

ヒューマン・モデリングのための顔情報計測ソフトウェアの開発

Face Measurement Software for Human Modeling

松本 吉央¹⁾ 開 一夫²⁾
Yoshio MATSUMOTO Kazuo HIRAKI

- 1) 奈良先端科学技術大学院大学 情報科学研究科 (〒630-0101 奈良県生駒市高山町 8916-5)
E-mail: yoshio@is.aist-nara.ac.jp
- 2) 東京大学大学院 総合文化研究科 広域科学専攻 (〒153-8902 東京都目黒区駒場 3-8-1)
E-mail: khiraki@idea.c.u-tokyo.ac.jp

ABSTRACT. The measurement of behaviors of a human is one of the major problems in Digital-Human research. The face of a human reflects its mental state such as the attention, the intention and the emotion. Thus the development of the technology for quantitative measurement of the facial information is very important for modeling a human. In this article, we describe the development of facial measurement software based on image processing technology. The key aspect of our system is the use of real-time stereo vision together with a simple algorithm which is suitable for real-time processing. We achieved a non-contact, passive, real-time, robust, accurate and compact measurement system for facial information such as the position and orientation of the face, gaze directions, blinking and lip motions. Many kinds of applications based on the developed software are presented to show the feasibility of the software.

1 背景

デジタル・ヒューマン技術において、人の行動を計測することは重要な課題である。特に人間の顔は、注意や意図、感情といった心理状態に応じて様々な変化を表出するため、その変化を定量的に計測する技術の確立は、ヒューマン・モデリングに欠かせない。近年、人間工学の分野において人体の計測技術が進み、人間の全身行動に関してはマークを取り付けることで非接触に各部の変位を計測することは可能になってきた。これに対して、人間の顔に関する情報を非接触で計測する技術はまだほとんど実用化されていない。

従来技術としては、視線方向を計測する装置として、頭部装着式のもの (Eye Camera) が開発され実用化されていた。また、頭の位置や姿勢を計測する装置として、リンク機構を用いた接触型のものや、磁気センサを用いたものがある。しかし、これらの頭部に装着する装置にはかなりの重量があったり、利用者に大きな拘束感、違和感を与えたりするため、人間の自然な状態での行動を計測することが難しいという欠点があった。この問題を解決するため、カメラを顔の前に置き、非接触で顔の位置・姿勢や視線を計測する方法が研究されてきた。従来研究で開発されたシステムは大きく以下の3つに分けられる。

- A. 顔のトラッキング
- B. 瞳の画像を用いた視線方向の計測
- C. 顔のトラッキング+瞳の画像を用いた視線方向の計測

A としては、Pentland ら [1]、Zelinsky ら [2] の研究がある。どちらも、単眼カメラを用いて顔の特徴点をトラッキングし、カルマンフィルタを利用して顔の位置・姿勢を検出している。また Zelinsky らはリアルタイム処理を実現しており、実際にロボットと人間のインタラクションの一つとしてジェスチャー認識に応用している [2]。しかし単眼のシステムでは顔の3次元の位置・姿勢を推定する処理

が複雑になる上、高い精度は期待できない。B としては、松田ら [3]、中野ら [4]、Clarke [5] のシステムが挙げられる。これらのシステムでは、目だけを拡大した画像を用いて顔の位置・姿勢は検出できないため、顔を動かす状況には対応できない。C としては、伴野ら [6] のシステムがある。このシステムではマーク付きの眼鏡をかけることで顔のトラッキング処理を簡単化した上で、視線方向の検出を行った。この研究ではステレオカメラを用いており、画像から直接3次元情報が得られるため位置・姿勢の推定アルゴリズムが簡単化でき精度も高いが、頭部に専用の眼鏡を装着する必要があり、非接触とはいえない。

2 目的

本プロジェクトは、頭部の動き、視線方向、瞬目、表情などを同時に非接触かつリアルタイムに計測するソフトウェアを開発し、人間の顔から得られる情報を記録するためのシステムを構築することにより、デジタル・ヒューマン研究における基本的な計測技術を確立することを目的としている。本ソフトウェアの開発は、大きく以下の4つのサブテーマから構成される。

- 1) 顔認識モジュールの開発
- 2) 顔計測モジュールの開発
- 3) 顔行動抽出モジュールの開発
- 4) 応用システムへの組み込みと評価

まず 1. においては画像中から顔を見つけて個人識別を行うことで、計測対象の人物を特定するためのモジュールを開発した。これは、2. の顔計測モジュールの初期化 (3次元顔モデルの選択) に当たる。次に 2. においては、顔の位置、姿勢、変形、視線方向、瞬目などを実時間で計測するモジュールの開発を行った。これらのデータは、顔に関する最も基本的な物理量である。次に 3. においては、2. で計測した物理量に基づき、注視対象、表情、ジェスチャなどを実時間で推定するモジュールを開発した。これらの情

報は、計測された物理量の組み合わせから抽出されるシンボル情報である。最後に 4. においては、1.~3. までで開発した一連のシステムを、認知心理実験システムやヒューマンインタフェースへ組み込むことで、その機能の確認、性能評価を行った。また、そこでの問題点をフィードバックすることでシステムの改良を行った。

なお、ソフトウェアの実装においては、

- リアルタイム処理の実現
- 様々なハードウェアプラットフォームでの実現

という 2 点を特に重視した。顔の位置・姿勢に関してはそれぞれ約 2mm, 2 度の精度で、また視線については約 5 度の精度で計測することができるようになり、その他に瞬き、口や眉の変形なども計測することができるようになった。また、顔の動きや特徴の位置変化から、簡単なジェスチャや表情の変化を認識することが可能となった。なお、本ソフトウェアを実装するハードウェア構成として、カメラについては、

