

ライティングスイッチフォトメトリー法を用いた顔の向きの判定

4 K-9

日置尋久, 佐治斉, 國井利泰

東京大学

1 はじめに

人の顔識別能力・表情認識能力を理解して計算機で実現しようとする研究が近年盛んに行なわれているが、まず問題となるのは、どのような方法で画像を入力するかということである。表情を判断するには、従来の表情研究で利用されている顔の主なパーツの形状はもちろん、額・頬などのしわの形状も重要な情報であり、空間的に密な情報が求められる。また人の表情の変化に追従できる程度に時間的に密な情報を得る必要がある。以上の点を考慮して、佐治らは、ライティングスイッチフォトメトリー(LSP)法という画像入力法を提案した[3]。LSP法は、光学的ステレオ法[4]に時間の概念を採用入れ、物体の動き、及び形状の変化を局所的な法線の変化としてとらえようとするものである。LSP法では、光学的ステレオ法と同様に単一カメラ・3光源のシステムにおいて、光の当たる方向を順次高速に切り替えて画像を撮る。微小時間での法線方向の変化が時間に比例すると仮定することにより、各時刻での法線場の再構成を可能にした。ここでは、表情解析をより安定に行なえるようにするため、顔の向きを自動的に判定するアルゴリズムを実現する。

2 ライティングスイッチフォトメトリー法

問題を簡単にするために、まずいくつかの仮定を設ける。1. 投影法 — 平行投影を仮定する。2. 光源 — 平行光を仮定する。これによって、顔面上の全ての点で入射角が同じであると見なすことができる。また、各光源は常に同じ強さの光を発生し、 Δt で順次切り替えられるものとする。3. 反射特性 — 顔面は完全拡散面であると仮定する。この仮定により、反射光 L を、入射光の強さ I 、反射係数 R 、入射角 i により、 $L = IR \cos i$ と表すことができる。

以上の仮定の元で、ある時刻 t (光源 S_2 に照らされているものとする)における、ある点での局所的な法線 N_t (単位ベクトル)は次のように求められる。

$$\frac{1}{3}(2L_{1,t-\Delta t} + L_{1,t+2\Delta t}) = I_{S_1} R_t N_t \cdot S_1$$

$$L_{2,t} = I_{S_2} R_t N_t \cdot S_2$$

$$\frac{1}{3}(L_{3,t-2\Delta t} + 2L_{3,t+\Delta t}) = I_{S_3} R_t N_t \cdot S_3$$

但し、 $S_k, k = 1, 2, 3$ は単位光源ベクトルを表す。

Detecting the Viewing Direction of a Face Using Lighting Switch Photometry

Hirohisa Hioki, Hitoshi Saji, and Toshiyasu L. Kunii
The University of Tokyo

3 顔面領域の抽出

各時刻の2次元顔面画像から、後の処理の対象となる顔面領域を抽出する。画像の大半を被写体が占めているものと仮定して、1. ノイズ除去、2. 閾値処理、3. 影領域の除去、4. 領域ラベリング、の4段階の処理で顔面領域のみを抽出する。

まずメディアンフィルタリングによって、ノイズの影響を軽減する。次に輝度値のグラディエントが大きい領域(境界領域)において、輝度値のヒストグラムを作る。こうして得られたヒストグラムは2つの山の間に谷があるような形になるので、判別閾値選定法[2]により、背景と被写体とを区別する適切な閾値を得ることができる。LSP法では、3つの光源の全てに照らされている領域においてのみ法線場の計算が有効であるので、画像の中で影になっている部分を除去する。1シーンに対応する画像を撮る間の被写体の動きを無視できるものとするれば、閾値処理が既に施されているため、影領域の除去は単に画像の論理積をとることによって行なわれる。最後に画像中のそれぞれの領域に固有のラベルをつけることにより、各領域の大きさを調べる。最初の仮定から、単に最大の領域をとれば、顔面領域を選び出すことができる。

4 顔の向きの判定

与えられた法線場から顔の向きを判定するのに、1. 形状的に安定であり、2. 顔の向きを的確に表す、という条件を満たす鼻領域を抽出し、そこでの法線の平均をとる、という方法を用いる。鼻領域を抽出するのに、顔の位置・向きに依存しないガウス曲率・平均曲率を法線場から計算する[1]。顔面領域の形状をガウス曲率(K)・平均曲率(H)によって、次のように分類する。

