

高性能な画像認識システムのための特徴量抽出手法

Robust feature-extraction methods for high-performance image-recognition systems

井上 俊明

Toshiaki Inoue

要旨 画像のテクスチャ構造に着目した特徴量であるLBP (Local Binary Pattern)は、輝度変化に頑強で計算量も少ないため、組み込み機器に適した高性能な画像認識への応用が期待できる。そこで本研究では、より軽量で識別精度の高いバリエーションのひとつであるCS-LBP (Center-symmetric LBP)のさらなる高精度化手法を提案する。輝度勾配のバイナリ符号化を基本とする従来のCS-LBPに対し、輝度の2次微分をバイナリ符号化し、さらにそのとり得る値の組み合わせについて画素の回転に相当する関係とみなせるものをグルーピングし、再符号化を行うことで、従来のCS-LBPに対して軽量化しつつ効果的に識別精度を向上できる。テクスチャや一般物体、風景などのさまざまな評価画像を用いてシミュレーションを行い、提案手法の有効性を確認した。

Summary LBP (Local Binary Pattern), suitable for the robust and low-cost texture classification, is expected to be applied to high-performance image-recognition systems in various areas or fields such as traffic, security, and medical care. In this paper, we describe discrimination improvements of CS-LBP (Center-symmetric LBP) which is one of LBP descriptors. In the conventional methods, features are extracted as binary codes of primary spatial differentials from pixel values, whereas in our new method, binary codes of secondary spatial differentials followed by re-encoding with rotation-invariance relation among all binary patterns is adopted. In the multi-class image-classification tests using our experimental system, the proposed method is effective in various datasets such as generic objects, indoor and outdoor scenes, as well as texture images.

キーワード: 画像認識, 特徴量, CS-LBP, 識別精度, 組み込み機器

1. はじめに

近年、自動車や歩行者を取り巻く交通環境における安全性の確保、IoTの進展に伴う新たな情報サービスの創出、その他インフラ管理、防犯、健康維持、医療診断の高度化など、さまざまな分野で画像情報の重要性が高まっている。特に自然画像から顕著な物体や部分的なシーンを抽出する画像認識は、これらの分野に共通する重要な要素技術のひとつとしてさらなる高性能化が期待されている⁽¹⁾。

画像認識のプロセスにおいて、一般的に特徴量抽出は初段に位置し、クラスタリングや分離境界生成などの後続処理に大きく影響することから、認識対象のアビアランス変化や撮影環境変動に頑強で高い識別精度をもつように設計することが効果的である。さらに組み込み機器への実装では、より計算量の少ない軽量なものが望まれる。

一般物体を対象とした特徴量の高精度化や軽量化については、これまでにも多くの提案がなされており、本来テクスチャの識別に適した特徴量として考案されたLBP⁽²⁾やその派生であるCS-LBP⁽³⁾も、代表的な手法として多く引用

されている⁽⁴⁾⁻⁽⁸⁾。そこで本研究では、CS-LBPをベースとした特徴量の高精度化手法を提案し、画像認識に応用してその有効性を評価する⁽⁹⁾。

以降、2章でCS-LBPの概要を述べるとともにその高精度化手法を提案する。3章でさまざまな画像を用いて提案手法の有効性を評価した結果を述べ、4章でまとめと今後の展望を述べる。

2. 特徴量抽出手法と高精度化

2.1 LBPおよびCS-LBPの概要

画像のテクスチャ構造に着目した局所特徴量として、LBP (Local Binary Pattern)やCS-LBP (Center-Symmetric Local Binary Pattern)がよく知られている⁽²⁾⁻⁽³⁾。LBPは、原理的にはスケーリングや回転などの幾何学的な変動に頑強ではないが、輝度変動に頑強で高速に計算できる特徴量として広く引用されている。さらに識別精度を改善するため、回転変動への対応や他の特徴量との組み合わせなどの工夫が多数提案されている⁽⁴⁾⁻⁽⁶⁾。一方、CS-LBPは、

