

Real-time Head Pose Estimation with Stereo Vision

石井, 繁範
九州大学システム情報科学研究院知能システム学部門

高木, 和也
九州大学システム情報科学研究院知能システム学部門

有田, 大作
九州大学システム情報科学研究院知能システム学部門

谷口, 倫一郎
九州大学システム情報科学研究院知能システム学部門

<http://hdl.handle.net/2324/5921>

出版情報：情報科学技術フォーラム, FIT 2003, 2003-09. Information Processing Society of Japan
バージョン：

権利関係：ここに掲載した著作物の利用に関する注意 本著作物の著作権は（社）情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。



I-077

ステレオカメラを用いた実時間頭部姿勢推定 Real-time Head Pose Estimation With Stereo Vision

石井 繁範^{†‡}
Shigenori Ishii

高木 和也[†]
Kazuya Takaki

有田 大作[†]
Daisaku Arita

谷口 倫一郎[†]
Rin-ichiro Taniguchi

1. はじめに

顔の向きは人の意図を反映している場合が多く、人と計算機間のインタラクションといった様々なアプリケーションや人間の行動理解にとって実時間での頭部姿勢推定は重要な技術である。そこで本稿では、利用者に負担を与えず、多くの人が一のシステムを利用できるように、円形分離度フィルタとテンプレートマッチングを用いて顔の特徴点検出を行い、ステレオ視により頭部姿勢推定を行う手法について提案する。

2. ステレオカメラを用いた頭部姿勢推定法

ステレオカメラを用いて頭部の姿勢を推定する為には、2つのカメラで同じ特徴点を複数個検出する必要がある。頭部姿勢推定にとって、ロバストで精度の高い顔の特徴点検出は不可欠である。瞳や鼻孔が顔の他の部分に比べて円形であり暗いという特徴を持つために検出が比較的容易である点と、口端が被験者に寄らず比較的安定に検出できる点により、本研究では検出する顔の特徴点を両瞳・鼻孔・両口端として、分離度とテンプレートマッチングを用いて顔の特徴点検出を行い、ステレオ視によりそれぞれの顔の特徴点の3次元位置を計算することで頭部姿勢推定を行う。

大まかな処理の流れは図1の通りであり、主な画像処理は左のカメラから得られる画像で行う。

2.1 顔領域の抽出

まず、図1における左画像において、背景差分により背景以外の領域を抽出する。その領域内において人の肌の色情報を用いて顔領域を抽出する。今回は、各画素の色相値を計算し、あらかじめ定義しておいた肌色の領域を顔領域とした。

2.2 口端位置の検出

あらかじめ用意しておいた口端のテンプレート画像を用いて、求められた顔領域の中で正規化相関によるテンプレートマッチングを行い、左右の口端位置を検出する。

2.3 瞳、鼻孔の特徴候補点抽出

瞳や鼻孔が他の顔の部分に比べ暗いかつ円形であるという特徴を用いて、瞳や鼻孔の特徴候補点を得るために、まず顔領域の中から暗領域を抽出する。

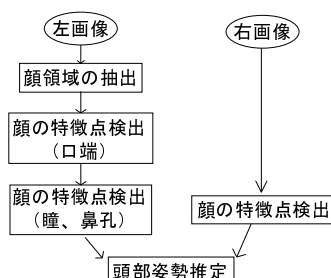


図1: 処理の流れ

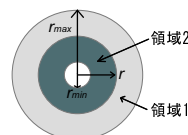


図2: 円形分離度フィルタ

次に、各暗領域の重心位置を計算し、その重心付近において、図2に示される2つの同心円領域からなる円形分離度フィルタを用いて分離度を計算する(式(1))[1]。これにより各暗領域の円形らしさが得られる。分離度を計算した結果がある閾値よりも大きい場合に、その場所を顔の特徴候補点とした。分離度 η は以下の式を用いて計算を行う。各画素の持つ分離度 η は、 r を $[r_{min}, r_{max}]$ の範囲で変化させた中の最大の値とし、各暗領域の分離度は、その重心位置付近の各画素の分離度の中で最大の値とする。

