

リアリティの本来の意味からすればコンピュータ内の世界も何らかの形で現実を反映したものであり、もしコンピュータ内に現実世界を完全に反映した世界を構築・提示できるならば、それはどこかの場所へ訪れることと何ら変わらない体験であると言える。しかし、現在の技術でそのようなクオリティを持った VR 世界の構築(複雑な世界全体をカバーする物理法則の実時間シミュレーション)と提示(高品位コンピュータグラフィックス画像の実時間提示)は到底不可能であり、何らかの違和感を持たざるを得ない。若年層ならば既にゲームなどを通してそうしたコンピュータ内の世界に慣れ親しんでおり、人とのコミュニケーションを行う場合に仮想空間を介することにさほど抵抗を感じないかもしれないが、高齢者に対して同様の事を期待するわけにはいかない。

本研究では、このような視点に立脚し、あくまで高齢者が現実の世界で積極的に活動することを支援することを目的とする。すなわち、加齢に伴い長時間の移動が苦痛に感じられるような場合にも、外界の世界を体験してコミュニケーションを行うことを支援する。このために、遠隔環境をそのまま自然な三次元空間として提示するシステムの開発を行う。

B. 研究方法

本研究の対象は、トレイグジスタンス技術の中でも人間に遠隔環境を自然な三次元空間として提示する、視覚提示技術である。従来、トレイグジスタンス視覚系は遠隔環境に存在するロボット頭部にステレオカメラを設置し、ユーザは頭部搭載型ディスプレイ(HMD: Head-Mounted Display)を装着して、ユーザの頭部運動に合わせてロボット頭部を制御することにより実現されていた。しかしながら、このような HMD を利用したシステムには、以下のような問題が存在する。

装着重量:画像提示装置と光学系を頭部に載せる必要があるため、軽量化には限度が存在する。

身体の消失:HMDは目の直前で視覚提示情報をすり替えるため、ロボットの腕や足などを完全に実装し実時間で制御しない限りは自己投射性(自分の身体がそこにあるような感覚)が得られず、空間に浮遊しているような感覚に囚われがちである。

酔いの誘発:遠隔ステレオカメラと HMD の組合せによるシステムは、ロボット頭部が人間の動きに対して遅れることなく追従し、実時間で提示画像が更新されることが必須条件である。しかし、現実には途中の通信などで遅延が生じるため、頭を動かすと世界が揺れるという現象が発生する。例えば、ユーザが静止した状態から首を右に振り、右を向いて止まるという状況を考える。このとき、理想的には首を振り始めた瞬間に提示される画像が左へ流れ、振り終えた瞬間に画像も静止するということが求められる。しかし、遅延が存在すると、首を振り始めても画像が静止状態のまま更新されないため、世界が一瞬頭部について回り、画像が流れ始めると運動している間世界は静止して感じられるが、右を向いて静止した後世界が左へ回るように知覚される。すなわち、運動開始と静止の瞬間に世界が揺れるように感じられるのである。こうした現象は HMD 装着者の酔いを引き起こすため、そのままでは使用に耐えない。

近年デバイスとしての HMD の改良が進み、高解像度かつ軽量の製品が市販されているが、それらは単なるパーソナルテレビとしてしか利用されていない。その根本的な原因は、上記のようなシステム全体としての視覚提示の時間的なクオリティが重要であるにもかかわらず、これに対する配慮が不足していること生じているものと考えられる。

一方で、VR 世界をコンピュータグラフィックス画像により提示する狭義の VR では、近年大型スクリーンを利用した CAVE などの投影型没入ディスプレイ (Immersive Projection Technology: IPT) が盛んに建設され、そのクオリティの高さに関する評価が定着しつつある。この評価に関しては、ユーザを包み込む映像視野角の大きさばかりが強調されがちであるが、前述の HMD における問題点の裏返しであると考えられる。すなわち、映像を表示するスクリーンが頭部に対してではなく周囲の空間に対して固定されているため、環境を表示する場合はユーザが運動を行っても映像を大幅に変更する必要がなく、結果としてシステムの時間遅れに対する許容度が高いことにあると考えられる。

このような視覚提示装置の特徴に基づき、本研究では HMD ではなく固定スクリーンを利用した実写画像によるレイグジスタンスシステムの構築を行う。しかしながら、実写ビデオ画像を実時間で提示し、なおかつ人間の運動に対応して正確な三次元空間を提示するシステムには、従来 HMD のみが利用されており、固定スクリーンを使用したシステムは存在しなかった。その原因は、スクリーンと人間の視点との位置関係が一般には非対称四角錐 (通常のカメラは上下左右対称) であり、かつ人間の運動によって変化するため、左右それぞれの視点から見た物体上の点をスクリーン上の正しい位置に表示することが困難であることに由来する。本研究では、この課題を解決することにより、固定スクリーンを利用した安定な視覚提示システムを構築して、人間にとって負担の少ない遠隔環境提示を実現することを目指す。

[倫理面への配慮]

本年度は被験者による実験は行わなかった。

C. 研究結果

前述のように、解決すべき課題は視点とスク

リーン面とで構成される非対称四角錐を実現すること、および人間の運動に対してその形状を実時間制御することである。前者に関しては、ある程度広角の画像を撮影しておき、スクリーン上にはその一部を切り捨てて表示することで実現できる。後者に関しては、画像の表示位置と大きさを実時間制御することで解決する。

