

全方位ヘッドマウントカメラによる指示対象物体の発見

間 下 以 大† 岩 井 儀 雄† 谷 内 田 正 彦†

人と人、或いは、人とコンピュータのコミュニケーションを広げるシステムとして、ウェアラブルカメラシステムの研究が注目を浴びている。本報告では、ウェアラブルカメラとして全方位ヘッドマウントカメラを用いた実時間指示動作認識システムと指示動作からの指示対象発見手法を提案する。実時間指示動作認識システムの特徴として、全方位ヘッドマウントカメラによるオクルージョン問題の解決、ランダムサンプリングと重点サンプリングによる高速化があげられる。指示対象発見では、基準点および基準面という概念と、線形モデルを用いて指示方向の決定手法について述べる。実験および考察では、指示動作認識システムの実装結果、指示対象発見に関する評価実験とその結果について述べる。

Indicated Object Detection from Head-Mounted Omnidirectional Camera

TOMOHIRO MASHITA,† YOSHIO IWAI† and MASAHIKO YACHIDA†

We propose a system for pointing gesture recognition and estimating pointing direction by using a vision sensor. By using random sampling and importance sampling, our method can track hands and estimate hand positions in real-time. By using the concepts of a cognitive origin and a reference plane, our system can also detect a direction to an indicated object. We use an omnidirectional vision sensor in order to cover the wide range of hand operations and movement of indicated objects. The camera is mounted on the head, which enables the system to be tolerant of the occlusion problem. The method for detecting an indicated object uses a linear model with the concepts of a cognitive origin and a reference plane.

1. はじめに

人間は、他者とコミュニケーションを行なうことによって他人の発想を自分の発想に取り入れ、そこから新たな発想を生み出すことができる。1人でいくら考えても解けない問題が、他者と相談したり、情報を交換することによって簡単に解けるといった経験は誰もが持っていることだろう。このことは、複数の人間が共通の場をもって刺激を与えあうことによって、新たな発想を生み出したり、創造性が活性化するという効果があることを示している。他者とのコミュニケーションの際に重要となるのが、どのような手段でどのような情報を取り出すことができるかということである。また、コミュニケーションを行なう共通の場も重要となる。

一般的なコミュニケーションの手段は、聞く、見る、話すといったものであり、これらを行なう場合は直接会って行なわれるか、遠く離れている場合には文字や音声のみといった限られた情報しか伝えることができない。

しかし、これらの手段以外の情報を他者に伝える、あるいは受け取ることができるようなシステムがあれば、さらに大きな効果が期待できるものとおもわれる。このような新たな発想を刺激する情報を取得、提示する技術として、視覚情報提示技術、身体情報センシング技術、身体情報認識技術があげられる。

このような技術で重要な要素となるのが、直感的インタフェースとリアルタイム性である。例えば、あるユーザの視野を他のユーザと共有することで危険予知トレーニングなどを行う発想支援システムにおいて、危険を予知したユーザが予知した危険を他のユーザに知らせる場合、危険を知らせる手段が直感的に理解しがたいものであったり、知らせた情報が既に遅れていて役に立たないものであっては意味がない。そこで、本報告では有効性が確認されている指示動作^{13),17)}を実時間認識する手法を提案する。また、自然な人の動きを妨げないために、認識システムには視覚センサを用いる。そして、その視覚センサにはユーザの視野と指示動作を行う腕を同時に撮影できる広視野角のヘッドマウントカメラが適している。

今回提案するシステムでは、単眼視による手法の問題点であったオクルージョン問題を、ヘッドマウントカメラを用いて天頂方向から対象人物の腕を撮影する

† 大阪大学大学院基礎工学研究科
Graduate School of Engineering Science, Osaka University

ことで緩和する。このとき、人物の腕の長さ、計算の基点となる肩位置を既知とすることで姿勢推定を行う。また、乱数を用いたランダムサンプリングと重点サンプリングによって、追跡対象の発見と追跡における処理を高速化し、リアルタイムでの指示動作認識を行う。そして、ヘッドマウントカメラに全方位視覚センサ HyperOmni Vision¹¹⁾を用いることにより、広範囲に及ぶ指示範囲に対応し、かつ指示対象を同時に撮影することが可能となる。また、全方位ヘッドマウントカメラの視野を他者に提示することで、視野の共有も可能となる。さらに、指示対象を発見する前処理として、指示動作の推定結果から指示方向を推定する手法を提案する。

