
 論 文

 ラプラシアン・ガウシアン符号とベクトル量子化ネットを用いる
 座標変換にロバストな高速画像照合法

 澁川 康裕[†] 黒木 秀一[†]

 A Fast Image Matching Method Robust to Coordinate Transformations Using Sign
 of Laplacian of Gaussian and Vector Quantization Net

 Yasuhiro FUCHIKAWA[†] and Shuichi KUROGI[†]

あらかし 対象画像がノイズ、明度変化、座標変換などを含む場合にも任意形状の検索画像の位置を高速に探索する画像照合法としてラプラシアン・ガウシアン (LOG; Laplacian of Gaussian) 符号とベクトル量子化 (VQ; Vector Quantization) ネットを用いる手法を提案する。本手法ではまず原画像から LOG 画像を生成する。これは高周波及び低周波のノイズを除去すると同時に、ラプラシアンが最低次の等方性微分演算子であるという性質を回転等を含む座標変換処理に利用するためである。次に任意形状の検索画像を探索するための窓処理を行う。これは通常のテンプレート照合のように探索したい任意形状の検索画像を含む方形領域を探索する場合、検索画像の周囲が遮へいなどにより変化したとき照合結果が影響を受けるのでこの影響を除くためである。以上の処理後、LOG 画像を 2 値に符号化した画像の相関を検索画像の画素数で正規化した正規化 LOG 符号相関を用いて対象画像と検索画像を FFT 演算で高速に照合する。更に VQ ネットを用いて小数個のテンプレート画像を生成し座標変換に対してロバストな画像照合を行う手法を提案する。本論文では提案手法の有効性、特に高速かつ効率的な画像照合が可能であることを実験により示す。

キーワード LOG 符号, テンプレートマッチング, ベクトル量子化, 高速画像照合

1. ま え が き

本論文ではノイズ、明度変化、座標変換などを含む対象画像から任意形状の検索画像の位置を高速に探索する画像照合法としてラプラシアン・ガウシアン (LOG; Laplacian of Gaussian) 符号 [1] とベクトル量子化 (VQ; Vector Quantization) ネット [2] を用いる手法を提案する。

対象画像の中から文字や記号などの特定のパターンを探索する手法として従来よりテンプレート画像照合すなわちテンプレートマッチングが広く用いられており様々な方法が提案されている [3]。画像の明度値を直接用いる正規化相関や残差二乗和 [4] は処理が簡明であり広く用いられているが、実際の使用環境では遮へいや明度変化等により対象画像が変化し照合性能が低下することがある [3]。そこで遮へいや明度変化に対し

てロバストな手法として、隣接画素の明度変化を表す増分符号を用いる増分符号相関 [5] や選択的正規化相関 [6] が提案されている。これらの手法では局所的な明度変化を用いることで局所的な遮へいや大域的な明度変化が照合結果に及ぼす影響を小さくしている。またブロック投票処理 [7] や標本化順位相関 [8] はテンプレート画像を複数の小領域に分割して各小領域ごとに照合処理を行うとともに遮へい等の例外部分を除くように統合処理することで遮へいや明度変化に対するロバスト性をもたせている。しかしこれらの手法では実際の応用においてしばしば生じる拡大縮小や回転等の座標変換に対するロバスト性があまり高くないことを示す結果が報告されている。すなわちブロック投票処理の拡大縮小と回転に対するロバスト性は増分符号相関よりも高いが正規化相関よりもわずかに高い程度であり [7]、標本化順位相関が用いる順位相関係数は正規化相関と同等以下となった [8] ことが報告されている。

一方、回転等の座標変換が施された多数のテンプレート (検索) 画像を効率的に探索する手法として、高

[†]九州工業大学工学部, 北九州市
 Faculty of Engineering, Kyushu Institute of Technology,
 Kitakyushu-shi, 804-8550 Japan

次元の検索画像を KL 変換 (Karhunen-Loeve Transform) により低次元の固有空間に投影した後, FFT (Fast Fourier Transform) により高速画像探索する手法が提案されている [9]. この手法の有効性はある範囲の回転を伴う画像に対する成功実験例を用いて示されているが, 一般には P 個の検索画像が $K (< P)$ 次元固有空間の近くにあり検索画像でない画像がその固有空間から遠くにある場合に適用でき, かつ K が P より十分小さいときのみ有効になる. これはこの手法が有効に適用できる画像や座標変換は限定的であることを意味するが, その適用可能範囲は明らかにされていない.

本論文では回転, 拡大縮小, 射影変換等の座標変換を伴う検索画像に適用でき, かつノイズや明度変化などにロバストな高速画像照合法を提案する. この手法ではまず原画像に LOG フィルタを施す. このフィルタはガウシアン関数の機能により原画像中の高周波の明度変化 (ノイズや局所的な明度変化等) を除去し, ラプラシアンの機能により低周波の明度変化 (バイアスや線形変化等) を除去する. 次に任意形状の検索画像を探索するための窓処理を行う. これは通常のテンプレート照合のように任意形状の検索画像を含む方形画像の探索を行うと対象画像中の検索画像の周囲が遮へいなどにより変化した場合にも照合結果が影響を受けるので, この影響を除くためである. 以上の処理後, LOG 画像の画素値の正負を ± 1 の 2 値に符号化した LOG 符号画像の相関を検索画像の画素数で正規化した正規化 LOG 符号相関 (2.3 参照) を用いると対象画像から検索画像を FFT 演算で高速に探索することができる. 最後に VQ ネットを用いて一般には無限個存在する座標変換画像の中から小数個の最適なテンプレート画像を生成することで座標変換に対してロバストな画像照合を行う.

ここで本手法で LOG 画像を用いる有効性と役割を要約しておく. (1) ラプラシアンは最低次の等方性微分演算子であり [1], 回転等の座標変換前後の正規化 LOG 符号相関の相似性 (3.2 参照) に寄与する. (2) LOG 画像あるいは LOG 符号画像のゼロ交差 (正と負の画素領域の境界) は原画像の主要な特徴 (対象物のエッジなど) を表す [1] ので LOG 符号画像を用いると原画像の特徴を保持しかつノイズや明度変化等にロバストな探索ができる. (3) LOG 符号画像は微少な座標変換に対してロバスト性をもつ (3.4 参照) ので必要とするテンプレート画像の個数を少なくすることが

できる. ところで LOG 画像の応用として LOG 画像を 3 値に符号化した画像を画像探索に用いる手法 [10] 等が提案され, その有効性が示されているが, 本手法のように座標変換に対するロバスト性に注目して検討したものはない.

以下, 本論文では LOG 符号, 窓処理, FFT 及び VQ を統合する手法を提案し, その有効性を実験結果を用いて示す. まず 2. で窓処理を含む正規化 LOG 符号相関とその高速処理について説明し, 次に 3. で VQ ネットを用いた座標変換画像の照合法について説明する. 最後に 4. で実験結果を示して検討するとともに本手法の有効性を示す.

