

単眼カメラ画像からの指先位置計測の一手法

木村 理恵子, 阪本 邦夫

島根大学総合理工学部 数理・情報システム学科

A Method of Finger Pointing Using a Single Camera

Rieko KIMURA and Kunio SAKAMOTO

*Department of Mathematics and Computer Science,
Interdisciplinary Faculty of Science and Engineering, Shimane University*

Abstract

A finger pointing system is described that can specify the position of pointed target. A 3-D position is generally detected and measured using a stereo viewing. This pair of stereo images allows us to obtain the 3-D information about the object. However it is possible to estimate the position of fingertip using a single image without stereo images. The length of arm is approximately calculated from stature and shoulder length, then the 3-D position of fingertip is estimated using the geometric model of human. We have developed a prototype system using a single camera and a personal computer. This system realizes the finger pointing such that user draws a picture on the floor in a virtual space. This paper proposes the measuring method of the position of finger using a single camera. We evaluated the result of specified positions by prototype system and made sure the performance required for practical use.

1. ま え が き

近年, バーチャルリアリティ(仮想現実感)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. これら各分野への応用を視野にいれて, 立体映像を表示する装置についても, 様々な表示方式の研究が行われている¹⁾. 立体表示装置を用いて立体映像システムを構築する場合, 観察者に対して立体映像の提示を行う方法には, 大きく分けて装着型と非装着型の装置を用いる方法があり, 多種の表示装置が提案されている. 装着型の立体映像表示装置とは, 立体ヘッドマウントディスプレイ装置などのようにメガネ型あるいはゴーグル型の装置を装着することにより, 立体映像を観察することができる表示装置で, メガネを着用するという煩わしさを伴うが, 容易に大画面表示を行うことが可能で, 観察者は没入感や臨場感を伴った立体映像を観察できることが特徴である²⁾. 一方, 非装着型の立体映像表示装置とは, 特殊なメガネを必要とすることなく立体映像を観察することができる表示装置で, レンチキュラレンズ方式あるいはパララックスバリア方式の立体ディスプレイ(立体テレビモニタ)や電子ホログラフィによる動画立体映像などがこれにあたる³⁾.

このように多様な立体表示装置が立体映像システムを構築する上で利用可能であるが, より普段の生活環境に近い状態で立体映像を観察するには, 立体ヘッドマウントディスプレイのようなゴーグル型の装置を装着する必要がある立体映像表示装置ではなく, 装着物の不要な非装着型の立体映像表示装置を用いたシステム構築が望ましいと考えられる. 非装着型の立体映像表示装置を用いて, 没入感や臨場感を伴った立体映像の提示を行うためには, 表示画面の大型化が必要不可欠であるが, 110インチの大画面で15~16名が一度に観察可能な立体映像システムが既に実用化されている⁴⁾.

筆者らは立体映像を用いた3Dワークスペースシステム⁵⁾⁻⁹⁾の構築を目指し研究を行っているが, 立体映像システムによって作り出された映像空間内で, 人物(操作者)の指示動作によりシステムが何らかの応答を行うようなインタラクションシステムを構築するためには, 人物や手の位置を検出することが必要となる. このような人間の動作のセンシングを行う技術にも, センサ装着型の方式と非装着型の方式が提案されている. センサ装着型方式では, 磁気式のセンサなどを用いて人間の動きを実時間でセンシングするモーションキャプチャシステムなどが既に実用化されており, 人間の動作を実時間で計測・認識を行うことにより, 3D-CG技術で生成され

たキャラクタの動作再生に利用されている。一方、コンピュータビジョンなどの研究の推移とともに、ビデオカメラの映像から3次元動作の計測を行う非装着型の方法も提案されてきている。非装着型の立体映像システムにより提示された映像に対するインタラクションを実現するような場合には、センシングシステムも非装着型とすることが望ましいといえる。

