

H-023

位相限定相関法に基づく一般画像照合アルゴリズムに関する検討 A Study on a Generic Image Matching Algorithm using Phase-Only Correlation

飯塚 智†
Satoshi Iitsuka

宮澤 一之†
Kazuyuki Miyazawa

青木 孝文†
Takafumi Aoki

1. はじめに

雑誌、商品パッケージ、ポスター、パブリックスクリーンなどの実世界のオブジェクトとオンラインコンテンツとの連携は、ユビキタスコンピューティングの有望なアプリケーションの1つである[1]。近年、カメラ付き携帯電話の普及に伴い、QRコードに代表されるようなビジュアルタグをカメラで撮影することで実世界オブジェクトとオンラインコンテンツとの連携を実現するサービスが広く普及しつつある。カメラ付き携帯電話を用いてビジュアルタグを撮影し、URLを表示させることでウェブサイトへのアクセスを容易にしたり、個人データを格納したタグを名刺に印刷し、携帯電話へのアドレス帳登録を容易にするなどといった応用がすでに実現されている。しかし、こうしたビジュアルタグの多くは、そもそも情報メディア上へ掲示されることを意図せずに設計されたため、掲示対象の美観を損なうだけでなく、一定のスペースを無意味に占有してしまうといった問題がある。

こうした問題に対し、本論文では、対象に特殊な情報を付加することなく、カメラを入力デバイスとして対象から任意の情報を取り出すことを考える。これを実現するために利用可能な技術的枠組みとしては、“一般画像認識”[2]–[4]が挙げられる。これは、制約のない実世界のシーンの画像に対して計算機がその中に含まれる物体を一般的な名称で認識することであり、40年以上にわたって盛んな研究が行われている。この枠組みを利用すれば、あらかじめ対象とそれに対応する情報をサーバなどに登録しておき、ユーザから送られてくるカメラ画像中にその対象が写っていればそれに対応する情報をユーザに返す、といった応用が可能となる。しかし、一般画像認識では、画像に生じる様々な曖昧性に対して頑健になるようにアルゴリズムが設計されていることがほとんどであり、例えばある程度似通った2つの対象を撮影した場合には同じ対象を撮影したものとみなされてしまう。そのため、ビジュアルタグの代用においては不向きな部分がある。

そこで、本論文では、実世界とサイバースペースとのリンクを登録された画像と撮影された画像とのマッチングによってシンプルに実現することを試みる。提案手法では、どのような画像に対しても曖昧性なくロバストなマッチングが可能な手法として、位相限定相関法 (Phase-Only Correlation: POC) [5], [6] を利用する。これは、画像の2次元離散フーリエ変換から得られる位相成分のみを用いて画像マッチングを行う手法であり、画像の輝度変化などの外乱に強く、多岐にわたる応用が実現されている。エッジやコーナーなど、画像中の特徴点に基づいてマッチングを行う手法とは異なり、明確な特徴が存在しないような画像でも正確にマッチングを行うことができる。また、位相成分の量

子化に対してマッチング性能の低下が起これにくいという特長を持つため、量子化した位相を登録データとすることで、画像をそのまま登録する場合に比べて登録容量を大幅に削減することが可能である。本論文では、実際に様々な物体を撮影することで構築した画像データベースを用いて提案手法の性能評価を行い、その有効性を示す。構築したデータベースに格納されている画像間には平行移動、回転、拡大縮小が生じているが、相似変換から逸脱するような変形はなるべく生じないように撮影している。さらに、位相成分の量子化レベルを変化させながら実験を行うことで、マッチング性能とのトレードオフについても議論する。

2. 位相限定相関法

ここでは、提案手法の基本技術であるPOCの定義について述べる。大きさ $N_1 \times N_2$ 画素の画像を $f(n_1, n_2)$ と $g(n_1, n_2)$ とする。ただし、定式化の便宜上、離散空間のインデックスを $n_1 = -M_1, \dots, M_1$ ($M_1 > 0$) および $n_2 = -M_2, \dots, M_2$ ($M_2 > 0$) とし、画像の大きさを $N_1 = 2M_1 + 1$ 画素および $N_2 = 2M_2 + 1$ 画素とする。これらの画像の2次元離散フーリエ変換 (2次元DFT) をそれぞれ $F(k_1, k_2)$ および $G(k_1, k_2)$ とし、次式で与える。

$$\begin{aligned} F(k_1, k_2) &= \sum_{n_1, n_2} f(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_F(k_1, k_2) e^{j\theta_F(k_1, k_2)} \end{aligned} \quad (1)$$

$$\begin{aligned} G(k_1, k_2) &= \sum_{n_1, n_2} g(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_G(k_1, k_2) e^{j\theta_G(k_1, k_2)} \end{aligned} \quad (2)$$

