このような良い性質を持つため、これまで、筋電位計測は義手などの入力インターフェースとして、また、人間工学における動作や疲労の解析手段等に用いられてきました。

しかし、筋電位だけを用いて全身の姿勢や動作を求めることは依然として困難な問題となっています。その原因の一つは、生身の人体のパラメータ、例えば、筋や腱の長さや、腱と骨との接合（それによるモーメントアーム等）を計測することの難しさです。また、関節の動きに対して、主に力を発生する筋肉（主動筋）、協力して順方向に働く筋肉（協働筋）、逆方向に作用する筋肉（拮抗筋）がある等、内部の動きに自由度があること、さらに、他の筋や骨に隠されている深層筋を計測するのが難しいことも理由としてあげられます。

このような問題に対処する一つの方法として、図1のように、外部から姿勢を計測し、その結果から逆運動学等を利用して各関節に作用しているトルクを求め、筋電位により推定される筋活性度と統合して筋張力を推定することが考えられます。現在、我々の研究室では、このようなモデルを用いた運動解析に取り組みつつあります。位置の選択性の高い電極やアレイ電極を用いた筋の活性状態の詳細な解析や、深層筋の活動推定などを用いてその精度を高めることも課題となっています。以下では、このような試みに先行して行ったリハビリテーション支援と把持行動の予測について紹介します。

2.2 肩のリハビリテーション運動の支援

リハビリテーションにおいて効果の高いトレーニングを行うためには、「正しい姿勢」で「適切な筋収縮」を発生させる必要があります。しかし、肩のように自由度の大きな機構では、正しいトレーニング動作を理解しにくく、また、一旦正しい動作を行っても、時間経過とともに所定の動作から外れ、「代償動作」（似ているにもかかわらず本来の目的を果たせていない別の動作）を行ってしまうことが問題になってきました。

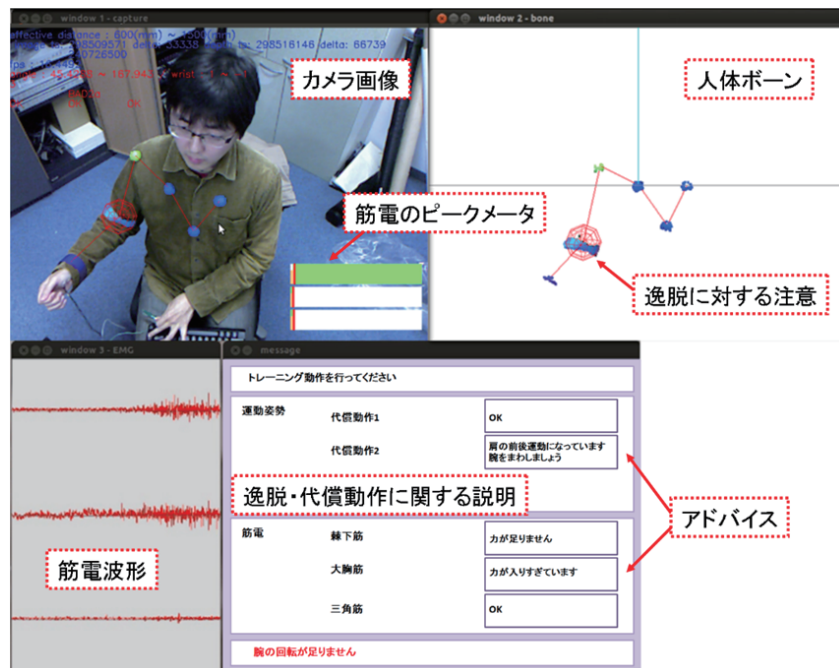


図2：情報提示の一例（点線の枠及び枠内の文字は本論文のために説明として付加したもので、利用者には提示されていない）：左上はRGBカメラ画像、右上は人体ボーン、左下は筋電位波形で上から棘下筋・大胸筋・三角筋後部、右下は文字情報