従来より非装着型のセンシング方式としては、2台以上のカメラを用いた受動型のステレオ画像計測、受動型のステレオ法に用いられるカメラ1台を光を発するプロジェクタに置き換えたアクティブステレオ法などが3次元画像計測法として提案されている¹⁰⁾。このように、測定対象の物体表面の色情報（映像情報）と距離情報を得るには、一般に2台以上のカメラが必要である。

操作者の意図には、ジェスチャーやコマンドといった手振りや音声などで伝達可能なもの以外に、対象を指し示したり、方向を示したりする「指差し」という動作が存在する。本論文では、この「指差し」動作をカメラ画像を用いて認識する問題を取り扱う。腕をまっすぐに伸ばした姿勢で「指差し」を行うことにより、1台のカメラ画像から指先位置の計測を行い、「指差し」の方向を決定できることを示し、大型スクリーン上に表示された臨場感ある映像や立体映像を指し示したり、「指差し」により方向を指示する場合に適用可能な、カメラ1台とパーソナルコンピュータで構成された計測システムの構築について報告する。

2. 指差しによるインタフェース

音声認識や画像を用いた認識などは、人間の意図を正確に伝達する上で重要な技術であり、ロボットとのコミュニケーションやコンピュータとのインタフェースを実現するため、様々な研究が行われている。この中で人間の行う「指差し」という動作は、方向を示したり、ヒトや物を指し示したりする場合に用いられる重要なコミュニケーション手段の1つである。本章では、人間の指差し動作をコンピュータに理解させる方法について簡単に述べる。

2.1 指差し方位の計測

「指差し」を含め、手や指を用いた手振りによる意思表示は、人間の重要なコミュニケーション手段の1つであり、ジェスチャーなどの手振りを認識して、人間と機械とのコミュニケーションを実現しようとする数多くの試みが報告されている。手振りを認識するためには、カメラ画像から手や指の関節などの各部位の位置を計測し、手や指の3次元姿勢を求めなくてはならない。この姿勢を決定する方法として、手指の3次元形状モデルを利用して姿勢を推定する方法¹¹⁾などが提案されており、手話認識など手振りをを用いるコミュニケーションの実現に応用されている。このように手や指の3次元姿勢が求ま

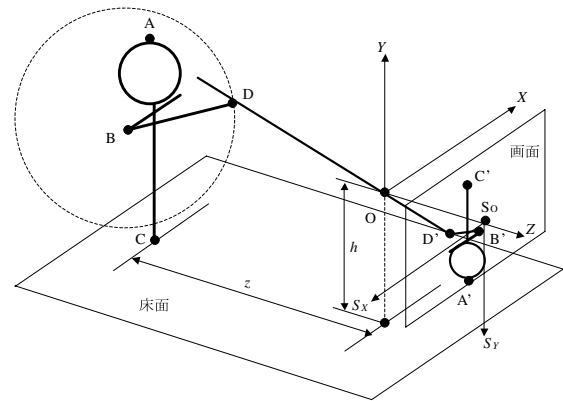


図1 指先位置の計測

ると、指差しの方向も検出できるが、「指」の方向が「指差し」の方向に一致するのは、指先と指差しの対象が近距離にある場合（「これ」という単語が付随して指差しを行う場合）である。

近距離の指差し以外に、「あれ」という単語が付随する遠距離の「指差し」がある。遠距離の指差しでは、指示方向を決定する際の基準となる指示基準点が必要となる。この指示基準点を決定するため、指差し動作のメカニズムを解明しようとする研究が行われている¹²⁾。指示基準点と指示動作点（指先）を結んだ直線上に指示対象点が存在するが、この指示基準点は、指示を行うヒトの姿勢に応じて変化するため、一意に決めることは困難である。そこで本研究では、視線の先に指示対象が存在するように、「指差し」は腕をまっすぐに伸ばした姿勢で行うこととし、両眼の中心点（視点）を指示基準点として、指差しの方向を求めることにした。

