

ポインティングデバイスとしての身体動作

久野義徳

埼玉大学情報システム工学科

kuno@cv.ics.saitama-u.ac.jp

概要: 人間は様々な身体動作で情報を発信できる。ここではそれを主としてポインティングデバイスとしてヒューマンインタフェースに用いることについて述べる。まず、ビデオカメラの画像から手の動きを認識することにより、手の動きでCG像やロボットを動かすことのできるシステムを紹介する。次に、行きたい方向に顔を向けることにより操縦できる知的車椅子について述べる。この車椅子では、周囲の歩行者の顔を見て、その顔の向きにより車椅子に気づいているか判断し、避け方を変える。最後に、この種のシステムに関して今後の課題を議論する。

Body Actions as Pointing Devices

Yoshinori Kuno

Department of Information and Computer Sciences, Saitama University

Abstract: Humans can transmit information through various body actions. This paper describes several systems using such body actions mainly as pointing devices for human interfaces. Firstly, we show systems that can recognize our hand motions from video images. We can thereby move computer graphic images or robots by our hand movements. Then, we introduce an intelligent wheelchair. We can turn it in a desired direction by turning our face in the direction. It changes the collision avoidance method with an approaching pedestrian by judging whether his/her noticing it from his/her face direction. Finally, we discuss issues concerning such human interfaces.

1 はじめに

人間同士のコミュニケーションでは言語の他に、視線、表情、ジェスチャなど多数の非言語的行動が重要な役割を果たしている。そこで、コンピュータと人間のコミュニケーションといえるヒューマンインタフェースに非言語的行動を利用しようという研究が盛んになっている [1]。人間の非言語的行動には無意識的・非意図的なものが多いが、コ

ンピュータへの意思伝達に使うという点から、現時点では、意識的・意図的な行動が対象になっていることが多い。すなわち、ジェスチャでコマンドを送ったり、手や視線、あるいは顔の向きで対象を指し示したり、対象の動きを操作するものである。これらは、ジェスチャによるコマンドを除けば、マウスなどによるポインティングにあたる。つまり、身体動作がポインティングデバイスとして使われている。これに関する先駆的な研究は MIT

のメディア研究所で開発された”Put-That-There”であろう [2]。ここでは磁気センサからの情報をもとに、操作対象や移動先を手で指し示すことができた。このようなことをビデオカメラの画像データから行うことができれば、装着物も不要で人間の行動を拘束しないものが実現できる可能性がある。そこで、多くの研究が進められている。ここでは、それに関して著者が関与してきた研究を紹介し、今後の課題を議論する。

2 ポインティングデバイスとしての手の動作

指でものをさしたり、手で対象の動きを示したりするのは、よく用いられる表現手段である。また、手でものの形や大きさを示すことも多い。これらは直接的なポインティングの動作であり、それをビデオカメラの画像から認識できれば、そのまま使いやすいインタフェースになると期待される。特に、マウスの動きが2次元平面上に限定されるのに対し、手は3次元空間で動かせる。したがって、3次元空間を対象とした場合に便利なインタフェースが実現できるのではないかと考えて、研究を進めてきた。

2.1 運動視差を利用したCG像の操作

コンピュータビジョンの応用としてヒューマンインタフェースが有望だと考えて、最初に行った研究である [3, 4]。同一空間上にない4点を画像上で追跡する。ここで、4点のうち1点を除いた3点で作る3角形を考え、4点目がその3角形上にあると仮定する。そして、ある時刻から次の時刻に3角形が動いたとき、4点目がどこに移動するか求める。実際には、4点目は3角形を作る平面上にないから、その画像上の位置は仮定により求めた位置と異なる。この差が運動視差である。これを利用することにより、安定に3次元運動が求められる。

手の上に4点を定めて追跡するのは困難なので、

4色の色球を手袋につけ、それを追跡することによりリアルタイムで動作するシステムを完成した。そして、1992年のデータショウに手の動きにより3次元CG像を操作できるシステムとして参考出品するなど、コンピュータビジョンの技術がヒューマンインタフェースに利用できることを示した。

