

1 はじめに

近年、カメラを用いた移動物体のリアルタイム追跡監視システムの開発が盛んに行われている。このようなシステムでは、移動物体を追跡すると同時に、カメラから得られた画像から、その物体の特定の部分、例えば動いている人間の顔など検出し、登録する、といった機能も重要である。特定の部分を検出できれば、その登録情報を用いて追跡を行うことも可能となる。

人間の顔領域を検出する従来手法としては、人間の肌色情報を用いて検出する手法や、顔のパターン情報を用いて検出する手法などが提案されている [1]。しかし、これらの手法では、検出に用いる情報をあらかじめシステムに入力しておく必要があるため、自然環境下での検出を行うことは困難である。また、入力した情報に適合する人間、例えば色情報によって検出を行う場合、特定の肌の色の人間のみ検出が可能となり、検出結果が入力された情報によって大きく左右されることが考えられる。

今回提案する手法は、人間の頭部を楕円体と仮定して、動画像中で移動している楕円体を検出することにより、顔領域を検出する手法である。

動画像中の移動物体を検出する従来手法として、あらかじめ取得された背景画像と差分をとる背景差分法 [2]、動画像中の隣接したフレーム間の差分をとるフレーム間差分法、画像全体でオプティカルフローを抽出し、空間的にまとまりをもって分布するフローを統合して移動物体を切り出すオプティカルフロー法 [3]、移動物体を含む領域をテンプレートとして登録し、以後毎回テンプレートと最も類似度の高い領域を検出するテンプレートマッチング法 [4] 等が提案されている。背景差分法では移動物体が動画像中で静止した場合でも、その物体を検出することが可能である、という利点がある。しかし、天候、時刻、照明などの環境変化に対して影響を受けやすく、背景画像を常に適切なものに更新しなければならない。オプティカルフロー法は、移動物体の移動量を検出することができるが、フローが必ずしも正確には抽出できるとは限らず、また計算量が非常に大きくなる。また、テンプレートマッチング法は、検出対象の見え方が変化すると、検出対象とテンプレートの設定位置との間にずれが生じ、検出対象を見失う場合がある。

本研究で適用したフレーム間差分法は、動画像中の

隣接するフレーム間で差分をとり、移動物体を検出する手法である。この手法は、隣接するフレームの対応する画素の差分をとり、その差分値が任意に設定した閾値を超えたならば、その画素領域を物体の移動が行われた領域 (以下、移動領域) と認識する。フレーム間差分は背景差分とは違い、時間的に微小な変化しか起こらない画像同士を比較するため、天候、時刻、照明のような環境要因を考慮しなくてもよい利点がある。また、フレーム間差分によって得られる情報を画像化したものをフレーム間差分画像 (以下、簡単に差分画像) という。本来、フレーム間差分は移動領域のみを検出するが、検出対象となる物体の既知である情報 (今回は楕円体) も積極的に用いることにより、よりロバストな検出が期待できる。

ただし、単一のカメラのみでは、人間の移動方向によっては楕円体の動きを検出できない場合が考えられる。この問題を解決する方法としては、複数のカメラから得られるアスペクト、すなわち見え方の異なる複数の差分画像を検出に用いる方法が考えられる。また、提案手法が投票アルゴリズムを基本としているため、この点においても複数のカメラを用いることによって、よりロバストな検出が期待できる。

本稿では、まず提案した顔領域検出手法の原理説明を行い、次にシミュレーション実験の結果について報告する。

2 提案手法

2.1 差分画像を用いた顔領域の検出

人間の頭部は、約 1 対 1 対 1.3 の楕円体と近似できることが知られている [5]。人間が動いているフレーム間で差分をとった場合、その差分画像には楕円体が移動したことによって生じた移動領域が存在する。この移動領域を検出することによって、人間の顔領域を検出することが可能となる。