2. 正規化 LOG 符号相関

2.1 画像の相関と FFT

以下では 2 次元画像 p を $p \triangleq p(x, y)$ または $p \triangleq p(\mathbf{x})$ のように表す. ここで $\mathbf{x} = (x, y)^T$ は画像中の位置を表す列ベクトルである. さて二つの整数 N と M に対し $N > M$ とし, 画素数がそれぞれ N^2 と M^2 の正方形画像 $p(\mathbf{x})$ ($\mathbf{x} \in I_p \triangleq \{0, 1, \dots, N-1\}^2$) と $q(\mathbf{x})$ ($\mathbf{x} \in I_q \triangleq \{0, 1, \dots, M-1\}^2$) の相互相関関数 (以下では単に相関関数ともいう) は

$$R(\mathbf{x}) \triangleq p * q \triangleq \sum_{\mathbf{i} \in I_q} p(\mathbf{x} + \mathbf{i})q(\mathbf{i}) \quad (1)$$

で定義されるが, これはフーリエ変換 F を用いて

$$p * q = F^{-1}(F^*(p)F(q)) \quad (2)$$

としても計算できる. ここで $F^*(p)$ は $F(p)$ の共役, F^{-1} はフーリエ逆変換を表す. したがって相関関数は FFT を 3 回用いて高速に計算できるが, このとき N と M のうち大きい N で計算し, N は 2 のべき乗とする. 式 (1) の右辺を直接計算するとその計算量は $O_{direct} = O(N^2M^2)$ であるが, 式 (2) 右辺の FFT での計算量は $O_{FFT} = O(3N^2 \log(N^2) + N^2)$ となる. 例えば, $N = 512$, $M = 64$ のとき FFT を用いた計算では直接計算するのと比べて計算量は約 1/75 となる.

実際の画像 $\tilde{p}(\mathbf{x})$ と $\tilde{q}(\mathbf{x})$ が方形画像であり各画素数 (横 \times 縦) が $\tilde{N}_1 \times \tilde{N}_2$ と $\tilde{M}_1 \times \tilde{M}_2$ のときは, \tilde{N}_i, \tilde{M}_i ($i = 1, 2$) 以上で最小の 2 のべき乗の整数を N とし, $N \times N$ の画像 $p(\mathbf{x})$ と $q(\mathbf{x})$ を次のように生成する. まず二つの窓領域を $W_p \triangleq \{0, 1, \dots, \tilde{N}_1 - 1\} \times \{0, 1, \dots, \tilde{N}_2 - 1\}$ 及び $W_q \triangleq \{0, 1, \dots, \tilde{M}_1 -$

$1\} \times \{0, 1, \dots, \tilde{M}_2 - 1\}$ とし, 次に窓関数 $w_p(\mathbf{x})$ と $w_q(\mathbf{x})$ ($\mathbf{x} \in \{0, 1, \dots, N - 1\}^2$) を

$$w_p(\mathbf{x}) \triangleq \begin{cases} 1 & \text{if } \mathbf{x} \in W_p \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$w_q(\mathbf{x}) \triangleq \begin{cases} 1 & \text{if } \mathbf{x} \in W_q \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

で定義して, $p = \tilde{p}w_p$ 及び $q = \tilde{q}w_q$ を生成する. ここで各窓領域内での各画素値は正, すなわち

$$p(\mathbf{x}) > 0 \quad \text{if } \mathbf{x} \in W_p \quad (5)$$

$$q(\mathbf{x}) > 0 \quad \text{if } \mathbf{x} \in W_q \quad (6)$$

であるとする.

2.2 LOG フィルタ

LOG フィルタは,

$$\nabla^2 G(\mathbf{x}) \triangleq \frac{-1}{\pi\sigma^4} \left(1 - \frac{\|\mathbf{x}\|^2}{2\sigma^2} \right) e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}} \quad (7)$$

で定義される. ここで $\mathbf{x} = (x, y)^T$ 及び $\|\mathbf{x}\|^2 = x^2 + y^2$ である. これはガウシアン関数 $G(\mathbf{x}) \triangleq (1/2\pi\sigma^2)e^{-\|\mathbf{x}\|^2/2\sigma^2}$ にラプラシアン $\nabla^2 \triangleq \partial^2/\partial x^2 + \partial^2/\partial y^2$ を施したものである. LOG フィルタをフーリエ変換すると,

$$F(\nabla^2 G) = \frac{-\|\mathbf{u}\|^2}{2\pi} e^{-\frac{\sigma^2\|\mathbf{u}\|^2}{2}} \quad (8)$$

となる. ここで $\mathbf{u} = (u, v)^T$ であり, u と v はそれぞれ x と y 方向の周波数を表す. この LOG フィルタは帯域通過フィルタであり, 低周波及び高周波のノイズを除去する性質をもつ (詳しくは 4.2 参照).

以下では画像 p に LOG フィルタを施して得られる画像を

$$p_{\text{LOG}}(\mathbf{x}) \triangleq p * \nabla^2 G \quad (9)$$

と表し, LOG 画像と呼ぶ.

2.3 LOG 符号を用いる任意形状画像の高速照合

本手法はテンプレート画像 $q(\mathbf{x})$ 内の任意形状の窓領域 $W_r \subseteq W_q$ 内の部分画像, すなわち

$$r(\mathbf{x}) \triangleq q(\mathbf{x})w_r(\mathbf{x}) \quad (10)$$

$$w_r(\mathbf{x}) \triangleq \begin{cases} 1 & \mathbf{x} \in W_r \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

で定義される $r(\mathbf{x})$ が正となる領域の画像を検索画像といい, この検索画像に最も照合する位置を対象画像 $p(\mathbf{x})$ の中から検出することを目的とする. この照合と位置検出の手法は様々考えられるが, 本手法ではまず $p(\mathbf{x})$ と $r(\mathbf{x})$ の LOG 符号画像, すなわち

$$p_{\text{SLOG}}(\mathbf{x}) \triangleq \text{sgn}_{w_p}(\mathbf{x})(p_{\text{LOG}}(\mathbf{x})) \quad (12)$$

$$r_{\text{SLOG}}(\mathbf{x}) \triangleq \text{sgn}_{w_r}(\mathbf{x})(q_{\text{LOG}}(\mathbf{x})) \quad (13)$$

を生成する. ここで

$$\text{sgn}_{w(\mathbf{x})}(f(\mathbf{x})) \triangleq \begin{cases} 1 & f(\mathbf{x}) \geq 0 \text{ and } w(\mathbf{x}) = 1 \\ -1 & f(\mathbf{x}) < 0 \text{ and } w(\mathbf{x}) = 1 \\ 0 & w(\mathbf{x}) = 0 \end{cases} \quad (14)$$

は画像 $f(\mathbf{x})$ の正負を符号化し窓関数 $w(\mathbf{x})$ を施した画像を表す. なおここで実際の LOG 画像の画素値が厳密に 0 になること, すなわち上式で $f(\mathbf{x}) = 0$ となることはほとんどなく, したがって上式の LOG 符号画像ともとの LOG 画像のゼロ交差 (正と負の画素領域の境界) はほとんど同じになる, すなわち LOG 画像のゼロ交差の情報は LOG 符号画像にそのまま受け継がれることを注意しておく.