- IEEE1394 高速度ステレオカメラ (80Hz)
- IEEE1394 高画質カメラ (30Hz)
- NTSC カメラ (30Hz)

の 2 種類から、PC については

- 高性能デスクトップパソコン
- 小型ノートパソコン

の 2 種類より、処理スピード、価格、実験環境などに応じて選択することができる。最も速いシステムでは 80Hz での処理を、また最も安価なシステムでは 10 万円程度でのシステム構築を、またポータブルなシステムでは 1.5kg 程度のシステム構築を可能とした。次に、開発した顔計測ソフトウェアを用いて、以下の応用システムを構築した。ヒューマン・モデリングの分野では、認知科学の分野で利用可能な心理実験における幼児の行動計測システム、自動車のドライバの車内/車外の注視行動を記録するシステム、の二つを構築した。また、インタフェース分野では、顔の動きによりポイントを操作する非接触型のコンピュータ・インタフェース、アイコンタクト、つまり視線が自分に向いていることを確認しながら対話する機能を持つロボットを開発した。

3 開発内容：アルゴリズム編

3.1 顔認識モジュール

まず、ステレオカメラから得られる画像中から、顔発見を行う機能を実装した。これには、顔全体を包含するテンプレート画像を用い、まず右カメラの画像から顔に似たパターンを検索し、次にその画像をテンプレートとして左画像から対応するパターンを検索する。これらのテンプレートマッチングには、明るさの変化（特に頭部の左右運動に起因するもの）に影響されにくいように、エッジ画像を用いる。また、顔全体を包含する低解像度のテンプレート画像を用いており、「基準顔」のモデル (Figure 1) を用いることで不特定のユーザの顔を発見することが可能である。このステレオ視の結果、カメラから一定距離内に任意のユーザがいる時に、顔領域を切り出すことが可能になった。

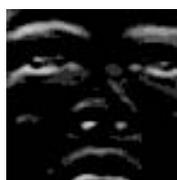


図 1: 顔発見に用いられる顔テンプレート

次に、検出された顔領域を用いて、顔認証を行うことでテンプレートを自動選択する機能を実装した。この機能には、商用の顔認証ソフトウェアである TrueFace を顔認証サーバとして用いる。TrueFace は Windows 上で動作するソフトウェアであるため、顔情報計測とは別の PC を用い、ソケット通信により検出した顔画像を送ることで認証を行うことにした。この部分に関しては、産業技術総合研究所の加賀美聡氏の協力のもと実装した。

また、検出された顔領域が登録していないユーザであった場合に、「基準顔」を元に目、口などの特徴をリアルタイムに画像中で探索し、自動的に 3 次元顔モデルを生成する機能を開発した。この機能により、顔の特徴をオペレータが手作業で指定せずに登録することができ、はじめてのユーザの計測を行う場合の利便性が大幅に向上した。Figure 2 は、一つの基準顔を用いてユーザの顔モデルを自動作成した時の結果の例である。



図 2: 顔モデルの自動構築の様子

3.2 顔計測モジュール

顔計測モジュールは、あらかじめ登録してある 3 次元顔モデルを用いて、顔に関する様々な物理量（顔の位置、姿勢、視線方向、口の開閉、瞬目など）をステレオカメラを利用して実時間で計測する機能を持つ。Figure 3 に、3 次元顔モデルの例を示す。顔モデルには、

- 顔発見用テンプレート画像
- 顔トラッキング用特徴テンプレート画像
- 顔トラッキング用特徴座標

が含まれている。

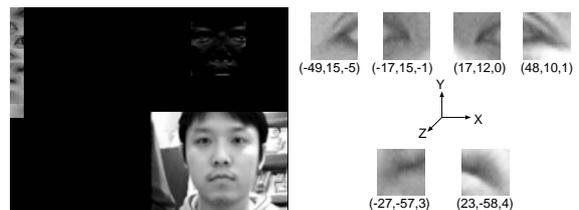


図 3: 顔モデル

顔情報計測モジュール全体の処理の流れを Figure 4 に示す。まず初期化として、顔認識部でも用いられる顔発見機能を用いる。そこで顔が発見された場合、顔モデルを用いて計測を開始する。計測においては、まず顔特徴を用いて顔のトラッキング（顔全体の位置・姿勢の推定）処理が行われ、その後、視線やその他の情報の推定処理が行われる。

ここではまず、顔トラッキング処理について述べる。3 次元顔モデルと 3 次元観測値の信頼性に基づく重み付けを考慮したモデルフィッティング処理により、高速に頭部の位置・姿勢 6 自由度を求めることができる。

