

適応学習型ビジョンへのアプローチ

大津 展之†

† 産業技術総合研究所 / 東京大学
〒 305-8568 茨城県つくば市梅園 1-1-1 中央第 2
E-mail: †otsu.n@aist.go.jp

あらまし 本稿(講演)では、これまで筆者が従事してきた適応学習型ビジョンへのアプローチと研究成果の概要について報告する。まず、パターン認識の一般的な枠組みに基づき、幾何学的な側面としての不変特徴抽出と統計的な側面としての判別特徴抽出の2段階の特徴抽出からなるアプローチを強調する。次に、その線に沿ってこれまで開発してきた、高次局所自己相関(HLAC)特徴抽出と多変量解析手法を組み合わせた汎用画像認識方式、およびその応用について紹介する。最後に、HLAC特徴抽出法を時空間へ拡張した立体高次局所自己相関(CHLAC)特徴抽出法に基づく汎用動画認識の最近の成果について報告する。

キーワード 視覚システム、パターン認識、特徴抽出、多変量解析

An Approach to Adaptive and Trainable Vision

Nobuyuki OTSU†

† National Institute of Advanced Industrial Science and Technology / University of Tokyo
Umezono 1-1-1 Central 2, Tsukuba-shi, Ibaraki-ken, 305-8568 Japan
E-mail: †otsu.n@aist.go.jp

Abstract This paper (invited talk) presents an outline of our approach and research results towards adaptive and trainable vision systems which the author has been devoted to. First, such an approach is emphasized that is based on the general framework of pattern recognition consisting of two stages of feature extraction; *invariant* feature extraction as the first *geometrical* aspect, and *discriminant* feature extraction as the second *statistical* aspect. Along the line of the general approach, a scheme of general-purpose adaptive and trainable image recognition comprising Higher-order Local Auto-Correlation (HLAC) feature extraction and multivariate analysis methods so far developed and its several applications are introduced. Finally, some recent researches using extended HLAC (CHLAC) features for several practical applications of motion recognition are presented.

Key words vision system, pattern recognition, feature extraction, multivariate analysis

1. はじめに

近年、犯罪やテロの増加に伴い、監視カメラによる映像監視の研究が盛んである。特に、監視カメラの知能化や省力化のためには、映像中の対象(たとえば人)およびその動きの認識や異常行動の検出を自動化する技術が必要であり、実用化のニーズは極めて高い。しかしながら、従来手法のほとんどは、動画から個々の動く物体を切り出し、あらかじめ用意したモデルに照らして対象の認識や動きの特徴を抽出する方式や、オブティカルフローにより動きを抽出する手法であり、計算量も多く、精度に限界があって、実用に足る認識性能が得られていない。

これに対して筆者らは、パターン認識における特徴抽出の理論的な視点から、適応学習能力を有する汎用な画像認識方式を

提案し様々な応用を行ってきた。本稿(講演)では、その概要について紹介する。

2. 適応学習型汎用画像認識方式

パターン認識では、一般に関数として表されるパターンから、認識に有効な何らかの特徴を抽出して認識を行う。

特徴抽出の一般的な枠組みとしては、幾何学的な側面としての「不変特徴抽出」と統計的な側面としての「判別特徴抽出」の2段階からなる特徴抽出が原理的に重要である[4]。

この枠組みを最も簡単な形で実践する方式として、高次局所自己相関(HLAC)特徴と多変量解析手法を組み合わせた適応学習型汎用画像認識方式を提案し[1],[2]、顔認識など様々な応用を行ってきた[3],[5],[6],[10]。さらに近年、HLAC特徴を動

画像の場合に拡張した立体 HLAC (CHLAC) に基づく汎用的で高速・高精度な「対象と動き」の特徴抽出方式を提案し[8]、動画画像認識への応用を行っている[9]、[10]。

動画は、2次元静止画像が時間軸に沿って並んだ3次元立体の数値データである。(動)画像認識の場合、対象の認識結果は本質的に画面内の対象の(時)空間的な位置に依らず不変である。これを「位置不変性」という。また、複数の対象がある場合、全体の特徴値がそれぞれ個別対象の特徴値の和になる「加法性」を持つと、以後の処理が容易となり認識精度が向上する。さらに特徴抽出法としては、計算量が少なく実時間処理が可能であることが望ましい。

積分特徴としてのHLAC/CHLACによる特徴抽出方式は、正にこれらの要請を全て満たす基本的で汎用的な「対象と動き」の特徴抽出方式となっており、対象のモデルや知識を一切必要とせず、対象の切り出しも不要である。この抽出データ(特徴ベクトル)を変量解析手法などの統計的手法で統合し、課題に有効な特徴をさらに適応学習的に抽出して認識することにより、(動)画像からの適応学習型の汎用認識方式が得られる。