構築したシステムを図 1 に示す。システムは、撮像部、画像処理部、および提示部から構成される。撮像部は、ステレオカメラ全体の位置を制御する平行移動機構、頭部回転運動に対応して左右の視点位置関係を制御しつつカメラの向きを一定に保つ姿勢保持機構から成る。姿勢保持は、観察者の頭部が回転してもスクリーン方向が常に一定であるために導入された仕組みである。一般に、空間中で人間の頭部など物体の運動を記述するには 6 自由度が必要であ



(a) 撮像部



(b) 映像提示部

図 1: 構築システム

るが、本方式ではカメラの姿勢を常にスクリーンに対して一定に保つため、カメラの回転 3 自由度は存在せず、カメラ 1 つあたりの制御自由度は 3 である。さらに左右のカメラの位置関係を指定するために 3 自由度必要であるが、左右の視点間距離（眼間距離）が一定という制約条件を設けると、結局 5 自由度制御できることが必要十分条件となる。これを、頭部全体の移動に対応する並進 3 自由度と、頭部回転による左右視点の位置制御の 2 自由度に分けてシステムを構成する。本年度は、椅子に着座した状態での利用を想定し、並進 3 自由度のうち上下方向の 1 自由度を省略して、並進 2 自由度とした。頭部回転運動に起因する視点位置制御は、左右振り向き動作(yaw)と首を傾げる動作(roll)に対応し、なおかつカメラの姿勢を一定に保つ機構を利用した。傾き動作(pitch)は、左右の視点が同時に上下方向に動くので、並進自由度で実現できる。ただし、本年度は前述の通り上下方向の並進運動を省略している。

画像処理部の役割は、ビデオ画像の位置とサイズを実時間で制御して出力することであり、本研究ではパーソナルコンピュータに装着したビデオ入出力インタフェース付きのグラフィックスボード上でハードウェア処理することにより実現した。本システムに限らず VR やトレイグジスタンスではシステム中の遅延を排除することが重要であるが、ここではグラフィックスボード上で処理を行うことにより、画像データの転送や CPU でのピクセル単位での処理などによる遅延要素をなくし、高速処理を実現した。

画像提示部で、単一のスクリーンに 2 台のプロジェクタを用いて左右用の画像を投影している。その際、人間の装着重量を軽減するため、円偏光フィルタによる左右画像の分離を行っている。

以上の構成により、HMD を使用したシステ

ムと比較して極めて安定な実空間の立体提示が可能となった。

D. 考察

遠隔実環境を立体提示する手法は、本研究のような実写カメラ画像をそのまま利用する方法の他に、遠隔世界を一旦三次元的に復元してコンピュータグラフィックスにより提示するというアプローチも考えられる。後者は近年活発に研究が進められており、Image-Based Modeling / Rendering (IBR) と呼ばれる。IBR を利用すると自由な視点からの画像を作成できるため、視覚提示装置の形態の対応範囲が広がるという利点があるが、反面処理が重くなり、余分な遅延を招く可能性もある。また、現在の技術では高精細な 3 次元形状復元を実時間で行うことは困難であり、復元過程でエラーが混入した場合、提示される世界に違和感を持つことは避けられない。その点、カメラで撮像したビデオ画像を一度も 3 次元復元することなく 2 次元画像のまま左右の目に提示するというアプローチは、3 次元への復元を人間自身に任せるという方針であり、機械に勝手な解釈をされるよりも自分自身で判断したいという人間重視の考え方に基づいている。

本研究で試作したシステムは、固定スクリーンを利用して生の実写ビデオ画像により遠隔空間の正確な 3 次元提示を行うという意味では世界に例を見ないものである。従来の HMD を利用したシステムでは世界揺れの現象を避けることは難しいが、このシステムでは見た目に派手な効果はないものの、安定した世界の提示が可能となった。立体画像についても、テーマパークで見ると誇張された立体画像は長時間体験すると疲労を覚えるが、我々が普段世界が立体的であると特別意識していないように、自然な立体画像は一見して立体とわからない程度のものであり、本システムで提示する

画像もそうになっている。

E. 結論

固定スクリーンを利用することにより、遠隔実環境の安定した 3 次元提示が可能であることがシステムの試作により実証された。本手法では従来の HMD を使用したシステムと異なり世界揺れの現象を生じることがなく、安心して利用することができるようになると期待される。

F. 研究発表

1. 論文発表

鄧 惟中, 関口大陸, 川上直樹, 柳田康幸, 奴久妻 章, 舘 暉: RCML: アールキューブ操作言語の開発, 日本バーチャルリアリティ学会論文誌, Vol. 6, No. 2, pp. 881–890, 2000.

2. 学会発表

Y. Yanagida, T. Maeda and S. Tachi: "A Method of Constructing a Telexistence Visual System Using Fixed Screens," *Proceedings of IEEE Virtual Reality 2000*, pp. 117–125, 2000.

柳田康幸, 舘 暉: 生活支援へ向けたテレイグジスタンスの基礎技術, 第 16 回ライフサポート学会講演予稿集, p. 40, 2000.