2.2 指先の3次元位置計測

「指差し」の方向を求めるには、指先と指示基準点（視点）の3次元座標の計算が必要となる。これらの3次元位置座標の計測には、多視点画像から空間中にある点の位置を求めて、指差し方位の決定する手法が報告されている¹²⁾。このように、3次元空間の位置情報を得るには、2台以上のカメラ画像を利用したステレオ画像計測を用いるのが一般的な方法であり、3次元空間中の任意の点の位置座標を求めることが可能である。しかし、

- ・腕をまっすぐに伸ばした姿勢で「指差し」を行うという「指差し」動作を行う場合の姿勢についての制約条件を利用するとともに、カメラ画像の撮影条件
- ・操作者は直立して床の上に立っている
- ・操作者の居る空間の床面は一樣に水平な平面である
- ・操作者の顔が写るように正面から全身像を撮影するという条件も満たしている場合、人間工学に基づく身体的特徴（身長と腕の比率など）を用いることにより、1台のカメラ画像による指先位置の計測が可能であり、画像上の測定点の位置から実空間での指先位置の座標を得る関係式の導出を行うことができる。なお、一般的な部

屋などの空間で指差しによるインタラクションを行う場合には、床面の条件は満たされており、また、視線の先に指示対象が存在するように、腕をまっすぐに伸ばした姿勢で「指差し」を行う動作は、操作者に不自由さを強いるものでもない。このように、任意の点の3次元空間座標の計測ではなく、操作者の「指差し」という動作、すなわち、指先位置のみの計測が可能であればよいという場合には、1台の計測用カメラのみで構成可能な簡易な計測システムを実現できる。この指先位置計測の幾何学的解析については3章で、計測処理システムについては4章で述べる。

3. 単眼カメラ画像による指先位置の計測

3.1 指先位置の計測

図1は、ピンホールカメラを用いて「指差し」を行っている操作者を撮影した場合について、実空間にいる人物とカメラの撮像面（カメラ映像の画面）に映った人物像を模式的に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。さらに、カメラ映像の画面中央の点S₀を原点とするスクリーン座標系を考え、これをS_X-S_Y座標系とする。また、人物の頭上の位置を点A、右肩の位置を点B、足元の位置を点C、指先の位置を点Dとし、これら各点に対応するカメラ映像の点をそれぞれA'、B'、C'、D'とする。このときカメラ映像上の点A'、B'、C'、D'のスクリーン座標を用いて、指先の位置である点Dの絶対座標を以下のようにして求める。

ピンホールカメラの焦点距離（焦点と撮像面間の距離）を $f(>0)$ とすると、点S₀の絶対座標は $(0, 0, f)$ と表される。また、点A'、B'、C'、D'のスクリーン座標をそれぞれ、 (sx_A, sy_A) 、 (sx_B, sy_B) 、 (sx_C, sy_C) 、 (sx_D, sy_D) とすると、点A'、B'、C'、D'の絶対座標はそれぞれ、 $(-sx_A, -sy_A, f)$ 、 $(-sx_B, -sy_B, f)$ 、 $(-sx_C, -sy_C, f)$ 、 $(-sx_D, -sy_D, f)$ となる。ここで、ピンホールカメラの焦点位置の床面からの高さを $h(>0)$ とすると、ベクトル $\vec{C'O}$ の各成分は $(sx_C, sy_C, -f)$ であるため、足元の位置である点Cの絶対座標 (x_C, y_C, z_C) は、直線C'C

$$X = sx_C \cdot t_C, \quad Y = sy_C \cdot t_C,$$

$$Z = -f \cdot t_C \quad (t_C \text{は媒介変数})$$

と床面 $Y = -h$ との交点として求めることができ、

$$x_C = -\frac{sx_C \cdot h}{sy_C}, \quad y_C = -h, \quad z_C = \frac{f \cdot h}{sy_C}$$

