

知っておきたいキーワード

Image/Video Signature

(正会員) 工藤大樹[†], (正会員) 西川博文[†]

[†]三菱電機株式会社 情報技術総合研究所

"Image/Video Signature" by Daiki Kudo and Hirofumi Nishikawa (Information Technology R&D Center, Mitsubishi Electric Corporation, Kamakura)

キーワード：MPEG-7, 画像同定, 映像同定, 画像特徴量, 画像検索, メタデータ

Image/Video Signatureとは

編集済みの画像を再編集したい、しかし、オリジナル画像がどこにあるかわからない、ということはありませんか？

画像共有サイトには、膨大な画像が蓄積されていますが、その中には、権利者の許可なく投稿、共有された違法な画像があります。これら膨大な画像から違法な画像を見つけるにはどのようにすればいいのでしょうか？

これらの問題解決に、画像同定技術のImage/Video Signatureが活用できます。画像同定とはどのように定義されるのかを、デジタルデータの一致不一致(同一性)判定との違いを示しながら説明します。

デジタルデータの一致を判断する方法として、SHA (Secure Hash Algorithm) に代表されるハッシュ関数があります。入力データが1ビットでも異なると、ハッシュ値はまったく異なる値を示し、効率的にデジタルデータの一致を確認できます。

しかし、画像データにおいては、 左 (Lena編集画像) に示すように、 回

転やモノクロ化などの編集・加工を施した場合、ハッシュ値は異なる値を示し、同一画像とは判定されません。一方で、人間の目には同一画像であることがわかります。上述したオリジナル画像の検索や、画像共有サイト上での

違法画像検索は、何らかの編集・加工が施された画像を対象にすることになります。このように、画像同定には通常のハッシュ関数では不可能な、編集や加工が施された画像に対するロバストな同一性判定が要求されます。 

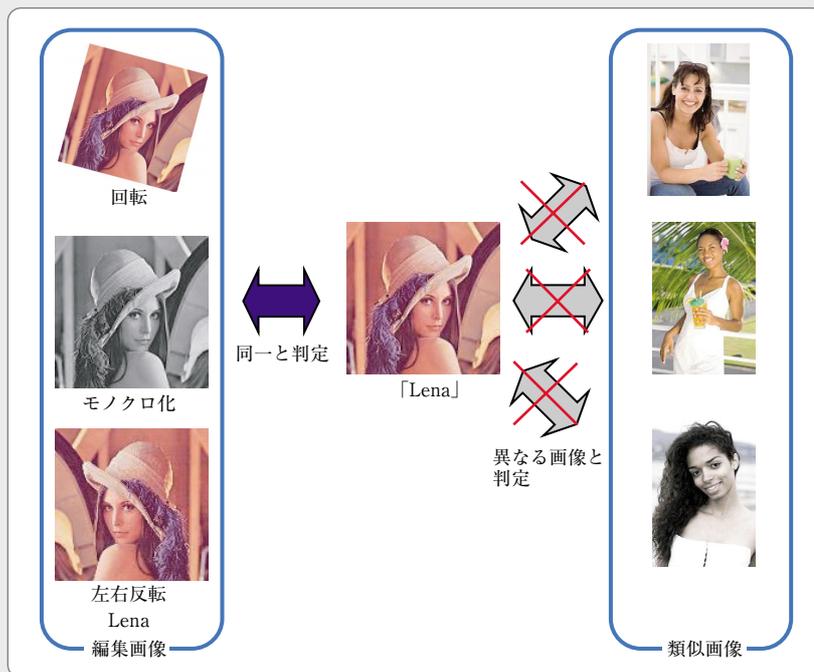


図1 画像の同定と類似画像検索

一方で、画像同定は類似画像検索技術とは異なります。例えば、図1右(類似画像)はいずれも女性の写真で、「Lena」の類似画像と言えるでしょう。しかし、画像同定では、これら