ただし、 $W_{N_1} = e^{-j\frac{2\pi}{N_1}}$ 、 $W_{N_2} = e^{-j\frac{2\pi}{N_2}}$ であり、 \sum_{n_1, n_2} は $\sum_{n_1=-M_1}^{M_1} \sum_{n_2=-M_2}^{M_2}$ を意味する。ここで、 $A_F(k_1, k_2)$ および $A_G(k_1, k_2)$ はそれぞれの画像の振幅成分、 $e^{j\theta_F(k_1, k_2)}$ および $e^{j\theta_G(k_1, k_2)}$ はそれぞれの画像の位相成分である。一般性を失うことなく離散周波数のインデックスを $k_1 = -M_1, \dots, M_1$ 、 $k_2 = -M_2, \dots, M_2$ とすることができる。 $F(k_1, k_2)$ と $G(k_1, k_2)$ の正規化相互パワースペクトル $R_{FG}(k_1, k_2)$ を次のように定義する。

$$\begin{aligned} R_{FG}(k_1, k_2) &= \frac{F(k_1, k_2) \overline{G(k_1, k_2)}}{|F(k_1, k_2) \overline{G(k_1, k_2)}|} \\ &= e^{j\theta(k_1, k_2)} \end{aligned} \quad (3)$$

ここで、 $\overline{G(k_1, k_2)}$ は $G(k_1, k_2)$ の複素共役である。また、 $\theta(k_1, k_2)$ は、 $\theta_F(k_1, k_2) - \theta_G(k_1, k_2)$ で与えられる。POC 関数 $r_{fg}(n_1, n_2)$ は、 $R_{FG}(k_1, k_2)$ の2次元離散

† 東北大学大学院情報科学研究科

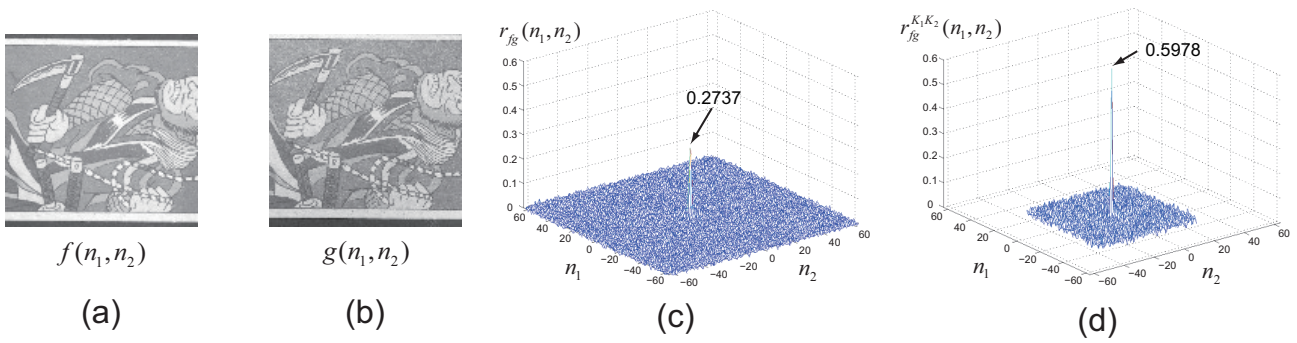


図 1: POC と BLPOC を使って同一物体同士をマッチングした例 : (a) 入力画像 $f(n_1, n_2)$, (b) 登録画像 $g(n_1, n_2)$, (c) POC 関数 $r_{fg}(n_1, n_2)$, (d) BLPOC 関数 $r_{fg}^{K_1K_2}(n_1, n_2)$ ($K_1/M_1 = K_2/M_2 = 0.5$)