G. 知的財産権の出願・登録状況

なし

マルチメディアを用いた高齢者支援システムの開発

分担する研究項目：ユーザインターフェースシステム，個人認証・セキュリティ技術

分担研究者 土居元紀 奈良先端科学技術大学院大学情報科学研究科助手

研究要旨：本研究ではマルチメディア技術を高齢者支援に応用する目的で，室内における人物の抽出および追跡について研究を行った。分散協調視覚の概念に基づき，複数のカメラを利用して人物を検出し，アクティブカメラにより頭部を追跡して顔の拡大画像の撮影を試みた。この結果，アクティブカメラによる自動撮影に成功した。同時に，撮影画像を個人認証に用いるための課題が明らかになった。

A. 研究目的

マルチメディア技術により高齢者の家庭生活を支援し，また，発作などの危険を早急に検知して重大な事故を未然に防ぐためには，24時間体制での自動監視が重要となる。また，同居人がいる場合は，個人の年齢や健康状態に応じてサービスの内容を適応させる必要がある。そのために，居住者を検出して追跡し，個人特徴を抽出して個人認証を行う必要がある。本報告では，監視画像から画像処理によって人物を検出し，頭部を自動追跡して拡大顔画像を撮影するシステムについて報告を行う。

B. 研究の方法

提案する監視カメラシステムは，入室者の個人の特徴を確実に得られる場所である，部屋の入り口に設置することにした。部屋の入り口を通過する人を検出し，その人の全身像と顔にズームした画像を撮影する。また，個人特徴として，顔画像だけでなく

身長および服の色を測定する。撮影された画像は画像ファイルとして記録されるが，このとき，個人特徴情報と撮影時刻が画像のヘッダに記録される。

システムは十分に早く応答し，さまざまな変動にも強いことが要求される。そのためアルゴリズムは計算コストが少ないことが求められる。また，分散協調視覚の概念を導入し，複数のカメラから得られる画像に対する画像処理プロセスを効果的に統合して，確実な顔画像撮影を行う。

以下に具体的な手法について述べる。

1. 顔の検出と個人特徴の抽出

本システムでは，顔にズームした画像を撮影するため，入り口を通過した人物の顔の位置を検出しなければならない。本システムでは，フレーム間差分により動物体を検出し，人物の頭部らしい部分を抽出することにした顔の検出は，二台の固定カメラを用いて行う。一台は人物の側方を撮影し，もう一台は人物の正面を撮影する。二台の

カメラで撮影されたそれぞれの画像における頭部の位置から、頭部の三次元的な位置が決定され、その位置を元に別に設置されたアクティブカメラが顔のズーム画像を撮影する。このアクティブカメラは、パン、チルト、ズーム等がコンピュータで制御できる。顔の位置検出および個人特徴の抽出処理の流れを以下に示す。

- ・ 入り口を通過する人物の検出 (図 1)
- ・ 人物の身長測定 (図 1)
- ・ 人物の顔位置の推定 (図 2)
- ・ 顔画像の撮影 (図 3)
- ・ 服装の色情報の取得 (図 2)

各処理には計算時間の少ない方法を用いており、実時間での実行が可能である。以下、各処理を詳しく説明する。

2. 入り口を通過する人物の検出

人物の側方を撮影するカメラは、ゲートセンサの役割を果たす。このカメラからの映像は、常に連続フレーム間差分をとり、二値化される。本稿では差分が生じた画素を黒画素とする。画像内の一定領域の黒画素数が閾値を超えた場合、カメラの前を人物が通過したと判定する。連続フレーム間差分を用いたのは、照明変動の影響をなるべく小さくするためである。

3. 人物の身長測定

身長推定も同様に、側方画像の差分画像において行われる。画像の横一列の画素における黒画素の合計数を、画像の下から上へと一列ずつ調べていき、合計数が閾値以下になった列の高さから暫定的な身長を計算する。実空間での身長の値は、顔位置推定処理において最終的に補正される。

4. 人物の顔位置の推定

顔の位置の推定は、正面画像をフレーム

間差分し二値化した画像において行われる。まず、側方画像における身長の高さは、正面画像における高さに変換される。ほぼ顔の大きさの長方形窓を、この高さで横方向に走査し、窓内の黒画素数が閾値を超えた時、長方形窓の位置を顔の位置とする。正面画像における顔の位置と側方画像における顔の位置から、実空間における顔の三次元的な位置を計算し、これに従い身長も補正される。

5. 顔画像の撮影

獲得された顔の三次元位置情報は、アクティブカメラのパン、チルト、ズームのパラメータに変換される。このパラメータによりアクティブカメラを制御し、顔のズーム画像を撮影する。そして、正面カメラで得られた全身像とともに記録する。

6. 服装の色の獲得

服装の色は正面画像から得られる。顔領域の下の領域から数点、色の値をRGB表色系で獲得し、各RGBの値それぞれについて平均する。これを服装の色の値とする。

[倫理面への配慮]

本年度は被験者による実験は行わなかった。

C. 研究結果

提案手法を試作システムに実装した。試作システムを図4に示す。本システムは、コンピュータ、二台の小型固定カメラ、アクティブカメラ、すべてのカメラからの画像をワークステーションに取り込むための四画面合成器より構成される。本システムは研究室の入り口に設置された。カメラは仕切り壁上に設置され、その視点の高さは165cmとなっている。

このシステムは顔画像を獲得し画像ファ

イルとして記録する。画像例を図5に示す。この画像ファイルはヘッダに身長や服装の色、画像を獲得した時間等の情報が書き込まれている。本システムが人物を検出してから顔画像を取得するまで約0.5秒である。