1.1 関連研究

1.1.1 ビジュアルトラッキング

これまでに視覚センサを用いた追跡手法は様々なものが提案されている。特に、人物などの肌色領域を追跡する場合には、肌色のモデル化が重要である。Yangら¹⁰⁾によって、肌色の分布に関する研究が為されている。それによると、3つの結論が得られている。第一に、人の肌色の分布は色空間において、特定の分布を示すという事、すなわち、その分布は特定の民族や個人には関係なく、これら特定の人物や民族といったものは肌色分布の一部を占めるに過ぎないという事、第二に、人物間の肌色の違いは、輝度の正規化によって軽減する事が可能であるという事、第三に、一定の照明条件下で、肌色の分布は、正規化色空間における多変量正規分布によって表現できるという事である。このことから、人物の手や顔のトラッキングに関する研究では追跡対象の色モデルを用いた手法が多く提案されている^{6),9)}。また、色モデルだけでなくエッジ検出等を行い楕円等の追跡対象の形状を組み合わせて用いる手法^{1),9)}もある。

逐次モンテカルロ法による手法^{4),9)}では、追跡対象に対する複数の仮説をもちながら追跡を行い、非線形運動についても適応できるといった利点がある。このような手法では、factor sampling と importance sampling により、画像全体や、ROI による予測領域に対して走査するよりも処理速度が速くなるという利点もある。

本システムでは、上述の逐次モンテカルロ法による手法を利用して、ランダムサンプリングと重点サンプリングという乱数を用いた手法で追跡を行なう。

1.1.2 姿勢推定

画像から人物の姿勢や動作を推定する手法はさまざまなものが提案されている。大きく分けると、単眼視による手法³⁾と複数のカメラによる手法¹⁶⁾とが挙げられる。

単眼視による手法は、システムの構成が単純になるという利点がある。しかし、オクルージョン問題がおこりやすい、奥行き情報が取得できないなどの欠点がある。

そのため、3次元の姿勢復元をおこなうには、既知パラメータを与えるなどしなくてはならない、リンクの長さなどを具体的に与えずに、他の拘束条件と評価関数を用いて、関節物体の姿勢を推定する手法⁵⁾もあるが、処理に時間がかかるという問題点がある。

複数のカメラによる手法では、複数視野により単眼視にくらべて視野が広く、多方向から観測するので、オクルージョンに対して強いという利点がある。また、視野は減少してしまうが、各カメラによる画像中の対応点をとることで、奥行き情報を取得することができ、3次元の姿勢復元が可能であるという利点がある。しかし、複数のカメラを用いるため、システムが複雑になる、複数の画像間での対応をとるための処理時間が長くなる、などの欠点がある。

本システムでは、実時間処理を目的とするため、高速処理が可能な単眼視によって姿勢推定を行なう。このとき、人物の腕の長さ、肩の位置を既知とすることで腕の姿勢推定を行う。

1.1.3 指示動作と指示対象の関係

人間の指示動作に関しては、神経科学等の分野で古くから研究されてきた^{2),7),8)}。また、近年ではヒューマンインタフェースなどの分野でも、指示動作は有効なインタフェースの一つとして研究されている^{12),14),15)}。

本システムでも直感的インタフェースの一つとして指示動作を考えている。これは、もしユーザの指示動作認識だけでなく指示対象をも発見することができれば、さらにコミュニケーションの手段が広がると考えられるからである。現在、指示動作と指示対象の関係については、手の届くような距離 (reaching distance) ではある程度明らかになってきているが、少し離れた距離 (walking distance) では明確な基準が得られていない。そこで、人物の指示動作から指示対象の方向を決定する手法を提案し、評価する。

2. モデル

2.1 カメラモデル-全方位視覚センサ

全方位視覚センサ HyperOmniVision¹¹⁾は鉛直下向きの双曲面鏡と上向きのカメラから構成される。双曲面鏡は式 1, 2 のように表され、二つの焦点 $(0, 0, c)$, $(0, 0, -c)$ を持ち、カメラは双曲面の中心軸と同軸に主点が双曲面鏡のその一方の焦点 $(0, 0, -c)$ にくるよう配置されている。このように配置することで、センサの周囲 360° を一度に撮像することができる (図 2)。また、図 1 に示すように双曲面鏡のミラー焦点に向かって入ってくる光は、双曲面の特性によりミラーに反射後、もう一方の焦点に向かう。すなわち、画像上での任意の写像点 $p(x, y)$ と俯角 α_C 、方位角 θ_C の関係は、式 3-6 から一意に決まる。これらの式から、全方位入力画像 (図 2) をミラー焦点を視点とした透視変換画像やパノラマ画像に容易に変換できるという特性