$$E = \sum_{i=0}^{N-1} w_i (Rx_i + t - y_i)^T (Rx_i + t - y_i)$$

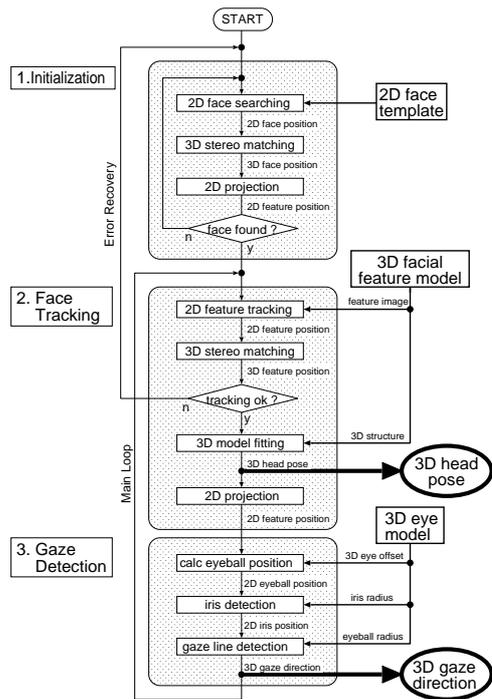


図 4: 顔計測モジュールの処理の流れ

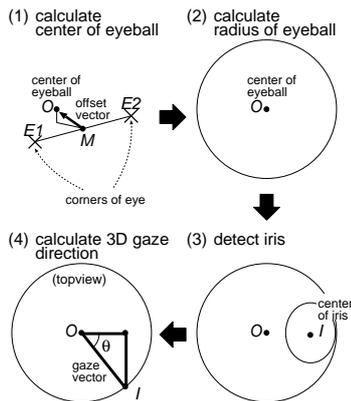


図 5: 視線のモデリング

ここで、 N は特徴の数、 x_i は顔モデル中のある特徴の位置ベクトル、 y_i は対応する特徴が観測された位置ベクトル、 w_i は観測の信頼性 ($0 \leq w_i \leq 1$)、 T, R は並進ベクトルおよび回転行列、そして E はフィッティングエラーである。最もよいフィッティングは、 E を最小化する T, R を求める問題に帰着でき、本ソフトウェアではこれを最急降下法を用いて解いている。

次に、目の虹彩の位置から視線方向を推定するモジュールについて述べる。顔トラッキングモジュールにより顔の位置・姿勢が計測できると、頭部に固定された眼球の中心位置を推定することができ、画像から得られる虹彩の中心位置と推定された眼球中心位置の関係より、水平・垂直の視線角度がわかるので、3次元ベクトルとして視線が得られる。この手順を Figure 5 に示す。また、眉、口の変形は、頭部位置・姿勢推定の後処理として別に特徴点追跡を行うことで、瞼はエッジ検出を行うことで、変形量を計測することができるようになった。

また、これらの計測結果は、以下の3つの方法で出力できるようにした。

- 画像と数値

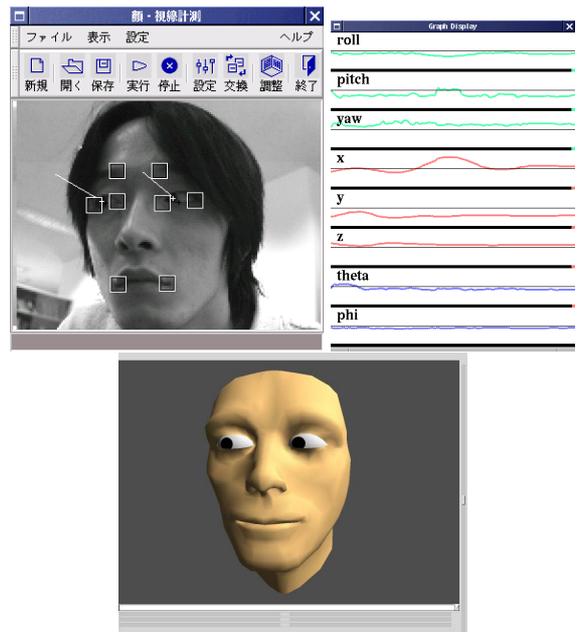


図 6: 顔計測モジュールの計測画面

- グラフ
- CG

その出力の様子を、Figure 6 に示す。Figure ??では、顔トラッキングに用いられる特徴を白い四角で、また推定された視線方向を直線で表示している。Figure ??では、顔の位置 (x, y, z) と向き (roll, pitch, yaw)、それに視線の水平・垂直の方向 (θ, ϕ) が表示されている。Figure ??では、これらのパラメータを用いてCGの顔画像を動かして表示しており、ユーザの顔の動きを真似しているような動きをすることになる。これら計測モジュールの画像処理は、約10msかかり、80Hzの高速度カメラのフレームレート内に収めることができた。

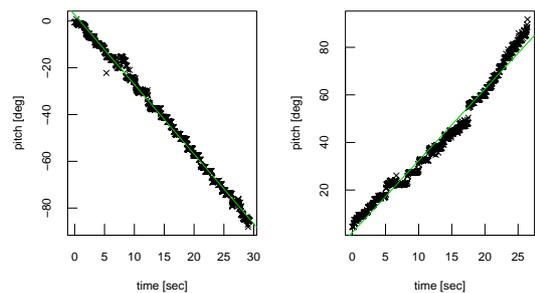


図 7: 頭部姿勢の精度評価実験の結果

次に計測精度に関してであるが、頭部位置・姿勢に関する精度評価実験の結果を Figure 7 に示す。ここでは定量的な計測を行うため、回転ステージに載せた人形の顔を用いて実験を行った。その結果、位置については誤差は約2mm、姿勢については誤差は約2度のいう結果が得られた。また、Figure 8 には、視線の評価実験の様子を示す。図中のボードには、10cm間隔でマーカが配置されており、約1m離れたところから被験者はボードのマーカを順番に見ながら、頭部位置・姿勢および視線方向の計測を行った。その結果、Figure 8 (下) のような視線の軌跡が得られ、順にマーカを見ていることが十分に判断できる。この軌跡

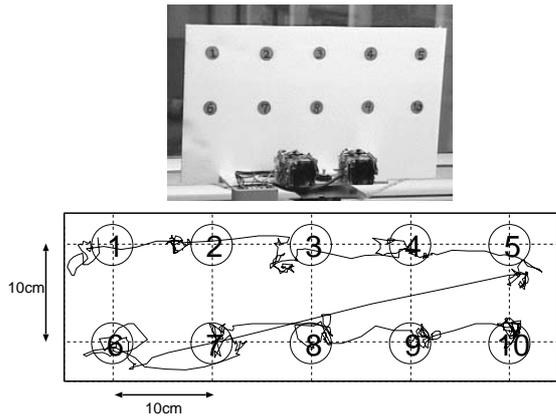


図 8: 視線方向の精度評価実験の結果

は、視線ベクトルとボードの交点を計算しプロットしたものである。このマーカの間隔(10cm)は、顔の位置から見ると約6度に相当する。この図から判断すると、視線方向は約5度の精度で推定できているといえる。