3 次元空間中にある大きさをもつ楕円体 E_1 を配置し、 E_1 に対して X-Y 平面方向に、距離 t だけ平行移動した楕円体 E_2 を配置する。ここで、以下の式 (1) によって表される領域 A は、楕円体が E_1 の位置から E_2 の位置に移動したときの時間微分領域である (以下、簡単に微分楕円体とする)。図 1 に、その様子を示す。

$$A = E_1 \cup E_2 - E_1 \cap E_2 \quad (1)$$

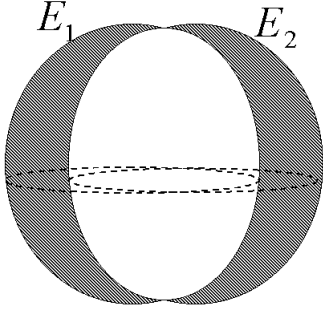


図 1: 微分楕円体

この微分楕円体 A を検出対象の 3 次元空間 (今回では対象物体の候補領域を包含する直方体) に等間隔に配置する。このときの微分楕円体の集合を B (ただし $B \ni A$) とする。

次に、カメラから得られる差分画像 D 中の、ある移動領域 p を 3 次元空間に逆投影する。さらに、逆投影した移動領域 p の射影直線 l_p と、微分楕円体 A との距離 d_p を計算する。図 2 にその様子を示す。

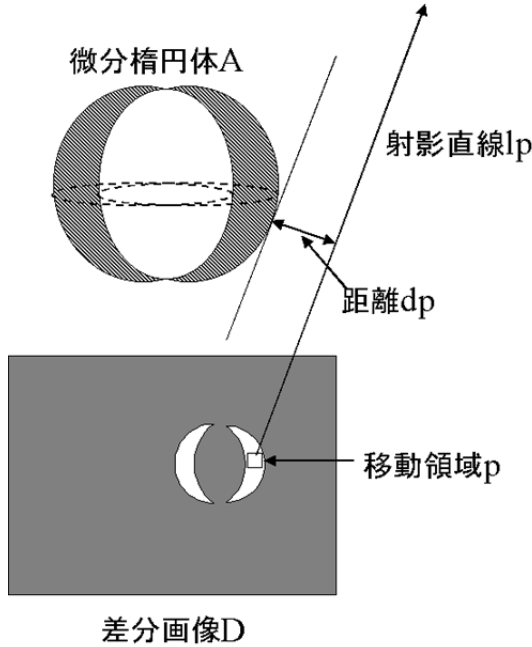


図 2: 微分楕円体と射影直線の距離

ここで、 l_p と微分楕円体 A の元となる楕円体 E_1, E_2

との最短距離をそれぞれ d_1, d_2 とすると、 d_p は次の条件で求められる。

$d_p = 0$: l_p が E_1 と交わる かつ l_p が E_2 と交わらない。または
 l_p が E_2 と交わる かつ l_p が E_1 と交わらない。または
 l_p が E_1 または E_2 と接する

$d_p = \min(d_1, d_2)$: それ以外の場合

D 中の全ての移動領域に対して、 A との距離を求め、その総和を差分画像 D と微分楕円体 A との距離 d_A と定義する。 D 中の全ての移動領域の集合を P (ただし $P \subset D, p \in P$) とすると、 d_A は次式 (2) によって得られる。

$$d_A = \sum_{p \in P} d_p \quad (2)$$

3 次元空間に複数配置された微分楕円体それぞれに対し距離 d_A を求め、その中で d_A が最小となる微分楕円体 A を検出結果とする。

$$d_{\min} = \min_{B \in A} d_A \quad (3)$$

式 (3) で、 $d_{\min}$ となる微分楕円体 A が、検出結果となる。

ここで、差分画像中の移動領域は、物体がある方向に移動したことによって生じるが、差分画像からは、どの方向に移動しているかを検出することはできない。そのため、微分楕円体が配置される位置 1 つに対して、X-Y 平面上水平に N 方向の移動を考慮した微分楕円体を配置する必要がある。ただし、本来全方向の移動を考慮する必要があるが、差分画像からは正負どちらの方向に移動しているかを検出できない。そこで、移動方向は全方向の半分、すなわち $N/2$ を考慮すればよいことになる。