このシステムを使って気づいたのは、回転運動の指示の際の問題である。3次元空間での並進運動は手の動きで簡単にできる。しかし、回転運動は、少しならよいが、大きく、例えば、対象物を何回転もぐるぐる回したいという場合に問題が生じる。このシステムでは、手の動きをそのまま対象物の動きとしたが、人間の手は、そのようにぐるぐるとは回せない。実際に、人間がそのような意図を伝えるときは、手を所定の方向に回しては戻すことを繰り返す。このシステムでは、そのような動作をされると、対象物は所定の方向に回転したり戻ったりを繰り返すことになる。そのときは、この問題を解決するために棒に球をつけたものを作り、それを人間が持って動かすことにした。実際に使ったところでは、棒の方が操作しやすかった。しかし、コンピュータビジョンを使ったインタフェースとしては、こういう補助具を使わずに済むようにしたい。そのためには、人間の動作を直接対象の動きにするのではなく、間に人間の意図を理解する処理が必要になる。しかし、この問題はまだ解決していない。

2.2 動きの基準

先のシステムではディスプレイの中のCG像が対象なので、使用者はだいたいディスプレイの方を向いている。しかし、ロボットを手の動きで操作する場合などを考えると、使用者の位置や向きが変わる可能性がある。一つの場所で動かないで使わなければならないシステムでは使いにくい。また、さらに複雑な問題は、使用者が空間をどのようにとらえて指示を出すのかということである。具体的にいうと「右」とか「左」を何を基準に考えるかということである。例えば、ロボットが近く

にいて見えているときに、指である方向をさし示した場合、実世界でその向きにロボットに行ってもらいたいと考えるのが普通であろう。この場合は、指の向きを世界座標で考える必要がある。しかし、ロボットが遠くにいて、ロボットに積まれたカメラの映像を見ながら指示を送る場合は、ロボット（その上に積まれたカメラ）を自分の身体と重ねあわせて、位置関係を考えて思われる。すなわち、ロボットを右に動かすときは、自分の身体を基準にして手を右に動かすようなことが自然だと思われる。

このような問題を解決するために複数視点画像からの不変量を用いた、CG 像や移動ロボットを操作するインタフェースを開発した [5, 6]。

ここで用いた複数視点不変量の原理は簡単である。3次元空間上に5点 $X_i, i \in \{0, \dots, 4\}$ があると仮定する (図 1)。それらのうち同一平面上にない4点を用いて、 X_0 を原点とする基底ベクトル

$$E_i = X_i - X_0 \quad (i \in \{1, 2, 3\}) \quad (1)$$

を設ける。この基底ベクトルを用いると、第5点 X_4 は α, β, γ を適当に選ぶことにより、次のように表すことができる。

$$X_4 = X_0 + \alpha E_1 + \beta E_2 + \gamma E_3 \quad (2)$$

カメラの投影を weak perspective と近似して、 $X_0 \dots X_4, E_1 \dots E_3$ を画像上に投影し、投影された座標をそれぞれ $x_0 \dots x_4, e_1 \dots e_3$ とする。weak perspective は線形変換なので、以上で述べたことは異なる位置から観測された2枚の画像それぞれについて式(2)と同様の関係が成り立つ。すなわち、次式が成立する(2枚目の画像に関するものについては、' を付けて示す)。

$$\left. \begin{aligned} x_4 - x_0 &= \alpha e_1 + \beta e_2 + \gamma e_3 \\ x'_4 - x'_0 &= \alpha e'_1 + \beta e'_2 + \gamma e'_3 \end{aligned} \right\} \quad (3)$$

式(3)では、それぞれが2次元ベクトルの方程式であるので、未知数3に対して、式の数4である。これを最小二乗法により解いて、基準座標系に対する3次元座標ベクトル $[\alpha \ \beta \ \gamma]^T$ を得ることができる。