となる。この結果から、操作者からカメラの焦点までの距離 $z(>0)$ は、点CのZ座標が負であることを考慮して、

$$z = -\frac{f \cdot h}{sy_C}$$

と表される。「指差し」を行っている操作者は、カメラに対し正対しているため、点A、B、Cは、 $Z = -z$ の同一平面上にあると仮定すると、ベクトル $\vec{A'O}$ の各成分は $(sx_A, sy_A, -f)$ であるため、頭上の位置である点Aの絶対座標 (x_A, y_A, z_A) は、直線A'A

$$X = sx_A \cdot t_A, \quad Y = sy_A \cdot t_A,$$

$$Z = -f \cdot t_A \quad (t_A \text{は媒介変数})$$

と平面 $Z = -z$ との交点として求めることができ、

$$x_A = \frac{sx_A \cdot z}{f}, \quad y_A = \frac{sy_A \cdot z}{f}, \quad z_A = -z$$

となる。また、右肩の位置である点Bの絶対座標 (x_B, y_B, z_B) も同様にして、

$$x_B = \frac{sx_B \cdot z}{f}, \quad y_B = \frac{sy_B \cdot z}{f}, \quad z_B = -z$$

と求まる。点Aと点CのY座標が求まったことにより、人物の身長 h の値を $|y_A - y_C|$ として得ることができる。また、ヒトの身長と腕などの各部分の長さの比は統計的に決まっており、腕を水平に広げた場合の指先間の長さは、ほぼ身長に等しいとされている。直線ACを体の中心線とすると、肩幅は $2 \cdot |x_A - x_B|$ で与えられるため、肩から指先までの腕の長さ $R(>0)$ は、

$$R = \frac{|y_A - y_C|}{2} - |x_A - x_B|$$

として求まり、身長 h の値を基に腕の長さ R を得ることができる。

ここでは、腕BDの長さ R として、指先の位置である点Dの絶対座標 (x_D, y_D, z_D) を求める。点Dの絶対座標は図1に示すように、直線D'Dと点Bを中心とする半径 R の球との交点として求まる。ベクトル $\vec{D'O}$ の各成分は $(sx_D, sy_D, -f)$ であるため、絶対座標系の原点を通る直線D'Dの方程式は、 $t\vec{D'O}$ のように係数 t を用いてベクトル表記される。また、係数 t を媒介変数として、 X, Y, Z の各成分を表記すると、

$$X = sx_D \cdot t, \quad Y = sy_D \cdot t, \quad Z = -f \cdot t$$

となる。一方、点Bを中心とする半径 R の球の方程式は、

$$(X - x_B)^2 + (Y - y_B)^2 + (Z + z)^2 = R^2$$

であるから、 X, Y, Z を代入して、 t についての2次方程式

$$(sx_D^2 + sy_D^2 + f^2)t^2 + 2(-sx_D \cdot x_B - sy_D \cdot y_B - f \cdot z)t + (x_B^2 + y_B^2 + z^2 - R^2) = 0$$

を得る。この2次方程式の解 t を得ることにより、点Dの絶対座標を求めることができる。ここで、

$$\alpha = sx_D^2 + sy_D^2 + f^2 \quad (> 0),$$

$$\beta = -sx_D \cdot x_B - sy_D \cdot y_B - f \cdot z,$$

$$\gamma = x_B^2 + y_B^2 + z^2 - R^2$$

とおくと, 2次方程式 $at^2 + 2\beta t + \gamma = 0$ の解は,

$$t = \frac{-\beta \pm \sqrt{\beta^2 - \alpha\gamma}}{\alpha}$$

と得られる.

一般に2次方程式の解は2個あり, ここでは直線 $D'D$ と点 B を中心とする半径 R の球との2交点(絶対座標系の原点に近い点と遠い点)に対応する. 図1に示すような状態では, 腕 BD の長さ R と操作者からカメラの焦点までの距離 z の関係は, 一般に $z > R$ であり, $\gamma > 0$ が成立する. また, 直線 $D'D$ の方程式のベクトル表記において, $t = 0$ のときは絶対座標系の原点を表しており, $Z < 0$ の領域に交点 D が存在するためには, $t > 0$ でなければならない. 従って, $\alpha > 0, \gamma > 0$ であり, かつ常に $t > 0$ であるためには, $\beta < 0$ でなければならない. このとき, 2次方程式の解 t のうち, 複号の負号を採ったとき与えられる解が, 絶対座標系の原点に近い点であり, 指先の位置である点 D となる. ゆえに, 点 D を与える媒介変数 t_D の値は,