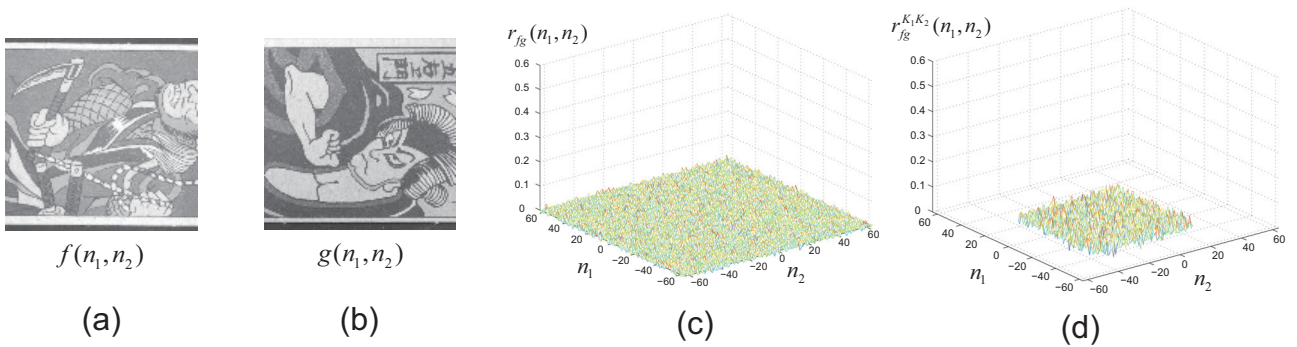


図 2: POC と BLPOC を使って異なる物体同士をマッチングした例 : (a) 入力画像 $f(n_1, n_2)$, (b) 登録画像 $g(n_1, n_2)$, (c) POC 関数 $r_{fg}(n_1, n_2)$, (d) BLPOC 関数 $r_{fg}^{K_1K_2}(n_1, n_2)$ ($K_1/M_1 = K_2/M_2 = 0.5$)

フーリエ逆変換 (2 次元 IDFT) として次のように表される .

$$r_{fg}(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{k_1, k_2} R_{FG}(k_1, k_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2} \quad (4)$$

ここで、 \sum_{k_1, k_2} は $\sum_{k_1=-M_1}^{M_1} \sum_{k_2=-M_2}^{M_2}$ を意味する . 2 つの画像が類似している場合、POC 関数 $r_{fg}(n_1, n_2)$ は、デルタ関数に近い極めて鋭いピークを有する . この相関ピークの高さは 2 つの画像の位相差スペクトルの線型性を表しており、位相差スペクトルが周波数に対して線形であれば、相関ピークの高さは 1 となる . 一方、異なる 2 つの画像間では、全くピークが現れない . POC 関数のピークの高さは、画像間の類似度を調べる指標となる . この特徴以外にも、POC 関数は、画像間の位置ずれ (平行移動) や明るさの変化、加算ノイズに対してロバストである [7] .

以下では、本論文で提案する手法で重要となる POC を用いた高精度画像マッチング [6] について概説する .

2.1 サブピクセル精度の平行移動量推定

POC を用いて高精度に平行移動量を推定するためには、(i) 離散データで得られた POC 関数への相関ピークモデルのフィッティング、(ii) 画像端での信号の不連続性を解消するための窓関数の適用、(iii) エイリアシングやノイズの影響を抑制するためのスペクトル重み付けが重要となる [8] . これらの高精度化手法を用いることによって、サブピクセル精度の平行移動量を求めることができる .

2.2 相似変換パラメータの推定

2 つの画像が相似変換の関係にある場合は、平行移動量を拡張することで、高精度に変換パラメータ (平行移動量、回転角度、拡大縮小率) を推定することが可能である [8], [9] . 画像の回転角度と拡大縮小率は、画像の振幅スペクトルの対数極座標変換 (Fourier-Mellin 変換) を利用して 2 次元の平行移動量に置き換えて推定する . まず、2 つの画像 $f(n_1, n_2)$ と $g(n_1, n_2)$ の 2 次元 DFT $F(k_1, k_2)$ と $G(k_1, k_2)$ をそれぞれ求める . 次に、それぞれの振幅スペクトル $|F(k_1, k_2)|$ と $|G(k_1, k_2)|$ を求める . 振幅スペクトルは回転と拡大縮