次にこれらの相関 $p_{\text{SLOG}} * r_{\text{SLOG}}$ を検索画像の画素数 $|W_r|$ で正規化した評価関数,

$$\begin{aligned} R_{\text{NCSLOG}}(\mathbf{x}) &\triangleq \frac{1}{|W_r|} (p_{\text{SLOG}} * r_{\text{SLOG}}) \\ &= \frac{1}{|W_r|} \sum_{\mathbf{i} \in W_r} p_{\text{SLOG}}(\mathbf{x} + \mathbf{i}) r_{\text{SLOG}}(\mathbf{i}) \end{aligned} \quad (15)$$

を正規化 LOG 符号相関 (NCSLOG; Normalized Correlation of Sign of LOG) と呼び, この値を最大にする位置 $\mathbf{x} = \mathbf{x}_d$ を検出することを提案する. この式よりこの評価関数が 1 回の相関演算, したがって 3 回の FFT 演算で高速に計算できることが分かる. この計算量はテンプレートマッチングの測度としてしばしば使用される次の正規化相関

$$\begin{aligned} R_{\text{nCSLOG}}(\mathbf{x}) &\triangleq \frac{\sum_{\mathbf{i} \in W_r} p_{\text{SLOG}}(\mathbf{x} + \mathbf{i}) r_{\text{SLOG}}(\mathbf{i})}{\sqrt{\sum_{\mathbf{i} \in W_r} p_{\text{SLOG}}(\mathbf{x} + \mathbf{i})^2} \sqrt{\sum_{\mathbf{i} \in W_r} r_{\text{SLOG}}(\mathbf{i})^2}} \end{aligned} \quad (16)$$

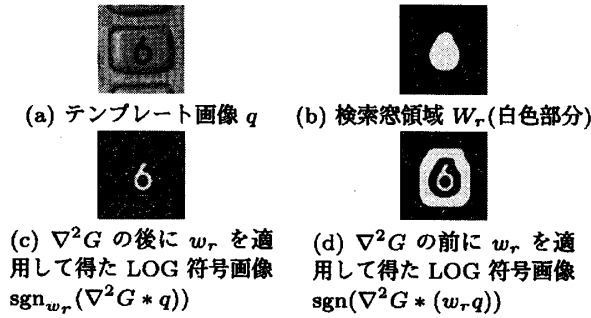

 図 1 $\nabla^2 G$ と w_r の適用順序の影響

 Fig. 1 Effect of application order of $\nabla^2 G$ and w_r . (a) Template image q . (b) Retrieval window W_r . (c) Sign of LOG image $\text{sgn}_{w_r}(\nabla^2 G * q)$ by applying w_r after $\nabla^2 G$. (d) Sign of LOG image $\text{sgn}(\nabla^2 G * (w_r q))$ by applying w_r before $\nabla^2 G$.

の計算量よりも少ないが、これらの評価値はほとんど同じになる。すなわちすべての $i \in W_r$ が $x + i \in W_p$ となる x について $ps_{LOG}(x + i)^2 = 1$ 及び $rs_{LOG}(i)^2 = 1$ となるので $R_{NCSLOG}(x) = R_{nCSLOG}(x)$ となる。また $i \in W_r$ かつ $x + i \notin W_p$ となる i と x について、すなわち検索画像中の位置 i が対象画像をはみ出す位置 $x + i$ にある場合には、 $ps_{LOG}(x + i) = 0$ となり $R_{NCSLOG}(x)$ は分子のみ、 $R_{nCSLOG}(x)$ は分母と分子が小さくなり、両者の値が異なるが、ここで例えば検索画像の 1 画素だけが対象画像の 1 画素と重なりかつそれらの符号が等しいとき $R_{NCSLOG}(x) = 1/|W_r|$ 及び $R_{nCSLOG}(x) = 1$ となり、この場合は本手法で用いる正規化の方が妥当であると考えられることを注意しておく。

またここで窓領域 W_q と W_r について説明しておく。まずテンプレート画像 q の窓領域 W_q は検索したい任意形状の画像よりも十分に大きい方形領域であるとし、一方、検索画像 r の窓領域 W_r は検索したい任意形状画像の少し外側を縁取ったものとする (図 1 (a) と (b) 参照)。更に W_r を表す窓関数 w_r は上述のように LOG フィルタを施した後の $q_{LOG} = (\nabla^2 G) * q$ に適用するが、これは LOG フィルタを施す前に w_r を適用すると、検索画像の縁に W_r の外の値 $w_r(x) = 0$ の影響が現れるので、その影響を除くためである (図 1 (c) と (d) 参照)。

3. 座標変換画像の照合

一般に画像をカメラで撮影する場合、拡大縮小、回転、射影変換などを伴うことが多い。以下、座標変換と照合関数の相似性について説明し、座標変換画像の探索法を提案する。

3.1 座標変換

以下、画像 $p(x)$ に座標変換 $h(x) = (x_h(x), y_h(x))^T$ を施した画像 $p(h(x))$ について考える。ここで $x_h(x)$ 及び $y_h(x)$ は $x = (x, y)^T$ の関数である。またこの変換は次の座標変換関数 $t_h(x)$ で表すこともできる。

$$p(h(x)) = t_h(x)(p(x)) \\ \triangleq \iint \delta(x_h(x, y) - u) \delta(y_h(x, y) - v) p(u, v) du dv \quad (17)$$

ここで $\delta(\cdot)$ はデルタ関数である。また本手法では主に次式で定義される射影変換について検討する。

$$h(x) = \frac{Ax}{b^T x + 1} \quad (18)$$

ここで

$$A \triangleq \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad b \triangleq \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad (19)$$

であり、 $a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2 \in \mathbb{R}$ である。これらの射影変換は、 $b = 0$ のときは線形変換であるが、それ以外の場合は非線形変換となる。

3.2 座標変換に対する照合関数の相似性

検索画像 $r(x)$ が座標変換 $h(x)$ を伴って対象画像 $p(x)$ に含まれる場合を考える。これを検討するための一つの尺度として次の関係を考える。すなわち $p(x)$ と $r(x)$ の相関を $R_{NCSLOG}(x)$ とし、一方 $p'(x) \triangleq p(h(x))$ と $r'(x) \triangleq r(h(x))$ の相関を $R'_{NCSLOG}(x)$ とするとき、それらに相似な関係

$$R'_{NCSLOG}(x) = t_h(x)(R_{NCSLOG}(x)) \quad (20)$$

が成立するとこの照合結果は座標変換に不変である (影響されない) といえる。更にこの相似関係が成立すれば座標変換に不変な検出が行われるといえる。

今、線形変換 $h(x) = Ax$ を考え、対象画像 $p(Ax)$ に LOG フィルタ $\nabla^2 G(Ax)$ を施した LOG 画像を連続系で表すと

$$p_{LOG}(Ax) \Big|_{\nabla^2 G(Ax)} \\ = \int p(Ax + Au) \nabla^2 G(Au) du \\ = |A|^{-1} \left[t_{Ax} \left(p_{LOG}(x) \Big|_{\nabla^2 G(x)} \right) \right] \quad (21)$$

となる。ここで $|A|$ は A の行列式であり、 $|A| > 0$ である。よって対象画像の LOG 符号画像は

$$p_{SLOG}(Ax) \Big|_{\nabla^2 G(Ax)} = t_{Ax} \left(p_{SLOG}(x) \Big|_{\nabla^2 G(x)} \right) \quad (22)$$

となる。この式は $p(Ax)$ の LOG 符号画像 (左辺) と $p(x)$ の LOG 符号画像を座標変換した画像 (右辺) が等しいことを表す。同様の関係は検索画像についても成立し、したがって相似関係式 (20) が成立する。

次に射影変換 $h(x) = Ax/(b^T x + 1)$ を考え、 $p(h(x))$ に $\nabla^2 G(h(x))$ を施した LOG 画像は

$$\begin{aligned} p_{LOG}(h(x)) \Big|_{\nabla^2 G(h(x))} \\ = \int p \left(\frac{Ax + Au}{b^T x + b^T u + 1} \right) \nabla^2 G \left(\frac{Au}{b^T u + 1} \right) du \end{aligned} \quad (23)$$