$$t_D = \frac{-\beta - \sqrt{\beta^2 - \alpha\gamma}}{\alpha}$$

であり, 指先の位置である点 D の絶対座標 (x_D, y_D, z_D) は, 媒介変数 t_D を用いて,

$$x_D = sx_D \cdot t_D, \quad y_D = sy_D \cdot t_D, \quad z_D = -f \cdot t_D$$

と求まる.

3.2 「指差し」位置の計測

前節のようにして, カメラ画像から指先の位置が求まると, 「指差し」位置は指先を通る視線の先の点として求めることができる. 図2において, 両眼の中心点を点 E , これに対応するカメラ映像上の点を E' とし, 点 E' のスクリーン座標を (sx_E, sy_E) とすると, 点 E' の絶対座標は $(-sx_E, -sy_E, f)$ となる. 点 E も点 A, B, C と同様に $Z = -z$ の同一平面上にあると仮定すると, ベクトル $\vec{E'O}$ の各成分は $(sx_E, sy_E, -f)$ であるため, 点 E の絶対座標 (x_E, y_E, z_E) は, 直線 $E'E$

$$X = sx_E \cdot t_E, \quad Y = sy_E \cdot t_E,$$

$$Z = -f \cdot t_E \quad (t_E \text{ は媒介変数})$$

と平面 $Z = -z$ との交点として求めることができ,

$$x_E = \frac{sx_E \cdot z}{f}, \quad y_E = \frac{sy_E \cdot z}{f}, \quad z_E = -z$$

となる. ここで, 操作者が床面上の点 G を指差しているとするとき, ベクトル \vec{ED} の各成分は $(x_D - x_E, y_D - y_E, z_D + z)$ であるため, 点 G の絶対座標 (x_G, y_G, z_G) は, 点 D を通る直線 EG

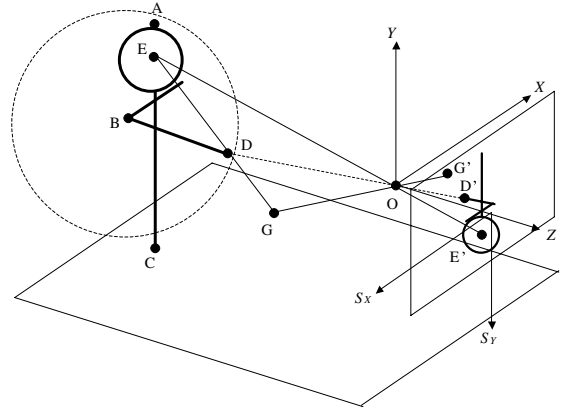


図2 「指差し」位置の計測

$$X = (x_D - x_E) \cdot t_G + x_D,$$

$$Y = (y_D - y_E) \cdot t_G + y_D,$$

$$Z = (z_D + z) \cdot t_G + z_D \quad (t_G \text{ は媒介変数})$$

と平面 $Y = -h$ との交点として求めることができ, 媒介変数 t_G の値を,

$$t_G = -\frac{h + y_D}{y_D - y_E}$$

として, x_G, y_G, z_G の値が求まる.

さらに, 床面上の点 G がカメラ映像に写る場合, 点 G のカメラ映像上で対応する点を点 G' とすると, 絶対座標系での撮像面の方程式は $z = f$ であり, ベクトル \vec{GO} の各成分は $(-x_G, -y_G, -z_G)$ であるため, 点 G' の絶対座標 (x'_G, y'_G, z'_G) は, 直線 GG'

$$X = -x_G \cdot t'_G, \quad Y = -y_G \cdot t'_G,$$

$$Z = -z_G \cdot t'_G \quad (t'_G \text{ は媒介変数})$$

と平面 $Z = f$ との交点として求めることができ,

$$x'_G = \frac{x_G \cdot f}{z_G}, \quad y'_G = \frac{y_G \cdot f}{z_G}, \quad z'_G = f$$

となる. さらに点 G' の絶対座標 (x'_G, y'_G, z'_G) をスクリーン座標系へ変換することにより, 点 G' のスクリーン座標 (sx_G, sy_G) は,

$$sx_G = -\frac{x_G \cdot f}{z_G}, \quad sy_G = -\frac{y_G \cdot f}{z_G}$$

と求まる.