原理的にはLBPに類似するが、画素位置の対称性を利用することでLBPに対して計算量を大幅に削減するとともに、平坦画像に対しても識別精度を改善した特徴量である⁽³⁾。

図1に、従来のLBPとCS-LBPの計算例を示す。LBPでは、注目画素 p_c の周囲8画素との輝度差をそれぞれ閾値0で2値化し、バイナリ重み付けをして得られた値(LBP値)をピンとし、画像領域にわたってヒストグラムを算出したもの(256次元のベクトル量)を基本特徴量とする。一方、CS-LBPでは、注目画素 p_c に関して点对称の位置関係にある周囲4画素ペアの輝度差の絶対値を、それぞれ任意の閾値 T で2値化し、バイナリ重み付けをして得られた値(CS-LBP値)をピンとし、画像領域にわたってヒストグラムを算出したもの(16次元のベクトル量)を基本特徴量とする。近年の応用では、テクスチャだけでなく歩行者や自動車などの一般的な物体の識別を目的とした汎用特徴量としての有効性も示されている⁽⁷⁾⁽⁸⁾。

そこで本研究では、画像認識への応用を目的として、従来のCS-LBPの識別精度をさらに改善する手法を提案し、その有効性を評価する。

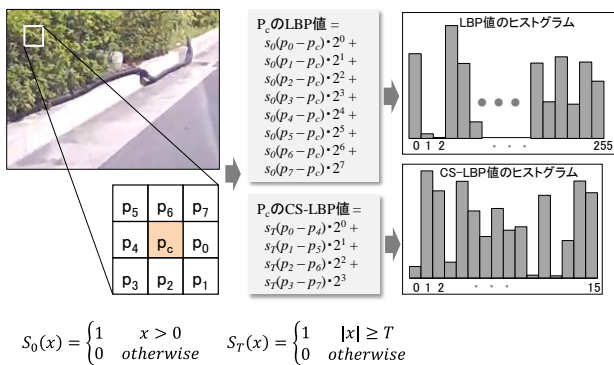


図1 従来のLBPとCS-LBPの計算例

2.2 回転不変性の導入による高精度化

従来のCS-LBPは、そのままでは画像の回転変動に頑強ではないが、画素単位で回転の位置関係を考慮した符号化を行うことで、識別精度の改善が期待できる。

図2に、回転不変性を導入したCS-LBP(回転不変CS-LBP)の原理を示す。従来のCS-LBP値がとり得る16値のパターンは、隣接8画素(点对称を考慮して4画素ペア)の2値状態(4ビット)に対応づけられる。ここで“ON”となる画素ペア(濃い色の画素で代表)を俯瞰して回転の位置関係にあるとみなせるパターンを纏めると、9つのグループに集約できる。そこで各グループに符号を再定義することで、幾何学的な厳密性はないが回転不変性を与えることができる。さらにとり得るパターン数、すなわち基本特徴量の次元数も16次元から9次元へ大幅に削減できる。

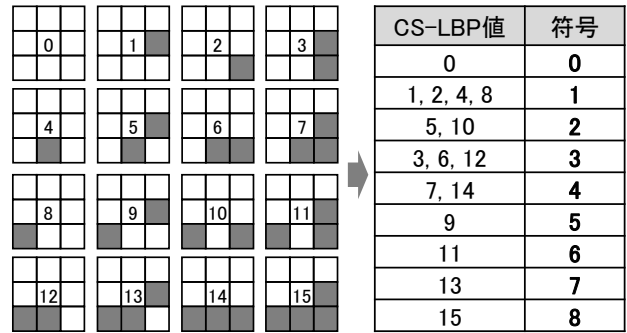


図2 回転不変性の導入

2.3 2次微分の導入による高精度化

従来のCS-LBPは、図1に示すように輝度に関する1次の空間微分を用いた特徴量であるが、式(1)に示すように、形式的に2次微分に置き換えることで、より複雑なテクスチャの特徴を捉えることが期待できる(2次微分CS-LBP)：

$$\Delta^+ CS-LBP = \sum_{i=0}^{(N/2)-1} s_T(p_i + p_{i+(N/2)} - 2 \cdot p_c) \cdot 2^i,$$

$$\Delta^- CS-LBP = \sum_{i=0}^{(N/2)-1} s_T(2 \cdot p_c - p_i - p_{i+(N/2)}) \cdot 2^i$$

