

知的光ビジョン

松岡 克典

人間のもつ高度な視覚機能をロボットなどの機器に組込むことができれば、人と同じ環境の中で、人の安全を確保しながら人と協調した作業が行える“夢のマシン”を開発することができる。しかし、そのような視覚機能を実現するためには、人が実際に活動している実世界の中で、時々刻々と変化する膨大な画像情報を高速に処理する必要がある。

例えば、視覚フィードバックを用いて環境に適応した動作や他のマシンとの協調作業を行うためには、サイクルタイムが1 ms以下のフィードバック速度が要求される。この速度は、現在のTVカメラのサイクルタイム(約33 ms)よりも1桁以上速い速度であり、新しい高速視覚情報処理システムが必要になる。

そこで2次元画像を局所走査することなく、並列的に処理することができる光情報処理技術を利用することにより、視覚機能に必要な画像処理の高速化を図ることができ、より高度な視覚機能を実現できる可能性がある。

このような光情報処理を用いた高速の視覚情報処理技術を「光ビジョン技術」と呼ぶことにし、ここでは筆者が進めている光ビジョンの研究を中心に紹介する。対象としているのは、人が活動している実世界シーンであり、その中から複数の対象物体を同時に検出することができる小型で高速の光情報処理システムの構築を目指している。

1. 相関光学系を用いたパターン認識

2枚の画像間の相関値は、互いの類似度を示してお

り、パターンを検出手段として利用することができる。相関光学系は、入力像と識別フィルター上のパターンとの相関演算をまさに光速で実行でき、高速のパターン検出システムとして期待されている。検出信号は、検出パターンが入力像の中に含まれている位置に生じる大きな相関ピークとして得られる。これを閾値処理して検出することにより、入力像に含まれる複数の対象物体を同時に検出することができる。

しかし、従来の相関光学系を用いたパターン認識では、単一の相関演算の結果から対象物体の検出を行うため、複雑な背景を含む実世界シーンに適用した場合、必ずしも良好な結果を得ることができなかった。

例えば、図1(a)に示す交差点のシーンから道路標識を検出する場合でも、従来法では十分な結果が得られない。図1(b), (c), (d)に示す有名な識別フィルター関数(マッチトフィルター¹⁾、SDF(synthetic discriminant function)フィルター²⁾、MINACE(minimum noise and correlation energy)フィルター³⁾を用いて検出した道路標識の位置は、それぞれ図1(e), (f), (g)の黒い部分になる。どの場合も、背景によって生じる偽信号が多く現れており、不十分な結果になっている。このように、単一の相関演算によって実行できるパターン検出能力には限界があった。

そのため、近年では、複数の相関演算を用いてパターン認識を行う光システムがいくつか提案されている⁴⁻⁸⁾。

2. 実世界シーンを対象にしたパターン認識

実世界を対象にしたパターン認識では、入力像の中に含まれる複数の対象物体がどこにあるかを瞬時に検出す

Optical vision system (1996年3月14日受理)
Katsunori MATSUOKA 大阪工業技術研究所光機能材料部情報光学研究室 (〒563 池田市緑が丘 1-8-31)

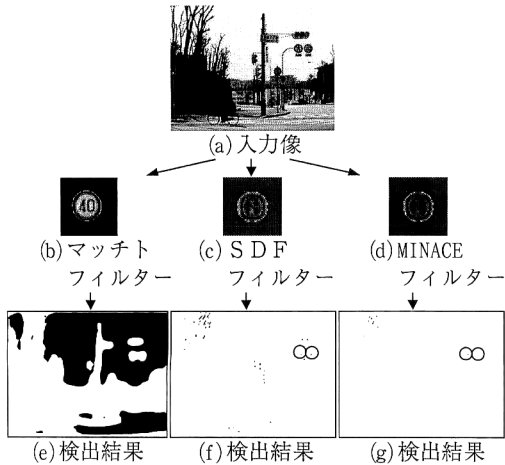


図1 相関演算を用いて実世界シーンから丸い道路標識を検出した結果。○印の位置に正しい検出信号が現れている。

ることが重要になる。しかし、それぞれの対象物体は、明るさ、大きさ、向きが異なっていたりするため、それらに影響されないロバストな処理が必要になる。しかも、背景や雑音などの未知情報を含んでおり、それらに対する高い耐性をもつ処理が必要となる。つまり次の4つの機能を同時に満足するパターン認識処理が必要になる。

