

モーメント特徴に基づく頭部姿勢推定手法

高松 亮 佐藤 誠

東京工業大学精密工学研究所

〒 226 横浜市緑区長津田町 4259

E-mail: rtakamat@pi.titech.ac.jp

あらまし 本報告では、使用者の頭頂部に装着された楕円形のマーカをビデオカメラで撮影し、得られたマーカの像から頭部の位置と方向を推定する手法を提案した。提案手法はビジョンベースのヒューマンインタフェースの構成要素として用いられることを意図したものである。すなわち、使用者の頭部姿勢の情報から、使用者の注意の方向や、首を縦方向や横方向に振ることで形成されるゼスチュアを推定することを想定し、使用者が首を縦および横にそれぞれ振った場合の頭部の回転角(2自由度)と、頭部位置の水平面への射影(2自由度)を使用者の頭部姿勢として推定する。実画像に基づく頭部姿勢推定実験を行なった結果、本手法の有効性が示された。

Head Posture Estimation Method Based on Moment Feature

Ryo TAKAMATSU and Makoto SATO

Precision and Intelligence Laboratory, Tokyo Institute of Technology

4259 Nagatsuta-cho, Yokohama-shi, Japan 226

E-mail: rtakamat@pi.titech.ac.jp

Abstract In this paper, we propose a method to estimate posture of human head, in which the position and the direction of the head are estimated from the image of singular marker put on the head of a user. As the feature, moments of image up to second degree are used. Posture is estimated by estimation system which consists of observation model based on local analysis and compensation function. Our method has the following characteristics. Firstly, it does not constrain user. Secondly, a user only have to put singular marker on the top of the head, and thus easily puts on and takes off the marker. Thirdly, because the features to represent the marker are moments up to second degree, we can expect considerable robustness to noise and less computational cost. The experimental results shows validity of our method.

1 はじめに

計算機の情報処理能力の高速化，大容量化にともない，ビデオカメラでとらえられた実時間画像を入力媒体とするヒューマンインタフェースが今後一般化するものと考えられている．画像情報を媒体とすることで，使用者の体の各部の姿勢や顔の表情などのノンバーバルな情報を，非接触な方法で入力することが可能になり，より効率の高くしかも自然なインタフェースを実現し得るものと期待される [1][2][3]．

使用者の発するノンバーバルな情報のうち最も基本的なものの一つに頭部姿勢がある．顔面の方向は視線方向とはほぼ一致すると考えられるため，頭部の向きから使用者の注意がどのような方向に向いているかをある程度推測することができる．また，人は首を縦方向や横方向に振ることによって「肯定」や「否定」などをあらわす基本的なゼスチュアを発生させることができる．

本報告では，ビジョンベースのヒューマンインタフェースを構成するための，頭部姿勢推定法を提案する．提案手法は，使用者の頭頂部に装着されたマーカをビデオカメラで撮影し，得られたマーカの像から頭部の位置と方向を推定するものである．

本手法では，楕円形のマーカを1個用いる．楕円の中心を頭部位置，長軸方向を頭部の方向とする．楕円を表現する特徴量として0次から2次までの画像モーメント [4] を用いる．モーメントは対象の中心位置や中心からの大きさといった対象の大局的な特徴を表現する量であるため，多数個のマーカや特徴点を抽出する手法 [5] でしばしば用いられる局所の特徴量と比較して雑音に対する頑健性が高いと考えられる [6]．また，マーカは頭頂部に1個装着するのみでよいとため，着脱が容易な形状にデザインすることが可能である．

提案手法による頭部姿勢の推定処理は2つの段階に分けることができる．すなわち，頭部の位置・方向と特徴ベクトルとの関係を記述した

数理モデルによる推定をまず行い，近似の導入に起因する推定誤差を双2次の多項式で表される補償関数で補償する．頭部姿勢の真の値と数理モデルによる推定値との関係は比較的単純なものであるため，少ないパラメータで補償関数を記述することができる．

なお，本報告では使用者の頭部姿勢の情報から，使用者の注意の方向や，首を縦方向や横方向に振ることで形成されるゼスチュアを推定することで，より高度なヒューマンインタフェースを構成することを想定し，使用者が首を縦および横にそれぞれ振った場合の頭部の回転角(2自由度)と，頭部位置の水平面への射影(2自由度)を使用者の頭部姿勢として推定するものとする．使用者は頭部の高さをほとんど変化させないものと仮定し，使用者は首を傾けることなくほぼ鉛直に保つものとする．