... (1)

ここで $\Delta^+ CS-LBP$ および $\Delta^- CS-LBP$ はそれぞれ2次微分の正係数および負係数、 s_T は図1に定義した閾値関数、 p_c は注目画素の輝度値、 N は p_c の周囲の画素数である(図1の例では8画素)。どちらの係数を用いるかは任意であり、図1に倣うと何れか一方を用いれば16次元、両方を用いれば32次元の基本特徴量が得られる。さらにこの基本特徴量に対して、前節で述べた手法をそのまま適用することで、回転不変性を導入でき、相乗的な識別精度の改善効果が期待できる(回転不変2次微分CS-LBP)。

2.4 識別精度の評価

さまざまなテクスチャを集めたデータセットであるTexture dataset⁽¹⁰⁾を評価画像として用い、提案手法を含む以下のバリエーションについて識別精度を比較評価した：

- 提案手法(4種)
 - 回転不変CS-LBP, 2次微分CS-LBP(正係数のみを用いた場合), および回転不変2次微分CS-LBP(正係数のみを用いた場合, 正係数と負係数の両方を用いた場合)
- 従来手法(2種)
 - CS-LBPおよび回転不変LBP⁽⁴⁾

データセットから17クラス×30枚の学習画像と17クラス×10枚のテスト画像をランダムに抽出したものを共通の評価画像とし、1)クラス毎に学習画像の平均特徴量を算出、

2) 平均特徴量とテスト画像の特徴量とのユークリッド距離が最小となるクラスに全テスト画像を分類, 3) テスト画像の正解クラスのラベルを用いて分類結果の正解率を算出, という手順を5回繰り返して正解率の平均値を採用した。

図3に, 各手法の識別精度を比較する。特徴量のベクトル長に関する識別精度の依存性も併せて評価するため, 評価画像を $(2^n \times 2^n)$ 個のブロックに分割し, 各ブロックの基本特徴量を連結したベクトルをその画像の特徴量と定義し, n を0から5まで変化させて識別精度をプロットした。いずれもベクトル長の増加に伴い識別精度は改善する傾向である。また, 従来のCS-LBPに対して回転不変性および2次微分を導入することでベクトル長によらず識別精度が改善され, 特に回転不変性の導入が効果的であることが分かった。

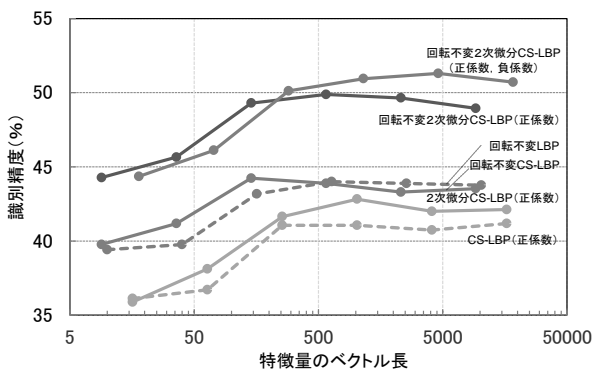


図3 識別精度の比較

3. 画像認識システムへの応用

3.1 認識アルゴリズムの概要

一般物体の認識手法として広く知られている bag of visual words (BoVW)に基づく画像表現と⁽¹¹⁾⁽¹²⁾, 多クラス画像の同時分類で効率が良いとされる部分空間法⁽¹³⁾, を用いた認識アルゴリズムをソフトウェアで実装し, 多クラス画像の識別を題材として提案手法の有効性を評価した。