高速度カメラを用いることから、ステレオ画像の解像度は通常の CCD カメラの解像度よりも低い 320x240 画素に制限されることになった。そこで、計測精度としては当初の目標よりもやや低い位置誤差約 2mm、姿勢誤差約 2度の計測結果となった。

3.3 顔行動抽出モジュール

顔計測モジュールで計測した物理量に基づき、ジェスチャ、表情、注視対象を実時間で推定する機能を開発した。

まず、ジェスチャ認識では、計測した顔の動きからジェスチャを認識するモジュールを開発した。本システムでは、一般に行われている画像からのジェスチャ認識と違い、精度良く頭部の6自由度の動きパラメータが計測できるため、高い認識性能が期待できる。ここでの認識にはジェスチャ認識で広く使われている連続 DP を利用し、“Yes”、“No”、“Look Left”、“Look Right” のようなジェスチャを認識できるようになった。次に、表情認識では、眉、口など特徴点の位置関係を用いて、典型的な表情変化の認識を実現した。認識する表情としては、“口を開けている”、“片目を閉じている”のような行動的なものを対象とした。これらの認識の様子を Figure 9 に示す。

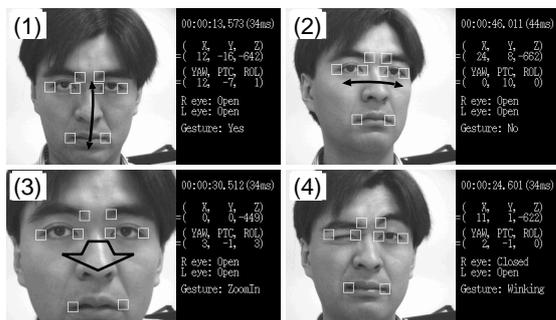


図 9: 顔行動抽出モジュールの処理の様子

次に、注視対象物の特定手順について説明する。はじめに、対象物の位置はカメラの原点からの3次元位置(グローバル座標系)で定義する。次にグローバル座標系で定義された対象物の位置をカメラ座標系に変換する。対象物は円形と仮定し、その大きさを円の直径を用いて定義する。対象物を定義すると、顔座標と視線ベクトルを用いた次式が成り立つかどうかを各対象について判定することで、対

象物を特定できる。

$$\tan^{-1}\left(\frac{L}{2\|P\|}\right) - \cos^{-1}\left(\frac{P \cdot G}{\|P\|\|G\|}\right) \geq 0$$

ここで、 P は顔からオブジェクトに対するベクトル、 G は視線ベクトル、また L はオブジェクトの直径である。

4 開発内容：実装編

4.1 ハードウェア

システムの実装にあたっては、様々なハードウェアプラットフォームでの実現を重視した。具体的には、

- IEEE1394 高速度ステレオカメラ (80Hz)
- IEEE1394 高画質カメラ (30Hz)
- NTSC カメラ (30Hz)

の2種類から、PC については

- 高性能デスクトップパソコン
- 小型ノートパソコン

の2種類から選んで組み合わせられるようにした。これらのシステムの形態は、処理スピード、価格、実験環境などに応じて選択することができる。最も速いシステムでは 80Hz での処理を、また最も安価なシステムでは 10 万円程度でのシステム構築を、またポータブルなシステムでは 1.5kg 程度でのシステム構築を可能とした。

また、これら任意のシステムの組み合わせを実現する上で、特に特徴的なこととして、フィールド多重化および CardBus キャプチャカードを用いることにした。

フィールド多重化 フィールド多重化とは、アナログレベルで2つのビデオ信号を多重化し、1つのビデオ信号にする技術である [7]。Figure 10 (上) にそれを用いた多重化デバイスのブロック図を示す。このデバイスには、外部同期のかかった2台のカメラからのビデオ信号が入力され、その2つの信号はアナログスイッチ IC を用いてフィールド毎に交互に出力される。この切替えは 60[Hz] という比較的低速で行われるため、実際のデバイスの実装は容易かつ安価な汎用部品だけを用いて実現できる。実際に作成したデバイスの写真を Figure 10 (下) に示す。このデバイスでは汎用部品だけを用いて片面に実装しているが、サイズは 5[cm] 四方に収まっている。本デバイスからの出力ビデオ信号の各フレームには、垂直方向の解像度が半分になったステレオ画像が含まれている。このようなアナログレベルでのビデオ信号の多重化の利点は、多重化された映像信号も NTSC ビデオ信号であるために、通常の NTSC ビデオ信号をキャプチャできるカードが1枚あればステレオ画像をキャプチャできることである。キャプチャカードを挿すスロットの数が少ない省スペース型 PC やノート型 PC では、フィールド多重化を行う利点は大きい。

CardBus キャプチャカード ノート PC で画像をキャプチャするには通常キャプチャデバイスが必要となるが、NTSC カメラをフレームレートでキャプチャできるデバイスは非常に限られている。我々は、日立 MSVCC03 という CardBus キャプチャカードを利用し、Linux 版ドライバを開発することで、顔情報計測ソフトウェアをノート PC で実行できるようにした。このカードを用いると最大 640x480 画素のカラー画像を、30fps で取り込むことができる。そのため、このカードとフィールド多重化技術を併用すると、通常のノート PC でステレオ画像をフレームレートで取り込むことが可能となる。ただし、このカードは既に生産中止となっており、入手は困難である。