試作システムについて、10人の被験者により評価実験を行ったところ、普通の歩く速さで入り口を通過した場合は、確実に拡大顔画像を取得できた。

D. 考察

被験者が早歩きでカメラに対しななめに横切るように歩いた場合は失敗が多かった。しかし、高齢者を対象にした場合、歩く速度はあまり早くないことが想定されるので、問題はないと思われる。また、服装の色が壁の色と近かったために、適切な差分画像が得られず、顔の検出に失敗した例もあった。

そして、身長測定の精度を評価するため、各被験者における身長の平均と標準偏差を計算した。ただし、顔検出に失敗した例のある被験者は除いた。標準偏差から測定誤差は平均約1.5cmだった。誤差の原因として歩行時の身体の上下動が考えられる。しかし、人物追跡やの手がかりとしては、許容される誤差であると思われる。

E. 結論

提案したシステムは、部屋の入り口において、入室者を検知し、顔の三次元的位置を取得し、その位置情報を元にアクティブカメラが顔にズームして顔画像を取得する。また、同時に身長と服装の色を測定する。入室者の全身像と顔画像は画像ファイルとして記録され、そのヘッダには身長等の身

体特徴情報が書き込まれている。このシステムの一連の処理は0.5秒程度で行われる。

実験の結果、通常の歩行においては安定した顔画像の獲得ができ、身長も1.5cmの誤差で測定が可能であった。

取得された顔画像についてみると、その顔の向き、表情等は多様で、個人認証に用いるにはより条件が絞り込まれた方が良いと思われる。そのために、カメラの方向に入室者が注意を向け、なるべく正面の顔画像を取得できる工夫が必要であると思われる。そのために、システムが擬人化され、システムからあいさつなど声をかけて注意を喚起する、コミュニケーション型のセキュリティシステムの実現を検討している。

また、入り口だけでなく、室内において全体を見渡して人物を抽出できるように、全方位視覚センサの導入を検討している。全方位視覚センサはカメラを鉛直上向きに置き、レンズの正面に双曲面ミラーを設置することによって、360度周囲の画像を一度に取得できるカメラである。外見を図6に、取得画像を図7に示す。現在リアルタイムで人物を追跡するシステムは完成しているが、多人数に対応できないなど、課題は多い。来年度は、全方位視覚センサを用いた人物監視システムについて研究を行う。

F. 研究発表

1. 土居 元紀,他:“セキュリティのための身体特徴抽出”,第50回パターン計測部会研究会資料,pp.1-6,2000.
2. 土居 元紀,他:“インテリジェントセキュリティカメラ”,映像情報メディア学会技術報告,Vol.25,No.22,pp.1-6,2000.