2.2 高速化手法

上述の手法では、射影直線と微分楕円体との距離を求める計算が膨大となると同時に、ノイズの多い差分画像からこのような厳密な処理を行うことは現実的で

はない。そこで、以下の手法により、簡略かつ高速化を行う。

- (1) 楕円体を座標変換により球体に変換する。
- (2) 透視投影を平行投影とする。
- (3) 3次元空間に配置する微分楕円体の個数を減らす。
ただし、候補の削減により、3次元空間中に配置する候補と、実際の対象物体との位置ずれが無視できなくなる。このため、検出結果には距離の和を用いる代わりに、距離の頻度を表すヒストグラムを用いる。

高速化手法 (1) は、楕円体と定義した人間の頭部を、画像の座標変換によって球体とする。球体に近似することにより、移動領域の射影直線と、微分楕円体との距離を求めるときの計算コストを小さくすることができる。座標変換の様子を図 3 に示す。

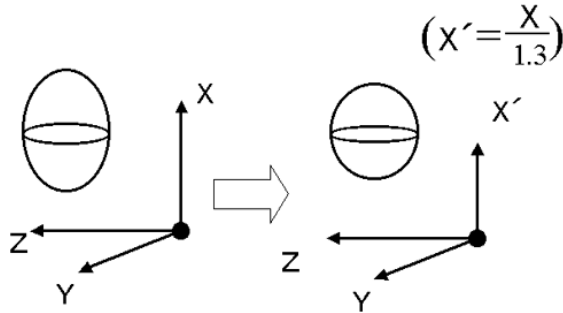


図 3: 高速化手法 (1)

また高速化手法 (2) は、本来、物体の画像面への投影は透視投影であるが、これを平行投影とする。平行投影により、物体の Z 軸方向の情報を無視することができるので、さらに計算コストを小さくすることができる。

高速化手法 (1)(2) より、上述の距離 d_p をより簡単な手法で求めることが可能となる。 D 中の移動領域 $p(x, y, z)$ とし、座標系はカメラ座標系とする。2つの球体 E'_1, E'_2 によって得られる微分楕円体を A とし、その2つの球体の中心をそれぞれ $S_1(x_1, y_1, z_1), S_2(x_2, y_2, z_2)$ とする。また、球体の半径を r とする。

移動領域 p を 3次元空間に逆投影し、その射影直線を l_p とする。このとき、 l_p と E'_1, E'_2 との距離 d_1, d_2 は、次式 (4) によって求められる。

$$\begin{aligned} d_1 &= \sqrt{(x - x_1)^2 + (y - y_1)^2} - r \\ d_2 &= \sqrt{(x - x_2)^2 + (y - y_2)^2} - r \end{aligned} \quad (4)$$

よって、 d_p は次式 (5) で求められる。

$$d_p = \begin{cases} 0, & \text{if } (d_1 - r)(d_2 - r) \leq 0 \\ \min(d_1, d_2), & \text{otherwise} \end{cases} \quad (5)$$

このように、射影直線 l_p と微分楕円体 A との距離は、高速化手法 (1)(2) によって容易に求めることが可能となる。

高速化手法 (3) は、対象物体の候補領域に配置する微分楕円体の個数を減らすことによって、計算コストを小さくすることができる。しかしこの場合、配置する微分楕円体の個数が、検出に対して十分に密でない可能性があるため、楕円体の組が実際の顔の位置に正確に配置されない可能性がある。よってその場合、差分画像との距離が最小となる微分楕円体を検出結果とすると、その検出結果が正確でない可能性が高くなる。そこで、検出結果を最小距離ではなく、距離の頻度のヒストグラムから求めることで、配置ずれを吸収する手法を導入する。