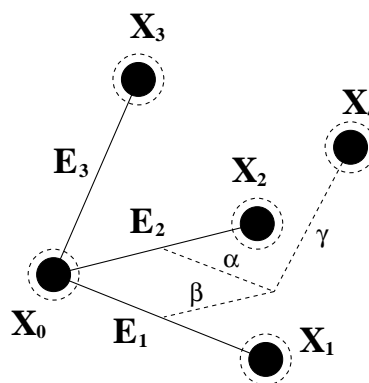


図 1: 複数視点不変量

同様に、3次元方向ベクトルを求める場合には、同じ座標系に対するもう1点 X_5 に対して式(3)を適用し、 $X_5 - X_4$ を方向ベクトルとすればよい。

これを利用してCG像やロボットを操作するインタフェースを開発した。この方法では、基準点が2台のカメラに写ってさえいればよい。したがって、機械的に精密なパンチルト機構でなく、簡単なもので特徴点を追跡すれば3次元情報が得られる。これで使用者の位置の制限の問題が解決できる。そして、基準点をシーンの中の固定物上にとれば、固定した世界座標で、使用者の身体の上にとれば、その人を中心に考えた座標系で空間情報が得られる。

このシステムでも、身体の上の特徴点を追跡するのは難しいので、図2のように身体の上に4つの球をつけて基準点とした。また、マーク付きの手袋をはめて手の特徴点とした。

しかし、実際には人体の上に基準点となる同一平面上にない4点をとるのは難しい。そこで、身体上には3点だけを基準点としてとり、4点目は仮想的に3点の作る平面の法線上にある点を考える方法を考案した[7]。これにより、使用者はさらに自由に動けるようになった。また、その利点を活かした応用として、液晶プロジェクタの表示画面内のCG像を動かすシステムを開発した。



図 2: ロボットを操作している様子

2.3 操作を意図した手の動きの選択

これまでに紹介したものは、手の動きに応じて対象物を動かすのが主な機能であったが、これに加えて、ものをつかんで置いたり、さらに両手で持って伸ばしたり縮めたりをジェスチャで行えるようにしたシステムを開発した [8]. しかし、これらのものには共通の欠点がある。それは、操作を意図しないで手を動かした場合にも、対象物を動かすジェスチャとみなされてしまう可能性があることである。したがって、操作を意図しないときは手を動かしてはいけないという、使いにくいものになってしまっている。そこで、これを解決するために他の非言語的行動の利用を検討した。考えたのは、対象物を操作しようというときはそれを見ているだろうということである。したがって、視線が求められれば、それが対象物に向いているときに手を動かしたときだけ、操作の意図があると考えればよいだろうということである。実際には視線の代用として顔の向きを求めて使用した [9, 10]. システムを開発して実験の結果、有効性を確認した。ただし、当然のことであるが、対象物の方を向いていても、それを操作する意図のない場合もある。これを解決するために、対象物を操作しようとしてじっと見ている場合と、画面をぼんやりと見ている場合を顔の向きの変化パターンから識別できないか検討している。しかし、現在までのと

ころ、まだ良好な結果は得られていない。

その他、手の動作を利用するものとして、スライドショーの操作を行うシステムについて検討した。ノート PC にカメラを搭載したものが発売されたとき、そのカメラを活用するものとして、手の動きにより、スライドを前に進めたり戻したりできるものを開発した [11]. さらに、液晶プロジェクタで投影した画面をカメラでとらえ、指示棒やレーザーポインタなどで画面をさしたときに、その位置情報を得られるようにした。これを用いてスライドの前後の他に、スライドの一覧を出して、その上で指示したものを表示することもできるようにした [12]. これは、直接的なポインティングデバイスとしての手の動作の利用といえる。

3 ポインティングデバイスとしての顔の向き

手のジェスチャの他にポインティングデバイスとして検討されているのは視線である。しかし、人間の場合、視線は手ほど直接的なポインティングの手段ではない。したがって、視線により細かいポインティングを行うのは手の不自由な人には有効だが、一般には使いにくい面があると思われる。そこで、概略の視線情報の有効な活用を検討している。その一つが、前章の最後に述べた、意図的に手を動かしているときの抽出である。なお、概略の視線情報ということなので、意図的部分の抽出のときと同様、実際には顔の向きを求めて利用している。