G. 知的所有権の取得状況

該当無し。

図1: 人物検出と身長推定

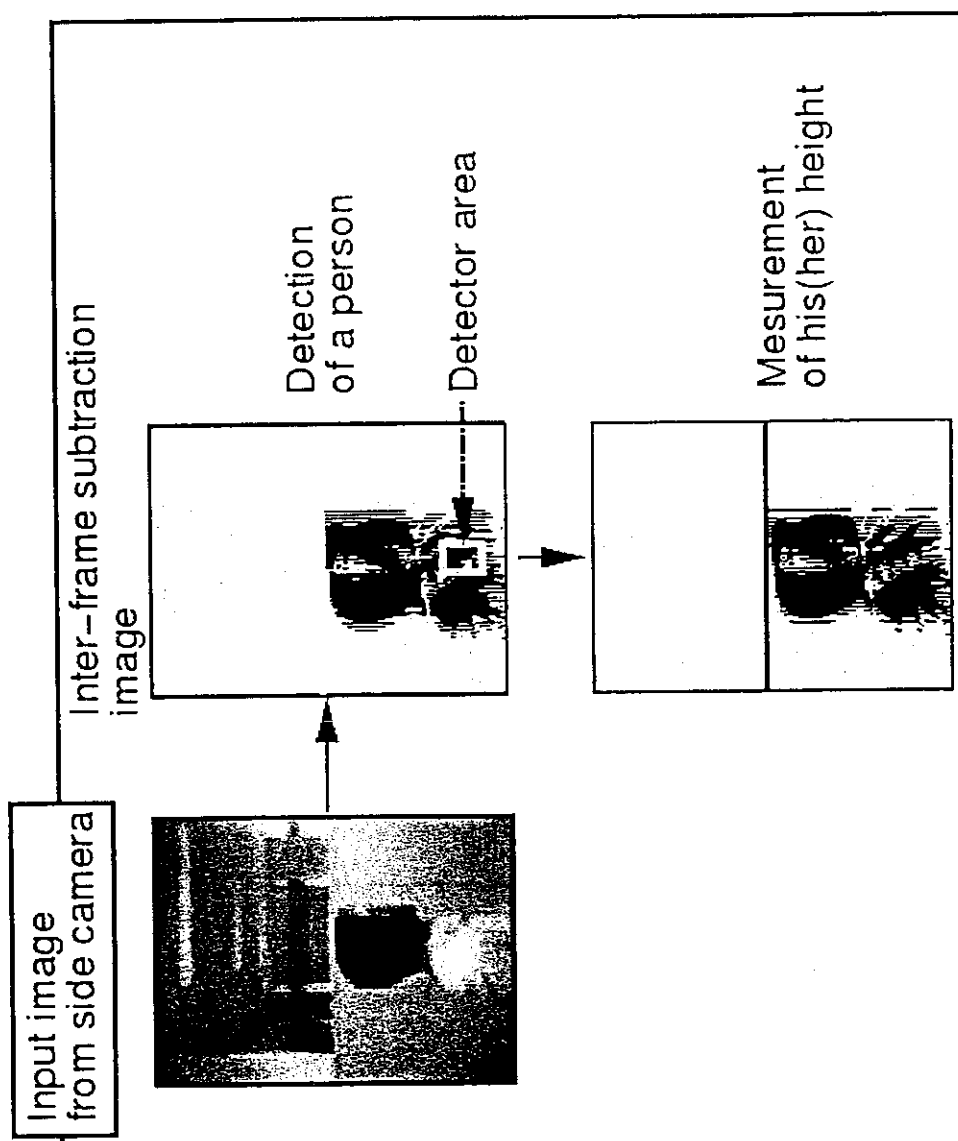


図2: 顔位置推定と服装の色の