このように指差しの方向を示すベクトル \vec{ED} は, 操作者の前方にあるスクリーンなど床面以外の任意の方向を指差している場合でも, カメラ画像から視点と指先の位置が求まれば算出することができるため, 実空間のワールド座標系とキャリブレーションを行い座標系の相対関係を導出しておけば, 実空間の位置(座標)を「指差し」により指示することもできる.

4. 指差しによるインタラクション

4.1 カメラ画像を用いた計測

1台のカメラ画像から指先位置の計測を行うためには、まず、カメラ画像から人物のいる領域を決定し、画像上の操作者の立ち位置から、実空間でのカメラから操作者までの距離や身長などを計測しなければならない。試作システムでは、顔と指先のカメラ画像上での位置により、顔領域と指先領域の区分けを容易に行えるように、計測の際にはカメラの高さなど設置位置を調整して固定した。文献13)の馬場らの提案による方法を用いて、カメラ画像から人物領域の抽出を行い、人物の存在する矩形領域を求める。この矩形がカメラ画像内に納まっている場合のみ、カメラ画像に操作者の全身像が写っていると判断し、操作者の身長をカメラ画像より算出する。

「指差し」を認識するためには、さらにカメラ画像から指先の抽出を行い、指先の位置を決定する必要がある。カメラ画像より肌色抽出画像を得ることで、顔、左手および右手の矩形領域を求める。カメラ画像の上部にある面積が最大の肌色領域が、顔の領域と考えられるため、この矩形領域の高さと人物領域の矩形の高さの比が、ヒトの身体的特徴に一致する場合に、顔領域として検出し、残りの肌色領域を左手および右手の領域とする。

さらに、右手による「指差し」認識を行う場合、操作者の身長および腕の長さが求まっているため、カメラ映像上での人物の頭上の位置を点A'、右肩の位置を点B'、足元の位置を点C' (図1参照)とし、点A'、B'、C'の座標をそれぞれ、 (sx_A, sy_A) 、 (sx_B, sy_B) 、 (sx_C, sy_C) とすると、右肩の位置を中心とする半径R'

$$R' = \frac{|sy_A - sy_C|}{2} - |sx_A - sx_B|$$

の円の内部に肌色領域が唯一存在する場合を、右手の領域として判断することで、左手の領域を画像から排除することができる。このように円内にある肌色領域を右手として認識することにより、カメラ画像上の「指差し」を行っている指先領域の位置から、右手の位置情報(空間座標)を得ることができる。なお、試作システムではインタラクションの結果を実空間へ投影しないため、実空間のワールド座標系とのキャリブレーションは行わず、計測された座標値はカメラの絶対座標系を用いて表すことにした。

4.2 インタラクションシステムの構成

本論文で提案する単眼カメラ画像による指先位置計測の有効性を確認するため、1台のカメラ画像を用いた指先位置の計測処理を、パーソナルコンピュータ上で動作するソフトウェアとして実装し、仮想空間での指差しによるインタラクションを実現する簡易なインタラクションシステムの構築を行った。試作システムは、CPUにIntel Celeron 1.0GHzを搭載したWindowsが動作する

表 1 USB カメラの仕様。

撮像素子	1/5 インチ CMOS イメージセンサー
有効画素数	10 万画素
焦点距離	20cm ~
画角	54 °
フレームレート	最大 16 フレーム/秒
インタフェース	USB インタフェース 1.1