$$\eta = \frac{\sigma_b^2}{\sigma_T^2} \quad (1)$$

$$\sigma_b^2 = n_1(\overline{P_1} - \overline{P_m})^2 + n_2(\overline{P_2} - \overline{P_m})^2 \quad (2)$$

$$\sigma_T^2 = \sum_{i=1}^N (P_i - \overline{P_m})^2 \quad (3)$$

ここで n_1 、 n_2 、 N はそれぞれ領域1、領域2、領域全体(n_1+n_2)の画素数、 σ_T は領域全体の全分散値、 σ_b は領域1、2の級間分散値、 P_i は位置 i の輝度値、 $\overline{P_1}$ 、 $\overline{P_2}$ 、 $\overline{P_m}$ はそれぞれ領域1、領域2、領域全体の平均輝度値である。

2.4 顔の特徴点の検出

まず左画像において、2.3節で得られた特徴候補点を次の3つの顔モデルに当てはめ、口端の位置や瞳の間の長さ、瞳と鼻孔の間の位置関係といった顔の幾何学的条件[2]を満たす特徴候補点の組み合わせの中で、分離度の総和が最も大きいものを瞳や鼻孔位置としている。分離度の総和を求める際に、前フレームでの特徴点位置から一定の範囲内にある特徴候補点に対して、その分離度に重みをつけることで、安定して特徴点位置を定める事が出来る。

モデル1 鼻孔と両瞳の位置関係を示す正面の顔モデル

モデル2 両瞳の位置関係を示すうつむき顔のモデル

モデル3 鼻孔と瞳の位置関係を示す横顔のモデル

なお、顔の幾何学的条件で瞳と鼻孔を検出できなかった場合、前フレームの特徴点画像との正規化相関によるテンプレートマッチングを行い瞳と鼻孔を検出する。

次に、左画像における顔の特徴点付近の画像をテンプレート画像として用意し、右画像において正規化相関によるテンプレートマッチングを行ない、右画像における瞳と鼻孔および口端の位置を検出する。

以上のような手法を用いて左右画像において顔の特徴点検出を行う。

[†]九州大学大学院システム情報科学府

[‡]現在、日本電気株式会社



(a) 左画像上の特徴点 (b) 右画像上の特徴点
図 3: 特徴点検出実験結果

2.5 頭部姿勢推定

これまで得られた左右画像における顔の各特徴点の 2 次元位置をもとに、ステレオ計算により各特徴点の 3 次元位置を計算する。3 次元位置が得られた顔の特徴点の中から 3 点を選び、その 3 点を通る 3 次元平面を求め、この平面の法線ベクトルを用いて頭部姿勢を推定する。本研究では、基準ベクトルをユーザが正面を向いているときの平面の法線ベクトルと定義し、この基準ベクトルと現フレームにおける平面の法線ベクトルとを用いて求めた、水平方向および垂直方向の顔の角度の変化量で頭部姿勢を表現する。

この比較を行う際、3 次元平面を求めた特徴点 3 点の組み合わせによって、得られる結果が異なる。そこで基準ベクトルを、あらかじめ特徴点 3 点の全ての組み合わせで求めておき、頭部姿勢の推定を行う際に、用いる特徴点の組み合わせによって基準ベクトルを変えることにより、特徴点の組み合わせに関わらずに頭部の姿勢を推定する事が出来るようにする。

3. 実験

本実験では Celeron1.3GHz の PC および 2 台の IEEE1394 カメラ (Sony DFW-V500) を用いた。その際カメラのフレームレートは 30fps とした。

3.1 顔の特徴点検出の実験

精度の良い顔の向き検出を行うためには、安定した顔の特徴点検出が必須である。そこで顔の特徴点検出の実験を行った。図 3 に左右画像における顔の特徴点の検出結果を示す。図 3 における十字は瞳、鼻孔といった各特徴点の検出位置を表す。ここで図 3 のように瞳や鼻孔上に十字が位置したとき、顔の特徴点を正しく検出できたとすると、顔の特徴点検出の精度は約 96%であった。