図4にアルゴリズムの全体構成を示す。学習では, まず画像を複数のブロックに分割し, 各ブロックから特徴量(基本特徴量を連結したベクトル)を抽出する。次に全画像の特徴量をクラスタリングして得られた量子化ベクトルを用い, 画像毎に特徴量を量子化してBoVW(量子化ヒストグラム)を算出する。最後に全画像の正解クラスラベルとBoVWから, クラス毎の部分空間を生成して学習結果とする。ここでは, クラスタリング手法として広く使われているk-meansを, 部分空間生成手法として比較的識別精度の高いとされるガウスクERNELを用いた非線形部分空間法を採用した。またBoVWはクラス数で次元数が決まるベクトル量であるが, ベクトル長を柔軟に制御して識別精度を最適化するため, 特徴量の抽出とは別の画像分割条(vw_1, vw_2, \dots)をパラメータとして設けた。

テストでは, まず画像から必要に応じて識別対象とする

領域(ウインドウ)を切り出す。次に学習と同様の画像分割条件で, ウインドウからの特徴量の抽出, およびBoVWの算出を経る。最後にBoVWがどのクラス部分空間に属するかを判定して識別結果とする。ここで画像を $(2^n \times 2^n)$ 個に分割し, n に関して全領域をテストすることで, 識別対象のスケールや並行移動の変動に対応した。また判定基準として, CLAFIC法(特徴空間の原点を起点として部分空間への射影量が最大となるクラスに分類する方法)と投影距離法(部分空間の原点を起点として部分空間への投影距離が最小となるクラスに分類する方法)があるが⁽¹³⁾, ここではより計算量の少ないCLAFIC法を採用した。

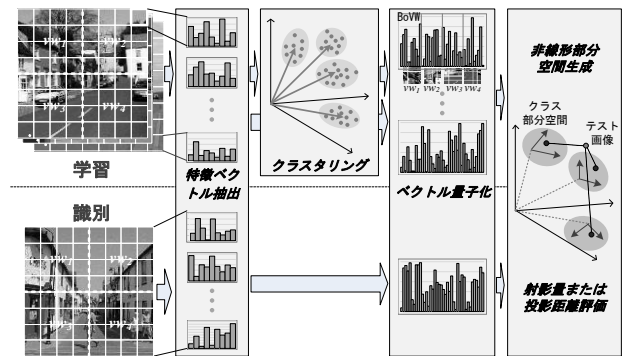


図4 認識アルゴリズムの概要

3.2 さまざまな画像への応用

テクスチャを用いた評価で識別精度が高かった, 回転不変CS-LBP, 正係数の回転不変2次微分CS-LBP, および回転不変LBPの3種に注目し, 前節で述べた認識アルゴリズムを用いて, さまざまな画像で提案手法の有効性を評価した。

● Texture datasetの例(テクスチャ構造)



● ALOI datasetの例(人工物体)



● Fifteen scene datasetの例(屋内外シーン)



● Event datasetの例(スポーツシーン)



図5 評価画像の例

図5に評価画像の例を示す。Texture datasetに加え、ALOI dataset(人工物体)⁽¹⁴⁾、Fifteen scene dataset(屋内外シーン)⁽¹⁰⁾、およびEvent dataset(スポーツシーン)⁽¹⁵⁾を用いた。図6に識別精度の評価結果を示す。評価画像毎に対応する学習画像とテスト画像の構成も図中に併せて示す。認識アルゴリズムに含まれる最適化パラメータ(式1のT, 画像分割数, ガウスクERNEL係数など)については、評価画像毎に予備実験を行って値を確定した。また識別精度については、学習画像とテスト画像をランダムに構成した5回の試行結果の平均値を採用した。

いずれの評価画像においても提案手法が優れており、従来手法に対して3.8%~28.9%の改善が得られた。特にFifteen scene datasetにおいて、2次微分の導入による精度改善が大きいことが分かった。一方で、回転不変性の導入を前提とした場合に、1次微分と2次微分の効果の差異は評価画像に依存するため、予備実験の結果を用いるなどで選択する必要があると考えられる。