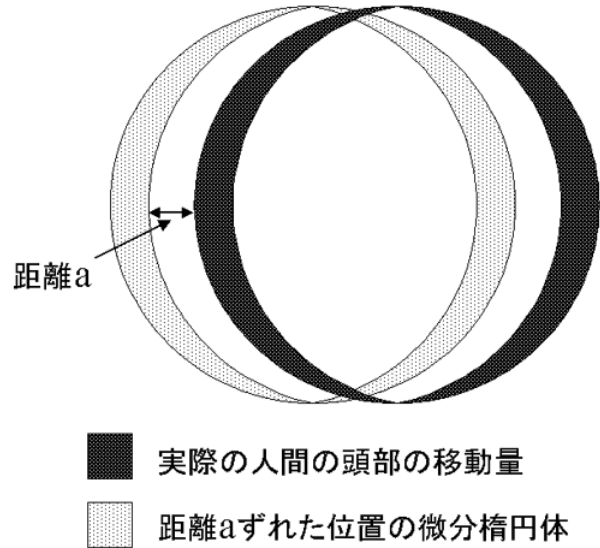


図 4: 距離 a だけずれた位置に配置された微分楕円体

図 4 に示すように、実際の顔の位置から距離 a だけずれた位置に配置された微分楕円体のヒストグラムは、以下の図 5 のように距離 a で頻度が非常に大きくなる。つまり、ヒストグラムで最大となる頻度を検出結果に

用いることによって、微分楕円体の配置ずれを吸収することが可能となる。

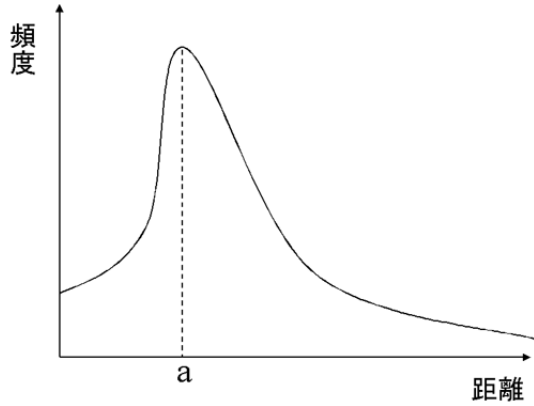


図 5: 距離 a ずれた位置に配置された微分楕円体のヒストグラム

差分画像 D 中の全ての移動領域の集合を P (ただし $P \subset D, p \in P$) とすると、微分楕円体 A のヒストグラムは、次式 (6) で求められる。

$$H_A(d_p) = H_A(d_p) + 1 \quad (p \in P) \quad (6)$$

配置する微分楕円体の集合を B (ただし、 $B \ni A$) とすると、全ヒストグラム中の最大頻度は、次式 (7) のようになる。

$$H_{\max} = \max_{A \in B} (\max H_A) \quad (7)$$

式 (7) で、 $H_{\max}$ となる微分楕円体 A を検出結果とする。

2.3 複数のカメラによる顔領域の検出

近年、ネットワークで結ばれた複数の視覚システム間の協調により、多様な状況・タスクに適応する分散協調視覚の研究が盛んに行われている [6][7]。これらの研究では、複数のカメラを用いて、それらの協調によって人物の顔を検出する研究も行われている [1]。

前節で述べた検出手法を、1つのカメラのみで行った場合、人間の移動方向がカメラの向きに対して平行の場合には、差分画像で得られる移動量が非常に小さくなってしまいうため、楕円体の動きを検出することが困難になる場合もある。しかし、前節で述べた検出ア

ルゴリズムは、容易に複数のシステムに拡張することが可能である。そこで次に、複数のカメラによって得られる、見え方の異なる複数の差分画像から、さらにロバストに顔領域を検出する手法を提案する。

例として、2つのカメラから得られる、2枚の差分画像を用いて顔領域を検出する手法を述べる。2つのカメラをそれぞれカメラ1、カメラ2とする。カメラ間の姿勢、位置の情報は既知であるとする。得られる差分画像をそれぞれ D_1, D_2 とすると、 D_1 中のある移動領域 p_1 を3次元空間に逆投影した場合の射影直線 l_1 と、ある位置に配置した微分楕円体 A との距離、また D_2 中のある移動領域 p_2 の射影直線 l_2 と A との距離は、カメラ間の姿勢、位置が既知であるため、前節の手法と同様に求めることができる。ただし、微分楕円体 A は、X-Y 平面上をある方向に水平に移動した楕円体によって得られるため、視点の異なる2つの画像面に投影される像は異なる。