小のみの影響を受けるため、平行移動の影響を排除することができる。自然画像では、そのエネルギーの大部分が低周波領域に集中し、高周波成分のエネルギーは相対的に小さいことが知られている。そのため、実際には、 $|F(k_1, k_2)|$ と $|G(k_1, k_2)|$ の代わりに、 $\log\{|F(k_1, k_2)| + 1\}$ と $\log\{|G(k_1, k_2)| + 1\}$ を用いて、高周波成分を強調しておく。強調された振幅スペクトルの対数極座標変換 (Fourier-Mellin 変換) を行い、 $F_{LP}(k'_1, k'_2)$ および $G_{LP}(k'_1, k'_2)$ を求める。ここで、 k'_1 と k'_2 は変換された画像のインデックスを表している。 $F_{LP}(k'_1, k'_2)$ と $G_{LP}(k'_1, k'_2)$ の間の水平方向の平行移動量は画像間の回転角度に対応し、垂直方向の平行移動量は画像間の拡大縮小率に対応している。そこで、 $F_{LP}(k'_1, k'_2)$ と $G_{LP}(k'_1, k'_2)$ の間の平行移動量を 2.1 で述べた方法により求め、画像の回転角度と拡大縮小率を算出する。

2.3 帯域制限位相限定相関関数による類似度評価

信頼性の低い高周波成分の影響を排除しつつ、画像の識別性能を向上させるために、帯域制限位相限定相関 (Band-Limited Phase-Only Correlation: BLPOC) 関数が提案されている [7]。BLPOC 関数は $R_{FG}(k_1, k_2)$ の 2 次元 IDFT のサイズ自体を、画像テクスチャの有効帯域に制限することによって、相関ピークのエネルギーを集中させ、画像の識別性能を向上させる。BLPOC 関数は次式で定義される。

$$r_{fg}^{K_1 K_2}(n_1, n_2) = \frac{1}{L_1 L_2} \sum_{k_1=-K_1}^{K_1} \sum_{k_2=-K_2}^{K_2} R_{FG}(k_1, k_2) \times W_{L_1}^{-k_1 n_1} W_{L_2}^{-k_2 n_2} \quad (5)$$

ここで、 $n_1 = -K_1, \dots, K_1, n_2 = -K_2, \dots, K_2, L_1 = 2K_1 + 1, L_2 = 2K_2 + 1$ である。BLPOC は、制限する周波数帯域の大きさに依存せず、常に相関ピークの最大値が 1 に正規化される。

実際に、POC 関数 $r_{fg}(n_1, n_2)$ と BLPOC 関数 $r_{fg}^{K_1 K_2}(n_1, n_2)$ を用いて同一物体同士のマッチングと異なる物体同士のマッチングを行った例を図 1 と図 2 に示す。POC 関数より BLPOC 関数を用いた方が高い識別性能を示すことがわかる。

3. 一般画像照合アルゴリズム

ここでは、POC を用いた一般画像マッチングアルゴリズムについて説明する。提案アルゴリズムは、(i) 回転、拡大縮小補正、(ii) マッチングスコア算出の 2 ステップから構成される。それぞれの処理について以下で述べる。さらに、位相量子化による登録データ容量削減の方法について説明する。

3.1 回転、拡大縮小補正

ここでは、登録画像と入力画像の間の回転角度、拡大縮小率を求め、補正する処理について説明する。正確な画像マッチングを行うためには、正確かつ高精度に画像間の回転角度と拡大縮小率を求める必要がある。

