

# Earth Mover's Distance を用いたロバスト画像照合における 特徴量の符号特性

Sign Characteristic of the Features for Robust Image Matching  
by Earth Mover's Distance

橋本 陽祐†

Yosuke Hashimoto

川上 肇‡

Hajimu KAWAKAMI

## 1 はじめに

画像照合では、照合対象の二つの画像に対して求めた特徴量間の相違度で画像間の違いを測る。その際、Earth Mover's Distance(EMD) は処理のロバスト化に有効であることが知られている [2]。EMD は一般化されたヒストグラム間に定義される距離であるため、EMD をテキスト識別に応用するのに、基底関数群で画像を展開し、基底関数を特徴量、展開係数の絶対値をその特徴量の頻度としてそれぞれ対応付ける方法も提案されている [1]。しかしながらこの方法では、展開係数の符号情報が有効に利用されないことが問題となる。そこで我々は、展開係数の符号によって基底空間を分割する方法を提案する。この方法を画像照合に応用して、照合法に備わるロバスト性と基底の分割方法との関係を調べたので報告する。

## 2 画像照合システム

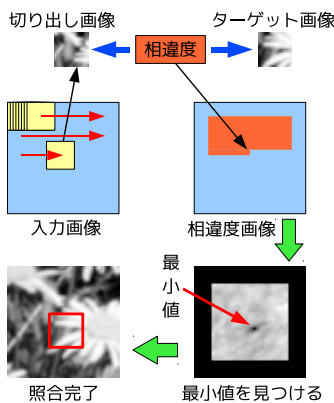


図1: 照合の手順

視野の中に部分画像として写るターゲットの位置を求める問題を画像照合と呼ぶ。この画像照合ではまず、比較的大きな入力画像からラスト順に部分画像を切り出し、その部分画像とターゲット画像との相違度を求め、この値が最小となる位置がターゲット画像と一致する部分であると考え、ここで画像照合では部分画像を認識する為に、ターゲット画像との相違度を測る処理が重要となる。その処理の為に、基底要素群  $\{Fe.k\}$  の線形

和  $\sum a_k Fe.k$  でこれらの画像を最小二乗近似し、成分群  $\{a_k\}$  を得る。これらのセットを画像の Signature とし、照合画像対のそれぞれに対応する Signature から EMD を計算し、画像間の相違度とする。

$$\text{画像} \approx a_1 Fe.1 + a_2 Fe.2 + \dots + a_k Fe.3 + \dots + a_n Fe.4$$

図2: 最小二乗近似

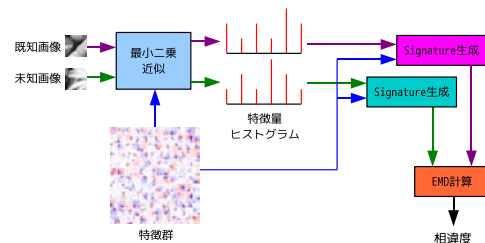


図3: 計算の全体像

## 3 符号情報の活用

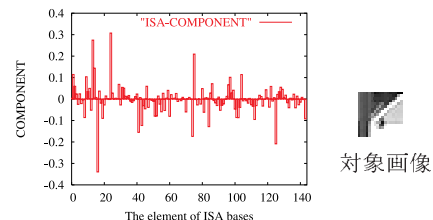


図4: 近似した各成分の例

画像を近似した際の各成分の符号は図4の様に正負に亘って一様に分布する。EMD で扱えるのは正の符号の頻度である。したがって EMD に応用する場合この成分を頻度と適切に対応づけた特徴量とすることが重要である。その方法を示す。

(方法1) 各成分を絶対値順にソートし、この値の大きな上位  $\alpha\%$  の成分を抽出し、基底を正負で2つに分割し、負の成分を絶対値に置き換える。各基底での成分の総和を1に正規化した後に、正負に分けて EMD を計算する。二つの EMD の値の和を相違度とする。

(方法2) 図5の様に値が負であった成分に対応する基底要素  $Fe.k$  を反転することで、成分  $a_k$  が正であったことにする。この値の大きな上位  $\alpha\%$  の成分を用いる。これらの総和を1に正規化した後に、反転した基底要素を

† 龍谷大学大学院 理工学研究科 電子情報学専攻  
‡ 龍谷大学 理工学部 電子情報学科

用いて EMD を計算する。

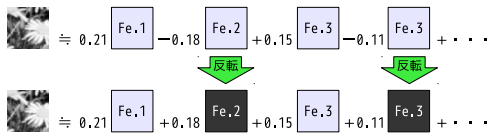


図 5: 負の成分の基底要素を反転

比較のため下記を従来法とする。  
 (従来法) 各成分を絶対値に置き換え、この値の大きな上位  $\alpha$  % の成分を用いる。それらの総和を 1 に正規化した後に EMD を計算する。