4.2 ソフトウェア

前章で述べたモジュール、特にその中でもソフトウェアの核となる顔計測モジュールを実装するにあたっては、

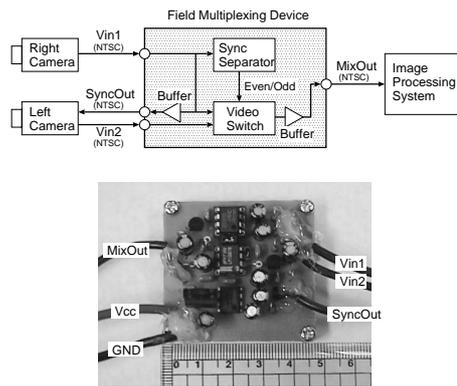


図 10: フィールド多重化回路



図 11: システム実装例 (左: デスクトップ PC+IEEE 高速ステレオカメラ, 右: ノート PC+NTSC カメラ)

リアルタイム処理の実現を重視した。リアルタイム処理の実現のためには、Intel Pentium 系プロセッサの MMX を利用した。全てソフトウェアにより画像処理を実現し、Pentium4 2.53GHz を用いた場合、

- 顔発見: 約 70ms
- 顔モデル自動作成: 約 20ms
- 顔トラッキング: 約 10ms

という処理速度が得られた。特に顔トラッキング (顔の位置・姿勢と視線方向の計測処理) を行うことができた。これにより、フレームレートが 80Hz の高速度カメラを用いてもコマ落ちすることなく処理が可能である。

また、様々なハードウェアでの実装および容易なメンテナンスを実現するために、ハードウェアの違いを吸収するような画像キャプチャライブラリを開発した。これは、Linux 上の共有ライブラリとして作られ、実行環境によって入れ替えるものである。

このために、ノートパソコン用キャプチャカードのデバイスドライバを開発し、またどのシステム構成でも計測ソフトウェアには変更せずに使えるように工夫した画像キャプチャライブラリを開発した。具体的には、

- NTSC カメラ用
- NTSC カメラ (フィールド多重化) 用
- IEEE1394 高性能カメラ用
- IEEE1394 高速度ステレオカメラ用

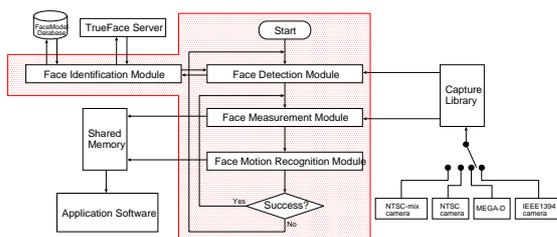


図 12: 開発したシステムの構成

の 4 種類のキャプチャライブラリを作り、そのうちの一つを libcapture.so というファイル名で置いて使う。顔情報計測側のプログラムは実行時にこのライブラリをロードすることになる。このようにしたこと顔情報計測側のプログラムは、実行プラットフォームによらずバイナリ互換とすることが可能になり、開発効率が向上し、また今後のプログラム配布における利便性、メンテナンス性の向上も期待できる。

以上、最終的に構築したシステムの構成を Figure 12 に示す。図中の網掛けになっている部分が顔情報計測システムの核となるソフトウェアである。また、計測した情報を利用するアプリケーションソフトウェアは別プロセスとして実行し、共有メモリ上に格納された計測結果を利用する。

5 開発内容: 応用システムの構築編

つぎに、開発した顔計測ソフトウェアを用いて、ヒューマン・モデリングのための計測・記録システムと、ヒューマンインタフェースの構築のためのユーザの計測・入力システムを開発した。

5.1 ヒューマン・モデリング分野

5.1.1 心理実験計測システム

応用システムの 1 つ目として、認知科学の分野で利用可能な、心理実験における人間の行動の非接触計測システムを開発した。従来の実験では、視線計測には頭部装着型のシステムを、また顔の動きを記録するにはビデオにより記録し後に手動で解析する手法が多く取られていたが、本システムを用いるとそれらの記録過程を自動化できる。

特に、乳幼児の実験には装着型の視線計測装置を用いることができなく、非接触型の顔・視線計測システムの実現が望まれていた。そこで、我々は開発した顔情報計測システムを乳幼児の実験に用いてみた。Figure 13 は、実験の様子を示す。被験者の乳幼児 (この場合は生後 12 ヶ月) は、テレビ画面に視覚的な刺激が写っているのを見ている。その映像の内容は人間とロボット、あるいは人間と人間の対話場面であり、その画面を注視している時間から、どの程度映像を理解できているかを調べる認知的な実験である (詳細は [8] を参照)。

乳幼児の顔は、映像への興味が低くなると左右に動くが、そのような場面でもかなりの部分で計測が可能であった。また、大きな動きにより計測が失敗した場面でも、顔が前を向いた瞬間に再び計測が始められるため、実験への支障はなかった。Figure 14 には、そのときの計測結果を示す。図の縦軸は顔の方向 (水平方向、正面が 0 度) であり、約 30 秒分の結果をグラフにしたものであるが、実際の実験はもっと長いもの (一人あたり 10 分程度) であり、計測データはファイルに保存されて後で解析に用いる。被験者への負担がまったくないため、このような乳幼児の認知心理実験においては特に実用的な計測システムである。



