単眼カメラ画像を用いた指差しポインタの構成

Finger Pointing Device Using Single Camera Image

正会員 阪本 邦夫†,

木村 理恵子†

Kunio Sakamoto† and Rieko Kimura†

Abstract A finger pointing system is described that can specify the position of the pointed target. We developed a prototype system using a single camera and a personal computer where the user points at a virtual 3D object floating outside the display. This paper proposes a method of measuring the spatial positions of the eye and finger using a single camera. We evaluated the results on specified positions obtained by the prototype system and ensured performance was adequate for practical use.

キーワード: インタフェース, インタラクション, 指差し, 立体映像, 3次元画像計測

1. ま え が き

近年, バーチャルリアリティの研究が盛んになり, 医療, 教育, CAD/CAM など様々な分野へ, 3次元画像技術が応用されるようになってきた¹⁾. 仮想的に作られた広い空間の中で物を掴んだり動かしたりする場合に適用可能な, 3次元的なジェスチャを高い精度で認識する方法²⁾などが提案されており, 手指を用いた仮想空間でのインタラクションが実現されている. これに対し, 立体テレビモニタ装置などにより比較的小さな空間に提示された立体映像へのインタラクションを実現するような場合には, 手の「形状」認識よりも, 指先の「位置」などの情報が有用であることが多く, また複雑な処理を必要としないため, リアルタイムなインタラクションを容易に実現することができる.

先に筆者らは, 指先を通る視線の先を指示対象とする「指差し」システムについて報告した³⁾. このシステムでは指示対象が実在の空間であるため, 操作者の立ち位置から視点や指先の位置を推定する方法を用いていた. しかし, 指先と視点位置の3次元空間座標が得られると「指差し」方位を決定できるため, 本論文で述べるモニタ装置で表示された映像への「指差し」による指示では, 視点や指先の位置をカメラ画像より直接計測する方法を用いる.

ビデオカメラを用いて任意の点の3次元位置情報を得るには, 2台以上のカメラを必要とするステレオ画像計測が

一般的で従来より用いられてきた⁴⁾. 本報告では, 1台のカメラ画像中の顔画像より視点位置の3次元位置計測を行う方法について述べ, モニタ映像への指示を行う指差しポインタについて報告する.

2. 単眼カメラ画像による3次元位置計測

2.1 3次元空間座標の計測

「指先」を利用して指差しポインティングデバイスを構成するためには, 指先以外に視点位置の3次元空間座標も求めなければならない. 3次元空間の直線上に並んだ3点を指標として用いることにより, 画像上の計測点から実空間での3次元座標の情報を得る関係式の導出を行うことができるため, 顔画像中の眼の特徴点(計測点)を用いて以下のように視点の3次元空間座標を得る.

図1は, ピンホールカメラを用いて操作者の顔を撮影した場合について, 実空間に存在する顔とカメラの撮像面に映った映像を図示したものである. 図1において, ピンホールカメラの焦点位置を原点Oとする絶対座標系を考え, 右手系でX-Y-Z軸をとる. 眼の特徴点の位置を点A, 点B, 点C, 点Dとし, これら各点に対応するカメラ画像上の点をそれぞれA', B', C', D'とする. また, 瞳の位置E_L, E_Rは, ABおよびCDの中点とし, 眼間距離(E_LE_R間の距離)をRとする. このとき特徴点間の距離AB, BC, CDの比をAB:BC:CD = k : 1 : kとし, $\vec{OA} = t_A \vec{OA'}$, $\vec{OB} = t_B \vec{OB'}$, $\vec{OC} = t_C \vec{OC'}$ (t_A, t_B, t_C は変数)とすると,

$$(k+1)\vec{OB} = \vec{OA} + k\vec{OC} \quad \text{より,}$$

$$t_B(k+1)\vec{OB'} = t_A\vec{OA'} + t_C k\vec{OC'} \quad (1)$$

$$|\vec{E_L E_R}| = |\vec{AC}| = R \quad (2)$$

2003年11月12日受付, 2004年1月28日再受付, 2004年3月11日採録

† 島根大学 総合理工学部

(〒690-8504 松江市西川津町 1060, TEL 0852-32-6100(代))

† Interdisciplinary Faculty of Science and Engineering, Shimane University

(1060 Nishikawatsu-cho, Matsue-shi, Shimane 690-8504, Japan)

図 1 3次元位置の計測
Measurement of the 3-D position.