は異なる画像と判定される必要があります。

ISO/MPEGでは画像同定を行うための特徴量の標準化作業を実施し、2009年に静止画像向けのImage

Signature¹⁾を、2010年に動画像向けのVideo Signature²⁾を策定しました。それぞれの特徴について紹介します。

Image Signature

静止画像の同定を行うための特徴量であるImage Signatureは、画像のトレース変換³⁾結果を1次元に縮退し、この1次元データをフーリエ変換したものです。

原画像とそれを45°回転した画像と、それぞれトレース変換・1次元縮退したデータ例を図2に示します。

45°回転画像から得られる1次元縮退データは、原画像から得られるデー

タから45°分のシフトで表現され、回転処理にロバストであることを示しています。

実際のImage Signature特徴量は、以下の二つから構成されます。

- ① Global Signature
- ② Local Signature

Global Signatureは、画像全体のトレース変換により得られる特徴量です。画像の拡大・縮小・回転などの基本的な画像処理に対してロバストな画像同定が可能です。特徴量間の単純な

ハミング距離をマッチング関数として利用することで、7,000万枚/秒という高速同定処理が可能です。

Local Signatureは、画像の特徴点周辺をトレース変換して得られる局所的な特徴量です。特徴点の例を図3に示します。Local Signatureによるマッチングでは、特徴量間のハミング距離の他、特徴点の幾何学的位置関係を利用しており、マッチング速度は10万枚/秒とGlobal Signatureに比べて劣ります。しかし、画像のクロッピングなど大幅な画像編集に対しても高いロバスト性を有しています。

なお、マッチング速度はいずれもIntel® Core2Duo 3.0GHzでの計測結果です。

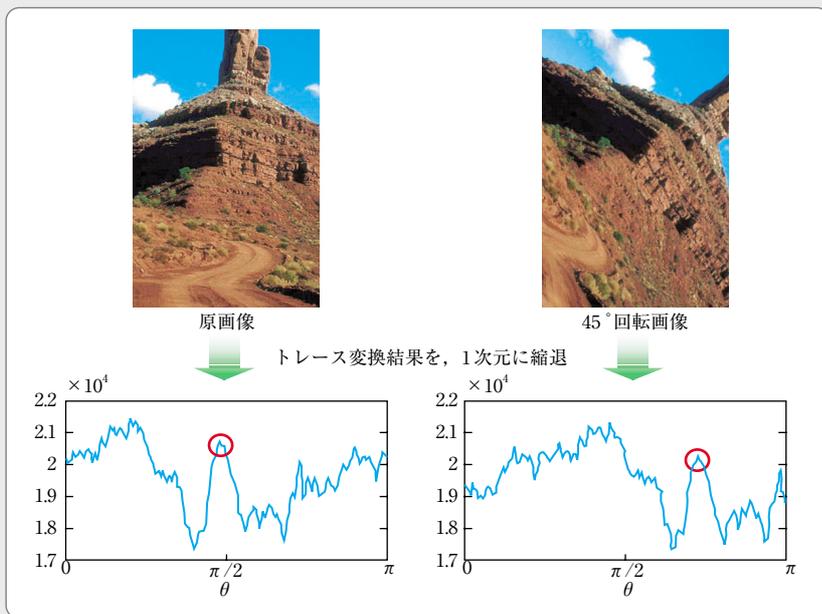


図2 回転画像と1次元縮退したトレース変換結果例

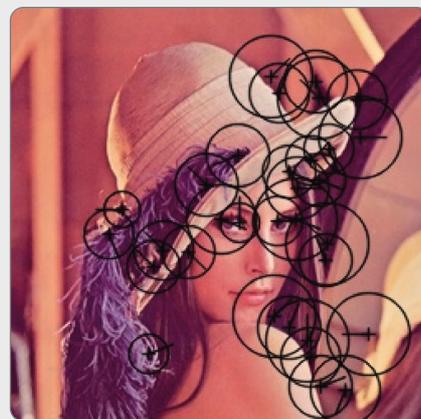


図3 Local Signature特徴点の例
(円の大きさは特徴量抽出範囲)

Video Signature

動画像の同定を行うVideo Signatureは、① Frame Signature, ② Word, ③ Bag-of-Words, の三つの特徴量から構成されます。

Frame Signatureは、映像中の全フ

レームに対して付与される特徴量であり、長時間の映像では演算処理が膨大になることから、Image Signatureに比べて簡易な抽出処理となっています。具体的には、画像フレームをさまざまなパターンの領域に分割し、各領域の輝度平均値や領域間の輝度差分値

から380次元の特徴ベクトルを取得します。図4に輝度平均値の取得領域の例を、図5に輝度差分取得領域の例を示します。得られた特徴ベクトルを予め規定されたテーブルにしたがって符号化することで、608ビットのFrame Signature特徴量を得ます。 ㊦