- (1) 対象物体の位置に依存しない処理。
- (2) 対象物体の明るさに依存しない処理。
- (3) 視線移動に伴う対象物体の大きさや形の変化に対して影響を受けない処理。
- (4) 背景・雑音に対して強靱な処理。

相関演算はシフト・インバリエントな処理であるため、(1)の機能を簡単に実現できるが、(2)～(4)の機能を同時に実現することは容易ではない。光を用いたパターン認識の分野では、これまで(3)の機能を実現するための研究が主に進められてきたが、(2)や(4)の機能については、あまり検討されてこなかった。

そこで、上記4つの機能を実現するための光学的手法として、①多重相関演算を用いる方法、②背景によって生じる偽信号を抑える識別フィルターの設計手法、③明るさの異なる対象物体の同時検出手法について検討を進めてきた。以下に、その内容について述べる。

2.1 多重相関光学系を用いたパターン検出

複数の相関光学系を用いて、1枚の入力シーンに含まれる異なる特徴パターンを複数同時に検出し、検出された特徴パターンの組合せから、入力像のパターン認識を

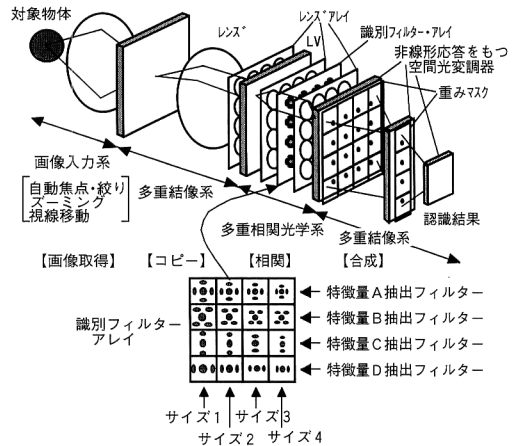


図2 多重相関光学系を用いたパターン認識システム。アレイ状に配置した複数の識別フィルターと入力像との相関演算を多重相関光学系で実行し、その結果を非線形な入出力応答をもつ空間光変調器上に重ね合わせることで、最終的な信号を得る。

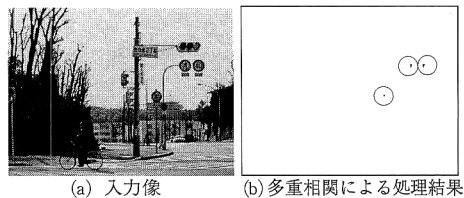


図3 多重相関による実世界シーンからの道路標識の検出結果。○印の位置に正しい検出信号が現れている。

行うことにより、優れたパターン認識能力を発揮できる。このようなパターン認識処理は、図2のような多重相関システムによって実行することができる。

例えば、図2の多重相関光学系で100種類の特徴パターンを検出できるとすると、各特徴パターンの組合せで分類できるパターンの種類は $2^{100} \approx 10^{30}$ 通りとなる。パターン認識に有効な組合せが全体の1%としても、 10^{28} 通りという膨大な種類のパターン識別能力を達成できることになる。対象物体の異なる形状や大きさも検出できるような特徴パターンの組合せを用意すれば、物体の変形や大きさの変化に対しても正しい認識ができるようになる。

図1と同じ実世界像から、多重相関を用いて道路標識を検出した計算機シミュレーション結果⁹⁾を図3に示す。カラー画像である入力像から濃淡画像成分と赤色画像成分を分離し、その各々に対して8種類の異なる大きさの道路標識を検出するSDFフィルター²⁾を適用した。

濃淡成分と赤色成分の処理結果の光学的 AND 演算を用いて、双方に検出信号が出力される場合を最終的な道路標識の検出信号とした。多重度が小さいにもかかわらず、背景からの偽信号の発生もなく、3つの道路標識を正しく検出できており、多重相関を用いたパターン認識の有効性が確認できる。

このような多重相関を用いたパターン認識では、どのような特徴パターンを、どのような組合せで、いくつ抽出すべきかが重要な課題になる。認識対象の代表的パターンを集めたトレーニング画像の主成分分析や因子分析を用いて特徴パターンを設計する方法を提案した⁹⁾。しかし現在のところ、それらの最適性を保証することはできておらず、今後の課題となっている。

2.2 背景に起因する偽信号の抑圧

実世界で得られる画像には、未知情報である背景や雑音が含まれている。ここでは、背景によって生じる偽信号の生成を抑える最適識別フィルターの設計方法について検討する⁸⁾。