となる。ここで $|b^T u| \ll 1$ とし、 $x' \triangleq Ax/(b^T x + 1)$ 及び $u' \triangleq Au/(b^T x + 1)$ とおくと $du' = |A|du/(b^T x + 1)$ となるので、

$$\begin{aligned} p_{LOG}(h(x)) \Big|_{\nabla^2 G(h(x))} \\ \simeq (b^T x + 1)|A|^{-1} \left(t_{h(x)} \left(p_{LOG}(x) \Big|_{\nabla^2 G_{\sigma'}} \right) \right) \end{aligned} \quad (24)$$

が得られる。ただし $\nabla^2 G_{\sigma'}$ は $\nabla^2 G$ の広がりを表す σ としてもとの σ の代わりに位置 x に依存する $\sigma' = \sigma/(b^T x + 1)$ を使用したものである。通常、 x の範囲は原点 $x = \mathbf{0}$ を中心とした $b^T x + 1 > 0$ の範囲を扱うので $(b^T x + 1)|A|^{-1} > 0$ となり、

$$p_{SLOG}(h(x)) \Big|_{\nabla^2 G(h(x))} \simeq t_{h(x)} \left(p_{SLOG}(x) \Big|_{\nabla^2 G_{\sigma'}} \right) \quad (25)$$

が得られる。同様の関係は検索画像にも成立するので、相似関係式 (20) が近似的に成立するといえるが、この結果の導出には $\nabla^2 G_{\sigma'}$ を使用する必要がある。一方、検索画像 $r(x)$ の窓領域 W_r の大きさが小さく、その領域で $|b^T x| \ll 1$ が成立するときは射影変換画像 $r(h(x))$ は線形変換画像 $r(Ax)$ で近似できるので、 $\nabla^2 G_{\sigma'}$ でなく $\nabla^2 G$ を使用しても相似関係式 (20) がほぼ成立する。ただしこの $r(h(x)) \simeq r(Ax)$ を $p(h(x))$ の中から検索すると式 (20) は任意の位置 x では成立せず、検索画像が存在する位置 $x = x_d$ において近似的に成立することになる。

以上の導出過程を見ると、射影変換 $h(x)$ を施した $\nabla^2 G(h(x))$ があれば対象画像 $p(h(x))$ から $p_{SLOG} \Big|_{\nabla^2 G(h(x))}$ が生成でき、更に $q(h(x))$ と $w_r(h(x))$ があれば $r_{SLOG} \Big|_{\nabla^2 G(h(x))}$ が生成でき、これらを用いて相似関係式 (20) をほぼ満足する位置 $h(x_d)$ を検出できることが分かる。そこで次の手法を提案する。

3.3 VQ ネットを用いた座標変換画像の照合

対象画像 $p(x)$ の中に座標変換 $h(x) \in H$ により変形された検索画像 $r(h(x))$ があり、その位置を検出することを考える。ここで座標変換集合 $H \triangleq \{h_j(x) \mid j = 0, 1, 2, \dots\}$ のパラメータ (例えば式 (19) の A や b) の範囲はあらかじめ分かっており、 H の要素は一般には無限個あるものとする。

以上の仮定のもとでは検索すべき画像は理論的には連続無限個あるが、実際にはそれらを量子化した有限個の検索画像を用いれば十分である。更にある座標変換 $h_j(x)$ に対して前節で示した画像照合を行うにはテンプレート画像 $q_j = q(h_j(x))$ 、窓関数 $w_j = w_r(h_j(x))$ 及び LOG フィルタ $v_j = \nabla^2 G(h_j(x))$ があればよい。そこで本手法では次のような VQ ネットを用いる。まずこの VQ ネットは N_c 個のユニットからなり、各ユニット i ($i = 1, 2, \dots, N_c$) はテンプレート画像 q_i 、検索窓 w_i 及び LOG フィルタ v_i を記憶ベクトルとして学習する。この学習は競合再初期化学習法 (CRL) [2] に基づく表 1 の学習法により行う。ここで CRL を用いる主な理由は以下のとおりである。まず VQ の代表的手法である K-means 法や LBG 法などのバッチ型 (例えば [2] 参照) では無限集合である H の中から有限個の訓練データの集合を適切に生成して学習させる必要があるが、オンライン型である CRL は H の中からランダムに訓練データを選出しながら学習できるという長所がある。また VQ の諸手法により求まる解は一般に初期値に依存した局所最適解になるが、CRL は量子化ベクトルの数が非常に多い場合の最適条件を用いて解の最適化を行う手法であり、その局所解からの回避能力や学習速度等の性能の高さが従来研究により示されている。

表 1 による学習後、VQ ネットの記憶 (q_i, w_i, v_i) を用いて式 (15) の $R_{NCSLOG}(x)$ を計算する。ここでテンプレート画像 q_i と対象画像 $p(x)$ に対して、検索画像の窓関数と LOG フィルタはそれぞれ w_i と v_i を用いる。すべての $i = 1, 2, \dots, N_c$ とすべての位置

表 1 CRL に基づく座標変換画像の学習

Table 1 CRL based learning of transformed images.

VQ ネットの第 i ユニットの画像 (ベクトル) $\mathbf{q}_i, \mathbf{w}_i, \mathbf{v}_i$ 及びスカラー値 d_i を記憶し, 以下の逐次学習を行う. まず各時刻 $k = 1, 2, \dots, K$ で座標変換 $\mathbf{h} \triangleq \mathbf{h}_{j(k)}(\mathbf{x}) \in H$ を生成する. ここで $j(k)$ は時刻 k で生成したランダムな座標変換の添字を表す. 更にこの座標変換 \mathbf{h} により変形したテンプレート画像を $\mathbf{q} \triangleq \mathbf{q}(\mathbf{h})$, 窓画像を $\mathbf{w} \triangleq \mathbf{w}_r(\mathbf{h})$ 及び LOG フィルタを $\mathbf{v} \triangleq \nabla^2 G(\mathbf{h})$ とし, 以下の処理を行う.

(1) 次の再初期化条件式 (RIC; reinitialization condition)

$$d_c > r_d \langle d_i \rangle \quad \text{and} \quad E < r_E \ln(N_c)$$

を調べる. ここで c は入力画像 \mathbf{q} に最も近いユニットの番号, $\langle d_i \rangle$ はすべての d_i の平均, $r_d (> 1)$ と $r_E (< 1)$ は定数とする. また,

$$E \triangleq - \sum_j \frac{d_j}{\sum_k d_k} \ln \left(\frac{d_j}{\sum_k d_k} \right)$$

は d_i のばらつきを表すエントロピーであり, すべての d_i が等しいとき最大値 $E_{\max} = \ln(N_c)$ をとる.

(2) 再初期化条件式 (RIC) が満たされるとき次の更新 (再初期化処理) を行う.

$$\mathbf{q}_i := \begin{cases} \mathbf{q} & \text{if } i = l \\ \mathbf{q}_i & \text{otherwise} \end{cases} \quad \mathbf{w}_i := \begin{cases} \mathbf{w} & \text{if } i = l \\ \mathbf{w}_i & \text{otherwise} \end{cases}$$

$$\mathbf{v}_i := \begin{cases} \mathbf{v} & \text{if } i = l \\ \mathbf{v}_i & \text{otherwise} \end{cases} \quad d_i := \begin{cases} \eta \langle d_i \rangle & \text{if } i = l \text{ or } i = c \\ \eta d_i & \text{otherwise} \end{cases}$$

ここで $:=$ は代入 (による更新) を表し, $\eta (< 1)$ は忘却係数を表す. また第 l ユニットの d_i の中の最小値 d_l をもつユニットとする.