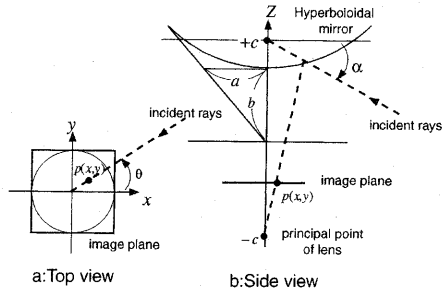


図1 HyperOmniVisionの光学系

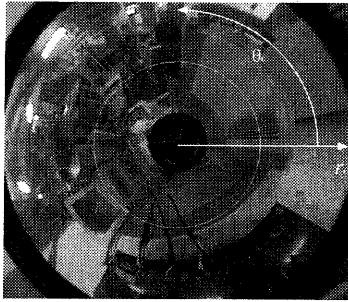


図2 入力画像

を持つ。

$$\frac{X^2 - Y^2}{a^2} - \frac{Z^2}{b^2} = -1 (Z > 0) \quad (1)$$

$$c = \sqrt{a^2 + b^2} \quad (2)$$

$$\tan \theta_C = \frac{Y}{X} = \frac{y}{x} \quad (3)$$

$$Z = \sqrt{X^2 + Y^2} \tan \alpha_C + c \quad (4)$$

$$\alpha_C = \tan^{-1} \frac{(b^2 + c^2) \sin \gamma - 2bc}{(b^2 + c^2) \cos \gamma} \quad (5)$$

$$\gamma = \tan^{-1} \frac{f}{\sqrt{x^2 + y^2}} \quad (6)$$

2.2 色モデル

本システムでは、カメラによって人物の腕の関節位置を観測することで腕の姿勢推定を行っている。本節では、腕の関節位置を観測するために用いる色モデルについて説明する。

1.1.1 節で述べたように、人物の肌色は色空間における正規分布(式7)で表現することが可能とされている。そこで、本システムでも手領域の追跡のための色評価に正規分布モデルを使用する。対象の色モデルを $C \triangleq (\mu, \Sigma)$ とすると、位置 x に対象のある確率密

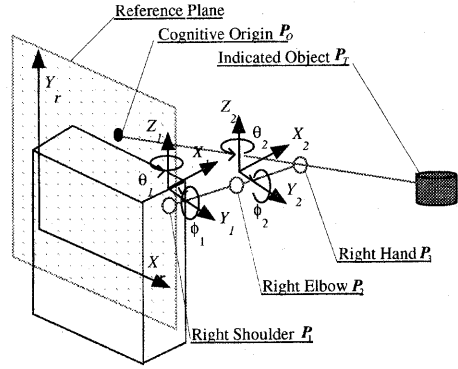


図3 関節物体モデル

度は、

$$p(x, C) = p(x; \mu, \Sigma) \triangleq \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (I(x) - \mu)^T \Sigma^{-1} (I(x) - \mu) \right] \quad (7)$$

となる。式7で、 μ は追跡対象の色空間中の平均ベクトルであり、 Σ は共分散行列、 $I(x)$ は入力となる画素の色である。また、3次元の色空間を用いているので $d = 3$ となる。

2.3 人体モデル

人体モデルは肩、肘、手を関節位置とした関節物体と仮定する。その自由度は、図3のように肩と肘の回転に各2自由度の計4自由度とする。

2.4 指示方向モデル

1.1.3 節で述べたように、人間の指示動作については古くから研究されているものの、明確な結論が得られていない。そこで、実空間中のある点と、指示動作を行っている手先位置 P_3 を結んだ直線上に指示対象物体 P_I があるとして、その点の位置が腕の姿勢によって変化するとした。今後この実空間中の点を基準点 P_O (図3 Cognitive Origin) と呼ぶ。

基準点 P_O の位置は人物の体を含む面上(図3 Reference Plane)を移動するものと仮定する。指示動作を腕をまっすぐに伸ばした状態、つまり $\theta_2 = 0, \phi_2 = 0$ とすると、指示動作の姿勢を表すパラメータは、肩における回転角 θ_1, ϕ_1 で表現される。この θ_1, ϕ_1 によって基準面上における基準点 P_O の位置 (X_r, Y_r) が決定するものとする。そして、 θ_1, ϕ_1 と X_r, Y_r の関係を線形モデルで仮定する(式8)。

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \theta & \phi & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta & \phi & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \end{bmatrix} \quad (8)$$

3. 指示動作推定, 指示方向推定手法およびアルゴリズム

本章では, ヘッドマウントカメラからの入力画像におけるトラッキングアルゴリズムと, トラッキングによって得られた情報をもとに, 人物の指示動作を推定するアルゴリズムについて述べる.