3.1 使用者の顔の向きの利用

使用者の顔の向きで操縦できる知的車椅子を開発した [13, 14, 15]. 使用者が行きたい方向を向けばそちらに回転する。これは意識的に行わなければならないが、目的の程度回転すると、使用者はほとんど無意識的に顔を正面に向ける。これが回転を止めることになる。自動車のハンドル操作では、回転の終了のために意識的に回転と逆方向

にハンドルを回すという操作が必要になるが，そのような意識的な操作の負担が軽減される．実際に病院でリハビリテーション中の車椅子利用者に乘っていただいたときも，事前の指示は「顔の向きに曲がります」というだけで，5人の被験者全員が操作でき，この点では有効性が確認できた．

知的車椅子の外観を図3に，構成を図4に示す．センサとしては，搭乗者を見るカメラと外部を見るカメラ，それに16個の超音波センサがある．搭乗者を見るカメラから図5に示すようにして顔の向きを求める．入力画像(a)から明るい領域を求め(b)，そこから雑音的な部分を除き顔領域を求める(c)．図中の縦線は顔領域の重心を通る直線である．最後に顔領域の中の目や口などの顔の特徴に対応する暗い部分を抽出し，それら全体の重心を求める(d)．図中の細い縦線はその重心を通る直線である．顔全体の重心と顔特徴の重心のずれが，概略の顔の向きに相当する情報になる．

このシステムでは顔の向きは，ほぼフレームレートで求められる．しかし，顔は常に少しは動いているので，顔の動きで車椅子を直接制御しては動きが安定しない．そこで，細かな動きには反応せず，意識的に顔を動かしたときにだけ反応するように，あるフレーム数のデータを平均して平滑化して使用する．しかし，どの程度，平滑化するかが問題である．回転を意図しない細かな動きに反

応しないようにするには平滑に用いるフレーム数を多くしてやればよいが，そうすると回転させようというときにも，なかなか回転を始めず，操作感が悪くなる．実験の結果，一つの固定値では難しいことがわかった．しかし使用者の感想から，

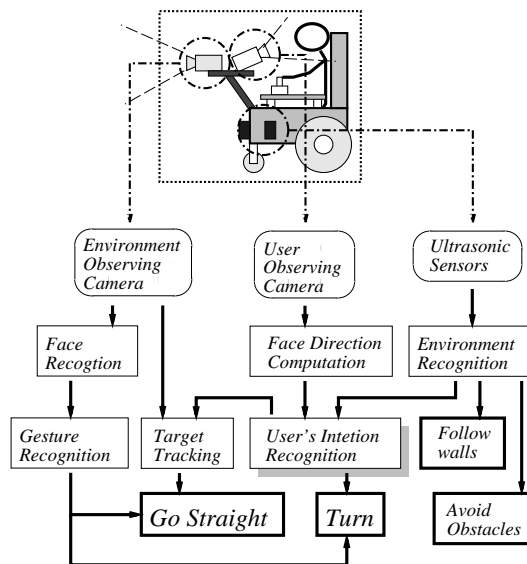


図 4: システム構成



(a) 入力画像

(b) 高強度領域



(c) 顔領域

(d) 顔特徴

図 5: 顔の向き計算



図 3: 知的車椅子の外観

左右に曲がる時は意識して顔を動かすので平滑の程度を大きくしても操作感はそれほど悪くないが、回転を止める時、すなわち、顔を正面に戻すときは、ほとんど無意識的な動作のためすばやく動かすので、そのときに反応が遅いと操作感が悪いということがわかった。そこで、左右へ曲がる時は平滑の程度を大きく、中央に戻すときは小さくすることにした。