☞ 予測符号化による特徴量の圧縮もオプションとして用意されています。

Wordは、40ビットで表現するFrame Signatureのサブセットです。

Bag-Of-WordsはWordのヒストグラムです。図6に示すように、45フレーム間隔で交互にヒストグラムを作成することで、長い映像からごく短くカットされた映像まで同定可能です。

静止画像の同定と異なり、動画の同定においては時間同期が必要になります。Video Signatureによるマッチングでは、次の二つの同期方法があります⁴⁾。

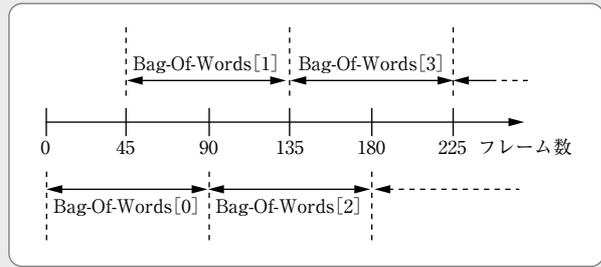


図6 Bag-Of-Wordsの抽出単位

(1) Bag-Of-Wordsで大まかな同期をとり、Frame Signatureマッチングで詳細な同期をとる方法

(2) Frame Signatureでキーフレームを選びキーフレーム同士を比較する方法

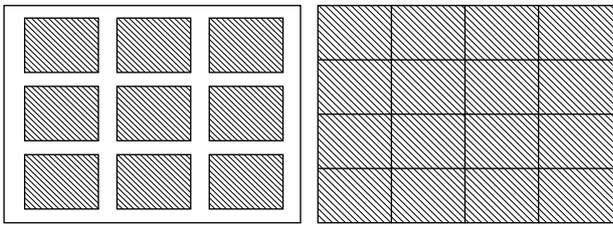


図4 Frame Signature抽出における輝度平均値算出領域の例
(斜線領域それぞれで平均値取得)

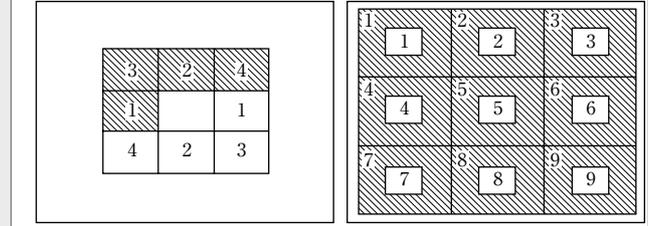


図5 Frame Signature抽出における輝度平均差分算出領域の例
(対応する番号の斜線領域と白抜き領域で差分値取得)

むすび

画像・映像を取り扱うパソコン、デジタルカメラ、スマートフォン等の性能向上、インターネットの爆発的な普及と通信回線の高速化・無線化により、「誰でも」、「いつでも」、「どこでも」大量の画像にアクセスすることが

可能となりました。これにより、画像検索技術全般に対する期待・要求が高まっています。

より一般的な物体や映像のシーン認識など、画像検索にはまだまだ多くの課題があり、今後、映像コンテンツ提供サービスを行う上での重要な技術といえます。

その中で、今回紹介したImage/Video Signatureは非常に高いロバスト性を有する画像同定技術であり、画像の著作権管理の他に、画像のフィルタリングやクラスタリング、効率的な画像アーカイブ構築への応用も期待されています。
(2012年4月27日受付)

参考文献

- 1) ISO/IEC 15938-3:2002/Amd 3:2009
- 2) ISO/IEC 15938-3:2002/Amd 4:2010
- 3) A. Kadyrov, M. Petrou, IEEE Trans. Pattern Analysis and Machine Intelligence, 23, 8, pp.811-828 (2001)
- 4) ISO/IEC TR 15938-8:2002/Amd 6:2011



くどう だいき
工藤 大樹 2001年、新潟大学工学部電気電子工学科卒業。2003年、東北大学大学院工学研究科電気・通信工学専攻博士前期課程修了。同年、三菱電機(株)に入社、画像符号化・画像処理・画像検索の研究に従事。正会員。



にしかわ ひろふみ
西川 博文 1989年、北海道大学工学部生体工学専攻修了。同年、三菱電機(株)に入社、画像符号化伝送・画像処理技術の研究開発、MPEG標準化活動に従事。正会員。