(3) 再初期化条件式 (RIC) が満たされないとき次の更新 (競合学習) を行う.

$$\mathbf{q}_i := \begin{cases} \mathbf{q}_i + \alpha (\mathbf{q} - \mathbf{q}_i) & \text{if } i = c \\ \mathbf{q}_i & \text{otherwise} \end{cases} \quad \mathbf{w}_i := \begin{cases} \mathbf{w}_i + \alpha (\mathbf{w} - \mathbf{w}_i) & \text{if } i = c \\ \mathbf{w}_i & \text{otherwise} \end{cases}$$

$$\mathbf{v}_i := \begin{cases} \mathbf{v}_i + \alpha (\mathbf{v} - \mathbf{v}_i) & \text{if } i = c \\ \mathbf{v}_i & \text{otherwise} \end{cases} \quad d_i := \begin{cases} \eta d_i + \|\mathbf{q} - \mathbf{q}_i\|^2 & \text{if } i = c \\ \eta d_i & \text{otherwise} \end{cases}$$

ここで $\alpha = 1 - E / \ln(N_c)$ はエントロピー E に依存した忘却係数である. なお $\mathbf{q}_i, \mathbf{w}_i, \mathbf{v}_i$ の初期値はゼロベクトル, d_i の初期値はゼロである.

$\mathbf{x} \in W_p$ に対して最も大きな $R_{NCSLOG}(\mathbf{x})$ の値を示す位置 $\mathbf{x} = \mathbf{x}_d$ を検出位置とする.

3.4 LOG 符号画像の座標変換に対するロバスト性

ここで LOG 符号画像が微小な座標変換に対してロバスト性をもつことを示す. まず 3.2 における議論

と同様に W_r 内で座標変換が $\mathbf{h}(\mathbf{x}) \simeq \mathbf{A}\mathbf{x}$ で線形近似できると仮定し, LOG 画像 $q_{LOG}(\mathbf{A}\mathbf{x}) \triangleq q(\mathbf{A}\mathbf{x}) * \nabla^2 G(\mathbf{A}\mathbf{x})$ と $q_{LOG}((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) \triangleq q((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) * \nabla^2 G((\mathbf{A} + \delta\mathbf{A})\mathbf{x})$ の違いを検討する. 画像 q が位置 $\mathbf{A}\mathbf{x}$ の近傍 (より詳しくいうと $\nabla^2 G((\mathbf{A} + \delta\mathbf{A})\mathbf{x})$ の影響範囲 (変換前の $\nabla^2 G(\mathbf{x})$ で半径 3σ 程度以内)) で 2 次近似できるとすると,

$$\begin{aligned} q_{LOG}((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) & \\ \simeq \frac{1}{|\mathbf{A} + \delta\mathbf{A}|} \int & \left(q(\mathbf{A}\mathbf{x}) + [\mathbf{q}^{(1)}(\mathbf{A}\mathbf{x})]^T (\delta\mathbf{A}\mathbf{x} + \mathbf{z}) \right. \\ & \left. + (\delta\mathbf{A}\mathbf{x} + \mathbf{z})^T [\mathbf{q}^{(2)}(\mathbf{A}\mathbf{x})] (\delta\mathbf{A}\mathbf{x} + \mathbf{z}) \right) \nabla^2 G(\mathbf{z}) d\mathbf{z} \end{aligned} \quad (26)$$

が得られる. ここで $\mathbf{q}^{(1)}(\mathbf{A}\mathbf{x})$ と $\mathbf{q}^{(2)}(\mathbf{A}\mathbf{x})$ はそれぞれ画像 $q(x_h, y_h) = q(\mathbf{h}(\mathbf{x})) \simeq q(\mathbf{A}\mathbf{x})$ の位置 $\mathbf{A}\mathbf{x}$ における 1 次偏微分ベクトル $\mathbf{q}^{(1)} = (\partial q / \partial x_h, \partial q / \partial y_h)^T$ と 2 次偏微分行列

$$\mathbf{q}^{(2)} = \begin{bmatrix} \frac{\partial^2 q}{\partial x_h^2} & \frac{\partial^2 q}{\partial x_h \partial y_h} \\ \frac{\partial^2 q}{\partial y_h \partial x_h} & \frac{\partial^2 q}{\partial y_h^2} \end{bmatrix} \quad (27)$$

である. 式 (26) の右辺において $\nabla^2 G(\mathbf{z})$ を \mathbf{z} の 1 次以下の項に施すと 0 になるので,

$$\begin{aligned} q_{LOG}((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) & \\ \simeq \frac{1}{|\mathbf{A} + \delta\mathbf{A}|} \int & \left(\mathbf{z}^T [\mathbf{q}^{(2)}(\mathbf{A}\mathbf{x})] \mathbf{z} \right) \nabla^2 G(\mathbf{z}) d\mathbf{z} \\ \simeq \frac{|\mathbf{A}|}{|\mathbf{A} + \delta\mathbf{A}|} & q_{LOG}(\mathbf{A}\mathbf{x}) \end{aligned} \quad (28)$$

が得られる. よって両辺を 2 値に符号化すると

$$r_{SLOG}((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) \simeq r_{SLOG}(\mathbf{A}\mathbf{x}) \quad (29)$$

が得られる. したがって画素値が画像上の位置の変化に対して 1 次や 2 次で緩やかに変化する部分が多い画像では原画像 $q((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) \simeq q(\mathbf{A}\mathbf{x}) + [\mathbf{q}^{(1)}(\mathbf{A}\mathbf{x})]^T \delta\mathbf{A}\mathbf{x} + \dots$ よりも LOG 符号画像 $r_{SLOG}((\mathbf{A} + \delta\mathbf{A})\mathbf{x})$ の方が $\delta\mathbf{A}$ の影響を受けにくいといえる. これは原画像よりも LOG 符号画像を用いる方が VQ ネットが必要とするユニット数を少なくできることを意味し, 後述の実験結果 (4.3 参照) もこれを示している.

また式 (29) は q が 2 次近似できる位置 $\mathbf{A}\mathbf{x}$ について導いたが, 実際にはより広い領域で成立する. すな

わち $q_{LOG}(Ax)$ と $q_{LOG}((A + \delta A)x)$ には相似関係が成立する (3.2 参照) ので, $q_{LOG}((A + \delta A)x)$ のゼロ交差 (正と負の画素領域の境界) は $q_{LOG}(Ax)$ のゼロ交差を δAx だけ移動したものになり, その移動部分以外では式 (29) が成立する. 更に σ が大きいほどゼロ交差の数が少ないので (例えば図 5 参照), δAx による変化は小さいといえる.

4. 実験と検討

以下, 本手法の有効性を実験結果を通して検討する. なお 4.1 と 4.2 では本手法の基本的性質を検討するため対象画像と検索画像は画素数がそれぞれ N^2 と M^2 の正方形の画像を用い, かつ VQ ネットを用いない場合を扱い, その後 4.3 で任意形状の検索画像と VQ ネットを用いる場合の検討を行う.