	$H < -\epsilon$	$ H \leq \epsilon$	$H > \epsilon$
$K > \epsilon$	凸領域	平面領域	凹領域
$ K \leq \epsilon$	平面領域	平面領域	平面領域
$K < -\epsilon$	鞍部	平面領域	鞍部

このような基準で顔面領域を分割した後、凸領域の中で顔面領域の7%以上20%以下の面積を占め、領域の重心が顔面領域の重心に最も近いものを鼻領域とする。ここで鼻領域の法線ベクトルを平均して、顔の向きとする。また鼻領域において、領域の各点の座標 (x_i, y_i) と重心の座標 (G_x, G_y) とから、共分散行列

$$M = \sum_i (x_i - G_x, y_i - G_y)(x_i - G_x, y_i - G_y)$$

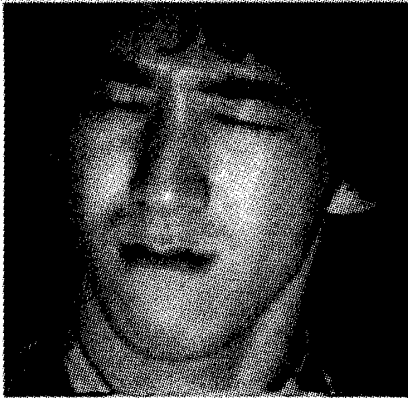


図 1: 顔面画像

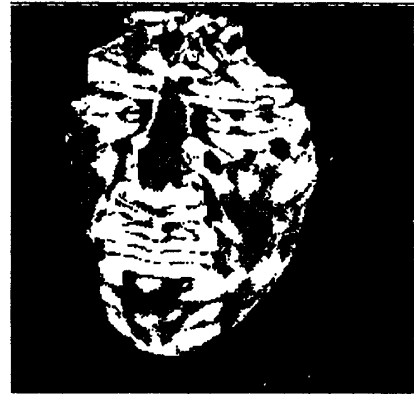


図 3: 領域分割



図 2: 法線場



図 4: 鼻領域の抽出

を計算して、固有ベクトルを算出し、この固有ベクトルと顔の方向から、鼻筋の方向とそれと直交する方向を決定する。

5 実験

図 1 の画像から顔面領域を抽出し、LSP 法により法線場を計算して、図 2 を得る。この法線場からガウス曲率・平均曲率を計算して、形状によって分割したのが、図 3 である。分割された領域から鼻領域を抽出し、鼻筋の方向とそれに直交する方向を示したのが図 4 である。この例では、顔の方向は以下のように計算された(但し座標系は、カメラの向いている方向を Z 軸の負の向きとして、右手系を構築するものとする)。

$$\begin{aligned} X: & (0.88, -0.05, 0.47) \\ Y: & (0.11, 0.99, -0.10) \\ Z: & (-0.46, 0.14, 0.87) \end{aligned}$$

6 結論

ここでは、時間変化する法線場をとらえる LSP 法を用いて、顔の向きを判定を行なった。今後は、この結果を利用して顔の向きに関する補正を行ない、位相幾何学的な特徴量を用いた表情解析を行なう予定である。なお、本研究は日産自動車(株)の協力のもとで行なわれたものである。

参考文献

- [1] Besl, P. J. and R. C. Jain, "Invariant Surface Characteristics for 3D Object Recognition in Range Images," *Computer Vision, Graphics and Image Processing*, vol. 33, 1986, pp. 33-80.
- [2] Otsu, N., "An Automatic Threshold Selection Method Based on Discriminant and Least Squares Criteria," *The Transaction of the Institute of Electronics and Communication Engineers of Japan*, vol. J 63-D, no. 4, 1980, pp. 349-356.
- [3] Saji, H., H. Hioki, Y. Shinagawa, K. Yoshida, and T. L. Kunii, "Extraction of 3D Shapes from the Moving Human Face Using Lighting Switch Photometry," in *Creating and Animating the Virtual World* (N. M. Thalmann and D. Thalmann, eds.), pp. 69-86, 1992.
- [4] Woodham, R. J., "Photometric Method for Determining Surface Orientation from Multiple Images," *Optical Engineering*, vol. 19, no. 1, 1980, pp. 139-144.