しかし、顔をゆっくりと動かしても、そちらに曲がるつもりでないこともある。例えば、壁にはつてあるポスターなどを走りながら見るような場合は、壁の方を向いていてもそちらに曲がりたいのではない。また、一般に何が近づいてきたら、そちらを見るのが普通である。この場合も、そちらに曲がりたくないわけではない。そこで、超音波センサのデータを用い、近くに物体がある場合は、その方向への顔の向きの平滑の程度を大きくするようにした。すなわち、少しそちらを向いただけでは曲がらないようにした。ただし、近くに物体がある場合にも、そちらに本当に行きたい場合もあるので、じっとそちらを向けば曲がるようになっている。

実験結果の詳細は省略するが、学生による予備実験と病院のリハビリテーション施設における車椅子利用者の実験で良好な結果を得ている。

3.2 周囲の人の顔の向きの利用

知的車椅子に関して、周囲の歩行者の顔の向きを利用して円滑な衝突回避を行う方法を検討している。人間は興味のあるものの方を頻繁に見る、すなわち顔を向けると考えられる。そこで、周囲の歩行者の顔の動きを調べ、車椅子の方をよく見てくれるか調べ、それに応じて車椅子の衝突回避動作を変更することを検討した [16, 17]。車椅子が多く走行する病院の廊下での観察と、いろいろなパターンに対する衝突回避実験の結果、車椅子に気づいていて車椅子より速く歩いている人に対する場合は、車椅子が少し手前で速度を落とし、歩行者に回避をうながすのが良いことがわかった。ま

た、車椅子に気づいていない人や、車椅子より遅く歩いている人に対しては、車椅子の方から避ける方がよい。もちろん、前者の場合でも、歩行者の方が避け始めなければ、車椅子の方が避けなければならない。

以上のことを次のようにして実現した。超音波センサで障害物を検出したら、それを人間と仮定して、標準的な人間の大きさと超音波センサの距離データから、カメラのズームレンズの焦点距離を変えて、人間だとしたら顔の周辺あたりになる画像データを入力する。その画像から肌色の検出と目などの顔特徴の検出を行い、顔部分が検出できるなら、その物体は人間だと判定する。そうでない場合は、車椅子は超音波センサのデータを基に障害物を回避する。これは静止障害物や後ろ向きの人間が対象になる。人間と判定した場合は、以後、顔領域の追跡を行う。そして、目の位置から顔の向きを求める。顔が一定の頻度以上で車椅子の方向を示すときは、車椅子に気づいていると判定する。また、超音波センサのデータから対象物の速度も計測する。この2つのデータから、前述のような基準で衝突回避の方法を選択する。

4 今後の課題

人間は全身から情報を発信しているといえる。その情報をとらえられれば、使いやすい機械が実現できると考えられる。しかし、特別な用途を除いては、既存の手段に対して優位なものはなかなかできないというのが現状であろう。コスト面の問題もあるが、その他の主な課題として以下の3つがあると考えて、研究を進めている。

1. 画像特徴の抽出と追跡の確実化

これはビジョンを使ったシステムに共通の課題だが、画像特徴の抽出や追跡をより確実に安定にする必要がある。ここで述べた研究でも、初期のものは処理速度の問題もあるが、マークを装着してこの問題を軽減している。もちろん、インタフェースとしては装着物が不要なものが望ましい。現状

では手の動きなどは肌色の部分を検出して追跡して求めることが多いが、長袖の服を着て、手が同じ生地 of 服の胴体の前を動く場合など、それを確実に追跡することは難しい。いくつかの運動モデルのどれかに従って動くなど、ある程度、手の運動パターンが限定できる場合には、その事前知識を利用することによって問題が解決できないか研究を進めている [18]。