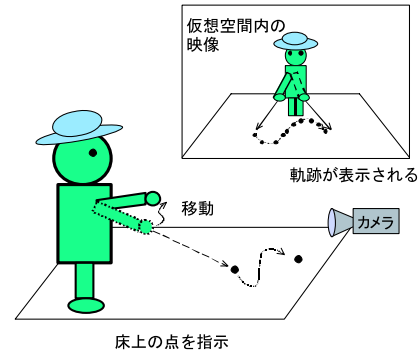


図 3 仮想空間でのインタラクション

IBM PC/AT 互換機を使用し、このコンピュータに表1に示すような仕様の一般に市販されているUSBカメラを1台接続して、カメラ映像の入力、画像認識、計測およびインタラクションの処理すべてを1台のコンピュータで行う簡易な構成とした。

試作システムでは、図3のように操作者が「指差し」を行い、指差した床面上に絵を描くというアプリケーションを動作させた。このアプリケーションは、指先を通る視線の先の床面上の点を「指差し」位置として、カメラ画像の上に点をプロットをしていくもので、コンピュータ内の仮想空間の床面上に絵を描くことができるため、操作者が行った「指差し」の位置を画面上で確認することができる。

使用したコンピュータは、高性能のスペックを備えたものではなく、カメラにより入力された縦352画素、横288画素の画像の処理も、画像処理ボードなどのハードウェアを用いずソフトウェアで行っているため、毎秒5フレーム程度の処理速度であるが、操作者が「指差し」により床面上に絵を描く動作を行ったとき、操作者の指示動作に追従して仮想空間内の床面上に絵が描かれており、簡易なインタラクションを実現するには十分な処理速度である。

このアプリケーションを用いて、操作者が行った指差しによる指示と、カメラ画像を用いた計測により得られる結果の比較を行ったが、これについては次節で述べる。

4.3 指差しによる指示

試作システムが動作する実空間の床面上に、直径160mmの円形のマーカを図4に示すように配置した。円形マーカは、カメラから3000mmと3500mm離れた位置に1000mm間隔で3列に配置されており、カメラから5000mm程度離れた任意の位置で操作者が、床面上の

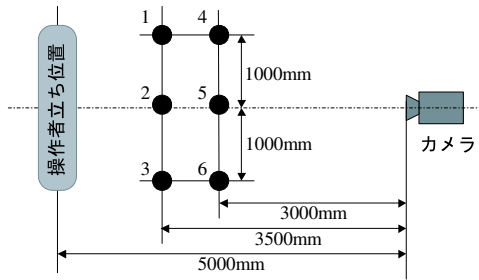


図4 マーカの配置

マーカを指差し、「指差し」による指示が床面の座標位置に、どのように対応しているかを確認した。図5は、被験者6名により床面上の6つのマーカを指し示したときの「指差し」方向の計測誤差を示したものである。図5に示すように、「指差し」方向の角度誤差は、操作者の左方に位置するマーカを指差した場合に大きくなるが、これは、右手で30°以上左方のマーカを指差すため、腕をまっすぐに伸ばして「指差し」を行う姿勢が保たれていないことによると考えられる。また、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。図6はカメラから5000mm離れた位置で床面近傍を「指差し」する動作を行った場合の指示位置の誤差を示したものである。「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、水平・垂直(X, Y軸)方向、奥行き方向ともに、カメラに近い位置で誤差が大きい。さらに奥行き方向については、操作者の近くでも誤差が大きくなるが、近くの床面を指差す場合には、指先は常に球面上を移動するため、カメラ画像上での指先の動きが小さくなり、サンプリング間隔が大きくなるためである。