相関光学系によるパターン認識では、入力像と識別フィルターとの相関パターンを閾値処理して、対象物体を検出する。従って、背景や雑音によって生じる相関ピークやサイドローブが閾値を越えると偽信号を生じる。そこで、得られる相関ピーク値の信号対雑音比 (SN 比; signal to noise ratio) と、得られる相関パターンのピーク対サイドローブ比 (PSR; peak to sidelobe ratio) の観点から検討を行った。

加算雑音 n を含む入力画像 g と識別フィルター f との相関強度 (相関値の絶対値の2乗) の期待値は、ベクトル表記で

$$\begin{aligned} \varepsilon[|f'(g+n)|^2] &= \varepsilon[|f'g|^2] + \varepsilon[|f'n|^2] + 2\varepsilon[(f'g)(f'n)], \\ &= |f'g|^2 + \sigma^2 |f|^2 \end{aligned} \quad (1)$$

と表される。ただし、添字 t はベクトルの転置を、演算子 $\varepsilon[]$ は期待値を表す。また、入力像と雑音は無相関で、雑音は平均値がゼロで分散が σ^2 の白色雑音と仮定した。式(1)の右辺第1項は、入力像と識別フィルターとの相関値であり、右辺第2項は雑音によって生じる偽りの信号である。つまり、雑音を含む入力画像と識別フィルターとの相関強度の SN 比は、用いる識別フィルターの総エネルギー $|f|^2$ が小さいほど良くなることからわかる。

入力像に含まれる雑音の分散値 σ^2 の推定値を変化させて MINACE フィルター³⁾ を設計し、実際の画像に適用して背景雑音に対する耐性について調べた。表1に、設計した MINACE フィルターの総エネルギー $|f|^2$ と

表1 雑音の分散 (σ^2) をパラメータとして設計した識別フィルターの特性

σ^2	総エネルギー $ f ^2$	PSR	検出された 画素数		PSR/ $ f ^2$ ($\times 10^{-3}$)
			θ	$\theta/2$	
0	34.23	0.94	551	4298	27.4
1	15.72	0.96	967	7349	61.0
10	11.44	0.96	83	722	83.9
50	10.90	0.57	483	2366	52.3
SDF	10.86	0.45	605	3388	41.4

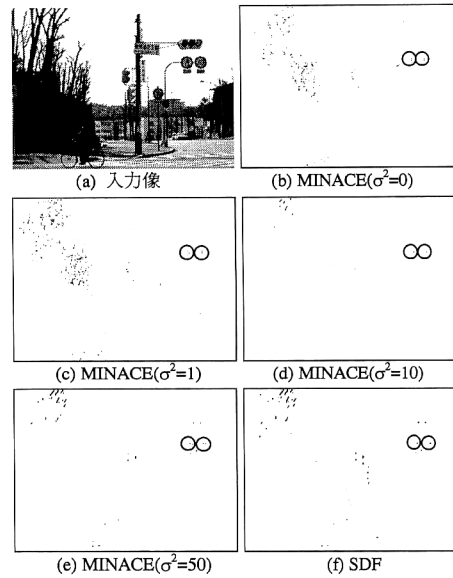


図4 雑音の分散 σ^2 をパラメータとして設計した識別フィルターによる道路標識の検出結果。

PSR の値を示す。PSR は、各トレーニング画像と識別フィルターとの相関像に対して、相関ピークを中心とする 10×10 画素の領域の内部の平均値とその外部の平均値の比として求めた。

図4(b)~(f)に、表1で示した各識別フィルターと図4(a)の入力像との相関演算の結果を示す。黒い領域が道路標識の検出信号を表す。表1の検出された画素数の欄には、閾値 θ を用いたときに検出された画素数と、閾値を $\theta/2$ にして検出された画素数をそれぞれ示した。この閾値 θ は、少なくとも2つの道路標識が検出される閾値の中で最大の値に設定し、背景から生じる偽信号が最も少ない状態を選んだ。図4の処理結果と表1の検出された画素数から、図4(d)の結果が最も偽信号の発生が少なく良好な結果と言える。