そのため、患者が正しくトレーニング動作を行う支援をするシステムを本学附属病院の理学療法士の協力を得て設計しました[10]。リハビリテーションの現場でよく行われるのは、医師や理学療法士が正しい動作、または、現在の状態可能な範囲で最も望ましいと考えられる動作を患者に指示し、それを患者が繰り返し再現する努力をすることです。そのため、このシステムでも、手本となる動作を記録しておき、利用者がそれを繰り返し再現することを手順としました。

繰り返しトレーニングを行う際の提示情報の例を図2に示します。このように計測結果を提示することにより、利用者が客観的に自分の状態を把握する補助を行うとともに、毎回の動作と基準動作との差を意識することができます。逸脱が起こっている場合には、該当箇所や理由・修正方法等が提示されるため、利用者は動作の修正を行いながら所定のトレーニング動作を継続することができます。

試作システムが概ね想定通りに動作していることを確認した上で、患者さんおよび理学療法士に試用してもらったところ、本システムの必要性や機能について、肯定的な意見を多くもらいました。現在の構成でも、理学療法士が患者のそばにいないと減らせない時間を減らせ、患者にとっても待ち時間が減ることが期待できるとの感想が得られています。また、トレーニング状態の履歴を残したり、過去のトレーニング履歴との比較を提示する機能を加えて欲しいとの要望がありました。上達の度合いを表示することによって患者の動機を高めたり、医師や理学療法士のために治療効果を可視化することなどが目的です。

2.3 把持行動の予測

いつ、どの物体を、どのように把持するのかを予測することが出来れば、ロボットや情報メディアが種々の支援をタイミング良く行うことが可能になります。例えば、把持対象に関する情報を遅延なく提供することや、危険物を掴まないように警告を行うこと、その他にも様々な支援が考えられます。

把持行動の典型的な時間的・空間的構造を図3に示します。まず最初に手の移動が始まり、それにやや遅れて手が開き始めます（これをプリシェイピングと呼びます）。様々な先行研究から、プリシェイピングの完了する時刻と物体に手が到達する時刻に一定の関係があることがわかっています。また、どのような形状の物体を掴もうとしているかによって、手の開き方や掴む際の閉じ方が異なります。

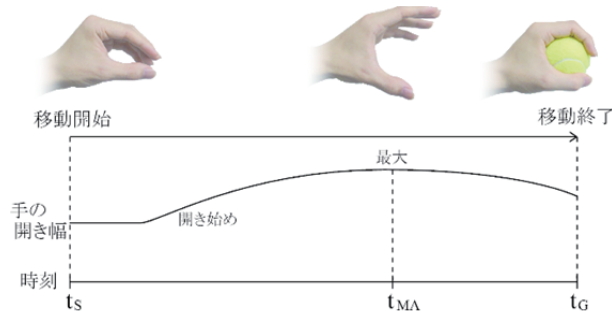


図3：把持行動の構造

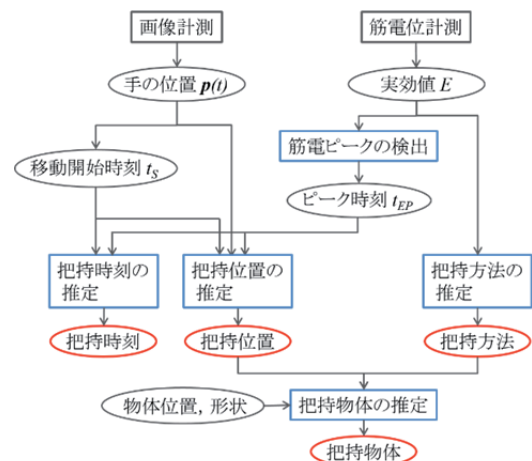


図4：把持予測のデータ計測・処理の流れ

このような知見を基に、把持の予測を行う手法を提案しました[5]。その方法は図4のようになっています。画像により手の位置を計測し、筋電計測により手を開く動作を検出するとともに手の形を認識します。この結果を統合して、把持時刻、位置、把持方法、把持物体を順に推定します。実験の詳細は

文献[5]をご覧くださいこととし、ここでは簡単な結果のみを紹介しします。手の移動が60%程度の時点で10～20%程度の誤差（移動距離や移動時間に対する誤差）で時刻や距離を推定でき、移動が完了する時点まで誤差が単調に減少します。また、掴み方（実験では4つの掴み方からの選択）は移動し始める時点でも70%程度の精度が得られ、移動が60%程度の時点では90%近い精度が得られています。