図 13: 幼児行動計測の様子

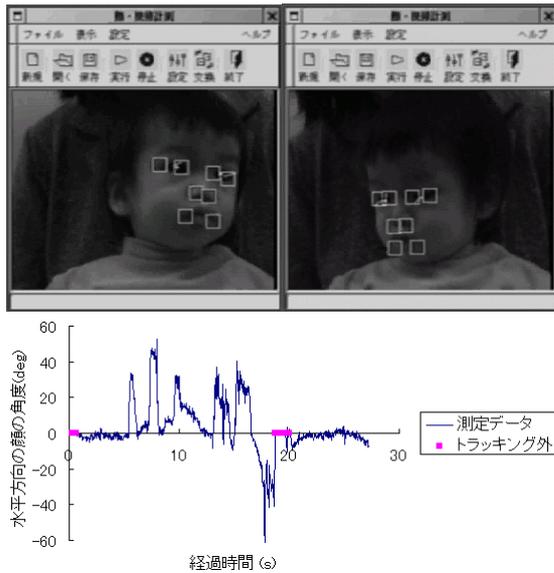


図 14: 幼児行動の計測結果

5.1.2 ドライバ行動計測システム

次に、自動車のドライバの挙動を計測・記録できるシステムを構築した [9]。そのカメラ、(c) は顔計測用カメラ、(d) は車外撮影用全方位カメラ、(e) は運転席の様子である。センサとしては、他に GPS、ジャイロセンサ、加速度センサ等が搭載されており、顔情報を含めた全ての計測データは同時にファイルに記録されるようになっている。

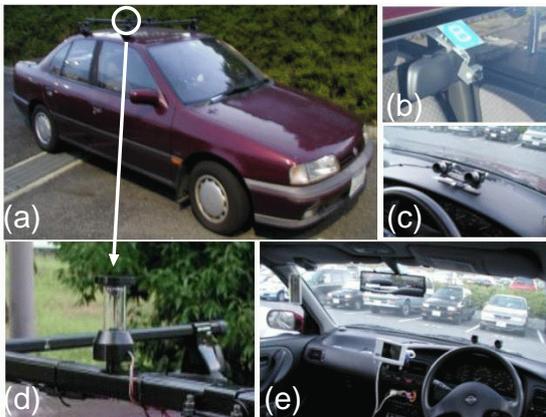


図 15: ドライバ行動計測システム

車内注視対象の認識 車内の対象物への注視時間を計測することによって、ドライバの行動推定が可能であることが知られている。そこで、運転席の配置をあらかじめモデリングしておくことで、メーター類、ミラーなどの車内の対象物を注視していることを検出することができるようにした。顔の位置・姿勢および視線方向などの情報は、カメラを中心とした座標系で計測される。車内の対象物は、Figure 15 に示すように座標を計測しておき、顔行動抽出モジュールの機能により注視対象が認識される。

車外注視の記録 車外に関しては、あらかじめ環境のモデルを構築しておくことはできない。そこで、外部環境の画像を撮影しておき、その画像上に視線方向や頭部方向を投影し、外部環境（例えば道路上）の注視対象を画像上で確認できるシステムを構築した。

外部画像の撮影には、周囲 360 度の撮影が可能な全方位視覚センサを用いることで、広い視野の画像を確保した。

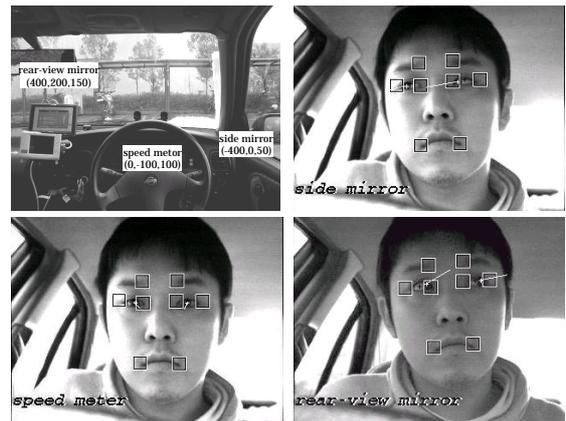


図 16: 車内対象物のモデリングと注視識別結果

その画像上に、ドライバの視線方向および顔正面方向をカーソルとして表示する。また、注視頻度の分布をヒストグラムとして表示できるようにし、どの部分にドライバの注意が集中するかを視覚的に表示できるようにした。

また、夜間はドライバの顔情報を計測するのが困難になるが、近赤外対応の CCD カメラを用い、赤外線投光器を設置することにより、Figure 15 に示すように夜間でもドライバの注視情報を計測できるようにした。

これらの車載計測システムの構築により、運転動作と視線情報の関係を解析することが可能になった。また、全ての情報はオンラインでも利用可能であるので、ドライバの運転支援を行うことも可能である。



図 17: 車外注視の記録ソフトウェア

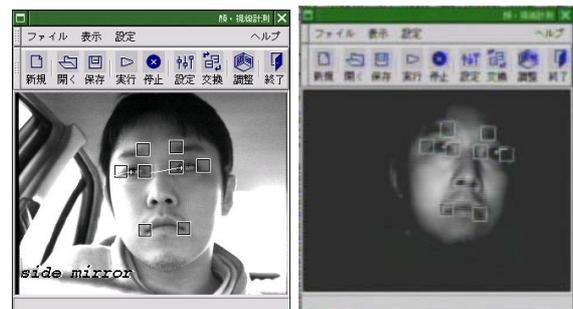


図 18: 昼/夜の走行実験の様子

5.2 インタフェース分野

5.2.1 ハンズフリー・コンピュータインタフェース

顔の動きによりポインタを操作する非接触型のコンピュータ・インタフェースの構築を行った。ここでは、顔情報計測の結果、顔の方向を用いて X Window 上のマウスを制御する。つまり、X Window 上でマウスで操作するソフトウェアなら何でも操作できることになる。視線で直接マウスの位置を指定するのではなく、ジョイスティックのように顔の向きを使って上下左右に動かす速度を与えることにした。これは、視線の精度がそれほど高くないためである。また、左右の目のウィンクで、左クリック、右クリックの動作をエミュレートした。現状では、マウスなど他のポインティングデバイスと比較して使い勝手がよいとは言えないが、手を使えない状況でも非接触でコンピュータを操作することが可能であるため、福祉分野などへの展開が期待できる。