獲得

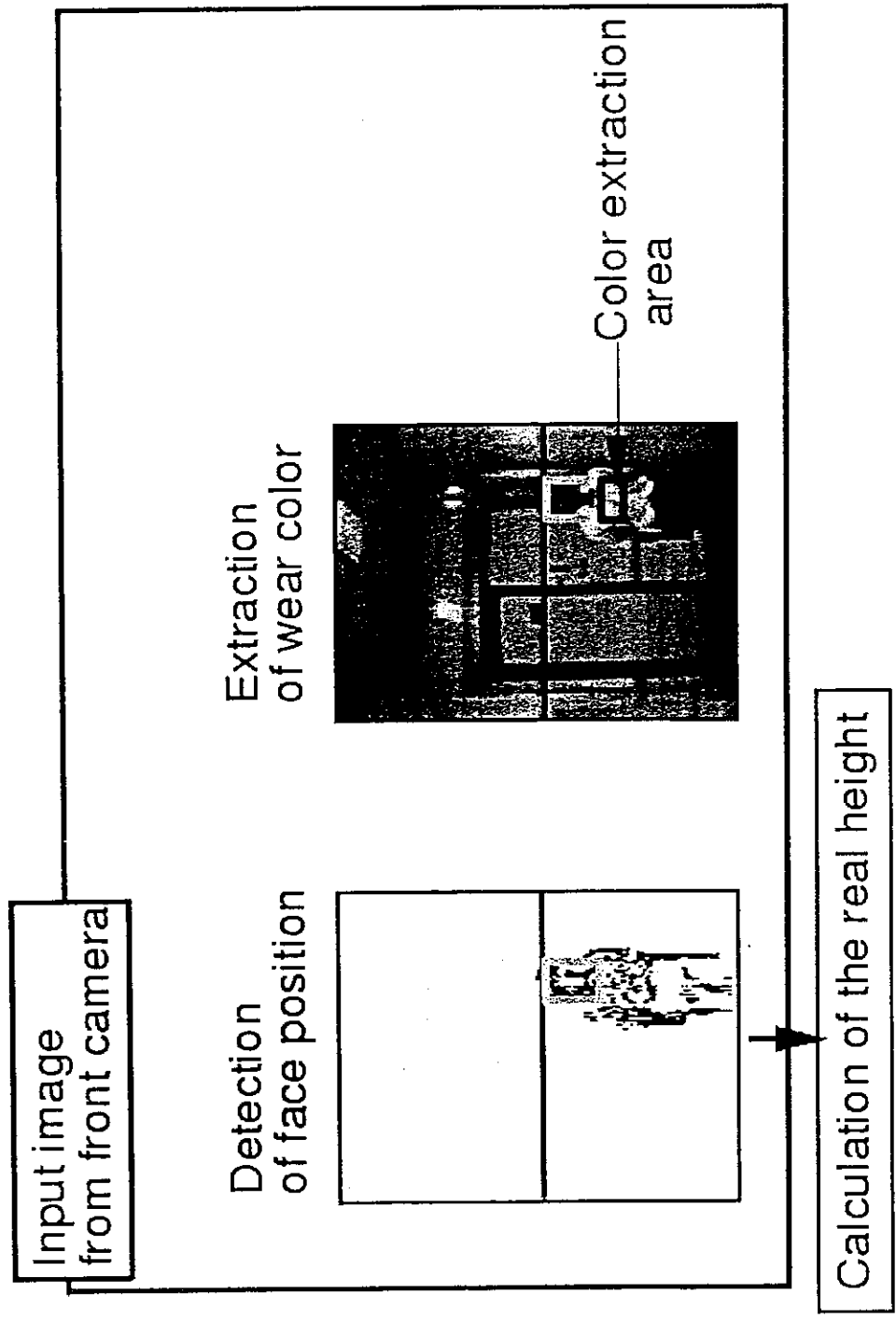
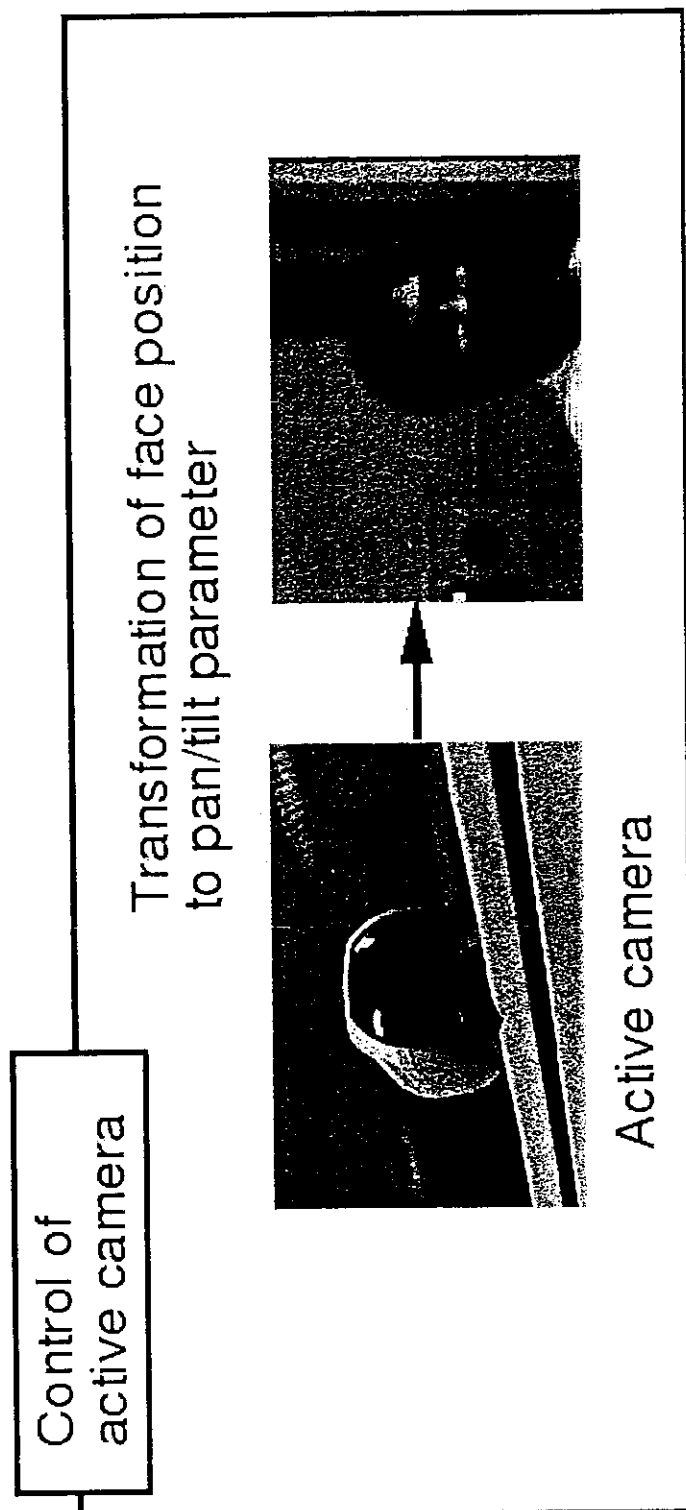