2 頭部姿勢推定法

2.1 観測モデル

本節では撮影されたマーカの像のモーメント特徴と，実空間でのマーカの位置・方向・大きさなどのパラメータとの関係を定式化した数理モデルである，観測モデルを構築する．

2.1.1 画像のモーメントと楕円形状

2次元濃淡画像を $f(x, y)$ で表す． $f(x, y)$ の (i, j) 次 ($i, j \geq 0$) のモーメントは次式のように表される．

$$\mu_{ij} = \iint x^i y^j f(x, y) dx dy \quad (1)$$

$i+j$ をモーメントの次数と呼ぶ．また，次式に示すような重心 $(x_c, y_c) = (\mu_{10}/\mu_{00}, \mu_{01}/\mu_{00})$ のまわりのモーメントを中心モーメントとよぶ．

$$m_{ij} = \iint (x-x_c)^i (y-y_c)^j f(x, y) dx dy \quad (2)$$

いま，内部の濃淡値が一定の楕円形状を観測して得られたモーメントを考える．楕円の中心

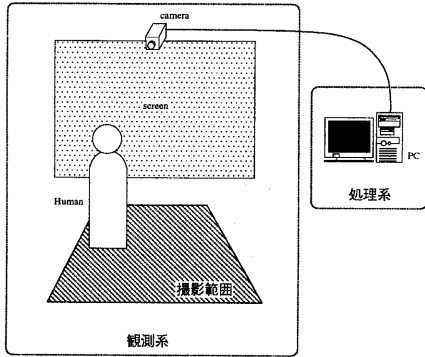


図 1: 頭部姿勢推定システム

(X_0, Y_0) は重心 $(\mu_{10}/\mu_{00}, \mu_{01}/\mu_{00})$ に等しい。また行列 M を、

$$M \equiv \begin{bmatrix} m_{20} & m_{11} \\ m_{11} & m_{02} \end{bmatrix}$$

と定義し、行列 M の固有値を $\lambda_1, \lambda_2 (\lambda_1 > \lambda_2)$ とおくと、楕円の長軸および短軸の長さ A, B は次式のように表わされる。

$$A = \sqrt{3\lambda_1/m_{00}}, \quad B = \sqrt{3\lambda_2/m_{00}}$$

また、楕円の長軸の方向は最大固有値を与える固有ベクトルの方向に等しい。

2.1.2 観測系の構成

図 1 に本手法に基づく頭部姿勢推定システムの全体像を示す。楕円形のマーカを頭部に装着した使用者をビデオカメラで撮影する。使用者は、ビデオカメラの視野にマーカが入っている範囲内で自由に動くことができる。

カメラとマーカの幾何学的配置を次のように仮定する。マーカの平均位置をワールド座標系 (x, y, z) の原点に置く。カメラは $(0, Y_0, -Z_0)$ の位置から光軸を原点に向けて設置するものとする。カメラが原点を見下ろす角度を θ_0 とすると次式が成立する (図 2)。

$$\cos \theta_0 = \frac{Z_0}{D_0}$$

$$\sin \theta_0 = \frac{Y_0}{D_0}$$

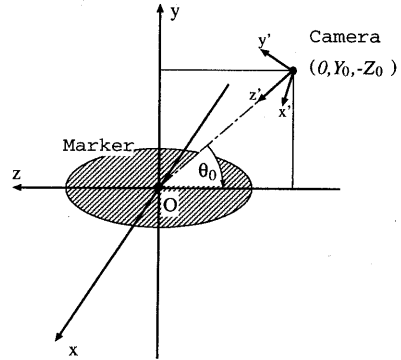


図 2: カメラとマーカの幾何学的配置

$$\text{ただし } D_0 = \sqrt{Y_0^2 + Z_0^2}$$

カメラ座標系 (x', y', z') を光軸が z' 軸に一致する右手系にとる。カメラ座標系とワールド座標系との間には次式のような関係が成立する。

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_0 & \sin \theta_0 \\ 0 & -\sin \theta_0 & \cos \theta_0 \end{bmatrix} \begin{bmatrix} x \\ y - Y_0 \\ z + Z_0 \end{bmatrix} \quad (3)$$