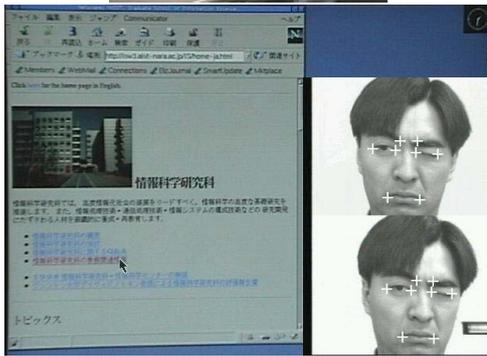


図 19: ハンズフリーマウスの操作の様子

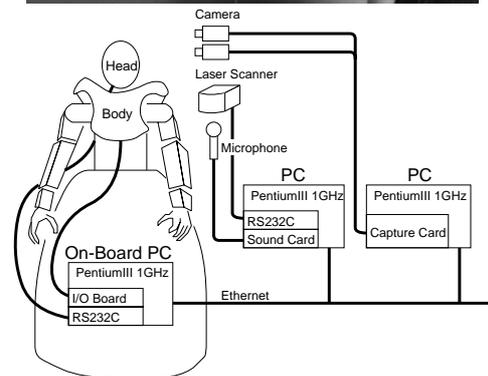
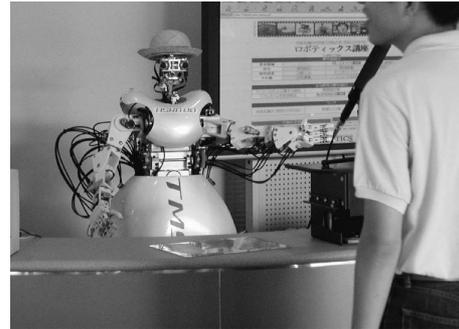


図 20: 対話ロボットシステム ASKA

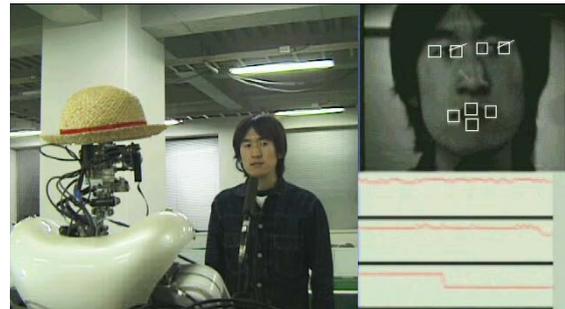


図 21: ロボットとの対話の様子

5.2.2 ロボット対話インタフェース

我々はこれまでに、人間型ロボットを用いた対話システムの研究を行ってきた [10]。そこでは、音声認識によりユーザの発話を理解し、ジェスチャを交えながら質問に答えるという学内の受付案内システムの開発を行っている。この対話システムに、顔情報計測ソフトウェアを組み込んだ。

従来のシステムでは、背景雑音により間違った反応をしてしまうことがあった。顔情報計測を用いると、ユーザの顔や視線、口の開閉を認識することができる。これにより、視線や顔が自分に向いていることを確認したら音声入力をスタートさせ、また口の動きが止まったら音声入力をストップさせ、背景雑音に反応しないようなシステムを実現した。ロボットがユーザの視線を意識しながらコミュニケーションを行うという意味で、ロボットと人間のある種の「アイコンタクト」が実現できたとも言える。

5.2.3 視線操縦型車いす

人間は通常、行きたい方向を見る。このことを利用して、視線や顔の方向に基づいて走行を行う車いすシステムを開発した [11]。Figure 15 に、開発したシステムの外観を示す。屋内用の電動車いす M-Smart (ミサワホーム製) をベースに作られている。この車いすは小型で、回転中心が座席の中心と一致しており小回りがきくという特徴をもっている。天井カメラと顔計測カメラが搭載されている。また、搭乗者の足もとにはレーザレンジファンダ LMS 200 (SICK 製) が搭載され、ロボット前方の距離

情報をスキャンすることができる。このロボットではノート型 PC (Pentium III 650MHz) をコントローラとして用いており、USB 経由の D/A 出力によりコントロールされる。このロボットは 2 時間以上 (ノート PC の駆動時間による制限) の走行が可能である。

車いすは、基本的にはユーザが見ている方向に進むが、レーザレンジファインダを用いた障害物回避機能と組み合わせることにより、壁のポスター等の方向を見ているもそこに衝突することはなく、(壁との角度により) 停止したり、壁に沿って滑らかに移動することができる。

また、本走行支援システムでは、顔と視線の方向を両方向同時に計測しており、顔の向いている方向と視線の示す方向が互いに異なる状態を判別できる。Figure 23 は、ジョイスティックによる手動操作において顔と視線の方向を計測した実験結果である。搭乗者が車椅子の操縦に集中している間 (150[s] ~ 340[s]) は、顔と視線の方向はほぼ一致していることが分かるが、搭乗者がわき見を始める (340[s] ~ 550[s]) と、顔と視線の方向が明らかに一致しなくなる。そこでこの実験結果から、顔と視線の方向の角度の違いが 15[deg] を超えた場合、“わき見状態”であると見なし、車椅子の速度を落としたり、車いすの旋回時にも速度を落とす。

