

D-12-96

## 人物のふるまいに着目した動的特徴量からの人物照合に関する一検討

A Study on Person Identification by Dynamic Characteristics of Human Behavior

守屋宏美\*1 河村圭\*2 渡辺裕\*2 富永英義\*1\*2

Hiromi MORIYA\*1 Kei KAWAMURA\*2 Hiroshi WATANABE\*2 Hideyoshi TOMINAGA\*1\*2

\*1 早稲田大学 理工学部 電子・情報通信学科

\*2 早稲田大学 大学院 国際情報通信研究科

\*1 Dept. of Elec. Info. and Comm. Eng., WASEDA Univ.

\*2 Graduate School of Global Info. and Tele. Studies, WASEDA Univ.

## 1 はじめに

近年、映像監視システムの需要が急速に拡大している。それに伴い、監視映像の効率的な蓄積、伝送、閲覧が重要な課題であり、長時間の監視映像から意味情報を抽出する技術が必須である。その技術一つに人物の照合が挙げられる[1]。

映像から人物を照合する要素技術は非常に多く研究されているが[2]、前処理として監視中の映像や蓄積された監視映像からの照合に適したシーン(フレーム群)をリアルタイム検出する研究は十分ではない。

本研究では、照合に適したシーンを選択をすることを目的とし、シーン選択における特徴量について検討を行った。

## 2 人物照合の従来手法

人物照合の手法は、顔認証、色情報識別、シルエット識別、歩行動作特性識別などの研究がある。しかし、それぞれ撮影条件への依存や男女差や年齢層までしか判別できないなど不十分な点があり、単独情報での照合には限界がある[2]。そこで、複数の情報を組合せによる人物照合が必要である。また、前述した4つの手法ではそれぞれ顔と全身の抽出が要求される。

## 3 主観評価実験

どのようなシーンがベストシーンであるのかを検証することを目的とし、映像中の人物の静的特徴量と動的特徴量の把握に必要なショット(フレームの位置)と時間(フレーム数)について、人物照合の従来手法から得られる仮説に基づく主観評価実験を行う。この実験は12の人物の画像を用い、15人の被験者が判定した。

## 3.1 ショットによる比較

人物の映る位置による参照のしやすさを比較するため、静止画を用いた実験を行う。従来手法において、人物の照合に使用する場合は顔画像認証の場合を除き、全身画像を使用している。

仮説 静的特徴量を得るには全身画像でなければならない。

上記の仮説に基づき、エントランスを想定する画像を使用する。上半身画像しか映らない撮影カメラに近いショットから壁の陰に隠れ半身しか映らない撮影カメラから遠いショットまでの範囲で段階的に4つの画像を選択する。実験は2つの同じ位置の静止画を比較し、そこに映る人物が同一人物であるか否かを判定する。

その結果、撮影カメラから遠く、壁に隠れた半身画像以外はすべての画像比較において、正しく判定された。よって、仮説は十分でない。

全身画像の場合、人物がより大きく写る方が判断し易かった。また、上半身画像においても正確に判定された。上半身画像であることで取得できる特徴量は減少するが、上半身画像に判断するのに有効な特徴量が認められる場合、識別に至る。人物の参照において、必要な条件は照合する人物が画像中にしめる割合(面積)の大きさであると考えられる。また、半身画像よりも全身画像の方が判断しやすいため、ベストショットは全身の写る面積最大のフレームとなる。

## 3.2 時間による比較

人物の映る時間による参照のしやすさを比較するため、時間(フレーム数)の異なる動画を用いた実験を行った。

人物の動的特徴を使用する場合の参照時間は、人物の照合を扱う研究において、1歩分を用いるか、2歩分を用いるか、統一された見解がみられない。

仮説 動的特徴量を得るには1周期分(2歩分)が適切である。

上記仮説に基づき、半周期(0.5秒=15フレーム)、1周期(1秒=30フレーム)、2周期(2秒=60フレーム)の3つの動画を選択した。歩く人物の1周期がほぼ1秒であることから、30フレームを基準とした。実験はそれぞれの動画の特徴を聞き、それが静的特徴量であるのか、動的特徴量であるのかを判定した。

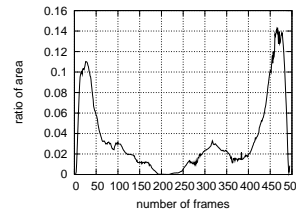


図1 面積比

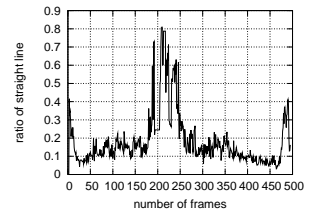


図2 直線長比



図3 462フレーム



図4 482フレーム

その結果、半周期分の動画では、静的特徴量のみで判定しており、1、2周期分の動画では、静的特徴量と共に動的特徴量が相次いだ。また、2周期分では、1周期分を用いた場合と全く同じ情報であった。よって、仮説は正しい。

人物の動的特徴量を使用するにおいて、半周期分の動画では不十分であり、1周期分以上のフレーム数が必要である。そこで、計算コストを考慮に入ると、参照するフレーム数は1周期分が適当である。

## 4 提案手法

人物の欠けるシーンは撮影映像視野から出る場合と建造物の陰に入る場合である。よって、オブジェクトの欠ける部分は直線で構成されることが多い。そこで、人物シルエット抽出を行い、画像全体に対する人物の面積比と抽出されたオブジェクトの周囲長に対する周囲長に含まれる直線の長さの比によって人物の照合に適したシーンを選択する手法を提案する。

ここでこの2つの要素によりの確なベストシーンを選択するため、それぞれのフレームがベストショットである可能性の度合いをベストショット度とし、2つの要素を組み合わせる。蓄積映像に対する高速検出時には2つのベストショット度の合わせたものの1周期分の平均が最も高い1周期をベストシーンとする。また、リアルタイム検出時にはベストショット度がある一定の値を超えた時点から1周期分をベストシーンとする。

## 5 実験と考察

固定カメラを用い撮影された映像中に映る人物の面積比と直線長比を求める。入力画像は主観評価実験に用いたものと同じものを使用し、人物抽出には既存の背景差分法を用いる。得られた2値画像を多角形近似することにより面積、および直線長を求める。

図1、図2に人物シルエット図から求めた各フレームに対する面積比、直線長比を示す。人物シルエットは精度良く検出されており、オブジェクト面積が大きいフレームではオブジェクトの欠落により直線長比が増加していることがわかる。また、図3、図4に面積がほぼ等しい異なる人物シルエット図を示す。このように、面積比だけでなく直線長比を用いることが有効であると言える。

## 6 まとめ

本稿では人物照合の従来手法と主観評価実験により抽出が必要な部位を明確にした。さらに、人物シルエットの面積比と直線長比を用いたシーン選択手法を提案し、実験により提案手法を有効性を確認した。

## 参考文献

- [1] Weiming Hu, et al., "A Survey on Visual Surveillance of Object Motion and Behaviors," IEEE Transaction on Cybernetics-Part C: Application and Reviews, Vol.34, No.3, August 2004.
- [2] 赤松茂, "コンピュータによる顔認識-サーベイ," 信学論(D-II), Vol.185-D-II, No.8, pp.2031-2046, August 1997.