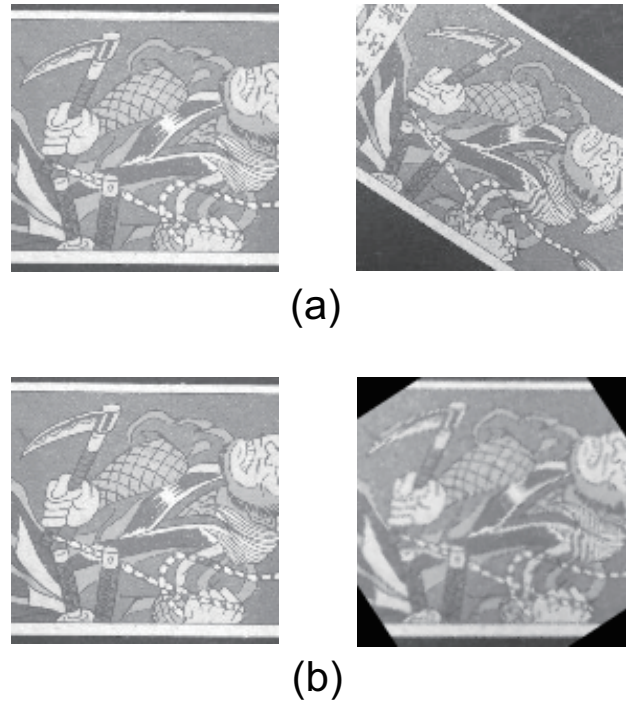


図 3: 回転、拡大縮小の補正: (a) 入力画像 (左図) と登録画像 (右図), (b) 回転、拡大縮小補正後の画像 (左図) と登録画像 (右図)

そこで、本論文では、2.2 で述べた POC に基づく高精度な相似変換パラメータ推定手法を用いる。算出された回転角度と拡大縮小率を用いて、登録画像に対して入力画像の回転角度と拡大縮小の補正を行う。図 3 に、回転と拡大縮小を補正した画像を示す。

3.2 マッチングスコア算出

信頼性の低い高周波成分の影響を排除しつつ、画像の識別性能を向上させるために、2.3 で述べた BLPOC 関数を利用してマッチングスコアを算出する。回転角度と拡大縮小が補正された画像間の BLPOC 関数を求め、相関ピークの最大値をマッチングスコアとする。本論文では、有効帯域を決定するための BLPOC 関数のパラメータは、 $K_1/M_1 = K_2/M_2 = 0.5$ ($M_1 = M_2 = 64, K_1 = K_2 = 32$) とした。

3.3 位相量子化による登録データ容量の削減

ここでは、登録データ容量の削減手法について述べる。具体的には、画像を 2 次元離散フーリエ変換して得られる位相を量子化することで登録データ容量を削減する。まず、原画像 (図 4 (a)) を 2 次元離散フーリエ変換して得られる振幅スペクトル (図 4 (b)) を対数極座標展開する (図 4 (c))。次に、対数極座標展開されたスペクトルと原画像をそれぞれ 2 次元離散フーリエ変換して得られる位相を計算する (図 4 (d) と (e))。周波数領域では、位相は原点对称な情報として得られるため、それぞれ第 1 象限と第 4 象限の位相のみを登録すればよい。位相は、 $-\pi$ から $+\pi$ までの値をとるが、この値を量子化しても位相限定相関法による画像マッ