が成立する。画像上において $A'B':B'C' = k':1$ とすると、

$$(k' + 1)\overrightarrow{OB'} = \overrightarrow{OA'} + k'\overrightarrow{OC'} \quad (3)$$

と表せる。式 (1), (3) において係数を比較することにより、

$$t_B = \beta t_A, \quad t_C = \gamma t_A$$

$$\text{ただし, } \beta = \frac{k' + 1}{k + 1}, \quad \gamma = \frac{k'}{k}$$

を得る。さらに式 (2) より、 $|\overrightarrow{AC}| = R$ の関係を用いると、

$$\begin{aligned} |\overrightarrow{AC}|^2 &= |\overrightarrow{OC} - \overrightarrow{OA}|^2 \\ &= t_C^2 |\overrightarrow{OC}|^2 - 2t_A t_C \overrightarrow{OA} \cdot \overrightarrow{OC} + t_A^2 |\overrightarrow{OA}|^2 = R^2 \end{aligned}$$

が得られ、 $t_C = \gamma t_A$ の関係を用いると、点 A の Z 座標は常に負であることより、

$$t_A = \frac{R}{\sqrt{(\gamma^2 |\overrightarrow{OC}|^2 - 2\gamma \overrightarrow{OA} \cdot \overrightarrow{OC} + |\overrightarrow{OA}|^2)}} \quad (> 0)$$

を得る。視点の位置を両眼の中央とすると、 $E_L E_R$ の中点、すなわち BC の中点が視点の位置であるから、 t_A , t_B , t_C の値が求まることにより、点 B', C' の座標値を用いて視点位置の 3 次元空間座標を一意に求めることができる。

2.2 計測点の検出

視点の 3 次元空間座標を得るためには、画像処理により顔画像中の眼の特徴点 (計測点) を検出しなければならない。まず、カメラ画像より HSV 表色系での閾値処理により、肌色 (顔) 領域と黒色 (髪) 領域を抽出する。肌色領域に隣接する髪候補領域中に一定以上の黒い領域を検出できれば、その肌色領域を顔と判断する。検出された顔領域中で、RGB 表色系で閾値処理を行い、眼の矩形領域を検出する。眼の矩形領域枠と瞳中心を通る直線との四つの交点を眼の特徴点 (計測点) として用い、視点位置の 3 次元空間座標を決定する。図 2 は、顔領域および眼領域の検出結果を示したものである。視点の座標位置および文献 5) の手法で得られた指先の 3 次元空間座標から、「指差し」の方位を決定して指差しポインタに利用するが、これについては次章で述べる。

図 2 顔および眼領域の検出
Detected regions.

図 3 モニタ座標系 $M_X-M_Y-M_Z$
 $M_X-M_Y-M_Z$ coordinates.

3. 指差しポインティングシステムの構成

3.1 モニタ座標系とのキャリブレーション

「指差し」を利用してテレビモニタで表示された物体を指し示すためには、指差しの方位と物体の位置関係を把握しなければならない。物体の表示はモニタ画面を基準にした座標系を用いて行われるため、図 3 に示すように、縦、横方向および前後方向に、それぞれ M_X , M_Y , M_Z 軸を設定し、これをモニタ座標系とする。指先の動きを撮影するカメラやテレビモニタの設置場所や方向は任意であるため、カメラ座標系とモニタ座標系は一般に異なり、計測を実施する前にキャリブレーションを行い、カメラ座標系とモニタ座標系との相対関係を求めておく必要がある。 M_X , M_Y 軸はモニタ画面上の平面にあり、 M_Z 軸は、この平面に直交するため、カメラ座標系でのモニタ画面 (平面) の位置を求めれば、モニタ座標系を決定することができる。

空間座標系において平面を決定するためには、平面上の 3 点の座標が必要であるが、これらの点をそれぞれ P_1 , P_2 , P_3 とする。一般にカメラとテレビモニタは、操作者と対面するように設置するため、モニタがカメラ映像に写りこむことはない。そこで、カメラで撮影可能な任意の 4 点の座標を用いて、モニタ画面上の点を決定する。図 3 に示すように、点 E_1 , F_1 および E_2 , F_2 は、点 P_1 の近傍を通る直線 $E_1 F_1$, $E_2 F_2$ 上の点であるとする。直線 $E_1 F_1$, $E_2 F_2$ の両方に直交する直線を考え、交点 P'_1 , P''_1 の中点を P_1 とする。交点 P'_1 , P''_1 は直線 $E_1 F_1$, $E_2 F_2$ 上の点であるため、

$$\overrightarrow{OP'_1} = \overrightarrow{OE_1} + s \overrightarrow{E_1 F_1}, \quad \overrightarrow{OP''_1} = \overrightarrow{OE_2} + t \overrightarrow{E_2 F_2}$$

(ただし、 s , t は変数) と表せる。ベクトル $\overrightarrow{P'_1 P''_1}$ は、

$$\begin{aligned}\overrightarrow{P_1'P_1''} &= \overrightarrow{OP_1''} - \overrightarrow{OP_1'} \\ &= \overrightarrow{E_1E_2} + t \overrightarrow{E_2F_2} - s \overrightarrow{E_1F_1}\end{aligned}$$