3.1 トラッキング

腕の姿勢推定を行うには, 各関節 P_i の入力画像における位置 p_i を推定する必要がある. この p_i は, 画像中における各関節の領域の重心位置と仮定する. このとき, 画像全域の探索によって各関節の発見と重心の推定を行っている処理時間が非常に大きくなってしまい, 実時間で処理が不可能になる. 本手法ではランダムサンプリングと重点サンプリングによって N 個のサンプリング点 $p_{sj} (j=1, 2, \dots, N)$ を分布させ, そのサンプリング点 p_{sj} の色情報を評価する事で追跡対象の重心推定を行う. これにより, 全画素を走査するより少ない画素数で重心推定が可能となり実時間で追跡が可能となる.

3.1.1 サンプリング

サンプリング点の決定は図4の様に, 極座標系で決定する. 極座標系を用いる事で, 画像中の位置による解像度の違いを緩和することができる.

ランダムサンプリング点 $p_R(r_R, \theta_R)$ の決定は, ランダムサンプリング領域の大きさを R_R, Θ_R , とし, サンプリング範囲に対するバイアスを R_B, Θ_B とすると, 一様乱数 $U(0, 1)$ を用いて

$$\begin{aligned} r_R &= R_B + R_R U(0, 1) \\ (R_B + R_R > r_R > R_B) \end{aligned} \quad (9)$$

$$\theta_R = \Theta_B + \Theta_R U(0, 1) \quad (10)$$

となる.

重点サンプリング点 $p_I(r_I, \theta_I)$ の決定は, 重点サンプリング領域の大きさを R_I, Θ_I , 追跡対象の予測位置を $G(r_G, \theta_G)$ とすると,

$$\begin{aligned} r_I &= r_G - \frac{R_I}{2} + R_I U(0, 1) \\ (R_B + R_R > r_I > R_B) \end{aligned} \quad (11)$$

$$\theta_I = \theta_G - \frac{\Theta_I}{2} + \Theta_I U(0, 1) \quad (12)$$

となる. 予測位置 $G(r_G, \theta_G)$ の決定は, 1フレーム前の予測位置 p_{n-1} を予測位置とする静止予測モデルとした.

極座標系によるサンプリング点 (r, θ) が決定したら,

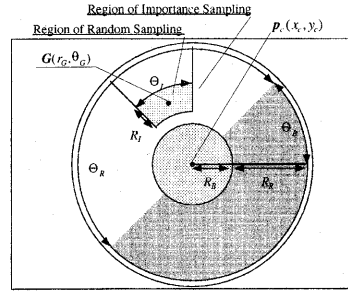


図4 サンプリング範囲

$$p_{sj} = \begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + r \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad (13)$$

として画像乗のサンプリング点 p_{sj} が決定する. ここで (x_c, y_c) は, 全方位画像の中心 p_c である (図4).

3.1.2 色評価

あるサンプリング点 p_{sj} に対する評価は, 2.2節で述べたように, 正規分布モデルを使用する. 実際の評価値はマハラノビス距離 r の2乗,

$$U_j = r^2 = (I(p_{sj} - \mu)^T \Sigma^{-1} (p_{sj} - \mu)) \quad (14)$$

によって行う. p_{sj} は, サンプリング点 p_{sj} の画素の色である.

この評価値 U_j をしきい値 K で2値化し (式15), 和をとることで面積とする (式16)

$$D_j = \begin{cases} 1 & U_j < K \\ 0 & U_j \geq K \end{cases} \quad (15)$$

$$S = \sum_{j=0}^N D_j \quad (16)$$

そして, 重心位置 p_i の推定は,

$$G = \frac{\sum_{j=0}^N D_j p_{sj}}{S} \quad (S \neq 0) \quad (17)$$

によって行われる.