図3: 顔画像の撮影



人物検出から顔画像取得まで0.5秒

図4: 実装システム

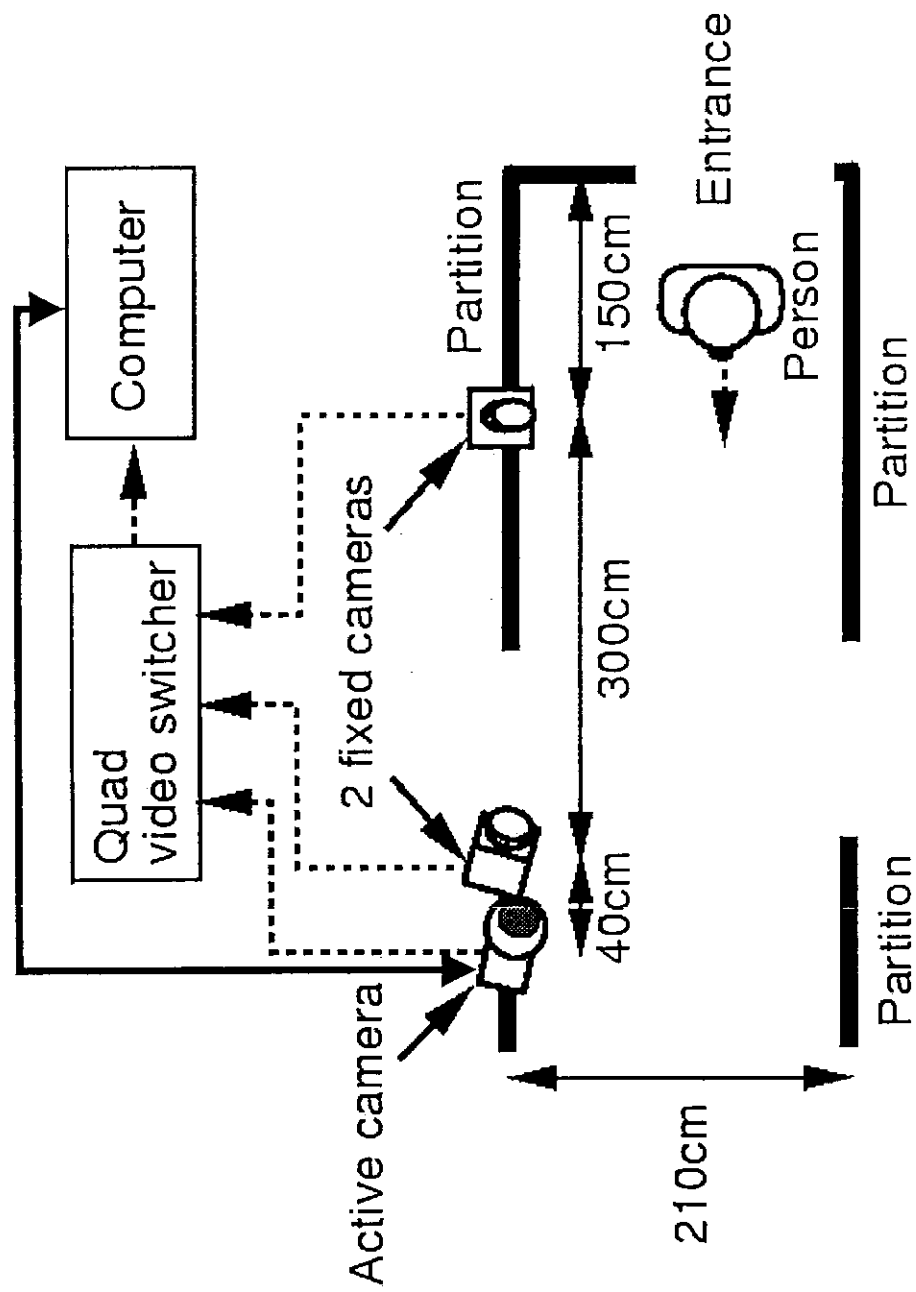


図5:出力画像



ヘッドに画像を獲得した日時と身長・服装の色が
書き込まれる

図6:HOVI外観

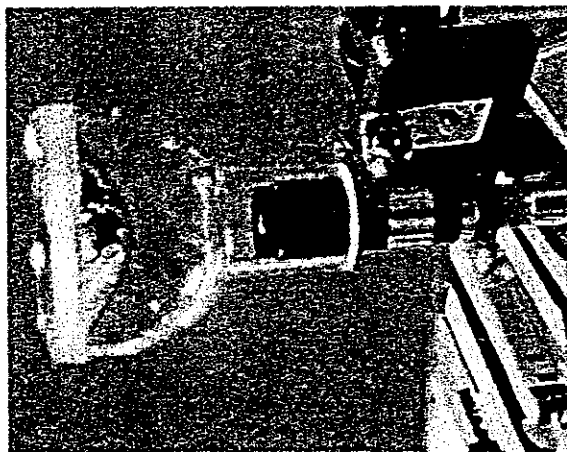
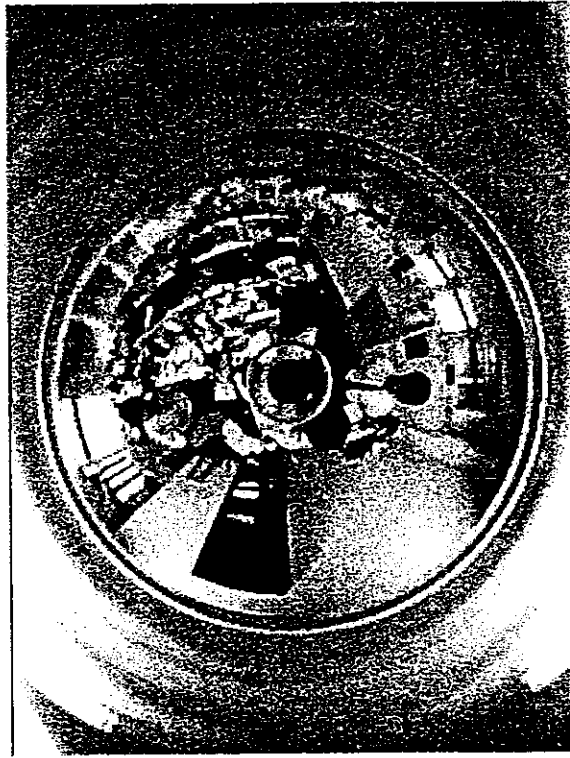
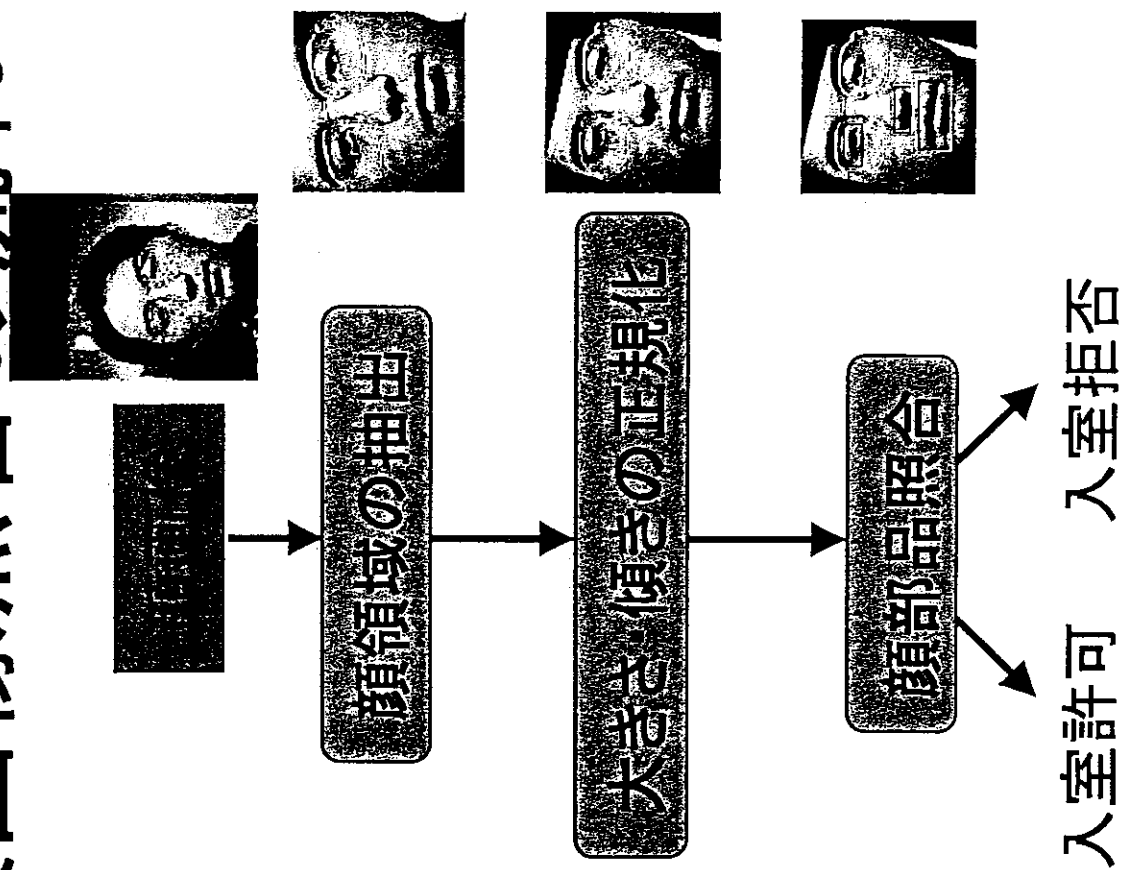