また、1枚の差分画像のみを用いて検出を行う場合、微分楕円体の位置ずれを考慮するため、距離の頻度のヒストグラムを用いた。しかし、2枚の差分画像を用いて検出する場合、距離のずれがそれぞれの差分画像に対して異なる可能性がある。このため、前節と同様の手法でヒストグラムを求め、その最大値を検出結果とすると、誤検出する可能性が高い。この問題を解決するために、配置した微分楕円体1体と2枚の差分画像ごとのヒストグラムを求める。そして、2つのヒストグラムの最大値の和を計算し、その値を微分楕円体の評価値とする。配置した全ての微分楕円体の中で、評価値が最大である微分楕円体を検出結果とする。

上述のように、複数のカメラシステムの協調により、異なる視点から撮影された差分画像を検出に用いることで、単一のカメラシステムでの検出における問題点を解決することが可能となる。また、前節で述べたように、本手法はヒストグラムによる検出、すなわち投票アルゴリズムにより検出を行う。よってこの点においても、複数のカメラを用いることによって、ロバストに顔領域を検出することができる。さらに、複数のカメラから得られる複数の差分画像を、協調によって全てのシステムが同様に保持するため、各システムで行う検出はすべて等価な処理によって行われる。よって、検出のための演算を、システム間で分散させることが可能となる。

3 実験

提案した顔領域検出アルゴリズムの性能を確認するために、2台のカメラを用いて実験を行った。実験は、あらかじめ人間が一人歩いている場合と二人歩いている場合の二通りの動画を記録しておき、その動画像を用いて提案した手法によって検出を行った。

提案した手法の有効性を示すために、上述した微分楕円体を用いる代わりに、楕円体1体を配置し、その楕円体と差分画像との距離を求めて検出を行う手法と、提案手法による検出を行った。さらに、それぞれの手法において、1枚の差分画像のみで検出を行う場合と、2枚の差分画像を使って検出を行う場合の検出率を比較した。カメラ1、2で得られる差分画像を差分画像1、2とする。検出の際、それぞれの画像の位置を補正するために、画像中の既知である4点による線形キャリブレーションを行った。

撮影は、蛍光灯で照明された比較的暗い廊下で行った。カメラから約4メートル離れた位置を移動する対象を撮影した。撮影に用いたカメラは、SONY EVI-G20であり、カメラ間の距離は約4メートル、カメラ光軸の交差角は約45°である。サンプリングは、360×240の256階調の白黒画像を秒間30フレームで10秒間行った。

検出に用いる微分楕円体の移動方向は、4方向の移動を考慮し、球体半径 r は12,14,16(pixels)(検出面上の大きさ)の3種類とした。球体の移動距離は、半径の3/10とした。また、直方体領域に配置する位置は、16×16×8とした。従って、合計6144の候補を配置した。撮影状況を図6に示す。

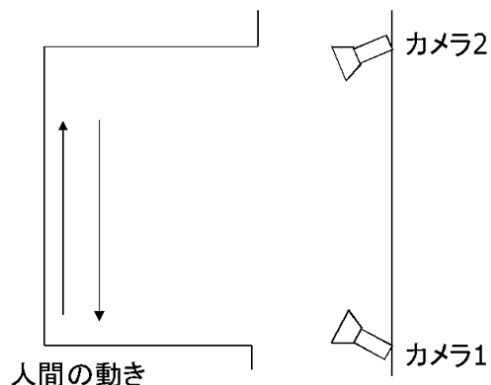


図 6: 撮影状況

図7に、人間が一人の場合の差分画像と検出結果の一例を示す。さらに、それぞれの検出手法における検出率を表1に示す。表1より、微分楕円体を用いて検出する手法の有効性と、2つのカメラを用いて検出する手法の有効性を確認できた。

表 1: 人間が一人の場合の検出率

有効フレーム 221	検出率 (%)			
	楕円体 1 体		微分楕円体	
配置する物体	楕円体 1 体	微分楕円体	楕円体 1 体	微分楕円体
検出に用いる画像	1 枚	両方	1 枚	両方
差分画像 1	71	76	69	88
差分画像 2	66	76	77	86

図8に、人間が二人の場合の差分画像と検出結果の一例を示す。また、それぞれの検出手法における検出率を表2に示す。表2より、人間が二人の場合においても、微分楕円体を用いて検出する手法の有効性と、2つのカメラを用いて検出する手法の有効性を確認できた。