3.2 腕の姿勢推定

このアルゴリズムでは, カメラ焦点から関節位置への単位方向ベクトル v_i をもとに, 肩座標 P_1 と各リンク (上腕, 下腕) の長さ L_i を既知として計算している. 全方位視覚センサの場合, 画像上における関節位置 (x_i, y_i) から v_i を求めるには, 式3の θ_C と, 式4の α_C によって, 以下の式で計算される.

$$v_i = \begin{bmatrix} \cos \theta_C \cos(-\alpha_C) \\ \sin \theta_C \cos(-\alpha_C) \\ \sin(-\alpha_C) \end{bmatrix} \quad (18)$$

図5に, カメラ焦点 (図5Focal Point) と各関節位置 (図5 P_i) 関係を示す. 各観測点 P_i への単位方向ベクトルを v_i , 奥行きを z_i とする. リンクで接続され

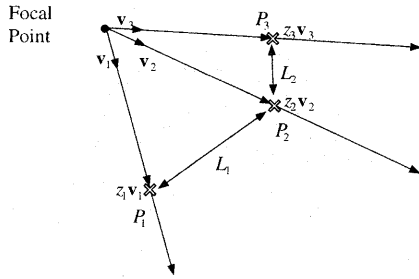


図5 単眼の場合の関係図

た2点 P_{i+1}, P_i 間の距離を L_i とする。ここで、 v_i, L_i は直接観測できるので既知であり、 z_i は未知であり求めるべきパラメータである。

以上のことから、次の制約条件が成り立つ。

$$\|z_{i+1}v_{i+1} - z_i v_i\| = L_i \quad (19)$$

または、観測誤差を考慮して

$$\|z_{i+1}v_{i+1} - z_i v_i\| = L_i + \epsilon_i \quad (20)$$

としても同様の議論が成り立つ。ここでは簡単のため $\epsilon_i = 0$ として話を進める。

式19を解析的に陽に解くと以下ようになる。

$$z_{i+1} = (v_{i+1}, v_i)z_i \pm \sqrt{[(v_{i+1}, v_i)^2 - 1]z_i^2 + L_i^2} \quad (21)$$

すなわち全てのリンクの姿勢が根元となる z_1 から相対的に決定できる(例えば $z_1 = 1$ として相対的に決定できる)。また、 z_1 の値が既知の場合には全ての姿勢が決定できる。

3.3 指示方向の推定

2.4 で述べた線形モデルのパラメータ $\alpha_1 \sim \alpha_6$ の学習について述べる。まず、 m 個の腕の姿勢 $\theta'_1, \dots, \theta'_m, \phi'_1, \dots, \phi'_m$ と、その指示対象の座標 $P_{T1}(X_{T1}, Y_{T1}, Z_{T1}), \dots, P_{Tm}(X_{Tm}, Y_{Tm}, Z_{Tm})$ があるとす。これを学習用データとする。それぞれの関係は式8から、

$$\underbrace{\begin{bmatrix} X_1 \\ Y_1 \\ \vdots \\ X_m \\ Y_m \end{bmatrix}}_B = \underbrace{\begin{bmatrix} \theta_1 & \phi_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_1 & \phi_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \theta_m & \phi_m & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_m & \phi_m & 1 \end{bmatrix}}_A \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \end{bmatrix}}_P \quad (22)$$

となり、線形モデルにおける誤差を最小にするパラメータ P は、

$$P = (A^T A)^{-1} A^T B \quad (23)$$

で求める事ができる。そして、指示動作の時に、 θ_1, ϕ_1 を求めれば、基準点 P_O が決定し、基準点と手先を通る直線上に指示対象 P_T が存在することになる。

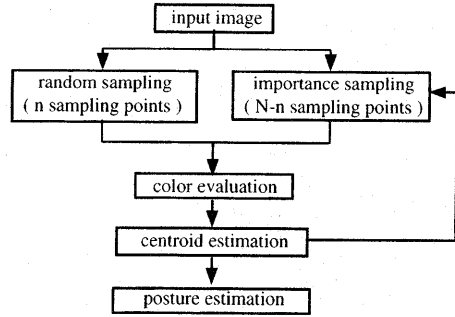


図6 システム概要

4. 実装システム

4.1 システム概要

システムの処理の流れを図6に示す。本システムでは、ランダムサンプリング点数 n と重点サンプリング $N-n$ の和 N を一定にすることで安定した速度で実行している。3.1で述べたように、重点サンプリングの予測は、前回に発見した追跡対象の重心位置 $p_{i,n-1}$ とし、サンプリング点の分布は一様分布とした。

4.2 システム構成

入力画像を撮影するカメラは、digital video camera VX2000(SONY)に全方位視覚センサミラーユニット、TOM-02-0013 HyperOmni Vision(末陰産業)を取り付けた。実験に使用した計算機は、ONYX2(R12000 300MHz)(SGI)であり、カメラ画像の取り込みから実験結果の出力までを行っている。画像取込時の画像サイズは 720×486 pixel だが、フィールド単位で処理しているため、一回の処理における画像サイズは 720×243 pixel となっている。実験では被験者の肘と手に異なる色マーカを装着して手の追跡を行った。追跡対象の色情報は、同じ環境で撮影した画像からマーカ部分のみを切り出して、マーカの色空間中における平均ベクトル μ 、共分散行列 Σ 、を計算したものをあらかじめ与えてある。また、人物の腕の長さ L_i も、あらかじめ与えてある。