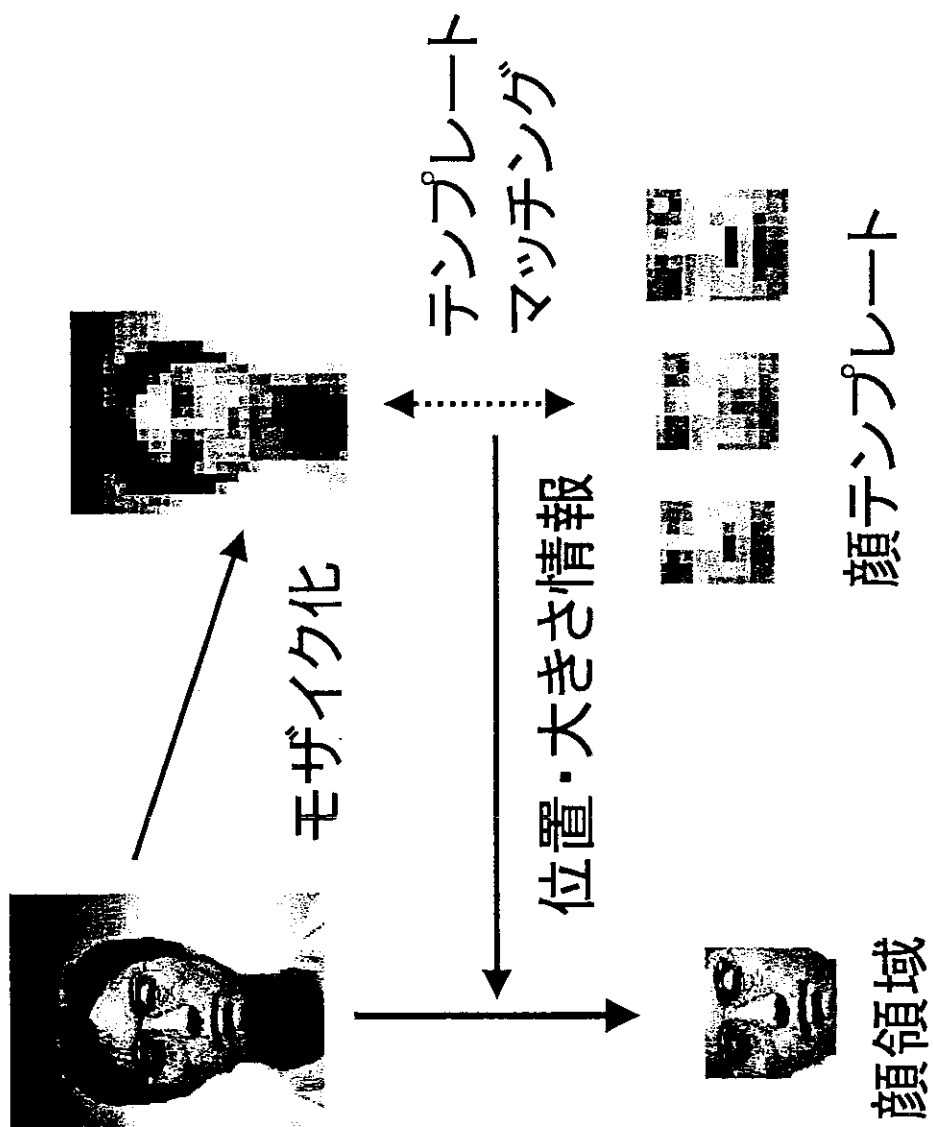
图7:HOVI取得画像



顔画像照合の流れ



顔領域抽出



目頭検出による 大きさおよび傾きの正規化

アフィン変換 → 厳密な基準点が求められる

目頭を基準としてアフィン変換

- 濃淡値の分布がはつきりしている
- 大きさ傾き変動に対し変化が小さい



□ 目頭領域

拡大

回転



階層的テンプレートマッチングで精度よく抽出できる

正規化処理