5. むすび

1台のカメラ画像から指先位置の計測を行う方法および指差しによるインタラクションを行うシステムについて述べた。本論文で述べた指先位置の計測方法を用いることにより、一般に市販されているUSBカメラを1台接続して、カメラ映像の入力、画像認識、計測およびインタラクションの処理すべてを1台のコンピュータで行う簡易な構成の「指差し」インタラクションシステムを構築することができた。また、指差しの精度を検討した結果、視野角40°でスクリーンに映像が表示されるような場合「指差し」の角度誤差は5°程度であり、画面領域を100程度に区分して「指差し」の指示が行えるため、画面上の複数の人物から指差しにより1名選ぶといった用途の指示システムなど、映像の中の物体単位の指示であれば十分に使用することができる。また、指先位置の計測に加えて「確定」や「キャンセル」といった指示をジェスチャー認識により行う機能を加えることにより、システムに操作者の意図を正確に伝達することができる。

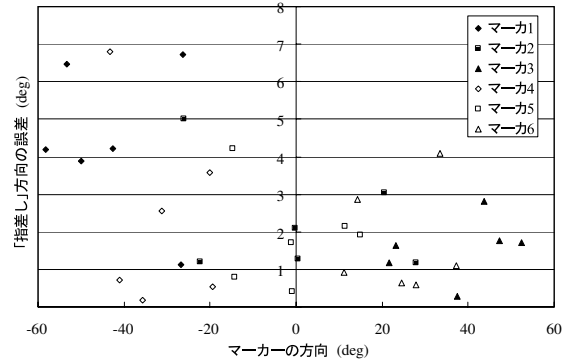


図5 「指差し」位置計測の評価結果

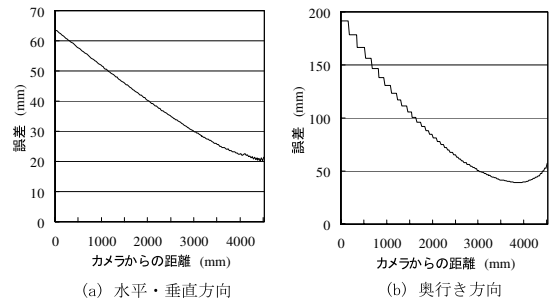


図6 「指差し」による位置計測の精度

と考えている。今後の課題としては、提案した計測手法とインタラクションシステムを改良し、立体映像表示技術を用いて空間に立体映像として表示された仮想物体への「指差し」によるインタラクションを行うシステムや、さまざまなコマンド入力のためのジェスチャー認識機能の追加などについて検討を行う予定である。

【参考文献】

- 1) 大越孝敬：“三次元画像工学”，朝倉書店（1991）
- 2) 泉武博監修：“3次元映像の基礎”，オーム社（1995）
- 3) 辻内順平編著：“ホログラフィックディスプレイ”，産業図書（1990）
- 4) 志水英二，岸本俊一：“ここまで来た立体映像技術”，工業調査会（2000）
- 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”，映情学誌，57，9，pp.1152 - 1154（2003）
- 6) 阪本邦夫，木村理恵子：“単眼カメラ画像を用いた指差しポインタの構成”，映情学誌，58，5，pp.721 - 723（2004）
- 7) 阪本邦夫，木村理恵子：“3Dワークスペース構築のためのシースルー立体表示ビューア”，映情学誌，58，7，pp.999 - 1001（2004）
- 8) 阪本邦夫，高木美和：“偏光パララックスバリア方式多視点立体ディスプレイの試作”，映情学誌，58，9，pp.1288 - 1290（2004）
- 9) 阪本邦夫，木村理恵子：“逆視領域のないパララックスバリア立体ディスプレイの一方式”，映情学誌，58，11，印刷中（2004）
- 10) 井口征士，佐藤宏介：“三次元画像計測”，昭晃堂（1990）
- 11) 島田伸敬，白井良明，久野 義徳：“三次元モデルを用いた二次元動画像からの手指姿勢の推定”，信学技報 PRU94-5，pp.25 - 32（1994）
- 12) 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”，情処学「コンピュータビジョンとイメージメディア」研資，No.133 - 017（2002）
- 13) 馬場，松尾，大橋，乃万，松尾，江島：“HeadFinder:単眼視動画像を用いた複数人追跡”，画像センシングシンポジウム 2001，pp363-368（2001）