3. 人間と機械・情報システムとの協調のデザイン

統制された環境下では、精度の良い画像認識システムが実現できるようになってきました。しかし、一般的な環境、例えば、人間がシステムを意識せずに自由に振る舞ったり、様々な未知の物体が持ち込まれたり、照明条件が任意に変化したりする等の悪条件下では、残念ながら100%に近い精度を得る目処は立っていません。

これを図5のようにモデル化してみます。認識対象や環境の状態を「状況」とし、それを空間（「状況空間」と呼ぶ）の一点で表すことにします。システム（認識プログラム）は環境や認識対象の状態に対する何らかの「前提条件」を仮定しており、それが満たされている（状況が図5のSに含まれる）場合には正しい認識結果が得られ、それを逸脱した場合には、結果は保証されません。従来の研究の多くは、この前提条件（S）をできるだけ一般的なものにし、多様な状況をカバーする方法を模索してきました（図5のBの矢印）。この努力により、パターン認識技術は大きな進歩を遂げてきましたが、人間が行動する環境を対象とする場合には、依然として想定外の状況が簡単に起こり得ます。

このような背景のもとで我々が取り組んできた手法は、人間が機械の欠点を補えるように機械の振る舞いやインタフェースをデザインすることです。人間が介在することは、問題を難しくするという悪い面だけでなく、人間によって状況を改善できる可能性があるという良い面も含んでいるからです。これにより図5の矢印Aに相当する「状況の改善」を得る方法を設計します。ここでポイントとなるのは、大きな負担なしにユーザが協力できるように、(1) システムの内部状態をわかりやすくユーザに伝えること、(2) 状況が悪い（システムが正しく動きにくい）時、それを改善する方法をユーザに伝えられること、等となります。また、それにより、(3) ユーザがシステムをうまく使いこなすコツを習得できること、が期待できます。このような方向性は、システムの精度自体を高める研究と相補的に機能し、人間にとって使いやすく、良い結果の得られるシステムを設計する方法論となることが期待できます。

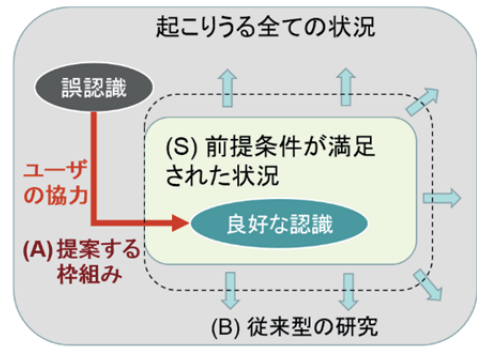


図5：前提条件と誤認識の関係

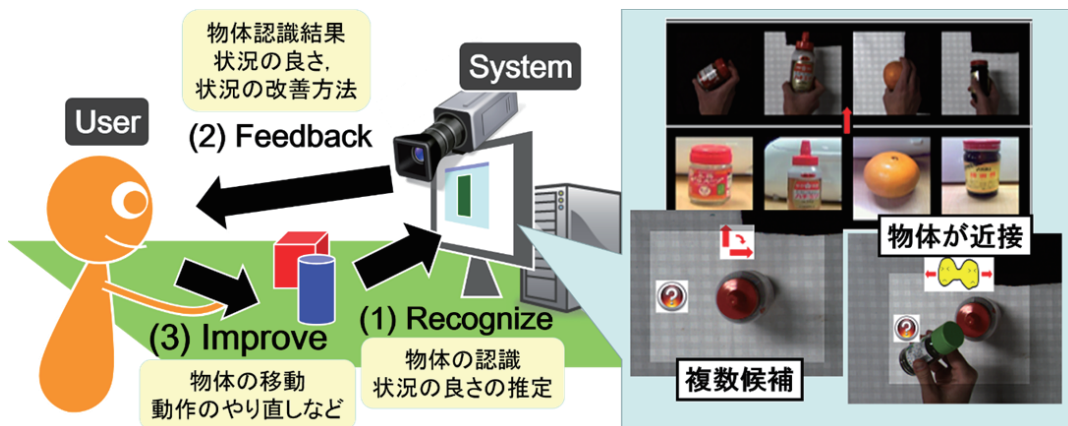


図6：物体認識をユーザと協調して行うシステム

このような考え方に基づいた2つの研究例を紹介します。一つは認識処理が失敗する悪状況に対応する機能を予めプログラムしておく方法であり、もう一つはそれを大量のデータから学習する場合です。