識別フィルターの総エネルギーが小さいほど、背景雑音の影響が小さく、PSR が大きいほど相関ピークのサ

イドローブによる偽信号の発生を抑えることができる。従って、小さな総エネルギーをもち大きなPSRを示す識別フィルタが、背景雑音を抑圧する意味で最適な識別フィルタになる。しかし一般的に、相関ピークのサイドローブを抑えると識別フィルタの総エネルギーは増大するため、両方の性質を同時に満足する識別フィルタを設計することは困難である。

そこで、背景から生じる偽信号の発生を抑制する評価指標として、識別フィルタのPSRと総エネルギーの比 ($PSR/|f|^2$) を考えることができる。図4で用いた各識別フィルタは、表1の最右欄の評価値をもち、評価値が大きな値をもつ識別フィルタが実験的にも背景雑音の抑圧に効果的であることが確認できる。

2.3 光学的適応閾値処理による多物体検出

実世界のシーンでは、明るい物体と暗い物体が混在していたり、暗い物体と明るい背景シーンが同一シーンの中に含まれている場合がある。従来の光システムでは、明るい物体に検出感度を合わせると暗い物体は検出できなくなったり、暗い物体に検出感度を合わせると背景からの偽信号が増大し、異なる明るさをもつ複数物体を同時に検出することは困難であった。

そこで、多重相関光学系と非線形処理を用いて、入力像の局所的な明るさに対応して検出感度を動的に変えることができる光学的適応閾値処理を用いたパターン認識を提案した¹⁰⁾。この方法では、入力像の明るい場所では明るい物体用の検出感度が、入力像の暗い場所では暗い物体用の検出感度が自動的に設定できる。

入力像 g の局所的明るさを求める相関フィルタを f_0 、対象物体を検出するための識別フィルタを f_1 としたとき、入力像との相関演算で得られる両者の相関強度の比は、入力像の明るさが変わっても一定である。そこで両者の比を一定値 θ で閾値処理することにより、入力像の明るさに依存しないパターン検出ができる。つまり、

$$|f_1 \star g|^2 / |f_0 \star g|^2 > \theta. \quad (2)$$

演算子 \star は相関演算を表す。式(2)は次式と等価である。

$$|f_1 \star g|^2 > \theta |f_0 \star g|^2 \quad (3)$$

つまり、識別フィルタ f_1 と入力像 g との相関パターンを、式(3)の右辺で表される入力像の局所的明るさを動的閾値として閾値処理することにより、明るさに依存しないパターン検出が実現できる。式(3)をそのまま光学系で実現することは困難であるが、次式のように変形すれば図5で示す光学系で実現できる。

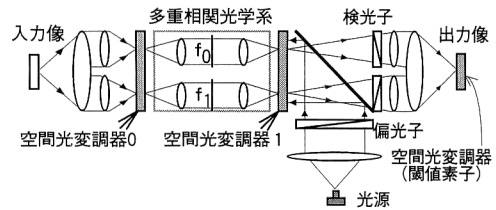


図5 光学的適応閾値処理を実行するパターン認識システム。

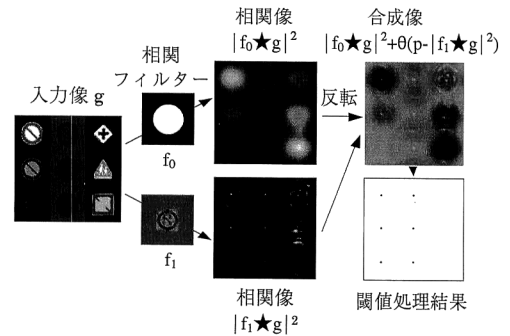


図6 光学的適応閾値処理の手続き。

$$|f_1 \star g|^2 + \theta(p - |f_0 \star g|^2) > \theta p \quad (4)$$

一定値 p は相関強度パターン $|f_0 \star g|^2$ の最大値を表し、パターン $(p - |f_0 \star g|^2)$ は相関パターン $|f_0 \star g|^2$ の明るさを反転した画像である。

図5の光学系では、多重結像系で複製した入力像 g を2つの相関光学系に入力して、相関フィルタ f_0 と識別フィルタ f_1 との相関パターンを空間光変調器1上のそれぞれ異なる場所に得る。相関フィルタ f_0 によって得られた相関パターンは明るさが反転した画像として、また識別フィルタ f_1 によって得られた相関パターンは明るさを反転しない画像として、それぞれ空間光変調器1から読み出す。空間光変調器1に偏光を用いた変調器を用いるならば、一方の読出し光に対する検光子の偏光面をもう一方の読出し光に対する検光子の偏光面より90度回転させることにより、一方のみを反転像として読み出すことができる。空間光変調器1から読み出された2つの相関強度パターンを、強誘電性液晶などを用いた閾値を有する空間光変調器2上に重畳して結像させることにより、式(4)の適応閾値処理が実現できる。この処理の流れを図6に示す。