4.1 処理時間の比較

本手法では FFT と LOG 符号を用いて計算量の低減を図っている. これらの効果を調べるため対象画像の画素数 N^2 を 256^2 及び 512^2 とし, それぞれに対して検索画像の画素数 $|W_r| = M^2$ を $20^2, 40^2, \dots, 100^2$ としして照合の計算時間を計測し, 図 2 の結果を得た. 図中の NCSLOG (FFT) と NCSLOG (direct) はそれぞれ FFT を用いる場合と用いない場合の正規化 LOG 符号相関, NCLOG (FFT) は正規化 LOG 相関, すなわち

$$R_{NCLOG}(x) \triangleq \frac{\sum_{i \in W_r} p_{LOG}(x+i)r_{LOG}(i)}{\sqrt{\sum_{i \in W_r} p_{LOG}(x+i)^2} \sqrt{\sum_{i \in W_r} r_{LOG}(i)^2}} \quad (30)$$

を FFT を用いて計算したもの, CC は正規化相関, 及び SCC は選択的正規化相関を表す. 更に図中の計算時間の単位は秒 [s] であり, 対象画像が与えられる前に行える処理の時間 (例えば式 (30) の $\sqrt{\sum_{i \in W_r} r_{LOG}(i)^2}$ の計算時間) は含めていない. 使用計算機の CPU は Athlon XP 2000+ (1667MHz), OS は Vine Linux 2.5 である.

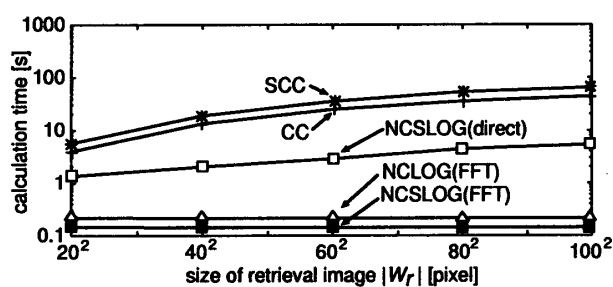
図より NCSLOG (FFT) は各サイズの対象画像と検索画像について他手法よりも高速であることが分かる. 例えば $|W_r| = 60^2$ のとき NCSLOG (FFT) は CC や SCC よりも $N^2 = 256^2$ と 512^2 の場合ともに約 173 倍高速である. また図より NCSLOG (FFT) と NCLOG (FFT) の計算時間は他手法と異なり $|W_r|$ には依存していないことが分かるが, これは FFT を用いる手法では検索画像を N^2 まで拡大して処理する (2.1 参照) ことから理解できる. 更に $N^2 = 512^2$ のときの NCSLOG (FFT) の計算時間は $N^2 = 256^2$ のときの FFT を用いない他手法よりも短いので, それらの手法で対象画像の大きさを 256^2 から 512^2 までの任意の値とするよりも NCSLOG (FFT) で 512^2 とした方が高速であることが分かる.

一方, NCSLOG (FFT) と NCLOG (FFT) の差は LOG 符号を用いる効果を表し, 図 2 では $N^2 = 256^2$ で約 1.5 倍, $N^2 = 512^2$ で約 2 倍高速である. これは NCLOG (FFT) では LOG 符号画像を生成するための符号演算は省略できるが非常に短く, 式 (30) の分母の $\sum_{i \in W_r} p_{LOG}(x+i)^2$ の計算量は分子の相関関数の計算量とほぼ同じになることから理解できる.

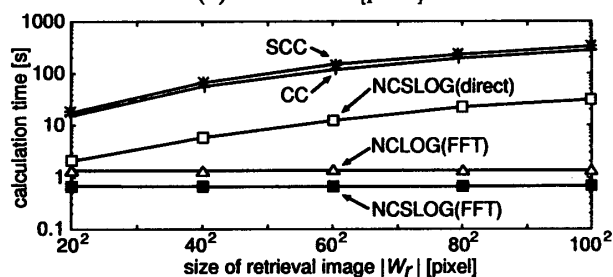
4.2 ノイズに対するロバスト性

2.2 で示したように LOG フィルタは帯域通過フィルタである. この空間周波数の帯域幅は $\sqrt{u^2 + v^2} = 0.87/\sigma \sim 2.0/\sigma$ [rad/pixel] (対応する空間周期は $\sqrt{x^2 + y^2} = 7.2\sigma \sim 3.1\sigma$ [pixel]) であり, この範囲外の低周波及び高周波のノイズ等は除去できる.

特に画像 $p(x)$ と $p'(x) \triangleq a p(x) + b^T x + c$ ($a > 0; a, c \in \mathbb{R}; b \in \mathbb{R}^{2 \times 1}$) を処理すると $p'_{LOG}(x) \triangleq (p'(x)) * (\nabla^2 G(x)) = a p_{LOG}(x)$ 及び $p'_{SLOG}(x) \triangleq$



(a) $N^2 = 256^2$ [pixel]



(b) $N^2 = 512^2$ [pixel]

図 2 計算時間の比較

Fig. 2 Comparison of calculation time.

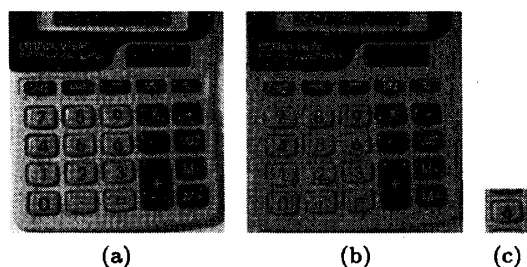


図 3 ノイズ付加実験で用いた画像：(a) 原画像 (512 × 512 [pixel]), (b) $\sigma_n = 30$ のノイズを付加した画像, (c) 検索画像 (100 × 100 [pixel])

Fig. 3 Images used in noise injection experiment: (a) original image, (b) noise injected image with $\sigma_n = 30$, (c) retrieval image.

表 2 原画像と付加ノイズの統計量
Table 2 Statistical quantities of original image and noise injected.

	原画像	付加ノイズ
明度の平均値	130	0
明度の標準偏差	75.4	0,5,10,15,20,25,30

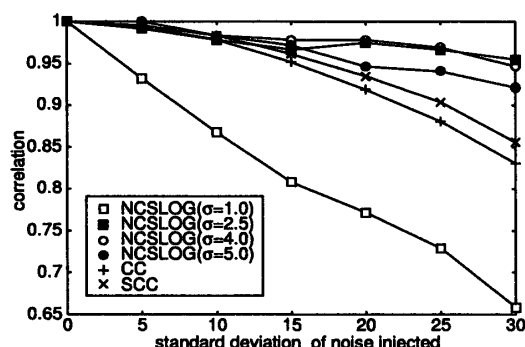


図 4 付加ノイズの標準偏差に対する相関値
Fig. 4 Correlation versus standard deviation of noise injected.