5.2節で述べる指示動作解析と指示対象発見に関する実験では、Polhemus社製磁気トラッカーFAS-TRACKを用いて計測を行った。この磁気トラッカーは、76cm以内であれば0.08cmRMSの静的精度で動作する。この磁気トラッカーは最大4つのレシーバを同時に用いることができ、各レシーバの3次元空間中における座標と姿勢を同時に取得することができる。

5. 実験および考察

5.1 指示動作推定システムの実装結果

実際の姿勢復元を実時間で行った様子を図7に示す。処理速度はおおよそ15Hzであった。

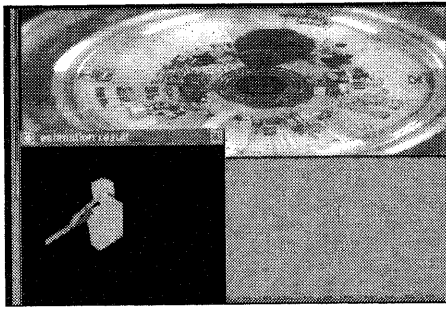


図8 実験の様子

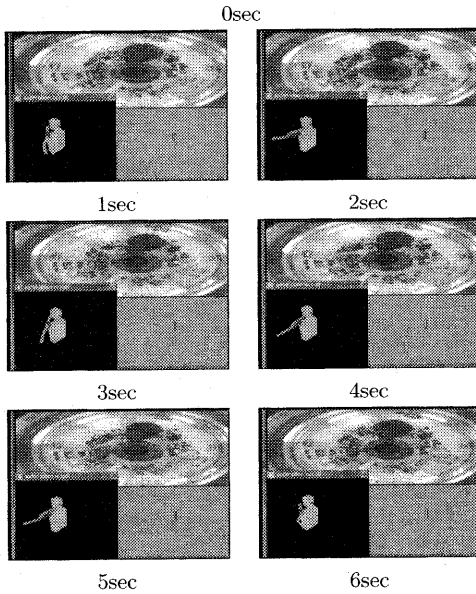


図7 実験結果

図7の各図における上段が入力画像であり、図7の左下が姿勢推定結果である。解析的に姿勢を推定するため上腕で2通り、下腕で上腕の各推定結果に対して2通りずつの解が得られている。この解の中から正しい姿勢を決定するのは本システムにおける指示動作認識の処理で行う予定である。

5.2 指示方向推定に関する評価実験

5.2.1 実験環境

指示対象発見には、まず2.4節、3.3節で述べた $\alpha_1 \sim \alpha_6$ のパラメータを学習するためのデータが必要になる。学習用データを取得した実験環境の様子図8に示す。被験者は図8のように、椅子に座った状態で指示してもらった。指示対象 P_T と指示対象 P_T と被験者の位置は図9のように設定し、指示対象 P_T は図9に加えて、高さについて4段階(床面より0cm,50cm,100cm,150cm)の4段階として、計80点を指示し、推定用データとした。さらに、推定用データとは異なる位置の指示対象45点を指示し、実

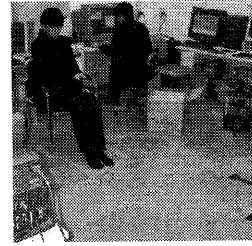


図8 実験の様子

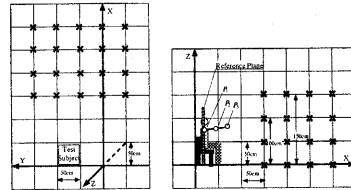


図9 指示対象と被験者の位置

験用データとした。本報告での評価実験はこの実験用データで行っている。姿勢の計測には、4.2節で述べた、Polhemus社製磁気トラッカーを用い、頭部、右肩、右肘、右手の4ヶ所にレーザを装着して姿勢を計測した。なお、今回の実験では個人差を考慮していないため、実験における学習パラメータ $\alpha_1 \sim \alpha_6$ は被験者毎に異なる。