2.1.3 楕円形状の観測モデル

楕円形のマーカと、それをカメラで撮影した像の関係をモデル化する。使用者は、楕円の長軸方向が顔面の前方向と一致するようにマーカを装着し、頭部を水平に保った状態で、カメラから下ろした鉛直線が床と交わる点の方向に頭部を向けているものとする。このとき、ワールド座標系での楕円の輪郭 (x_0, y_0, z_0) は媒介変数 α を用いて次のように表される。

$$x_0 = b \sin \alpha$$

$$y_0 = 0$$

$$z_0 = a \cos \alpha$$

使用者が頭部を移動・回転させたとする。移動ベクトルを (x_1, y_1, z_1) 、回転角をベクトル $\theta = (\theta_x, \theta_y, \theta_z)$ と表す。回転を各軸周りの回転の変換を表す関数、

$$R(\theta) = R_x(\theta_x)R_y(\theta_y)R_z(\theta_z)$$

を用いて表すと、マーカの楕円は以下のように表すことができる。

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R(\theta) \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \quad (4)$$

ここで、移動および回転と、マーカの大きさはカメラとの距離 D_0 に比較して十分に小さいと仮定すると、上式は以下のように近似できる。

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \simeq \begin{bmatrix} 1 & \theta_z & \theta_y \\ -\theta_z & 1 & \theta_x \\ -\theta_y & -\theta_x & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \quad (5)$$

カメラの焦点距離を f とすると、カメラ座標系内の点 (x', y', z') は投影面上の点 (X, Y) に次式で表されるように射影される。

$$X = f \frac{x'}{z'}, \quad Y = f \frac{y'}{z'}$$

カメラとマーカとの距離はマーカの大きさに比べて十分に大きいので、マーカの像はカメラの撮像面上に平行投影されたものとなる。すなわち、楕円の z' 座標を近似的に、

$$\begin{aligned} z' &\simeq -\sin \theta_0 (y_1 - Y_0) + \cos \theta_0 (z_1 + z_0) \\ &= D_0 + d_1 \end{aligned}$$

$$\text{ただし } d_1 = -\sin \theta_0 y_1 + \cos \theta_0 z_1$$

と表すことができる。ここで d_1 は (y_1, z_1) のマーカの移動の光軸方向の成分である。

いま、

$$k = \frac{f}{D_0 + d_1}$$

とおくと、

$$X = k(x_0 + \theta_z y_0 + \theta_y z_0 + x_1)$$

$$\begin{aligned} Y &= k(\cos \theta_0 (-\theta_z x_0 + y_0 + \theta_x z_0 + y_1 - Y_0) \\ &\quad + \sin \theta_0 (-\theta_y x_0 - \theta_x y_0 + z_0 + z_1 + Z_0)) \end{aligned}$$

楕円の式を代入して、かつ

$$w_1 = \cos \theta_0 y_1 + \sin \theta_0 z_1$$

$$W_0 = \cos \theta_0 Y_0 + \sin \theta_0 Z_0$$

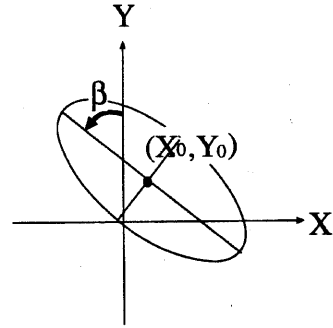


図 3: 投影面上の楕円形状

とおく。 w_1 よび W_0 はそれぞれ (y_1, z_1) および $(Y_0, -Z_0)$ が光軸と直交する方向の成分である。このとき、

$$\begin{cases} X = k(b \sin \alpha + \theta_y a \cos \alpha + x_1) \\ Y = k(-(\cos \theta_0 \theta_z + \sin \theta_0 \theta_y) b \sin \alpha \\ \quad + (\cos \theta_0 \theta_x + \sin \theta_0) a \cos \alpha + w_1 - W_0) \end{cases} \quad (6)$$