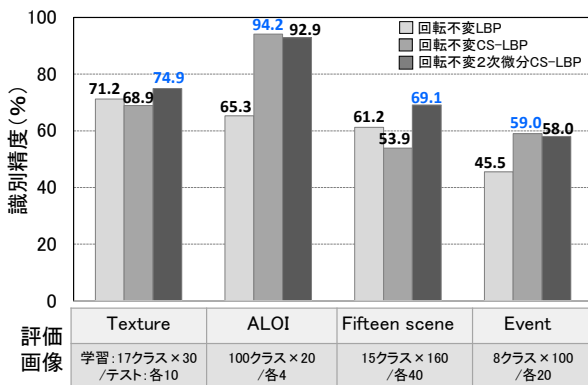


図6 さまざまな画像での評価結果

3.3 車載カメラを用いた路上工事検出への応用

実用的なシーン検出を題材として提案手法の有効性を確認するため、車載カメラで撮影した画像から路上工事領域を検出する初期的な実験を行った。

図7に実験結果をまとめる。学習画像として工事領域と非工事領域を切り出した画像各140枚を用い、テスト画像として未知の工事領域を含む画像3枚(a)を用いた。工事領域のみを検出したいため、縦と横をそれぞれ1分割、2分割、および4分割したウインドウを定義し、各ウインドウで算出したクラス毎の射影量を積算して得られた射影量のマップ(累積射影量)を利用した。(b)は工事クラスの累積射影量を可視化したものである。工事領域は、たとえば工事クラスの累積射影量を適切に正規化、閾値処理することで、(c)の白枠のように得られる。図の上段と中段は比較的良好に検出できている例、下段は中央分離帯付近で誤検出が含まれる例である。このような誤検出領域をネガティブサンプルとして学習画像に加えていくことで、さらに検出精度は改善されると考えられる。認識アルゴリズムに含まれるパラ

メータを調節することで、さらなる最適化の余地は残されているが、おおよその工事領域は検出されており、提案手法の有効性が示唆されたと考える。

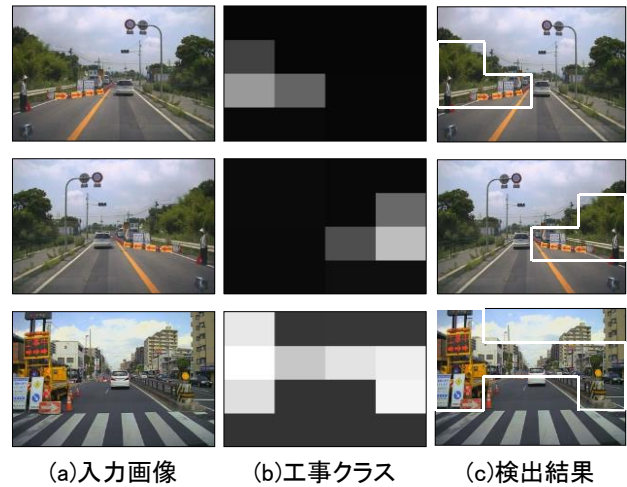


図7 路上工事領域の検出結果

4. まとめ

本研究では、輝度変動に頑強で計算量の少ない特徴量のひとつであるCS-LBPに着目し、回転不変性および2次微分の導入による高精度化手法を提案した。

LBPを含む従来手法のバリエーションに対して、テキストチャでの識別精度を効果的に改善できることを確認した。また提案手法による特徴量と、BoVWに基づく画像表現、およびガウスクERNEL関数を用いた非線形部分空間法、とを組み合わせることで多クラス画像識別に応用し、テキストチャだけでなく人工物体や屋内外シーンなどの一般的な評価画像でも有効であることを確認した。さらに車載カメラ画像からの路上工事領域の検出を題材とし、初期的な実験を行って提案手法の有効性を示唆した。