2. インタラクションによる失敗の回復

前項の研究が進み、処理がかなり確実に became としてもビジョンによるシステムを必ず成功するようにするのは難しい。そこで、失敗を回復する機構を導入する必要がある。ヒューマンインタフェースでは、機械の相手として人間がいるのだから、この失敗回復に人間を活用できないか検討している。しかし、機械の使用者は一般にビジョンの専門家ではない。そこで、どのようにすれば人間に負担にならない形で、その人が機械がどういう失敗をしているか知り、そしてそれを回復するための有益な情報を与えられるかが問題になる。これについては、機械の方が音声で何が分かって何が分からないのかという現状を伝え、それに対する人間の音声やジェスチャによる反応を認識することにより、問題を解決する方法を検討している [19, 20]。また、ジェスチャ認識では、未知のジェスチャをされたり、学習したものでも環境の変化により認識が確かでない場合がある。この場合、ジェスチャの意味を推定し、その結果に基づく行動 (推定できないときは可能な行動の中から適当に選択) を少し示し、それに対する人間の反応を見て推定が正しいかどうかを判定するロボット用のインタフェースも検討している [21]。

3. 新しい応用分野の開拓

ビジョン以外の手段であまり問題なくできているものをビジョンでやっても勝ち目は少ない。2次元的なポインティングなら、普通の状況ではマウスやタッチパネルで十分だろう。また、わざとらしい大げさなジェスチャを覚えなければならないようなものより、音声を利用する方が実用的で

あろう (音声とジェスチャを相補的・協調的に使うのは有望で検討している [19, 20])。

それではビジョンの特徴を活かせる応用は何であろうか。これに対してはまだ結論がない。ビジョンでは他のセンサと違い多くの情報が得られる可能性がある。しかし、それらを確実に得られるかという問題が多い。一方で、コンピュータには視覚認識は難しいが、人間は非常に得意なので、人間と機械のコミュニケーション手段として画像が使えるような応用は有望と思われる。ポインティングではないが、テレビのリモコンなどをどこに置いたか忘れてしまったような場合に対して、ものを置いたと思われる行動をした時点の画像を蓄積して、必要に応じてそれを一覧表示することで、人間を助けるシステムを開発している [22]。

もう一つ応用を考える上で新しいものが出るかもしれないと思っているのは無意識的・非意図的な行動の利用である。人間の非言語的行動には多くのものがあるが、その多くは、特定の意図をもって意識的に行うというものではなく、無意識的・非意図的である。そして、そのように無意識的・非意図的なのに、その人に関する重要な情報を伝達している。2.3 や 3.2 で述べた研究は、これに関する初期的な研究である。実際に、どういう状況でどのようなものを見れば、有用なシステムが作れるかは今後の課題である。

5 おわりに

ポインティングデバイスとしての身体動作ということで、手のジェスチャと顔の向きをコンピュータビジョンの技術で求め利用する研究について述べた。この種のヒューマンインタフェースでは高速に反応してくれなければ使いものにならない。実験システムを作り、想定したような動作は確認できたが、実装の面では速度や安定性に関して何とか使えるレベルにしたという程度である。したがって、本当に他の手段に比べて有効かどうかの定量的な評価の段階までには至っていない。前章で述べた課題に加え、この点も今後の課題である。

