

複数カメラの観測統合による広範囲な顔・視線計測

Face and gaze tracking system for wider range scene using integrated multi-camera observation

浅野 慧† 末永 剛† 竹村 憲太郎† 松本 吉央† 高松 淳† 小笠原 司†

Satoshi Asano, Tsuyoshi Suenaga, Kentaro Takemura, Jun Takamatsu, Yoshio Matsumoto, Tsukasa Ogasawara

1. はじめに

顔向きや視線を計測するシステムは、従来より心理学的実験や知的インターフェース分野での応用を目指して研究されてきた。また、手軽に計測を実現する手法も研究されており、近年においては非接触かつ安価な可視光カメラを用いたシステムが多数実現されており、^{[1][2]}我々も単眼カメラを用いた3次元顔・視線計測システムの開発を行ってきた^[3]。しかし、従来のシステムでは、計測範囲がカメラ1台の視野角に限定されていた。このため、計測範囲が机やモニタの前に制限され、ユーザの大きな動きへの対応や、広い空間での利用が困難であるという問題が存在していた。

本研究では、この問題を解決する手法として、複数のカメラから得られた特徴点の情報を統合し、頭部の3次元位置・姿勢と視線ベクトルを計測することで、広範囲の領域を計測する手法を提案し、その実装と評価実験の結果について述べる。

2. 複数カメラ観測統合と動き推定

単眼顔情報計測システム^[3]は、(1)詳細な3次元モデルを作成する顔モデル生成、(2)微分画像を用いた顔発見、(3)モデルの単位動きベクトル場と観測の実動きベクトル場とを比較し動き推定を行う6自由度動き推定という順序で、顔の位置推定を実現している。提案手法では、このうち6自由度動き推定において、複数カメラから得られた特徴点の動きベクトルを使用する。

2.1 単位動きベクトル場の生成

3次元顔モデルを3次元空間中で6軸方向に微小に動かし、それぞれの動きを2次元平面に投影する。これにより、顔が各軸方向に単位量動いた際に、特徴点を取りうる2次元的な動きを単位動きベクトル場として表現することができる。ここでは単位動きベクトル場として、 n 個の特徴点について m 台のカメラに対応する特徴点を記述する。X, Y, Z軸方向の並進に関する動きとしてベクトル場

$$\begin{cases} \mathbf{x}_t = (u_{xt1c1}, v_{xt1c1}, u_{xt2c1}, v_{xt2c1}, \dots, v_{xtncm})^T \\ \mathbf{y}_t = (u_{yt1c1}, v_{yt1c1}, u_{yt2c1}, v_{yt2c1}, \dots, v_{ytncm})^T \\ \mathbf{z}_t = (u_{zt1c1}, v_{zt1c1}, u_{zt2c1}, v_{zt2c1}, \dots, v_{ztncm})^T \end{cases}$$

を生成する。ここで u_{xt1c1} は、顔モデルがX軸方向へ並進した際、1番目のカメラで撮影した画像上での特徴点のX軸方向変位を表す。また、以下のようにX, Y, Z軸まわ

りの回転に関する3つの単位動きベクトル場を生成する。

$$\begin{cases} \mathbf{x}_r = (u_{xr1c1}, v_{xr1c1}, u_{xr2c1}, v_{xr2c1}, \dots, v_{xrnem})^T \\ \mathbf{y}_r = (u_{yr1c1}, v_{yr1c1}, u_{yr2c1}, v_{yr2c1}, \dots, v_{yrenem})^T \\ \mathbf{z}_r = (u_{zr1c1}, v_{zr1c1}, u_{zr2c1}, v_{zr2c1}, \dots, v_{zrenem})^T \end{cases}$$

実動きベクトル場として、 n 個ある各特徴点について、 m 台のカメラからオプティカルフロー \mathbf{f} を求める。

$$\mathbf{f} = (u_{f1c1}, v_{f1c1}, u_{f2c1}, v_{f2c1}, \dots, v_{fnem})^T$$

2.2 単位動きベクトル場と実動きベクトル場の比較
3次元的な動きの推定には、3次元顔モデルを2次元平面に投影した画像と、入力画像におけるそれぞれの特徴点の配置が近づくように、3次元顔モデルをフィッティングする必要がある。任意の微小な顔の動きによる実動きベクトル場は、6つの単位動きベクトル場の線形和で表現できる。

$$\mathbf{f} = a_1\mathbf{x}_t + a_2\mathbf{y}_t + a_3\mathbf{z}_t + a_4\mathbf{x}_r + a_5\mathbf{y}_r + a_6\mathbf{z}_r$$

このときの各単位ベクトル場の係数 a_1, \dots, a_6 が各自由度の移動量に対応する。これを展開することで、特徴点数 $n \times$ カメラ数 $m \times 2$ 個の連立一次方程式を得る。 $n \times m > 3$ の際に連立方程式は冗長になるため、最小二乗法を用いて各係数を求める。この際には各特徴点の相関値を重みとして考慮した重み付き最小二乗法を用いる。

$$\mathbf{V} = (\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t, \mathbf{x}_r, \mathbf{y}_r, \mathbf{z}_r), \mathbf{a} = (a_1, a_2, \dots, a_6)^T$$

