

カメラと人物との間の距離に依存しないグラフ構造特徴量の一検討

A Graph Structure Proposal of Distant Invariant Features between Camera and People

並木重哲
Shigeaki NAMIKI

亀田能成
Yoshinari KAMEDA

大田友一
Yuichi OHTA

筑波大学 大学院システム情報工学研究科
Graduate School of Systems and Information Engineering, University Tsukuba

1 まえがき

人物行動識別の分野におけるカメラ映像のデータセットの多くは、人物の全身が十分映り込むように、一定の大きさで映るものである。一方で、オフィスのように小規模の室内環境においては、カメラと人物との間の距離によって、映像中の人物の見かけの大きさが変化しやすい。たとえば、カメラと人物との間の距離が大きい場合、人物の見え方は小さくなり、カメラと人物との間の距離が小さい場合、人物の見え方は大きくなり、一部の隠れが生じうる。本研究では、カメラと人物との間の距離にロバストに人物行動の識別を行うことを最終的な目標とし、本稿ではそのために利用する特徴量について検討する。

スケール変化にロバストで、かつ物体の見た目の記述力の高い特徴量の一例として SIFT[1] が挙げられる。しかし、人物の見かけの変化が大きいこと、動きに関する情報が必要であることから別の特徴量が必要である。そこで本研究では特徴点のオプティカルフローを人物の動きベクトルとして利用する。特徴点の検出位置は、人物の見かけの大きさや、服のしわ、照明環境などによってフレーム毎にある程度異なる。しかし、検出位置は異なっても、特徴的な動きベクトルは各フレームにおいても検出されると考えられる。

動きベクトルをベースとする人物行動の特徴量として、映像中の人物領域に検出された全ての特徴点のグラフ構造を定義する。ノードは各特徴点に対応し、オプティカルフローの向きと大きさを持つ。エッジは接続するノード同士の向きと大きさにより定義される。ある人物行動同士の比較を行う際には、各グラフ構造の全てのエッジの類似性を考える。これは、部分的なオクルージョンに対してロバストである必要があるためである。

2 動きベクトル

人物領域から特徴点を検出するために、統計的背景差分を行う。こうして算出された人物領域に対する特徴点の算出と、連続 2 フレームからのオプティカルフローの算出手法として、特徴点の算出と追跡を同時に行う KLTtracker[2] を用いる。ここで、人物領域になるべく一定数の特徴点検出されるよう、パラメータを調節する。算出されたオプティカルフローを、人物の行動を反映した動きベクトルとして扱う。

3 動きベクトルのグラフ構造

人物領域に算出された特徴点をまとめてグラフ構造として扱う。ある 2 フレームで得られた特徴点の数を n と

して、その特徴点集合から得られるグラフ G を考える。頂点 i についてオプティカルフローが得られるが、それを $\mathbf{v}_i = (\theta_i, s_i)$ で表す。ここで θ_i は画像中の方向、 s_i はその大きさを表す。また、辺 ij を表す特徴量ベクトルについては $\mathbf{e}_{ij} = (\theta_i, \theta_j, k_i)$ で定義する。ただし $s_i \leq s_j$ のとき $k_i = \frac{s_i}{s_j}$ 、 $s_j < s_i$ のとき $k_i = \frac{s_j}{s_i}$ とする。ここで ij は全ての考えうる組み合わせ ($1 \leq i, j \leq n, i \neq j$) とする。

この 2 フレームと別の 2 フレームについて同様にグラフ \hat{G} (頂点数 \hat{n}) を求める。

$n < \hat{n}$ としたときの G と \hat{G} の辺の一つの辺の対応付け p によるグラフ間の類似度を $Diff_p(G, \hat{G})$ として、

$$Diff_p(G, \hat{G}) = \sum_{n}^{n} f(\mathbf{e}_{ij}, \mathbf{e}_{ij}) \quad (1)$$

ただし一つの ij に対しては高々一つしか ij を割り当てないものとする。また、関数 f は \mathbf{e}_{ij} と \mathbf{e}_{ij} の類似度を評価し、その類似度が一定未満だと 0 を返す関数であり、上限は 1 である。 $n < \hat{n}$ としたとき、 G と \hat{G} の辺の対応付けについては $N = {}_n C_2 P_n C_2$ 通り考えられる。その最大を考え、

$$Diff(G, \hat{G}) = \max_N Diff_p(G, \hat{G}) \quad (2)$$

を G と \hat{G} の類似度とする。実際には N は大きな数になると考えられるので、計算過程ではフローベクトルの大きさなどを考慮して近似によってその値を求める。

4 おわりに

本稿では、カメラと人物との間の距離にロバストな特徴量として、映像中の人物領域における動きベクトルをベースとしたグラフ構造を提案した。今後の課題は、人物とカメラとの間の距離、人物の見え方、照明条件など様々な条件において本手法の有効性を検討することである。

参考文献

- [1] David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, (2004), pp. 91-110.
- [2] Bruce D. Lucas and Takeo Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," International Joint Conference on Artificial Intelligence, (1981), pp. 674-679.