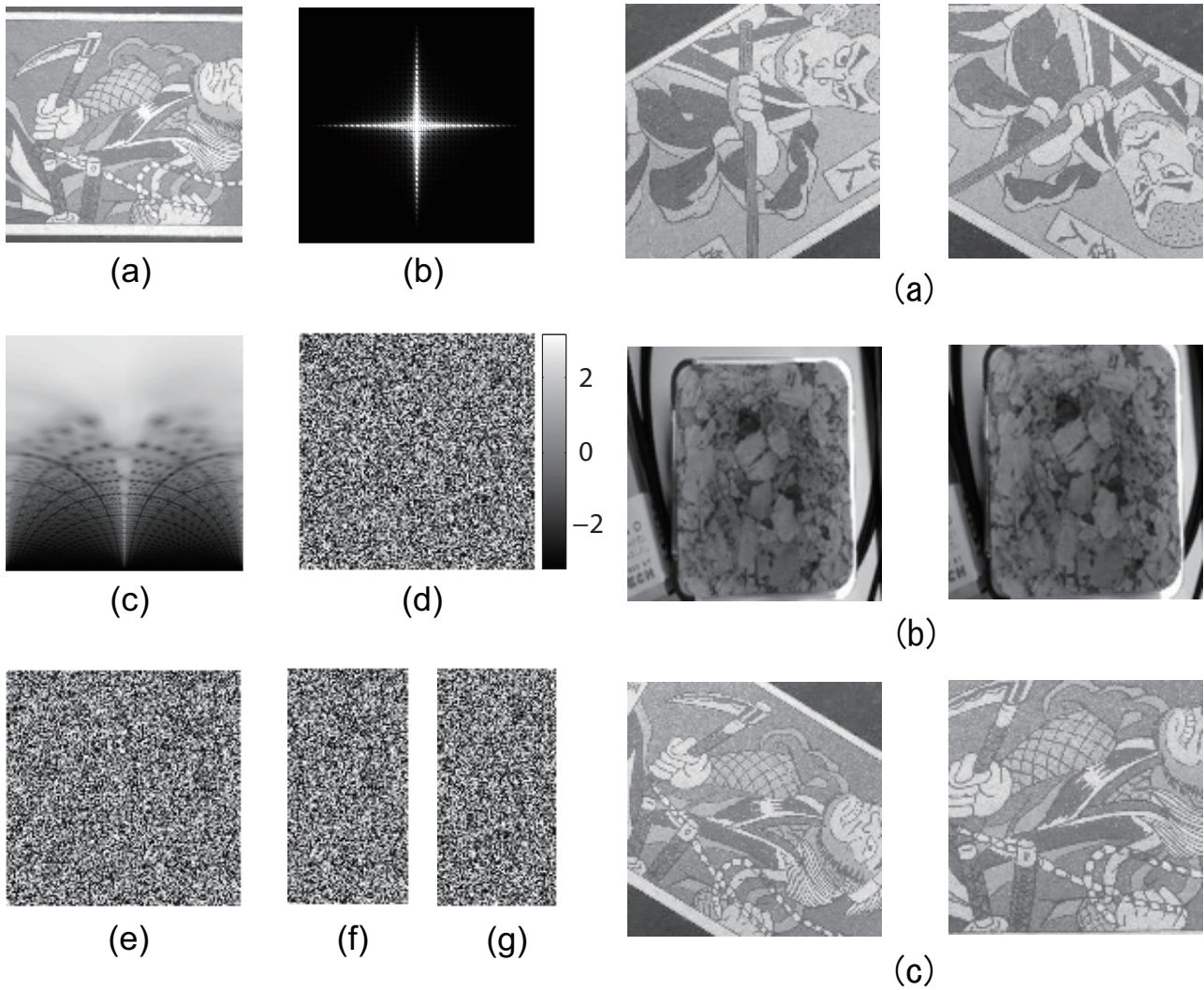


図 4: 登録データ生成: (a) 原画像, (b) 原画像の振幅スペクトル, (c) 対数極座標展開後のスペクトル, (d) 対数極座標後のスペクトルの位相成分, (e) 原画像の位相成分, (f) 登録データ (回転, 拡大縮小の補正に利用), (g) 登録データ (マッチングスコア算出に利用)

チングの性能は大きな影響を受けないことが報告されている [10], [11]. そこで, それぞれの位相の第 1 象限と第 4 象限, かつ低域から縦軸方向に K_1 , 横軸方向に K_2 までの位相を量子化したものを登録データとして利用する (図 4 (f) と (g)). 画像の回転, 拡大縮小の補正するために, 対数極座標展開後のスペクトルの位相成分を利用する. また, マッチングスコアを算出するために, 原画像の位相成分を利用する.

4. 実験・考察

提案アルゴリズムの性能を評価するために, 一般画像データベースを用いて性能評価実験を行った. 今回の実験では, 一般に市販されているコンパクトデジタルカメラを用いて様々な物体を撮影した画像データベースを作成した. このデータベースには, 119 個の物体を