表 2: 人間が二人の場合の検出率

有効フレーム 217	検出率 (%)			
	楕円体 1 体		微分楕円体	
配置する物体	楕円体 1 体	微分楕円体	楕円体 1 体	微分楕円体
検出に用いる画像	1 枚	両方	1 枚	両方
差分画像 1	58	61	61	63
差分画像 2	60	64	71	74

提案する手法は、差分画像を用いて顔領域の検出を行っている。この手法の問題点として、移動している物体の色と、背景の色が非常に近い場合、フレーム間差分では移動領域が正確に検出されないことが挙げられる。よってこの場合において検出を行うと、誤検出する可能性がある。この場合の誤検出結果を図9に示す。

また、複数の人間が存在する状況で、人間同士の距離が配置する微分楕円体の大きさとほぼ等しい場合、人間の間の領域を誤検出するという問題もある。この場合の誤検出結果を図10に示す。



図 7: 人間が一人の場合の差分画像と検出結果



図 8: 人間が二人の場合の差分画像と検出結果



図 9: 顔領域の誤検出 1

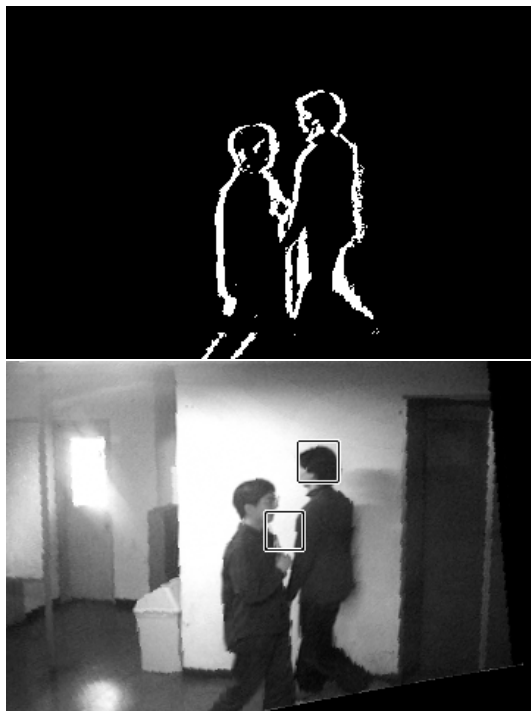


図 10: 顔領域の誤検出 2

4 まとめ

複数のカメラによる、協調型顔領域検出アルゴリズムを提案した。また、2 台のカメラによって得られた動画像を用いて検出のシミュレーション実験を行い、提案手法の有効性を確認することができた。今後は、実験で明らかになった課題に関する検討を進める予定である。

謝辞

この研究の一部は、文部省科学研究費補助金 (課題番号 11680354), および財団法人日東学術振興財団の助成を得た。

参考文献

- [1] 尺長 健:”自然環境下での顔画像の分散協調登録”, 分散協調視覚による動的 3 次元状況理解 '99/3 pp.151-pp.164
- [2] 山田 浩正 伊藤 渡 上田 博唯:”背景差分法における波の語検出抑制法の検討”, 電子情報通信学会技術研究報告 PRMU98-109 '98/10 pp.151-pp.164
- [3] 長井, 久野, 白井:”複雑変動背景化における移動物体の検出”, 信学論,D-II,Vol.J80-D-II,No.5, '97 pp.1086-1095
- [4] 羽下 哲司 藤原 秀人 鷺見 和彦:”首振り,ズームカメラを用いたトラッキング型侵入監視システム”, 電子情報通信学会技術研究報告 PRMU99-67 '99/9 pp.23-pp.30
- [5] Stan Birchfield:”Elliptical Head Tracking Using Intensity Gradients and Color Histograms”, IEEE Conference on Computer Vision and Pattern Recognition, Santa Barbara, California '98/6
- [6] 松山 隆司 和田 俊和 丸山 昌之:”能動視覚エージェントによる移動対象の協調的追跡”, 画像の認識・理解シンポジウム I(MIRU'98) '98/7 pp.365-pp.370
- [7] 中澤 篤志 加藤 博一 井口 征士:”分散カメラシステムによる人物の追跡”, 画像の認識・理解シンポジウム II(MIRU'98) '98/7 pp.1-pp.6