であり、 $\overrightarrow{P_1'P_1''} \cdot \overrightarrow{E_1F_1} = 0$ 、 $\overrightarrow{P_1'P_1''} \cdot \overrightarrow{E_2F_2} = 0$ より、

$$A \begin{pmatrix} s \\ t \end{pmatrix} = \vec{b} \quad (4)$$

として、 s 、 t についての連立方程式を得る。ただし、

$$A = \begin{pmatrix} |\overrightarrow{E_1F_1}|^2 & -\overrightarrow{E_1F_1} \cdot \overrightarrow{E_2F_2} \\ \overrightarrow{E_1F_1} \cdot \overrightarrow{E_2F_2} & -|\overrightarrow{E_2F_2}|^2 \end{pmatrix}$$

$$\vec{b} = \begin{pmatrix} \overrightarrow{E_1F_1} \cdot \overrightarrow{E_1E_2} \\ \overrightarrow{E_1E_2} \cdot \overrightarrow{E_2F_2} \end{pmatrix}$$

とおいた。式 (4) より s 、 t を求めることにより、点 P_1 の座標を決定することができる。同様に点 P_2 、 P_3 の座標を得ることにより、カメラ座標系でのモニタ画面（平面）の位置を求めることができる。

3.2 指差しポインタ

前節で述べたキャリブレーションを行うことにより、カメラ座標系とモニタ座標系との相対関係が求まるため、視点位置と指先の位置から決定される「指差し」方位とモニタに表示された物体との位置関係から、操作者が「指差し」により指し示した対象物体を特定することができる。この原理を用いて、「指差し」により画面内の映像オブジェクトを指示するポインティングシステムを構築した。試作システムは、Windows が動作する IBM PC/AT 互換機と市販の USB カメラ（対角画角 54° ）を 1 台使用する構成とし、カメラにより入力された縦 352 画素、横 288 画素の画像から、特徴点（測定点）を抽出する画像処理部、指先、視点位置の画像計測および「指差し」対象を決定する計測部をコンピュータ上で動作するソフトウェアとして実装した。

この試作システムを用いて、操作者がモニタの前方 600mm の位置でモニタ画面上を「指差し」する動作を行っている場合の、視点位置、指先位置および指示対象位置の計測結果を図 4 に示す。また視点および指先位置が、それぞれ 600mm、400mm カメラから離れている場合の指示動作を仮定して、理論上生じる計測誤差と視点から指示対象点までの距離との関係を、図 5 に示した。デジタル画像を用いた計測では、3 次元空間がサンプリングされ量子化誤差が生じる。「指差し」の方向が同じ場合、操作者から遠くなるにつれて量子化誤差が大きくなるため、操作者から離れたモニタ画面に近い位置で誤差が大きく、また水平・垂直方向に比べ、奥行き方向の空間分解能が粗くなる。

4. む す び

モニタ画面の映像オブジェクトに対する「指差し」による指示を実現するための、指差しポインタシステムについて述べた。このシステムは、モニタ映像に対する「指差し」動作の認識を実現するものであるため、立体映像へのイン

図 4 3 次元位置計測の結果
Result of position measuring.

図 5 「指差し」位置の計測精度
Accuracy of measuring position of finger pointing.

タラクション以外に、コンピュータインタラクションを実現する入力デバイスとしても使用できるのではないかと考えている。今後はコマンド入力のためのジェスチャー認識機能の追加について検討を行い、「指先」によるインタラクションを行う立体映像システムの構築を行う予定である。

〔文 献〕

- 1) 泉武博監修：“3 次元映像の基礎”，オーム社（1995）
- 2) 島田，白井，久野，三浦：“形状の個性差に適応する精密なジェスチャー推定”，インタラクション 98, pp.5-12（1998）
- 3) 阪本邦夫，木村理恵子：“単眼カメラ画像からの指先位置計測の一手法—指差しによる仮想空間でのインタラクション—”，3 次元画像コンファレンス 2003, pp.213-216（2003）
- 4) 井口征士，佐藤宏介：“三次元画像計測”，昭晃堂（1990）
- 5) 阪本邦夫，高木美和：“立体映像へのインタラクションのための 3 次元位置計測の一手法”，映像学誌，57, 9, pp.1152-1154（2003）

さかもと くにお
阪本 邦夫 1994 年，姫路工業大学工学部電子工学科卒業。1999 年，大阪市立大学大学院工学研究科電気工学専攻後期博士課程修了。同年，秋田県立大学システム科学技術学部助手。2002 年，島根大学総合理工学部助教。仮想環境システム，画像計測，画像処理，ネットワーク映像メディア，3 次元画像表示の研究に従事。博士（工学）。正会員。

きむら りえこ
木村理恵子 2003 年，島根大学総合理工学部数理・情報システム学科卒業。現在，同大大学院総合理工学研究科数理・情報システム学専攻博士前期課程において，仮想空間におけるインタラクションに関する研究に従事。