となる。

一方、 (X, Y) 投影面上で長軸 A 、短軸 B 、中心 (X_0, Y_0) 、 X 軸となす回転角 β の楕円形状が観測されたものとする(図 3)。このとき次式が成り立つ。

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix} \begin{bmatrix} B \sin \alpha \\ A \cos \alpha \end{bmatrix} + \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} \quad (7)$$

式 (6)、式 (7) を比較することにより、次の関係式が得られる。

$$B \cos \beta = kb$$

$$-A \sin \beta = k \theta_y a$$

$$B \sin \beta = -k(\cos \theta_0 \theta_z + \sin \theta_0 \theta_y) b$$

$$A \cos \beta = k(\cos \theta_0 \theta_x + \sin \theta_0) a$$

$$X_0 = kx_1$$

$$Y_0 = k(w_1 - W_0)$$

上式より頭部位置 (x_1, y_1, z_1) および頭部方向 $(\theta_x, \theta_y, \theta_z)$ の推定式を次のように求めることができる。

$$x_1 = \frac{b}{B \cos \beta} X_0$$

$$y_1 = \left(\frac{bY_0}{B \cos \beta} + W_0 \right) \cos \theta_0 - \left(\frac{bf}{B \cos \beta} - D_0 \right) \sin \theta_0$$

$$z_1 = \left(\frac{bf}{B \cos \beta} - D_0 \right) \cos \theta_0 + \left(\frac{bY_0}{B \cos \beta} + W_0 \right) \sin \theta_0$$

$$\theta_x = \left(\frac{R_0}{r_0} - \sin \theta_0 \right) \frac{1}{\cos \theta_0}$$

$$\theta_y = -\frac{R_0}{r_0} \tan \beta$$

$$\theta_z = \left(\frac{R_0}{r_0} \sin \theta_0 - 1 \right) \frac{\tan \beta}{\cos \theta_0}$$

$$\text{ただし } R_0 = \frac{A}{B}, \quad r_0 = \frac{a}{b}$$

2.2 補償関数

前節で導かれた観測モデルによる頭部姿勢の推定式は局所解析に基くものであるため、解析の中心から離れた範囲では推定値に誤差が生じる。そこで本節ではこの推定誤差を補償する補償関数を導入する。

頭部姿勢の真の値と数理モデルによる推定値との関係は比較的単純なものであるため、補償関数は低次の多項式で表現することにする。具体的には、マーカの位置の補償に双1次の多項式を、マーカの方向の補償に双2次の多項式をそれぞれ用いる。

2.2.1 マーカの位置の補償

観測モデルによって推定されたマーカの位置 (x_1, z_1) を補償関数 $g_x(x_1, z_1)$ および $g_z(x_1, z_1)$ を用いて次式のように補償する。

$$\hat{x}_1 = g_x(x_1, y_1), \quad \hat{z}_1 = g_z(x_1, y_1)$$

ここで、 (\hat{x}_1, \hat{z}_1) は補償されたマーカの位置である。補償関数としては次式に表されるような双1次の多項式を用いる。

$$g_x(x_1, y_1) = p_{00} + p_{10}x_1 + p_{01}y_1 + p_{11}x_1y_1$$

$$g_z(x_1, y_1) = q_{00} + q_{10}x_1 + q_{01}y_1 + q_{11}x_1y_1$$

係数 $p_{ij}, q_{ij} (i, j = 0, 1)$ は、あらかじめ推定位置 (x_1, z_1) と真の位置との組を補償関数を設計するためのデータとして用意し、最小2乗法によって求める。

2.2.2 マーカの回転角の補償

マーカの回転角の推定値 (θ_x, θ_y) の持つ推定誤差は使用者の頭部位置に依存して定まる。ある頭部位置において (θ_x, θ_y) を補償する補償関数を次式のように表す。

$$\hat{\theta}_x = g_{\theta_x}(\theta_x, \theta_y), \quad \hat{\theta}_y = g_{\theta_y}(\theta_x, \theta_y)$$