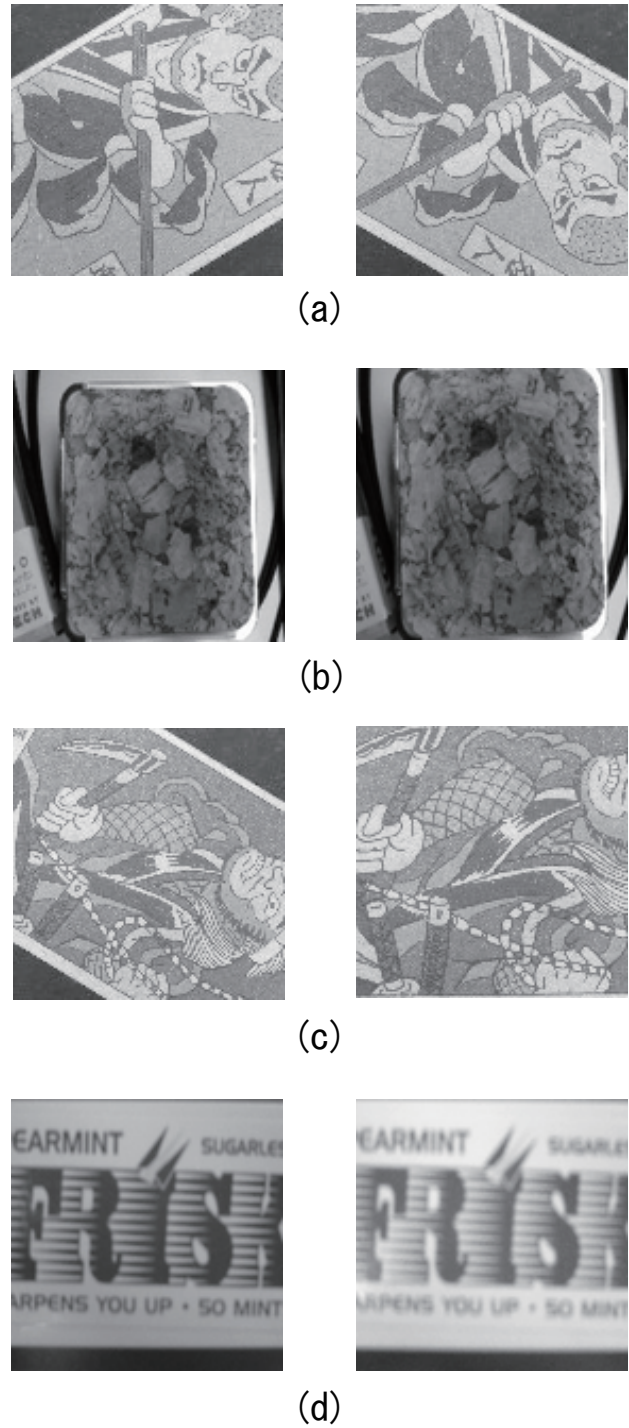


図 5: データベースに格納されている画像の例: (a) 回転の含まれる画像ペア, (b) 拡大縮小が含まれる画像ペア, (c) 回転と拡大縮小が含まれる画像ペア, (d) 明るさの異なる画像ペア

異なるタイミングで 10 回撮影することにより得られた計 1,190 枚 (119 物体 × 10 回) の画像が格納されている. 10 回の撮影では得られた画像間には平行移動, 回転, 拡大縮小が生じているが, 相似変換から逸脱するような変形はなるべく生じないように撮影している. 図

5 に、データベースに格納されている画像の例を示す。

提案アルゴリズムの照合性能を評価するために、照合アルゴリズムの性能評価に一般的によく用いられている EER (Equal Error Rate) を用いる [12]。EER とは、FNMR (False Non-Match Rate) と FMR (False Match Rate) が等しくなるときのエラー率であり、EER が小さいほど性能が高いことを意味する。まず、FNMR を求めるために、同一物体同士におけるすべての組み合わせ ${}_{10}C_2 \times 119 = 5,355$ ペアについてマッチングスコアを求める。続いて、FMR を求めるために、異なる物体同士におけるすべての組み合わせ ${}_{1190}C_2 - 5355 = 702,100$ ペアについてマッチングスコアを求める。そして、EER により照合性能を評価する。

本論文では、登録画像の振幅スペクトルを対数極座標展開した後のスペクトルの位相成分と原画像の位相成分をそれぞれ 2 ビット、4 ビット、6 ビット、8 ビットと変化させたときの性能を調べる。評価する性能は、(i) 照合性能、(ii) 登録容量、(iii) 処理時間の 3 つである。実験結果を表 1 に示す。

提案アルゴリズムの EER は、位相の量子化レベルにより変化する。提案アルゴリズムにおいて、EER が最小となるのは原画像の位相の量子化レベルを 4 ビット、原画像の振幅スペクトルを対数極座標展開した後のスペクトルの位相の量子化レベルを 6 ビットとしたときであり、3.577% である。提案アルゴリズムにおいて、画像を登録データとした場合の登録データ容量は 16,384 [byte] である。処理時間は、Intel Core 2 Duo E6850 (3.00 GHz \times 2) 上の MATLAB 7.2.0 で評価した。提案アルゴリズムにおいて、1 回のマッチングに要する時間は、0.060 秒である。以上により、提案アルゴリズムは、照合性能を大きく低下させることなく、位相成分を量子化することで登録容量を削減することができ、かつ、マッチングに要する時間も高速であるため、実用的な照合アルゴリズムである。