$\text{sgn}_{w_p}(\mathbf{x})(p'_{LOG}(\mathbf{x})) = p_{SLOG}(\mathbf{x})$ が得られるが、これは LOG 符号画像においては低周波ノイズである明度バイアス c 、線形明度変化 $\mathbf{b}^T \mathbf{x}$ 及び明度スケール a の影響は完全に除去できることを意味する。一方、統計的に無相関な部分を除いて相関値を計算する選択的正規化相関は正規化相関よりもノイズに強いと考えられるが一般にその能力は明らかでない。

各手法のノイズに対するロバスト性を検討するため、図 3(a) の原画像にノイズを加えた対象画像 (図 3(b) 参照) に対し、原画像から切り出した数字 “6” を含む検索画像 (図 3(c)) を用いて (すなわちこの画像領域全体を窓領域として) 各手法による照合実験を行った。付加ノイズ (正規ノイズ) 及び原画像の統

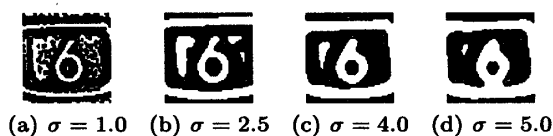


図 5 異なる σ に対する LOG 符号画像
Fig. 5 Sign of LOG images for different σ .

計量を表 2 に示すが、すべての場合で照合は成功し、各照合位置での相関値は図 4 のようになった。この図から NCSLOG の相関値は LOG フィルタの係数が $\sigma = 2.5, 4.0, 5.0$ のとき CC や SCC よりもノイズの影響を受けにくいことが分かる。しかし $\sigma = 5.0$ のときは $\sigma = 2.5$ や 4.0 のときよりもノイズの影響を受けている。これは $\sigma = 5.0$ のときの LOG フィルタが検索画像を特徴づける周波数の一部も除去したからであると考えられる。この現象は検索画像の LOG 符号画像 (図 5) を見ると、 $\sigma = 1.0, 2.5, 4.0$ と大きくなるにつれて検索画像を構成する線が太くなり高周波ノイズを除去できるようになるが、 $\sigma = 5.0$ では線が太すぎて検索画像の主要な特徴の一つである “6” の外側のエッジの一部が消失していることから理解できる。またこの結果は LOG 符号画像で原画像の特徴を保持するには σ をうまく選ぶ必要があることを意味している。

4.3 座標変換画像の照合

本手法は座標変換された画像を VQ ネットを用いて照合する機能と任意形状の検索画像を照合する機能をもつ。これらの効果を調べるため以下の二つの実験を行った。

4.3.1 実験 1

まず前述の図 3(a) と (c) の画像をそれぞれ対象画像とテンプレート画像の原画像とし、テンプレート画像の “6” を縁どった任意形状窓と “6” を含む方形窓を用意した。次にロール角 θ_R 、ピッチ角 θ_P 、ヨー角 θ_Y をそれぞれ $-10 + 20i/3$ [deg] ($i = 0, 1, 2, 3$)、拡大率 γ を $0.8 + 0.2i$ ($i = 0, 1, 2$) とした次式の射影変換を原対象画像に施した 192 枚の画像を対象画像として照合実験を行った：射影変換は

$$\mathbf{h}(\mathbf{x}) = \frac{\gamma \begin{bmatrix} C_{RCY} - S_{P}S_{R}S_{Y} & C_{Y}S_{R} + C_{R}S_{P}S_{Y} \\ -C_{P}S_{R} & C_{P}C_{R} \end{bmatrix} \mathbf{x}}{\beta(C_{Y}S_{P}S_{R}, C_{R}S_{Y})\mathbf{x} + 1} \quad (31)$$

とし、 $C_R \triangleq \cos \theta_R$, $C_P \triangleq \cos \theta_P$, $C_Y \triangleq \cos \theta_Y$, $S_R \triangleq \sin \theta_R$, $S_P \triangleq \sin \theta_P$, $S_Y \triangleq \sin \theta_Y$ であり、 β はカメラ

ラと対象物の距離 $L \approx 20$ [cm] の影響を表すパラメータであり、対象物の実際の幅 $W \approx 10$ [cm] と画像上の幅 $w \approx 500$ [pixel] を用いて $\beta \triangleq W/(wL) = 0.001$ となった。この射影変換を用いると検索画像内の位置 $\mathbf{x} = (x, y)^T$ の $|x|$ と $|y|$ の範囲はそれぞれ 50 未満、したがって $|\mathbf{b}^T \mathbf{x}| = |\beta(c_{YSPSR}, c_{RSY})\mathbf{x}| < 0.01$ となり 3.2 で述べた条件 $|\mathbf{b}^T \mathbf{x}| \ll 1$ は満たされていると考えられる。

VQ ネットの学習は $\theta_R, \theta_P, \theta_Y$ をそれぞれ $-10 \sim 10$ [deg], γ を $0.8 \sim 1.2$ の範囲の射影変換を用いて表 1 の手順で 10,000 回行った。ここで LOG フィルタの係数 σ は前述の実験で検索画像の特徴をうまく保持すると考えられる $\sigma = 2.5$ を用いた。ユニット数が 5 個の VQ ネットの学習後の荷重ベクトルを図 6 に示す。図中の \mathbf{w}_i と \mathbf{w}'_i はそれぞれ任意形状窓 W_r と方形窓 W'_r を学習した結果であり、これらの荷重ベクトルを用いて、各手法で上記の 192 枚の対象画像について照合したときの成功率を表 3 に示す。ここで照合位置が正しい位置から半径 10 画素以内であれば成功とした。表より提案手法の NCSLOG はすべて正しい位置を検出できたが、CC と SCC は正しく検出できない場合があったことが分かる。なおユニット数は全手法について順次 1 個ずつ増やしていき、照合成功率が初めて 100% となった 5 個の場合を示している。なお、任意形状窓を使用した CC と SCC が照合成功率 100% となるにはそれぞれ 6 個のユニット数を必要とした。以上の結果は 3.4 で述べたように LOG 符号画像の座標変換に対するロバスト性により、他手法より少ない

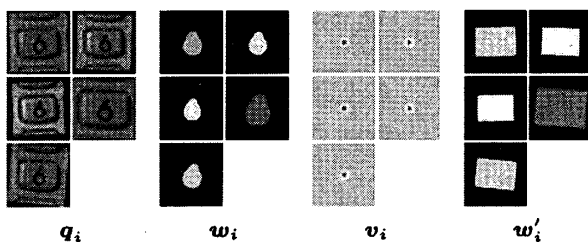


図 6 実験 1 の学習後の荷重ベクトル

Fig. 6 Weight vectors after the learning for the experiment 1.

表 3 実験 1 の照合成功率

Table 3 Correct matching rate for the experiment 1.

	任意形状窓 W_r	方形窓 W'_r
NCSLOG	100%	100%
CC	99.5%	99.5%
SCC	96.9%	97.9%

ユニット数で 100% の照合成功率が達成できたことを表していると考えられる。一方、任意形状窓と方形窓の照合成功率の差はほとんどないが、これはこれら両窓内の画像はもとの対象画像から切り出したものであり、両画像が対象画像に含まれることから理解できる。

4.4 実験 2

この実験では次の照合課題を通して方形でない検索画像を用いる有効性について検討した。まず電子情報通信学会パターン認識・メディア理解研究会が 1998 年に主催したアルゴリズムコンテストで用いられた図 7(a) と (b) の画像をそれぞれ背景画像とテンプレート画像とし、更にそれらを合成するために当該研究会が提供している「例題作成プログラム」を使用して回転角 $\theta_R = -10 + 4i$ ($i = 0, 1, 2, 3, 4, 5$) 及び拡大率 $\gamma = 0.9 + 0.1i$ ($i = 0, 1, 2$) でテンプレート画像を変換して背景画像と合成した合計 18 枚の画像を対象画像 (図 7(c) 参照) として照合実験を行った。なお上記コンテストではカラー画像が用いられたが、本実験ではグレー画像に変換して使用した。