3.2 頭部姿勢推定実験

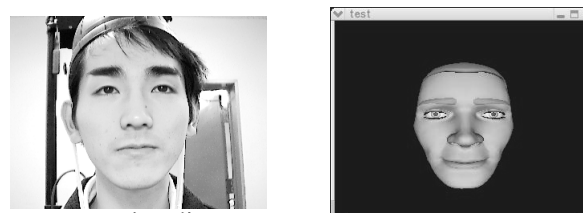
提案手法による頭部姿勢推定精度を調べる為に、頭部に装着した超音波センサから得られる結果との比較を行った。表 1 にそれらの比較結果をまとめる。

まず、センサからの測定結果と提案した手法による推定結果との誤差の最大値は、水平方向・垂直方向共に大きくなっているが、この理由として、一般に人が顔を動かす際には顔よりも先に瞳を動かしてしまい、実際の顔の動きよりも瞳の動きが大きくなってしまふ点が考えられる。しかし、それらの誤差の平均値は、水平方向で 2 度、垂直方向で 4 度であるという結果が得られたことから、提案手法で比較的精度よく頭部姿勢推定が行われているといえる。さらに、頭部姿勢推定に要した 1 フレーム当たりの平均時間は約 35 ミリ秒であったことから、処理時間は十分早く、実時間頭部姿勢推定を行う事が出来るといえる。

次に、提案手法によって推定された頭部姿勢を 3D モデルで再現したものを図 4~6 に示す。各図の (a) はカ

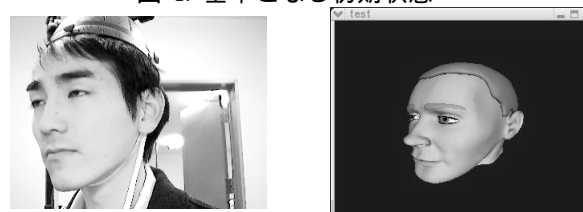
表 1: 本手法の頭部姿勢推定誤差

顔の回転方向	水平方向	垂直方向
最大誤差 (°)	14	8
平均誤差 (°)	4	2



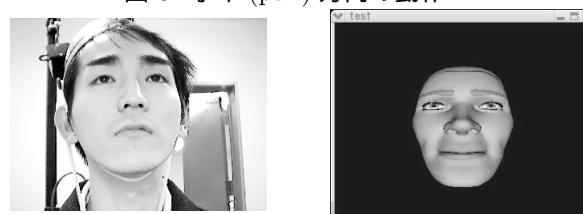
(a) 左画像 (b) 再現画像

図 4: 基準となる初期状態



(a) 左画像 (b) 再現画像

図 5: 水平 (pan) 方向の動作



(a) 左画像 (b) 再現画像

図 6: 垂直 (tilt) 方向の動作

メラから得られた左画像、(b) はこのときに推定された頭部姿勢を 3D モデルで再現したものである。図 4 は基準となる頭部姿勢の初期画像であり、図 5 は水平 (pan) 方向に、図 6 は垂直 (tilt) 方向に頭部を動かした場合の様子である。これらの図からも、被験者の頭部姿勢に対して比較的正確な頭部姿勢推定が出来ているといえる。

4. まとめ

本稿では、ステレオカメラを用いた顔の向き検出手法について提案し、提案した手法を用いて顔の向き検出実験を行ない、ほぼ正確に顔の向きが検出できていることを示した。今後は、照明条件の変化やオクルージョンに対してよりロバストな顔の向き検出手法について検討を行う必要がある。

参考文献

- [1] 福井和宏, 山口修, “特徴形状とパターン照合の組み合わせによる顔特徴点抽出”, 電子情報通信学会論文誌 D-II, vol.J80-DII, no.8, pp.2170-2177, 1997.
- [2] Y.ARAKI, N.SHIMADA and Y.SHIRAI. “Face Detection and Face Direction Estimation Using Color and Shape Features.” International Conference on Machine Automation, pp.351-356, 2000.