具体的には、補償関数として次式に示されるような双2次の多項式を用いる。

$$g_{\theta_x} = r_{00} + r_{10}\theta_x + r_{01}\theta_y +$$

$$r_{20}\theta_x^2 + r_{11}\theta_x\theta_y + r_{02}\theta_y^2 +$$

$$r_{21}\theta_x^2\theta_y + r_{12}\theta_x\theta_y^2 + r_{22}\theta_x^2\theta_y^2$$

$$g_{\theta_y} = s_{00} + s_{10}\theta_x + s_{01}\theta_y +$$

$$s_{20}\theta_x^2 + s_{11}\theta_x\theta_y + s_{02}\theta_y^2 +$$

$$s_{21}\theta_x^2\theta_y + s_{12}\theta_x\theta_y^2 + s_{22}\theta_x^2\theta_y^2$$

係数 $p_{ij}, q_{ij} (i, j = 0, 1, 2)$ の決定は位置の補償関数の場合と同様に、あらかじめ推定値 (θ_x, θ_y) と真の値との組を補償関数を設計するためのデータとして用意し、最小2乗法を用いて行う。

3 頭部姿勢推定実験

本報告で提案した頭部姿勢推定手法の評価を行うために、室内のいくつかの位置からスクリーン上の点にマーカを向け、ビデオカメラで撮影した画像データを用いて推定実験を行った(図4)。

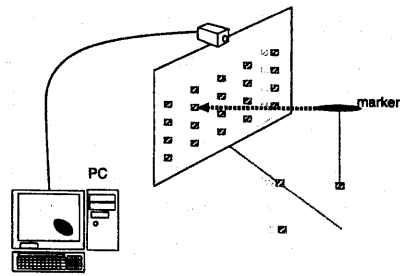


図4: 頭部姿勢推定実験

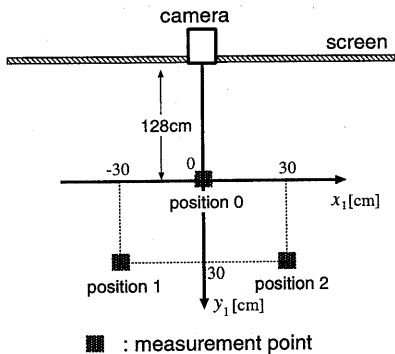


図 5: 測定点の位置

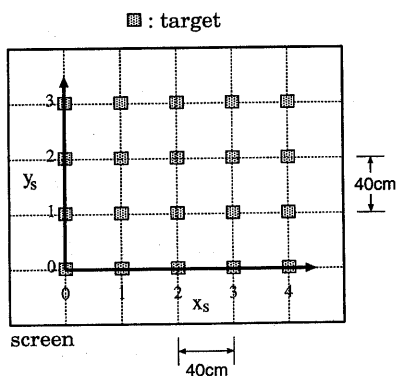


図 6: スクリーン上の視点位置

画像データの撮影条件は以下の通りである。床の上に定められた計3箇所の測定点(図5)において、マーカーの中心を床からの高さ160[cm]の位置に固定し、マーカーの長軸を図6に示すようなスクリーン上の各点に向ける。マーカーは長軸が28.1[cm]、短軸が12.5[cm]の、内部が赤色の楕円形である。スクリーンの上部に固定されたビデオカメラで撮影されたカラー画像に対して、ルックアップテーブルを用いた簡単な色フィルタで赤色を強調する処理を行い、入力データを作成した。図7に、地点0において撮影されたマーカーの画像を例として示す。

これらの画像のモーメントを求め、観測モデルを用いてマーカーの位置推定を行なった結果と、補償関数によって補償された推定位置と

(0,3)	(1,3)	(2,3)	(3,3)	(4,3)
(0,2)	(1,2)	(2,2)	(3,2)	(4,2)
(0,1)	(1,1)	(2,1)	(3,1)	(4,1)
(0,0)	(1,0)	(2,0)	(3,0)	(4,0)

図 7: マーカー画像(各座標値は図6の (x_s, y_s) に相当)

を図8に示す。図より、補償関数を用いることにより推定誤差が減少し、マーカーの真の位置 $(x_1, z_1) = (0, 0), (-30, 30), (30, 30)$ により近い推定位置が得られていることがわかる。なお、観測モデルによる推定位置のうち地点0、すなわち $(x_1, z_1) = (0, 0)$ の結果は他の地点のものと比較して推定誤差が小さい。これは観測モデルを導く際に用いた局所解析を地点0のまわりで行なったためである。

