

擬似乱数と LBP による輝度変換攻撃された電子透かし画像検出 Extraction of Intensity Conversion Attacked Digital Watermarking Image Using Pseudo-Random Pattern and LBP

国松 亜蘭†
Aran Kunimatsu

今村 幸祐‡
Kousuke Imamura

黒田 英夫§
Hideo Kuroda

藤村 誠¶
Makoto Fujimura

1. はじめに

近年、各種デジタルコンテンツの不正複製による著作権侵害が深刻な社会問題となっており、電子透かしをはじめとする対策技術が盛んに検討されている。この一つとして、インターネット上の不正複製画像を効率的に検出するために、画素値ヒストグラムの相関係数を用いた画像検索法[1]を提案した。この手法は、ガンマ変換等の輝度変換攻撃に対応していなかったため、新たに Local Binary Pattern (LBP) [2]ヒストグラムを用いた輝度変換攻撃画像の検出法を導入する。提案法では、LBP 算出時に画像領域に応じて制限された擬似乱数を加えることで、画像全体の構造特徴を反映した LBP ヒストグラムを作成する。これにより、輝度変換攻撃画像の高精度な検出が可能となることを示す。

2. LBP ヒストグラムと相関係数

LBP とは図 1 のように、注目画素を閾値として近傍画素を二値化し、ビットパターンを重み付けにより数値化することで、局所的なパターンを特徴量化する手法である。

22	47	40	0	1	0	1	2	4
63	45	39	1	0	0	128	8	8
50	31	58	1	0	1	64	32	16
Pixel Value			Binary Pattern			Weight		

LBP Pattern = 01001011, LBP Value = 75

図 1 LBP (Local Binary Pattern)

この LBP を対象画像の各画素に対して求め、その正規化ヒストグラムの相関係数を 2 枚の画像の類似評価に用いる。相関係数の算出式を式 (1) に示す。ここで、 x_i y_i は正規化された LBP 値 i の生起率を示し、 \bar{x} \bar{y} はそれぞれの相対平均を表す。ガンマ変換等の輝度変換の多くでは、局所的な画素値の大小関係が変化しないため、LBP ヒストグラムの相関係数も変化し難い。

$$\rho = \frac{\sum_{i=0}^{255} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=0}^{255} (x_i - \bar{x})^2 \sum_{i=0}^{255} (y_i - \bar{y})^2}} \quad (1)$$

† 金沢大学大学院自然科学研究科
Graduate School of Natural Science and Technology,
Kanazawa University

‡ 金沢大学理工研究域
Institute of Science and Engineering,
Kanazawa University

§ FPT 大学計算機基礎
Department of Computing Fundamentals, FPT University

¶ 長崎大学工学研究科
Graduate School of Engineering, Nagasaki University

2.1 LBP ヒストグラムの高尖度対策

LBP 値の出現頻度には大きな偏りが発生し易く、画像の類似性が低くとも相関係数が高くなる傾向がある。これは、特定の LBP 値が極端に高い頻度で発生することに起因し、例えば、イラストのような背景の変化が少ない画像が挙げられる。この対策として、LBP 算出時に注目画素の 8 近傍に、一様生起率で整数値を取る擬似乱数を加える (図 2)。これにより、画像の特徴を保持したまま LBP ヒストグラムの高尖度部分を分散することができる。提案法では、擬似乱数を $r_i \in \{-2, -1, 0, 1, 2, 3\}$ の範囲とする。

22	47	40	r_0	r_1	r_2	+	$r_i \in \{-2, -1, 0, 1, 2, 3\}$
63	45	39	r_7		r_3		
50	31	58	r_6	r_5	r_4		
Pixel Value			Pseudo-Random Sequence				

図 2 擬似乱数の付与

2.2 擬似乱数の制限による構造情報の付加

LBP は局所的なビットパターンであり、その統計的特徴を表す LBP ヒストグラムでは、各ビットパターンが画像のどの位置に存在するかという構造的特徴は考慮されない。つまり、図 3 に示すような 2 つの画像は LBP ヒストグラムにおいて区別ができない。その結果、LBP ヒストグラムによる検出では、過検出が多く発生してしまう。



図 3 統計的特徴が等しく画像構造のみが異なる画像例

そこで、高尖度対策のために与える擬似乱数に対して、画像中の領域位置に応じて特定の特徴を与える。この結果、画像の平坦な領域のパターンは、画像中の位置に応じて特定の範囲の LBP に分散することになる。つまり、平坦領域の画像中の位置が異なる画像は、LBP ヒストグラムでは同一画像として検出され難くなる。具体的には、画像中の領域位置に応じて、8 近傍に与える擬似乱数の一部を乱数の範囲を超えた定数に固定する方法を用いる。

0*0*	*****	**1***0*	r_0	r_1	r_2	r_3	r_4	r_5	r_6	r_7
*****	*****	*****	}	* : random						
*****	*****	*****		1 : max ($r_i = 4$)						
0*1*	*****	**1***1*		0 : min ($r_i = -3$)						

図 4 画像中の領域位置によって制限された擬似乱数

本論文では図 4 に示すように、画像を 3×3 の 9 領域に分割して、四隅の領域についてそれぞれ異なる制限を付与している。図中の各領域内の表記は加える擬似乱数列 {r₀ r₁ r₂ r₃ r₄ r₅ r₆ r₇} を示し、* が [乱数]、0 が [定数-3]、1 が [定数 4] となることを表している。このように一部のビットを定数とすることで、平坦な領域の LBP が擬似乱数によって分散する範囲が領域毎に限定される。結果として、画像の構造を反映した LBP ヒストグラムを得ることが期待できる。