#### 4 実験

RGB 色空間で変動が混入する画像を変動の無い画像と照合する処理を、変動の大きさを変化させて行うことにより、方法 1 および方法 2 を用いた場合の画像照合法に備わるロバスト性を測る実験を行った。

##### 4.1 方法

(1) 基底の抽出: まず  $153 \times 153$  画素の画像を 25 枚用意し、画素値の分布が平均 0・分散 1 となるように各画像を正規化した後、無作為に選んだ 400ヶ所の位置で  $17 \times 17$  画素の部分画像を各画像から切り出す。これら合計  $10^4$  個の部分画像群をサンプル画像として独立部分空間分析 (ISA) に基づき ISA 基底を抽出する。そこから最大値ノルム最小化法を用いて情報統合基底  $W$  を生成する [3]。図 6 は生成した  $W$  の要素である。(2) ターゲット画像:

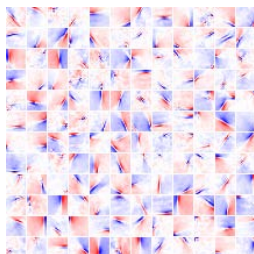


図 6:  $W$  の要素

図 7 に例示する各原画像 (大きさ  $51 \times 51$  画素で 8bit の明度画像、10 種類) から大きさ  $17 \times 17$  画素の部分画像 1 枚を切り出しターゲット画像とした。ターゲット画像の例を図 8 に示す。(3) 変動が混入する画像: 最大振幅  $z$  の変動を乱数に基づき生成し、それを原画像群に RGB 色空間で加算して入力画像群を合成した。この  $z$  を変動強度と呼ぶ。 $z = 170$  の明度画像例を図 9 に示す。(4) 照合: ターゲット画像  $P_j$  と同じ大きさの部分画像  $Q_k$  を入力画像群の全ての位置で切り出し、方法 1 と方法 2 の EMD を用いてターゲット画像との相違度  $M_{jk}$  を測る。このとき各基底は  $\alpha = 70\%$  として生成する。変動の初期値を変えてこの照合を 10 回繰り返した。(5) 評価: 上記の繰り返し毎に照合法に備わる識別力

$$C = \frac{\text{異種間での } M_{jk} \text{ の最小値}}{\text{同種間での } M_{jk} \text{ の最大値}} \quad (1)$$



図 7: 原画像の例 (10 種中の 5 種)



図 8: ターゲット画像の例



図 9: 入力画像 ( $z=170$ ) の例

を算出する。この  $C$  は 1 未満のとき単純な閾値処理で誤認識が生じることを表すので、 $C - 1$  の分布に対する負方向の安定性を測る指標

$$R_M = \frac{m_{C_M} - 1}{\sigma_{C_M}} \quad (2)$$

で照合法に備わるロバスト性を評価する。ここで  $m_{C_M}$  は  $C$  が平均以下である識別力  $C$  の平均値、 $\sigma_{C_M}$  はそれらの標準偏差である。

#### 5 結果と考察

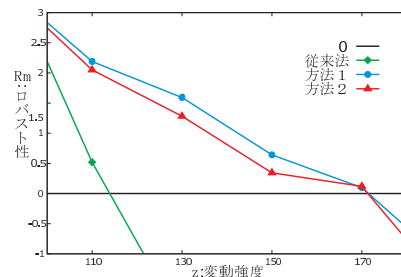


図 10: 各方法でのロバスト性

図 10 に変動強度に対する提案方の  $R_M$  を示す。この結果を評価するために従来法の結果も併せて同時に示す。同図より、提案方の  $R_M$  が負に転ずる  $z$  の値は 170 付近であるのに対して従来法では  $z = 110$  付近である事が分かる。これにより提案法はより大きな変動に対して正しく照合できる可能性が備わる事が分かる。

#### 6 むすび

特徴量の符号情報を適切に導入することにより EMD をロバスト化する方法を示した。今後の課題は 2 つの方法の解析である。

#### 参考文献

- [1] Yossi Rubner, Carlo Tomasi, "PERCEPTUAL METRICS FOR IMAGE DATABASE NAVIGATION", Kluwer Academic Publishers, 2001, p.43.
- [2] Richard Szeliski, "Computer Vision", Springer, 2011, p.613.
- [3] 橋本 陽祐, 川上 肇 "均等分割型 EMD によるロバスト画像照合" Fit2010, H-014.