まず与えられたテンプレート画像をそのまま用いると提案手法以外での照合は全く成功しなかった。これは与えられたテンプレート画像中のキャラクタ周囲の画素値は 255 であるが、キャラクタ及び対象画像中のキャラクタ周囲の画素値のほとんどは 10 から 50 前後の値であり、テンプレート画像をそのまま用いて CC や SCC を適用すると画素値が 255 をとる部分が相関値に占める割合が非常に大きくなることが主要因であることが分かった。そこでテンプレート画像のキャラクタ以外の部分の画素値 255 をキャラクタ部分の画素値の平均値で置き換えて新しいテンプレート画像を作成して以下の比較実験を行った。

まず回転角 θ_R を $-10 \sim 10$ [deg], 拡大率 γ を $0.9 \sim 1.1$ の範囲内の値として新しいテンプレート画像を変

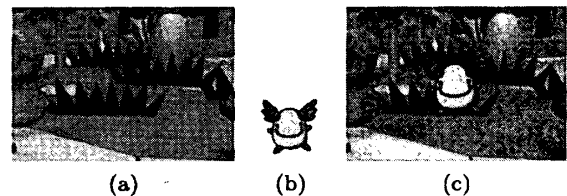


図 7 実験 2 で用いた画像: (a) 背景画像 (352 × 240 [pixel]), (b) テンプレート画像 (128 × 128 [pixel]), (c) 対象画像の例

Fig. 7 Images used in the experiment 2: (a) background image, (b) template image, and (c) an example of source images.

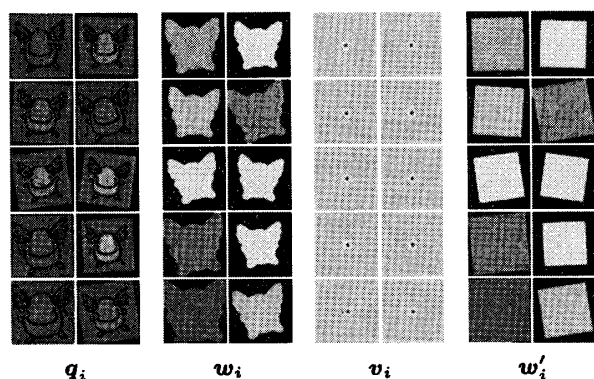


図 8 実験 2 の学習後の荷重ベクトル

Fig. 8 Weight vectors after the learning for the experiment 2.

表 4 実験 2 の照合成功率

Table 4 Correct matching rate for the experiment 2.

	任意形状窓 W_r	方形窓 W'_r
NCSLOG	100%	50.0%
CC	77.8%	38.9%
SCC	72.2%	27.8%

換した画像について VQ ネットの学習を 10,000 回行った。ユニット数が 10 個の VQ ネットの学習後の荷重ベクトルを図 8 に示す。図中の w_i と w'_i はそれぞれ任意形状窓 W_r と方形窓 W'_r を学習した結果を表す。これらの荷重ベクトルを用いて上記の対象画像 18 枚について照合したときの成功率を図 7 に示す。ここで実験 1 と同様に検出位置が正しい位置から半径 10 画素以内であれば成功とした。またユニット数は全手法で順次 1 個ずつ増やしていき、照合成功率が初めて 100% となった 10 個の場合を示している。表 4 より任意形状窓を使う NCSLOG が 100% の照合成功率を達成しており、LOG 符号画像と任意形状窓を使う NCSLOG の有効性が理解できる。一方、任意形状窓を使う CC で照合成功率 100% とするにはユニット数が 14 個必要であり、SCC では 20 個のユニットでも 88.9% の照合成功率しか得られなかった。なおこの実験結果では NCSLOG は CC や SCC よりもかなり良い結果となったが、前節の実験 1 では NCSLOG と CC との差はわずかであった。これは LOG 符号画像の座標変換に対するロバスト性は原画像の性質に依存する (3.4 参照) ことに起因していると考えられる。

5. むすび

本論文ではノイズ、明度変化、座標変化等を含む対象画像から任意形状の検索画像を高速に探索する画像

照合法として正規化 LOG 符号相関とベクトル量子化ネットを用いる手法を提案した。正規化 LOG 符号相関は FFT を用いて高速に計算でき、従来の正規化相関等よりも計算時間を大幅に低減できることを示した。また LOG 符号を使用することでノイズや明度変化の影響を低減できることを示し、比較実験によって正規化相関や選択的正規化相関よりもノイズに対するロバスト性が高いことを確認した。更に VQ ネットでテンプレート画像、窓画像及び LOG フィルタを学習して使用する方法について検討し、他手法よりも少ないユニット数で座標変換画像を探索できることを示した。

文 献

- [1] D. Marr, Vision: A Computational Investigation into the Human Representation and Processing of Visual Information, W.H.Freedman Company, 1982.
- [2] 西田 健, 黒木秀一, 佐伯知則, “再初期化法を用いた適応ベクトル量子化,” 信学論 (D-II), vol.J84-D-II, no.7, pp.1503-1511, July 2001.
- [3] 金子俊一, “実世界マシンビジョンのためのロバスト画像照合技術,” 電学論 (C), vol.121, no.5, pp.830-833, 2001.
- [4] D.I. Barnea and H.F. Silverman, “A class of algorithms for fast digital image registration,” IEEE Trans. Comput., vol.C-21, no.2, pp.179-186, 1972.
- [5] 村瀬一郎, 金子俊一, “増分符号相関によるロバスト画像照合,” 信学論 (D-II), vol.J83-D-II, no.5, pp.1323-1331, May 2000.
- [6] 佐藤雄隆, 金子俊一, 五十嵐悟, “選択的正規化相関によるロバスト画像照合,” 電学論 (C), vol.121, no.4, pp.800-807, 2001.
- [7] 齊藤文彦, “ブロック照合投票処理を用いた速へいに強い画像マッチング,” 信学論 (D-II), vol.J84-D-II, no.10, pp.2270-2279, Oct. 2001.
- [8] 流郷達人, 宮本 敦, 金子俊一, 五十嵐悟, “標準化順位相関に基づくロバスト画像照合,” 画電学誌, vol.31, no.3, pp.363-369, 2002.
- [9] M. Uenohara and T. Kanade, “Use of Fourier and Karhunen-Loeve decomposition for fast pattern matching with a large set of templates,” IEEE Trans. Pattern Anal. Mach. Intell., vol.19, no.8, pp.891-898, 1997.
- [10] K. Sumi, M. Hashimoto, and H. Okuda, “Three-level broad-edge matching based real time robot vision,” Proc. International Conference on Robotics and Automation, pp.1416-1422, IEEE, 1995.

(平成 15 年 5 月 13 日受付, 16 年 1 月 5 日再受付)



刈川 康裕 (学生員)

平 13 九工大・工・機械知能工卒. 平 15 同大学院博士前期課程了. 現在, 同大学院博士後期課程に在学中. 主にニューラルネットによるパターン認識の研究に従事. 日本神経回路学会会員.



黒木 秀一 (正員)

昭 55 九工大・工・電気卒. 昭 60 東工大大学院博士課程了. 同年より九工大・制御・助手を経て, 平 3 同大・助教授. 工博. 主にニューラルネットの研究に従事. 日本神経回路学会, 計測自動制御学会等各会員.