実際に、図 3 で示した画像例に対して相関係数を計算した結果を表 1 に示す。擬似乱数を制限した LBP ヒストグラムでは明らかに相関係数が低くなり、構造の異なる画像を区別できていることがわかる。

表 1 統計的特徴が等しく画像構造のみが異なる画像での擬似乱数の付加による相関係数の変化

Image A、Image B 間の相関係数値			
画素値ヒストグラム	L B P ヒストグラム		
	乱数なし	高尖度対策	制限疑似乱数
1.0000	1.0000	0.9988	0.9088

3. 電子透かし画像検索アルゴリズム

我々が提案した 2 段階画像検索法は、前段検索とメイン検索の 2 段階で構成される。前段検索では、所有画像と類似した不正複製画像候補を抽出する。次に、不正複製画像候補をメイン検索に入力し、実際に電子透かしの抽出と、不正複製画像の判定を行う。

本論文の、ヒストグラム間の相関係数を用いた類似画像検索は前段検索にあたり、そのアルゴリズムを図 5 に示す。所有する画像の各ヒストグラムをあらかじめ用意することで処理の高速化を図っており、入力画像から作成した画素値ヒストグラム、LBP ヒストグラムとの相関係数を算出して、設定した閾値以上であれば類似画像と判定する。そうでなければ類似画像とは判定せず、次の検索対象画像の評価に移る。Th_o、Th_c はそれぞれ原画像、圧縮画像の画素値ヒストグラムの相関係数における閾値を示しており、値は 0.95 に設定した。Th_{LBP} は LBP ヒストグラムの相関係数における閾値を示しており、値は 0.98 に設定した。

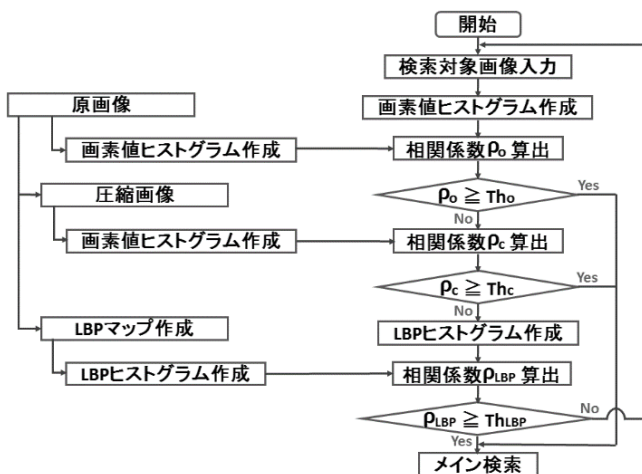


図 5 ヒストグラム間の相関係数を用いた画像検索法

4. シミュレーション実験

提案する画像検索法の、輝度変換攻撃耐性を検証する。評価基準とシミュレーション条件を以下に示す。

●評価基準

$$\text{Recall} = \frac{tp}{fn + tp} \quad (2)$$

$$\text{Precision} = \frac{tp}{tp + fp} \quad (3)$$

$$\text{Error rate} = \frac{fn + fp}{U} \quad (4)$$

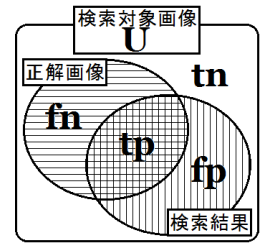


図 6 画像検索のモデル

●シミュレーション条件

原画像：MIRFLICKR No.1~No.10000 (10000 枚)

輝度変換攻撃画像：10000×11 = 110000 枚

(γ = 0.50, 0.60, 0.70, 0.80, 0.90, 1.00, 1.11, 1.25, 1.43, 1.67, 2.00)

$$\text{ガンマ変換 } I_{\text{output}} = 255 \times \left(\frac{I_{\text{input}}}{255} \right)^{1/\gamma}$$

各ヒストグラムに対する検出精度の評価を表 2 に示す。画素値ヒストグラムではほとんど検出できなかった輝度変換攻撃画像が、乱数なしの LBP ヒストグラムを導入することで検出されることがわかる。また、高尖度対策を行うことで、Recall、Precision、Error rate のすべてが改善した。しかし、過検出が多く、Precision、Error rate は良好とはいえない。そこで、加える擬似乱数を制限した場合は、Recall がやや低下しているが、Precision は上昇するという結果が得られ、Error rate についても改善した。

表 2 輝度変換攻撃画像の検出結果

		Recall	Precision	Error rate
画素値ヒストグラム		0.2250	0.4455	0.0078
LBP ヒスト グラム	乱数なし	0.9350	0.0555	0.0680
	高尖度対策	0.9775	0.1324	0.0509
	制限疑似乱数	0.9716	0.2477	0.0324

以上より、LBP 算出時に擬似乱数に制限を加えることで、輝度変換攻撃画像が検出可能となると共に、過検出をコントロールし、より目的に合った画像検出が可能となった。

5. まとめ

インターネット上の不正複製画像を検出する画像検索法における輝度変換攻撃画像の検出法を提案した。その中で、LBP 算出時に制限された擬似乱数を加えることで、提案する画像検索の精度向上を実現した。

今後、この結果を踏まえて、動的な閾値設定を行い、最終的な前段検索システムを設計していく予定である。

謝辞

本研究の一部は科研費(24560466)の助成を受けたものである。

参考文献

[1] K. Imamura, H. Kuroda and M. Fujimura, "Image Content Detection Method Using Correlation Coefficient between Pixel Value Histograms", Springer CCIS 260, pp.1-9, 2011.
 [2] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on featured distribution," Pattern Recognition, vol. 29, no. 1, pp. 51-59, 1996.