5.2.2 基準面の妥当性

推定用データから得られた、指示直線の様子を図10、基準面における基準点の位置と腕の姿勢の関係を11,12に示す。図11,12で基準面の位置は、肩の平均位置を通る $X-Z$ 平面に平行な面を基準面とした。図10,11,12における各 X, Y, Z 軸の原点は、図9と一致している。

図10は、指示対象の高さ毎に指示直線を引き、 $X-Y$ 平面に写像したものである。この図10から、写像した直線が収束する点が基準面上にあり、この点が肩と頭の間を移動していることが分かる。

X_r と ϕ_1 の関係を図11に、 Y_r と θ の関係を図12に示す。図11から、指示動作時における上下の腕の動きと、基準面における水平方向の移動が線形な関係であることが読み取れる。また、図12から、指示動作時における左右の腕の動きと、基準面における垂直方向の移動が線形な関係であることが読み取れる。

このことから、2.4で述べた、線形モデルは有効であると思われる。

5.2.3 評価実験

提案手法、肩基準、頭部基準における指示方向推定精度の比較実験の結果を表1に示す。この実験では、図13に示すように、手から指示対象への方向を真値とし、基準点から手先への方向を推定値とした、誤差 e

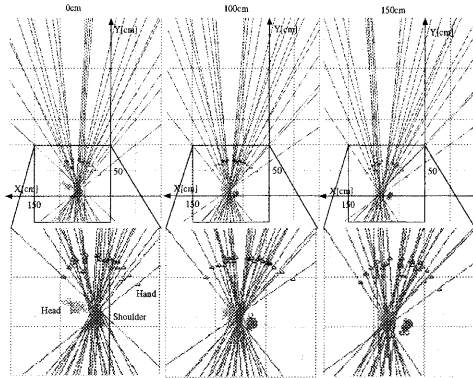


図 10 指示直線の様子

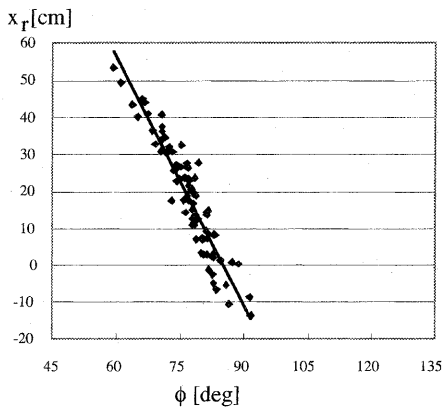


図 11 x_r と θ の関係

表 1 提案手法、肩基準、頭部基準の比較

絶対誤差平均 [deg]	提案手法	肩基準	頭部基準
被験者 1	3.23	13.36	15.67
被験者 2	2.51	14.92	22.52
被験者 3	2.76	18.05	20.28
被験者 4	4.16	15.57	16.59
被験者 5	3.85	13.25	18.85

は

$$e = \left| \cos^{-1} \frac{(P_T - P_3, P_3 - P_O)}{|P_T - P_3| |P_3 - P_O|} \right| \quad (24)$$

となる。基準点は提案手法によって推定された基準点の場合、肩を基準点とした場合、頭部を基準点とした場合としている。表 1 から、本手法が今まで用いられてきた肩、或いは頭部を基準点とした場合よりも精度よく方向推定が行えていることがわかる。

同様の環境で被験者に物体（段ボール箱）を指示してもらい、その様子を全方位視覚センサで撮影した。そして、その全方位画像に指示直線を射影した。その全方位画像を図 14 に示す。推定した指示直線上に指示対象は存在しないものの、誤差の範囲を考慮するこ