とおくと

$$\mathbf{f} = \mathbf{V}\mathbf{a}$$

となり、それぞれの特徴点の相関値行列 \mathbf{W} を加えた重み付き擬逆行列 $\mathbf{V}^+ = (\mathbf{V}^T\mathbf{W}^{-1}\mathbf{V})^{-1}\mathbf{V}^T\mathbf{W}^{-1}$ を用いて

$$\mathbf{a} = \mathbf{V}^+\mathbf{f}$$

となり、ベクトル \mathbf{a} を用いて6自由度の移動量が求まる。

2.3 実装

提案手法の実装を図1に示す。本システムの特徴としては、個々の特徴点トラッキング等、カメラ側の処理の並列化が可能なる点、そして計測を統合する際に画像ではなく特徴点の実動きベクトルのみを転送する点が挙げられる。これにより、カメラ数の大幅な増加にも十分に対応できるシステムであると言える。また、複数カメラの計測可能な特徴点全てを6自由度動き推定に利用できるため、特徴点が複数のカメラに写り、実動きベクトルが冗長になる場合には精度向上が期待できる。

†: 奈良先端科学技術大学院大学, Nara Institute of Science and Technology

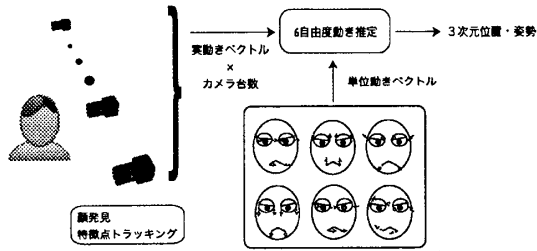
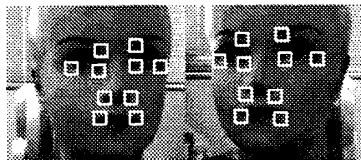
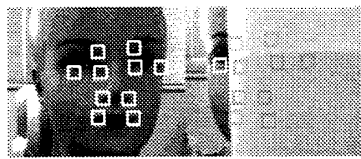


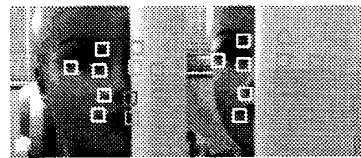
図1 提案手法の実装



(a) 通常状態



(b) 片カメラへのオクルージョン



(c) 両カメラへのオクルージョン

図2 2台カメラによる計測

3. 精度評価実験

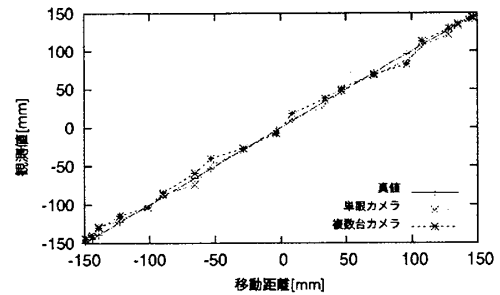
3.1 実験の概要

本論文で提案した複数台カメラを用いた顔の動き推定手法の評価を行うため、精度評価実験を行った。定量評価を行うため、汎用ロボットアームの手先にマネキンの頭部を装着し、位置・姿勢を各軸方向に変化させ、顔の動きの推定を行った。また、特徴点が冗長になった場合の精度向上を確認するため、従来の単眼カメラを用いた顔計測システムとの比較も行った。

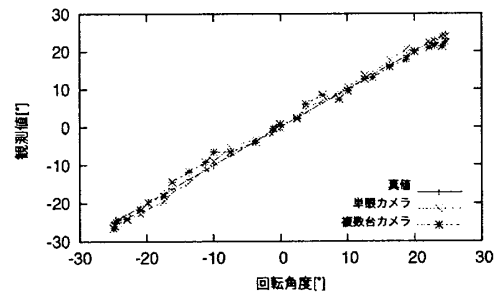
3.2 実験結果

図2は、2個のカメラを用いてリアルタイムに顔の動きを計測した結果である。図2(a)から従来同様に顔の向き、および視線ベクトルが抽出できている様子が確認できる。図中の矩形は各特徴領域を表し、相関値が70%以上の場合は白の矩形、それ以下の場合は相関値の低下に伴い、濃い矩形で貢献度を示している。(b)(c)においては、特徴点に対する重みを考慮し、相関値の下がった特徴点の貢献度を下げることによって、遮蔽に対してもロバストにトラッキングを行っている様子が確認できる。

従来の単眼カメラと提案手法の計測結果として、X軸



(a) X軸方向の並進



(b) X軸方向の回転

図3 実験結果

方向への並進および回転の計測結果を図3に示す。また全軸の計測誤差は、単眼の場合には平均7.86mm, 3.72°, 提案手法の場合には平均5.35mm, 3.48°であった。これらの比較結果より、提案手法においても従来の手法と同等の精度が得られたことがわかる。

4. おわりに

本論文では、複数台のカメラを用いて、特徴点ベクトル場を生成し、モデルと比較することで、より広範囲を計測可能にする手法を提案した。また、システムの実装および精度評価実験を行い、この有効性を確認した。今後は3台以上のカメラを用いた実験を行い、視野範囲の拡大の有効性を確認していく予定である。

参考文献

- [1] F. Davoine, F. Dornaika: "Head and Facial Animation Tracking Using Appearance-Adaptive Models and Particle Filters. Real-Time Vision for Human Computer Interaction", (G.Kisachanin, V.Pavlovic and T.S. Huang, eds.), Springer Verlag, chapter8, pp.121-140, 2005.
- [2] H. Yamazoe, A. Utsumi, T. Yonezawa, S. Abe: "Remote gaze estimation with a single camera based on facial-feature tracking without special calibration action", Proceedings of the 2008 symposium on Eye Tracking Research & Application Symposium, pp.245-250, 2008.
- [3] 松本 吉央, 佐々尾 直樹, 小笠原 司: "単眼カメラによる顔トラッキング-3次元顔形状に基づく6自由度動き推定", 第6回計測自動制御学会システムインテグレーション部門講演会, 3I4-1, 2005.