一方、従来のCS-LBPに対して2次微分を導入することで、ノイズの影響を受けやすくなる懸念がある。これについては大域的あるいは局所的にノイズ量を推定し、閾値Tに反映するなどの対策が考えられる。また本研究で用いた多クラス画像識別の枠組みで実用性を高めるためには、より大規模なデータセットを用いた多様なアピランスの学習、オンライン学習の導入、事前知識その他補助的な情報の活用、などを候補とした高精度化が課題となる。さらにモバイル機器への組み込みでは、専用ハードウェア処理などによる低電力高効率化やリアルタイム処理なども実装課題として加わる。

動画を含む自然画像から、任意に物体や部分的なシーンを高精度かつ高速に識別する技術は、冒頭で述べたさまざまな分野への応用で今後も重要性を増すと考えられ、本提案に限らずより優れた手法を提案していきたい。

参考文献

- (1) 佐藤, 大友(編), 「画像認識技術の実用化への取り組み」, 情報処理学会誌, Vol. 51, No. 12, 2012年12月.
- (2) T. Ojala, M. Pietikäinen and D. Harwood, “A comparative study of texture measures with classification based on featured distributions”, *Pattern Recognition*, Vol. 29, Issue 1, pp. 51-59, Jan. 1996.
- (3) M. Heikkiä, M. Pietikäinen, and C. Schmid, “Description of Interest Region with Center-Symmetric Local Binary Pattern”, *Indian Conference on Computer Vision, Graphics and Image Processing(ICVGIP'06)*, *Lecture Notes in Computer Science(LNCS)* 4338, pp.58-69, 2006.
- (4) T. Ojala, M. Pietikäinen and T. Mäenpää, “Gray Scale Rotation Invariant Classification with Local Binary Patterns”, *European Conference on Computer Vision(ECCV'00)*, *LNCS* 1842, pp.404-420, 2003.
- (5) T. Ahonen, J. Matas, C. He and M. Pietikäinen, “Rotation Invariant Image Description with Local Binary Pattern Histogram Fourier Features”, *Proceedings of Scandinavian Conference on Image Analysis (SCIA'09)*, pp.61-70, 2009.
- (6) 寺島, 喜田, 「勾配情報を用いたLocal Binary Patternの改良」, *DEIM Forum 2014予稿集*, F5-4, 2014年3月.
- (7) A. Satpathy, X. Jiang and H. L. Eng, “Human detection using Discriminative and Robust Local Binary Pattern”, *IEEE International Conference on Acoustics, Speech, and Signal Processing(ICASSP'13)*, 13-4, May, 2013.
- (8) Y. Zheng, C. Shen, R. Hartley and X. Huang, “Pyramid Center-Symmetric Local Binary/Trinary Patterns for Effective Pedestrian Detection”, *Asian Conference on Computer Vision(ACCV'11)*, *LNCS*, pp.281-292, 2011.
- (9) 井上, 「CS-LBP特徴量を用いた多クラス画像識別の高精度化に関する一検討」, *電子情報通信学会技術研究報告*, Vol. 116, No. 89, pp. 1-6, 2016年6月13日.
- (10) http://www-cvr.ai.uiuc.edu/ponce_grp/data
- (11) G. Csurka, C. Christopher, R. Dance, L. Fan, J. Willamowski, and C. Bray, “Visual Categorization with Bags of Keypoint”, *Proceedings of ECCV Workshop on Statistical Learning in Computer Vision*, pp.59-74, 2004.
- (12) 永橋, 伊原, 藤吉, 「画像分類におけるBag-of-featuresによる識別に有効な特徴量の傾向」, *情報処理学会研究報告*, 2009-CVIM-169, No.3, pp.1-8, 2009年11月19日.
- (13) 前田, 村瀬, 「カーネル非線形部分空間法によるパターン認識」, *電子情報通信学会論文誌, D-II*, Vol. J82-D-II, No.4, pp.600-612, 1999年4月.
- (14) <http://aloi.science.uva.nl/>
- (15) http://vision.stanford.edu/lijiiali/event_dataset/

著者紹介

井上 俊明 (いのうえ としあき)

研究開発部に所属。LSI回路・アーキテクチャ技術の研究, プラズマディスプレイ製品向け信号処理LSIの開発を経て, 現在, 映像処理技術の研究に従事。