3.1 物体認識への適用

上記の考え方を物体認識システムへ適用した例を図6に示します[4]。ここでは人間（ユーザ）が系に含まれており、(1) 物体の認識と状況の良さの認識、(2) ユーザへの情報呈示と状況改善の提案、(3) 利用者によるシーンの改善や問題自体の変更、を含むループ系が構成されます。系の動作は、(1)～(3)を繰り返すことによってユーザを満足させることが目的です。ここで、悪い状況とその改善策としては、表1のものを用意しました。これらは因果関係や事前知識に基づいて人手で構成したものであり、さらに多様な悪い状況の検出方法や有効な改善策を収集することは今後の課題となっています。

表1：悪状況とその改善策

i	認識の失敗	ユーザに示される改善策
1	認識結果として複数候補が残存している	物体の向きを変える
2	認識結果の候補がない	同上
3	認識結果の候補が2つ、かつ悪状況が検出されない	正解を教示するインタフェースを提供
k	認識困難な状況	ユーザに示される改善策
1	物体の移動が速い	一度動きを止める
2	複数物体が近接している	近接している物体を離す
3	物体色が背景色と類似している	異なる色を持つ背景に移動させる
4	物体が鏡面反射を起こしている	照明条件が良くなるように向きを変える/移動させる
5	物体が暗すぎる	同上

簡単なシステムを構成し、ユーザが一定時間以内にシステムに正しい認識をさせられるかどうかを試した結果、全般的に、認識精度と正しい認識までにかかる時間の両方の面で、性能が向上しました。ユーザがシステムに協力する（システムの状況改善要求に応える）時間を含めてもトータルの時間が短縮されたことは興味深い結果と言えます。これは主に、ユーザがあてのまない試行錯誤をする時間が減ったことに起因します。ただし、依然として100%の認識精度とはなっておらず、今後より多様な状況に対応する必要があります。

3.2 ジェスチャインタフェースへの適用

ジェスチャインタフェースは直感的で使いやすいインタフェースを目指して開発されてきました。しかし、実現されているジェスチャインタフェースは必ずしも使いやすいものではなく、かえってユーザのストレスを溜める場合も多いのが現状です。人間の振る舞いや周囲環境の多様性にシステムの能力が追いついていないことが元々の原因ですが、失敗が起こった際のシステムの振る舞いをユーザが理解・予測できないこともそのストレスを大きくしています。

その一つの解決方法として、図5の状況空間を実際に構成し、その空間の中で状況を良い方向へ向かわせるインタ

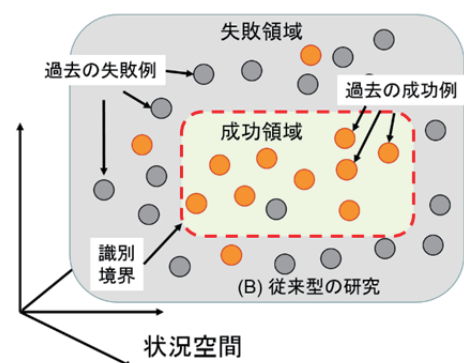


図7：成功領域と失敗領域を大量データから推定する

フェースを設計した例を紹介します。状況空間の次元（軸）として、システムが用いているアルゴリズムの前提条件を用いれば良いこととなりますが、複雑な認識処理に対してこのような指標を設定することは困難です。そのため、前提条件そのものではなく、前提条件の充足に深く関与する指標を冗長性を許して選ぶことにします。次に、図7のように状況空間の中で良好な認識が行われる領域とそうでない領域を区別します。その方法として、正解データ付きの多数のデータを利用し、成功例が大半を占める領域（以下、成功領域と呼ぶ）とそうでない領域（失敗領域と呼ぶ）の間に識別境界を設定します。また、この識別境界からどれだけ成功領域側に入っているかを状況の良さ（状況スコアと呼ぶ）と定義します。システムが状況スコア、および、状況スコアを良くする方法をユーザに提示できれば、ユーザがシステムに協力できることとなります。ただし、良い状況へ向かう方法は無数にあり、その目標点・経路の選び方、その提示方法が問題となります。例えば、以下の2つが考えられます。