3. いくつかの応用例

以下、ここでは、最近の成果である動画画像認識への応用例をいくつか述べる。

3.1 Gait 認識

本方式(CHLAC+判別分析+k-NN 識別)を、米国の HumanID プログラムの一環として NIST が取り纏めている、71 人の gait (歩様) から個人を認識するテストデータ(HumanID Gait Challenge Dataset) に対して適用した結果、従来の上位 5 位までの手法を大幅に上回る世界最高性能であることが実証された(図 1)。特に歩きの条件が異なる難しい課題(Probe: D-G)に対して、従来手法に比べて高い識別率をあげている。

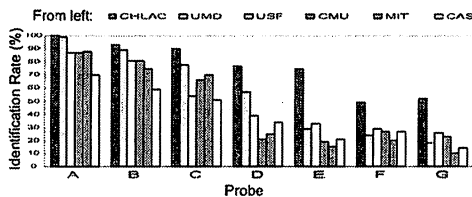


図 1 Comparison of gait recognition methods

3.2 異常検出

CHLAC は、画像中に複数の対象が有る場合、それぞれの対象の特徴の和が全体の特徴となる加法性を持つので、通常(正常)動作の特徴ベクトルは、特徴空間(251次元)のある部分空間(通常動作部分空間 S_N) に分布することになる。従って、異常行動は、常時学習(教師無し)の PCA) によって得られる S_N からの逸脱(その距離を異常値)として、直ちに高速かつ高精度に検出・認識される[9]。複数人の場合でも異常値検出力は同じである(図 2)。対象のモデルや知識は不要であり、異常検出の様々な課題に応用できる。

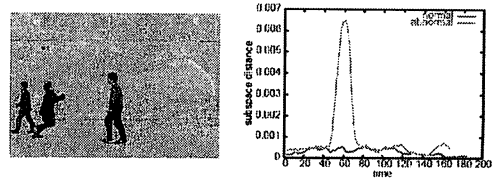


図 2 Example of abnormality detection.

3.3 移動体追跡

画面を分割し HLAC 特徴の加法性を利用することにより、複数移動体の同時自動追跡も可能である[7]。追跡には、対象の形情報に加えて色情報も重要な要素となるので、HLAC 特徴をカラー画像に対応できるように拡張してある。画像レベルでの相関(照合)に基づく従来手法と異なり、移動対象の特徴レベルでの同定・認識に基づくため、移動体の切り出しも不要で計算量も少なく、実時間での頑健な追跡が可能である。

4. おわりに

本方式は、対象に関するモデルや事前知識を一切必要とせず、対象およびその動きを判別特徴空間の点として区別する。さらに対象の切り出しも不要であり、計算量も一定で少ないため、実時間処理が可能である。そのため、極めて汎用的な「対象と動き」の認識方式となっている。対象の認識のみならず、異常行動の検出や、特定した対象の追跡にも適用できるため、ニーズの高い知的防犯カメラなど、セキュリティ分野における自動(無人)映像監視を始め、ロボットの視覚など、広くコンピュータビジョンの研究開発と実用化に大きく貢献すると期待される。

文 献

- [1] 大津, 島田, 森: “N 次自己相関マスクによる図形の特徴抽出,” 電子通信学会技術報告, PRL-78, 31, 1978.
- [2] N. Otsu, T. Kurita: “A new Scheme for Practical Flexible and Intelligent Vision Systems,” Proc. IAPR Workshop on Computer Vision (MVA1988), 431-435, 1988.
- [3] T. Kurita, N. Otsu: “Face Recognition Method using Higher Order Local Autocorrelation and Multivariate Analysis,” Proc. 11th ICPR, 213-216, 1992.
- [4] 大津, 栗田, 関田: 行動計量学シリーズ第 12 巻「パターン認識: 理論と応用」, 朝倉書店, 1996.
- [5] Y. Shinohara, N. Otsu: “Facial Expression Recognition using Fisher Weight Maps,” Proc. 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition, 499-504, 2004.
- [6] T. Ishihara, N. Otsu: “Gesture Recognition using Auto-regressive Coefficients of Higher-Order Local Auto-Correlation Features,” *ibid.*, 583-588, 2004.
- [7] 河合正明: 実環境動画画像からの移動物体のロバストな追跡, 東京大学工学部 卒業論文, 2004.
- [8] T. Kobayashi, N. Otsu: “Action and Simultaneous Multiple-Person Identification using Cubic Higher-order Local Auto-Correlation,” Proc. 17th ICPR, 741-744, 2004.
- [9] T. Nanri, N. Otsu: Unsupervised Abnormality Detection in Video Surveillance, Proc. IAPR Conf. on Machine Vision Applications, 574-577, 2005.
- [10] N. Otsu: Towards Flexible and Intelligent Vision Systems - From Thresholding to CHLAC, *ibid.*, 430-439, 2005.