この車いすシステムは、屋内・屋外を含めて様々な場所

においてデモを行っており、顔情報計測のための画像処理の安定性の高さを示している。



図 22: 視線操縦型車いす

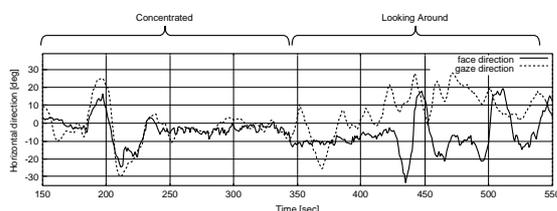


図 23: 集中度の認識

6 開発成果の特徴

本プロジェクトで開発した顔情報計測ソフトウェアは、従来開発されていたシステムと比べて以下の特徴を持つ。

- 非接触、非侵襲。
- 顔が比較的大きく動いても計測が可能。
- 顔の位置・姿勢も（別のセンサを組み合わせることなく）視線と同時に計測できる。
- 完全に受動的なシステムであり、視線の計測に近赤外線照射を用いない。
- 自動で顔モデルを構築することが可能。
- 瞬き、口の開閉など顔の特徴の変化や、ジェスチャも計測可能。
- 特別なハードウェアを用いず、汎用 PC だけ処理を実現しているので安価。
- 用途（処理速度、精度、システムのサイズなど）に応じて様々な PC およびカメラを選択できる。
- 共有メモリおよびソケット通信によるデータ出力機能を持ち、アプリケーションの構築が容易。

そのため、本プロジェクトで開発した応用システムも含めて、下記のような幅広い応用分野での利用が可能である。

- コンピュータインタフェース
- ロボットインタフェース
- 福祉機器（例．手を使わない機器操作）
- 自動車安全システム（例．居眠り、わき見の検出）
- 人間工学（例．人のモデリング、人に優しいデザインのための指標）
- 心理学（例．視覚刺激に対する反応の記録、解析）
- 画像伝送（例．顔画像のパラメータ抽出による圧縮）
- アミューズメント（例．遠隔対戦ゲーム）

7 今後の課題・展望

顔計測モジュールに関しては、視線推定のための眼球モデルの構築を自動で行うソフトウェアを作成する計画であったが、時間的に至らなかった。実用化にあたり必要性は高いと考えているので、プロジェクト終了後に引き続き開発する。また、顔行動抽出モジュールに関しては、表情の時系列の変化まで含めたジェスチャの認識には至らなかった。単純な特徴点のトラッキングするだけでは難しく、より詳細な画像処理が必要であることが分かったので、今後画像処理手法の検討を行っていく予定である。さらに、画像処理のサブピクセル化等の手法を用いて、視線方向推定の精度を向上させていきたい。

また、本プロジェクト開発後には、さらに幅広い応用分野においての研究協力を募り、応用システムのプロトタイプの開発および評価実験を進め、ソフトウェアの完成度を向上させる。また、ソフトウェアの市場調査を行い、近い将来的、ヒューマン・モデリングのための計測システム単体として、あるいはシステムを組み込んだ応用システム（インタフェース）として実用化することを目指している。

8 参加企業及び機関

奈良先端科学技術大学院大学、東京大学および松江精機工業株式会社。

参考文献

- [1] A.Azarbayejani, T.Starner, B.Horowitz, and A.Pentland. Visually controlled graphics. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 15, No. 6, pp. 602-605, 1993.
- [2] A.Zelinsky and J.Heinzmann. Real-time Visual Recognition of Facial Gestures for Human Computer Interaction. In Proc. of Second Int. Conf. on Automatic Face and Gesture Recognition, pp. 351-356, 1996.
- [3] 松田圭司, 永見武司, 山根茂. 高速楕円近似による汎用視線位置計測システム. 第 10 回生体工学シンポジウム論文集, pp. 401-404, 1995.
- [4] 中野倫明, 坂口靖雄, 山本新. 屋外環境における視線検出の一方式. 電子情報通信学会論文誌 (D-II), Vol. J79-D-II, No. 9, pp. 1638-1642, 1996.
- [5] Andrew H.Clarke. Current trends in eye movement measurement techniques. In W.H.Zangemeister, H.S.Stiehl, and C.Freksa, editors, Visual Attention and Cognition, pp. 347-364. Elsevier Science B.V., 1996.
- [6] 伴野明, 岸野文郎, 小林幸雄. 瞳孔の抽出処理と頭部の動きを許容する視線検出装置の試作. 電子情報通信学会論文誌 (D-II), Vol. J76-D-II, No. 3, pp. 636-646, 1993.
- [7] Y. Matsutomo, T. Shibata, K. Sakai, M. Inaba, and H. Inoue. Real-time Color Stereo Vision System for a Mobile Robot based on Field Multiplexing. In Proc. of IEEE Int. Conf. on Robotics and Automation, pp. 1934-1939, 1997.
- [8] 有田亜希子, 開一夫, 神田崇行, 石黒浩. ロボットは話相手になれる? : インタラクティブロボットに対する乳児の認識. 情報処理学会関西支部 環境知能研究会予稿集, 2002.
- [9] Kentaro Takemura, Jun'ichi Ido, Yoshio Matsumoto, and Tsukasa Ogasawara. Drive Monitoring System Based on Non-Contact Measurement System of Driver's Focus of Visual Attention. In Proc. of IEEE Int. Symposium on Intelligent Vehicle, 2003 (to appear).
- [10] Ryuichi Nishimura, Takashi Uchida, Akinobu Lee, Hiroshi Saruwatari, Kiyohiro Shikano, and Yoshio Matsumoto. ASKA: Receptionist Robot System with Speech Dialogue System. In Proc. of 2002 IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS'02), pp. 1314-1319, 2002.
- [11] 松本吉央, 後藤健志, 谷裕章, 怡土順一, 紙弘和, 伊野智行, 小笠原司. 搭乗型ガイドロボット watson の開発. 第 8 回ロボティクスシンポジウム予稿集, 2003 (to appear).