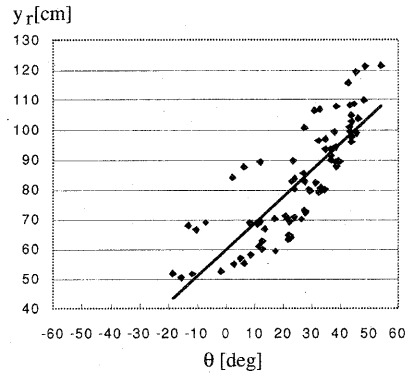


図 12 y_r と θ_1 の関係

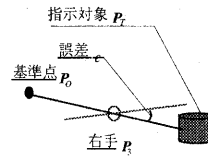


図 13 指示方向推定における誤差

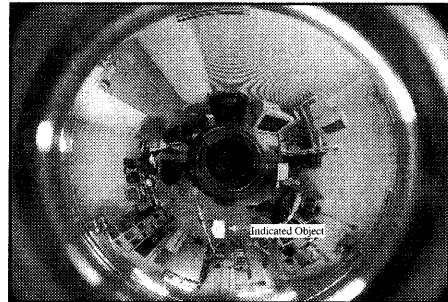


図 14 指示動作と指示直線

とで指示対象の発見が行えるものと思われる。

6. おわりに

本報告では、全方位ヘッドマウントカメラによる実時間指示動作推定システムと、指示動作からの指示対象発見手法を提案した。

提案システムでは、およそ 15Hz の処理速度で、人物の腕の姿勢を推定する事ができた。この処理速度はコミュニケーションなどに用いるには十分であると思われる。

指示動作と指示対象の関係に関しては、人物の指示動作から指示対象のある方向を決定する手法を提案した。この手法と、あらかじめ登録された指示対象となりうる物体の画像のマッチングを組み合わせることで指示対象物体の発見ができるものと思われる。

今後の課題は、指示方向推定を全方ヘッドマウントカメラシステムに実装し、指示対象発見まで行えるシステムの構築と、指示方向推定においては、視線あるいは頭部の向きや、人物の体全体の姿勢を考慮した指示方向推定手法の構築である。

謝 辞

本研究の一部は、日本学術振興会未来開拓学術研究推進事業 (JSPS-RFTF99P01404)、文部科学省科学研究費補助金 (重点領域研究・奨励研究) の補助を受けた。

参 考 文 献

- 1) S.T. Birchfield. Elliptical head tracking using intensity gradients and color histograms. In *CVPR98*, pp. 232-237, 1998.
- 2) W. Russell Brain. Visual disorientation with spatial reference to lesions of the right cerebral hemisphere. *Brain*, Vol. 64, pp. 244-272, 1941.
- 3) Chounhong Pan and Songde Ma. 3d motion estimation of human by genetic algorithm. In *ICPR00*, Vol. 1, pp. 159-163, 2000.
- 4) Michael Isard and Andrew Blake. Icondensation: Unifying low-level and high-level tracking in a atochastic framework. In *ECCV*, Vol. 1, pp. 893-908, 1998.
- 5) C. Pan and S. Ma. 3d motion estimation of human by genetic algorithm. *ICPR00*, Vol. 2, pp. 159-163, 2000.
- 6) L. Sigal, S. Sclaroff, and V. Athitsos. Estimation and prediction of evolving color distributions for skin segmentation under varying illumination. In *CVPR00*, pp. 152-159, 2000.
- 7) JohnF. Soechting and Maratha Flanders. Sensorimotor representations for pointing to targets in three-dimensional space. *Journal of Neurophysiology*, No. 62, pp. 582-594, 1989.
- 8) Michael E. Woodin and Alan Allport. Independent reference frames in human spatial memory: Body-centered and environment-centered coding in near and far space. *Memory & Cognition*, Vol. 26, pp. 1109-1116, 1998.
- 9) Y. Wu and T. S. Huang. A co-inference approach to robust visual tracking. In *ICCV01*, Vol. 2, pp. 26-33, 2001.
- 10) J. Yang, L. Weier, and A. Waibel. Skin-color modeling and adaptation. In *Asian Conference on Computer Vision (ACCV'98)*, pp. 687-694, 1998.
- 11) 山澤一誠, 八木康, 谷内田正彦. 移動ロボットのナビゲーションのための全方位視覚センサ hyperomni vision の提案. 信学会論文誌 D-II, Vol. j79-D-II, No. 5, pp. 698-707, 1996.
- 12) 福本雅朗, 間瀬健二, 末永康仁. Finger-pointer: 画像処理を用いた仮想指示棒. HC 91-12, 電子情報通信学会, 91.
- 13) 林健太郎, 久野義徳, 白井良明. ユーザ中心と世界固定の視点による空間指示動作の解釈とその応用. PRU 95-93, 電子情報通信学会, 1995.
- 14) 田中宏一, 和田俊和, 松山隆司. 3次元人体形状計測に基づく指差し動作の解析. *CVIM 133-17*, 情報処理学会, 2002.
- 15) 吉田千里, 乾敏郎. ポインティング動作に基づく空間認知特性の検討. HCS 2000-09, 電情報通信学会, 2000.
- 16) 中井博章, 山田和弘, 阿部亨. ステレオ動画像による他関節物体の3次元運動の推定. In *MIRU00*, 第1巻, pp. 155-163, 2000.
- 17) 宮里勉, 岸野文郎. 臨場感通信会議における仮想空間を介した指差し指示動作知覚の評価. 電子情報通信学会, Vol. j80-D-II, No. 5, pp. 1221-1230, 1997.