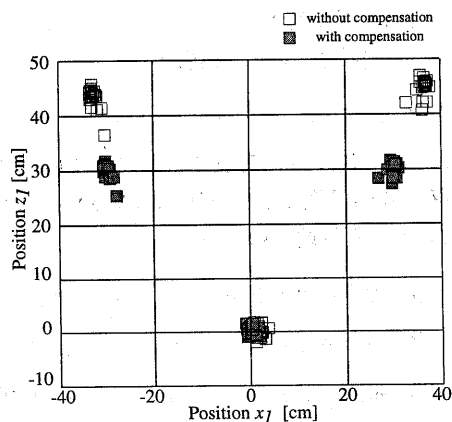


図 8: マーカー位置の推定結果(□:観測モデルによる推定位置, ■:補償関数による補償結果)

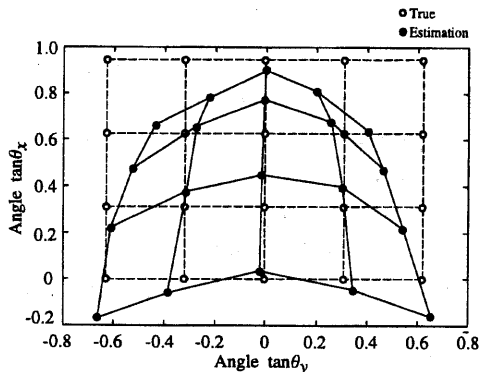


図 9: 観測モデルによるマーカー方向の推定結果 (地点 0)(○:真の値, ●:推定位置)

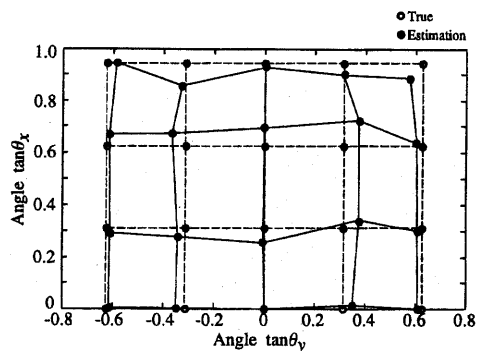


図 11: 補償関数によるマーカー方向の補償結果 (地点 0)(○:真の値, ●:推定位置)

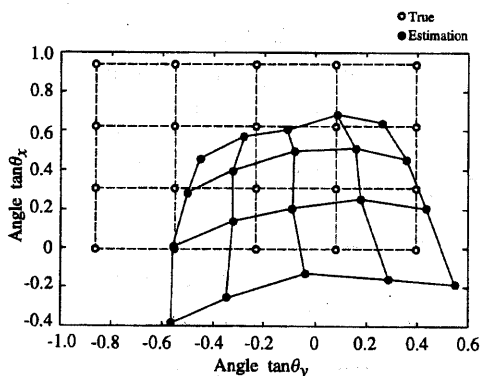


図 10: 観測モデルによるマーカー方向の推定結果 (地点 1)(○:真の値, ●:推定位置)

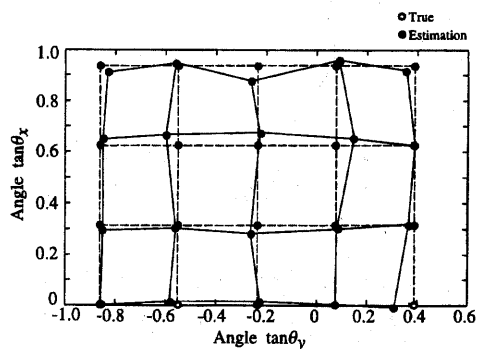


図 12: 補償関数によるマーカー方向の補償結果 (地点 1)(○:真の値, ●:推定位置)

観測モデルによる頭部の向き (θ_x, θ_y) の推定値を図 9, 図 10 に示す。地点 2 における推定結果は図 10 とほぼ鏡像の関係にあるので、ここでは省略する。図より推定された頭部方向同士を結んだ格子のトポロジーは、真値同士を結んだ格子のトポロジーを良好に保存していることがわかる。また、原点近傍の推定精度が高く、原点から離れるにしたがって誤差が大きくなっている。また、地点 0 に比較して地点 1 における推定誤差が大きくなっている。これらは、頭部位置の推定の場合と同様に、局所解析を原点まわりで行なったことが理由である。