本研究の一部は科学研究費補助金 (07650492, 09221217, 0955080, 12650249) による。

参考文献

- [1] 黒川隆夫, “ノンバーバルインタフェース,” オーム社, 1994.
- [2] R. A. Bolt, “Put-That-There,” *Computer Graphics*, Vol. 14, No.3, pp.262-270, 1980.
- [3] R. Cipolla, Y. Okamoto, and Y. Kuno, “Robust structure from motion using motion parallax,” *Proc. IEEE 4th International Conference on Computer Vision*, pp.374-382, 1993.
- [4] 岡本恭一, ロベルト チボラ, 風間 久, 久野義徳, “定性的運動認識を用いたヒューマンインタフェースシステム,” *電子情報通信学会論文誌*, Vol.J76-D-II, No.8, pp.1813-1821, 1993.
- [5] Y. Kuno, K. Hayashi, K.H. Jo, and Y. Shirai, “Human-robot interface using uncalibrated stereo vision,” *Proc. 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.525-530, 1995.
- [6] K.H. Jo, K. Hayashi, Y. Kuno, and Y. Shirai, “Vision-based human interface system with world-fixed and human-centered frames using multiple view invariance,” *IEICE Trans. Inf. & Syst.*, Vol.E79-D, No.6, pp.799-808, 1996.
- [7] 林 健太郎, 久野義徳, 白井良明, “ユーザの位置の拘束のないジェスチャによるヒューマンインタフェース,” *情報処理学会論文誌*, Vol.40, No.2, pp.556-566, 1999.
- [8] K.H. Jo, Y. Kuno, and Y. Shirai, “Manipulative hand gesture recognition using task knowledge for human computer interaction,” *Proc. 3rd IEEE International Conference on Face and Gesture Recognition*, pp.468-473, 1998.
- [9] 石山智之, 久野義徳, 島田伸敬, 白井良明, “視線情報による選択的ジェスチャ認識に基づくヒューマンインタフェース,” 第4回画像センシングシンポジウム講演論文集, pp.175-178, 1998.
- [10] Y. Kuno, T. Ishiyama, S. Nakanishi, and Y. Shirai, “Combining observations of intentional and unintentional behaviors for human-computer interaction,” *Proc. CHI 99*, pp.238-245, 1999.
- [11] 島田伸敬, 村嶋照久, 久野義徳, 白井良明, “プレゼンテーション補助のためのジェスチャインタフェース,” 第5回画像センシングシンポジウム講演論文集, pp.67-70, 1999.
- [12] 古川大輔, 島田伸敬, 久野義徳, 白井良明, “ジェスチャによるプレゼンテーション支援システム,” *インタラクシオン 2000 論文集*, pp.53-54, 2000.
- [13] Y. Adachi, Y. Kuno, N. Shimada, and Y. Shirai, “Intelligent wheelchair using visual information on human faces,” *Proc. 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.354-359, 1998.
- [14] 足立佳久, 中西 知, 久野義徳, 島田伸敬, 白井良明, “顔の視覚情報処理を用いた知的車椅子,” *日本ロボット学会誌*, Vol.17, No.4, pp.423-431, 1999.
- [15] S. Nakanishi, Y. Kuno, and Y. Shirai, “Robotic wheelchair based on observations of both user and environment,” *Proc. 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.912-917, 1999.
- [16] Y. Kuno, Y. Murakami, N. Shimada, and Y. Shirai, “Intelligent wheelchair observing the faces of both user and pedestrians,” *Preprints of IFAC Workshop on Mobile Robot Technology*, pp.232-237, 2001.
- [17] 村上佳史, 久野義徳, 島田伸敬, 白井良明, “知的車椅子のための歩行者の顔の観察に基づく衝突回避,” *日本ロボット学会誌 (採録決定)*.
- [18] 鄭 文皓, 久野義徳, 島田伸敬, 白井良明, “Coupled Switching Linear Model を利用した複雑なジェスチャー認識,” *情報処理学会研究報告*, Vol. 2000, No.106 (2000-CVIM-124), pp. 1-7, 2000.
- [19] T. Takahashi, S. Nakanishi, Y. Kuno, and Y. Shirai, “Human-robot interface by verbal and nonverbal communication,” *Proc. 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.924-929, 1998.
- [20] S. Cheng, Y. Kuno, N. Shimada, and Y. Shirai, “Human-robot interface based on speech understanding assisted by vision,” T. Tan, Y. Shi, and W. Gao (Eds.), *Advances in Multimodal Interfaces - ICMI 2000, Lecture Notes in Computer Science 1948*, Springer, pp.16-23, 2000.
- [21] 村嶋照久, 久野義徳, 島田伸敬, 白井良明, “人間と機械のインタラクシオンを通じたジェスチャの理解と学習,” *日本ロボット学会誌*, Vol.18, No.4, pp.590-599, 2000.
- [22] 川崎広一, 久野義徳, “行動認識を用いた記憶支援システム,” 第7回画像センシングシンポジウム講演論文集, pp.351-356, 2001.