- シンボル型フィードバック：ユーザが状況の改善のために行える行動を予めリストアップし、それらの行動を行った場合の状況スコアを推定する。改善されたスコアが良く、ユーザの負担が小さいものを選択する。
- 事例型フィードバック：蓄積された過去のデータから、成功領域内にあり、現在の状況にできるだけ近いものを選択する。



図8:ユーザの協力を促すための表示例（どのようにユーザが動けば認識がより正確になるかを伝える）。左：ディスプレイ上のコンテンツを手で操作するインターフェース（ユーザに画像で情報提示できることを想定）、中：シンボル型フィードバック（手のひらをこちらに向けてるように矢印などで指示）、右：事例型フィードバック（手を少しあげた過去のサンプルを例示）

その簡単な実装例を図8にあげます。ここで対象としているジェスチャは2種類の手形状で、どちらでもない場合を合わせて、3カテゴリの識別問題となっています。正解付きのサンプルはそれぞれの手形状について2万例を与え、成功領域と失敗領域の識別境界を求めるためにサポートベクターマシン（SVM）を用いています。簡単な例ではありますが、2種類の提示のどちらによってもユーザが協力できること、それによって状況スコアが増加する（認識率が上がる）ことが確認できています。また、上記の2つの方法を比べると、事例型フィードバックの方が協力を要する時間が長く、ユーザに少し多くの負担をかけていること推測されますが、一回の協力による状況スコアの増加は大きい傾向が見られました。ユーザが許容できる負担や慣れに応じて使い分ける必要性が示唆されています。

このように、大量のデータからシステムの特性を解析し、人間を含んだ系としての総合的な性能を向上させていく方法論を構築することが重要だと我々は考えています。

4. おわりに

本稿では、人間にとって安心して使えるロボットや情報メディアを実現するための人間のセンシング

や機械システムとの協調のデザインについて紹介しました。必要とされる機能を十分な精度で提供するためには、今後一層の研究を必要としています。脳科学の最新の成果を利用した生体計測や人間の特性のモデル化などを取り込んで、今後さらに発展するとともに、奥の深い分野となることが期待できます。

参考文献

- [1] M. Ozeki, S. Maeda, K. Obata, Y. Nakamura, *Multimedia Tools and Applications*, 44 (3), pp. 433-448, 2009
- [2] 青山秀紀、尾関基行、中村裕一、電子情報通信学会論文誌、J92-D (8), pp. 1317-1328, 2009
- [3] K. Ishikawa, M. Toda, S. Sakurazawa, J. Akita, K. Kondo, Y. Nakamura, *IEEE/ACIS International Conference on Computer and Information Science (ICIS 2010)*, pp. 37-42, 2010
- [4] 近藤一晃、西谷英之、中村裕一、電子情報通信学会論文誌、Vol. J94-D (8), pp. 1206-1215, 2011
- [5] 青山秀紀 他、電子情報通信学会論文誌、Vol. J95-D (3), pp. 527-538, 2012
- [6] Zhiwen Yu, Zhiyong Yu, X. Zhou, C. Becker, and Y. Nakamura, *IEEE Trans. on Knowledge and Data Engineering*, 24 (4), pp. 759-768, 2012
- [7] 中村裕一、近藤一晃、小泉敬寛、システム制御情報学会誌、56 (2), pp. 90-95, 2012
- [8] Zhiwen Yu, X. Zhou, and Y. Nakamura, *IEEE Pervasive Computing*, 12 (2), pp. 68-75, 2013
- [9] 中村裕一、情報の科学と技術、63 (2), pp. 57-62, 2013
- [10] 朝倉僚 他、電子情報通信学会論文誌、Vol. J97-D (1), pp. 50-61, 2014