5. むすび

本論文では、位相限定相関法 (Phase-Only Correlation: POC) を用いた一般画像照合アルゴリズムを提案した。提案アルゴリズムは、画像を 2 次元離散フーリエ変換して得られる位相成分を用いることで、明るさの変化やノイズなどに対してロバストな画像マッチングが可能である。また、位相成分を量子化することで認証性能の低下を防ぎつつ、データベースに登録する容量を削減することができる。本論文では、実際にさまざまな物体を撮影することで構築した画像データベースを用いて提案手法の有効性を確認した。今後は、より一般的な射影変形などに対応したアルゴリズムを検討することや一般画像マッチングを利用したアプリケーションを開発する予定である。

参考文献

[1] J. Rekimoto and K. Nagao: "The world through the computer: computer augmented interaction with real world environments", Proc. UIST'95, pp. 29-36 (1995).

表 1: 実験結果

原画像の位相の量子化レベル [bit]	対数極座標展開後のスペクトルの位相の量子化レベル [bit]	EER [%]	容量 [byte]
2	2	3.906	1,024
2	4	3.744	1,536
2	6	3.617	2,048
2	8	3.698	2,560
4	2	3.827	1,536
4	4	3.701	2,048
4	6	3.577	2,560
4	8	3.661	3,072
6	2	3.784	2,048
6	4	3.720	2,560
6	6	3.585	3,072
6	8	3.621	3,584
8	2	3.783	2,560
8	4	3.734	3,072
8	6	3.595	3,584
8	8	3.660	4,096

- [2] 柳井啓司: "一般物体認識の現状と今後", コンピュータビジョン・イメージメディア, 48, SIG16(CVIM19), pp. 1-24 (2007).
- [3] P. Singla and P. Domingos: "Object identification with attribute-mediated dependences", In Proceedings of the 9th European Conference on Principles and Practice of knowledge Discovery in Database (PAKDD-2005), pp. 297-308 (2005).
- [4] S. Tejada, C. Knoblock and S. Minton: "Learning domain-independent string transformation weights for high accuracy object identification", In Proceedings of the 8th European Conference on Principles and Practice of knowledge Discovery in Database (PAKDD-2002), pp. 350-359 (2002).
- [5] C. D. Kuglin and D. C. Hines: "The phase correlation image alignment method", Proc. Int. Conf. Cybernetics and Society, pp. 163-165 (1975).
- [6] 青木孝文, 伊藤康一, 柴原琢磨, 長嶋聖: "位相限定相関法に基づく高精度マシンビジョン—ピクセル分解能の壁を越える画像センシング技術を目指して—", IEICE Fundamentals Review, 1, 1, pp. 30-40 (2007).
- [7] K. Ito, H. Nakajima, K. Kobayashi, T. Aoki and T. Higuchi: "A fingerprint matching algorithm using phase-only correlation", IEICE Trans. Fundamentals, **E87-A**, 3, pp. 682-691 (2004).

- [8] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi and K. Kobayashi: “High-accuracy image registration based on phase-only correlation and its experimental evaluation”, Proc. IEEE Int. Symp. Intelligent Signal Processing and Communication Systems, pp. 86–90 (2002).
- [9] Q. Chen, M. Defrise and F. Deconinck: “Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition”, IEEE Trans. Pattern Anal. Machine Intell., **16**, 12, pp. 1156–1168 (1994).
- [10] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi and H. Nakajima: “An effective approach for iris recognition using phase-based image matching”, IEEE Trans. Pattern Anal. Machine Intell., **30**, 10, pp. 1741–1756 (2008).
- [11] S. Iitsuka, K. Ito and T. Aoki: “A practical palmprint recognition algorithm using phase information”, Proc. 19th Int. Conf. Pattern Recognition (2008).
- [12] J. Wayman, A. Jain, D. Maltoni and D. Maio: “Biometric Systems”, Springer (2005).