各地点ごとに求められた補償関数を用いて推定値を補償した結果を図 11, 図 12 に示す。頭部位置の場合と同様に地点 2 についての結果は省略する。図 9, 図 10 に比較して、特に局所解析の導入に起因する誤差が増加する、原点から離れた領域での誤差が減少していることから、双 2 次の多項式による補償関数の記述が有効であることがわかる。

補正関数による推定値の補正後の、推定誤差の絶対値の平均はマーカーの位置に関して $1.01[cm]$ 、方向に関して $0.0363[rad](2.08[deg])$ であり、高い推定精度が得られた。

4 おわりに

本報告では、使用者の頭頂部に装着された楕円形のマーカをビデオカメラで撮影し、得られたマーカの像から頭部の位置と方向を推定する手法を提案した。提案手法はビジョンベースのヒューマンインタフェースの構成要素として用いられることを意図したものである。すなわち、使用者の頭部姿勢の情報から、使用者の注意の方向や、首を縦方向や横方向に振ることで形成されるゼスチュアを推定することを想定し、使用者が首を縦および横にそれぞれ振った場合の頭部の回転角(2自由度)と、頭部位置の水平面への射影(2自由度)を使用者の頭部姿勢として推定する。

提案された手法は以下のような特徴を持っている。

- 特徴量として対象の大局的な特徴を表現し、雑音に対し頑健な低次の画像モーメントを用いる。
- マーカは頭頂部に1個装着するのみでよい。ため、着脱が容易な形状にデザインすることが可能である。
- 頭部姿勢を推定する方法として、局所解析に基づく数理モデルによる推定をまず行い、近似誤差を補償関数で補償するという相補的アプローチをとる。

補正関数による推定値の補正の結果、マーカの位置および方向に対して高い推定精度が得られた。

今後論じられるべき問題の一つに、頭部の方向に関する補償関数 $g_{\theta_x}, g_{\theta_y}$ の頭部位置 (x_1, z_1) に対する依存性がある。本報告では室内の3箇所地点毎に補償関数 $g_{\theta_x}, g_{\theta_y}$ をあらかじめ求め、各々の地点における頭部方向の推定値を補償することで評価を行った。しかし、実際には使用者の頭部はさまざまな位置をとる。そこで、本報告のように離散的な地点で求められた補償関数同士を、頭部位置をパラメータとして補間

することで、任意の頭部位置 (x_1, z_1) に対する補償関数を合成することが考えられる。

また、本手法を複数台のカメラによるシステムに適用することで、より推定精度を高めると同時に頭部姿勢を推定可能な空間をより広くすることも考えられる。

今後は、これらの課題を解決しつつ、本手法に基づくヒューマンインタフェース環境を構築する予定である。

参考文献

- [1] M.Fukumoto, K.Mase and Y.Suenaga: "REAL-TIME DETECTION OF POINTING ACTIONS FOR A GLOVE-FREE INTERFACE", Proc. IAPR Workshop on Machine Vision Applications, pp. 473-476 (1992)
- [2] J.J.Kuch and T.S.Huang: "Virtual Gun A Vision Based Human Computer Interface Using the Human Hand", Proc. IAPR Workshop on Machine Vision Applications, pp. 196-199 (1994)
- [3] 天谷賢治, トム カルバート, クリステイ マッケンジー: "小数マーカによる手と指姿勢の認識法", 信学技報, PRU95-186, pp.63-69 (1995)
- [4] M-K.Hu: "Visual pattern recognition by moment invariants", IRE Trans. on Information Theory, IT-8, 2, pp. 179-187 (1962).
- [5] 辺見一男, 井上和夫: "ディスプレイ上の像を把握操作できる3次元インタフェース — マーカ追跡による方法 —", 信学論, Vol.J74-DII, No.9, pp. 1264-1272
- [6] 高松 亮, 佐藤 誠, 奈良 裕, 河原田 弘: "局所モーメントによる運動物体の追跡に関する基礎的研究", 信学論, Vol.J79-A, No.2, pp. 399-407 (1996).