この適応閾値処理を用いて、明るさの異なる複数の道路標識の同時検出を行った計算機シミュレーション結果を図7に示す。識別フィルタ f_1 と相関フィルタ f_0 は図6に示すものを用いた。この識別フィルタ f_1 は、

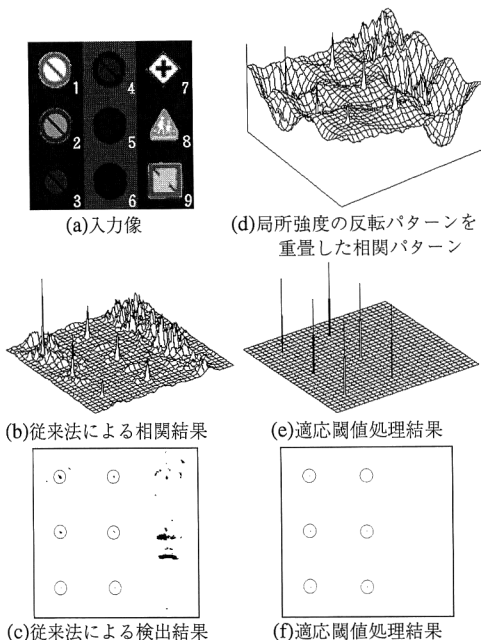


図7 明るさの異なる複数物体の同時検出。

丸い道路標識のみを検出するように作製した SDF フィルター²⁾である。図7(a)の入力像は、明るさの異なる9つの道路標識を含んでおり、それぞれの相対強度は図中の番号順に3, 2, 1, 1, 0.8, 0.6, 2.5, 3, 3である。この入力像と識別フィルター f_1 の相関強度パターンは図7(b)のようになる。暗い検出パターンの相関ピーク値は、検出パターン以外の道路標識によって出力される相関値よりも小さいため、一定の閾値処理では、図7(c)のように正しい検出が行えない。図7(b)の相関強度パターンに局所的明るさの反転像を加えたパターンは図7(d)のようになり、適応閾値処理によって図7(e)のように丸い道路標識のみを正しく検出することができるようになる。

2種類の実世界のシーンに対して適応閾値処理を用いて道路標識の検出を行った結果を図8に示す。図8(b), (e)は、それぞれ(a), (d)の入力像に対して適応閾値処理によって得られた検出結果であり、図8(c), (f)は、それぞれ(a), (d)の入力像から従来法である SDF フィルターを用いて検出した結果である。図8(b)では3つの道路標識すべてを検出できている。また図8(e)では、4つの道路標識のうち2つを検出できている。しかも両者ともに、従来法に比べて、背景に起因する偽信号の発生が抑えられている。この結果より、適応閾値処理を用いたパターン検出処理の有効性が確認できる。

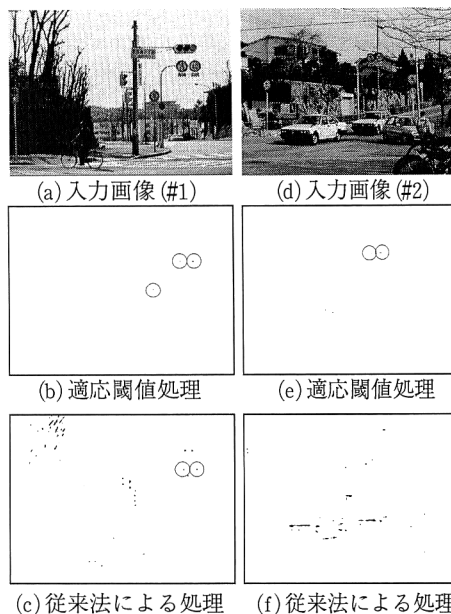


図8 適応閾値処理による実世界シーンからの道路標識の検出。○印の位置に正しい検出信号が現れている。

3. まとめ

実世界を対象にした光学的なパターン認識技術について、筆者の研究を紹介した。実世界のシーンを対象とした場合、未知情報である背景に対して偽信号を生じないロバストな処理が必要であり、多重相関演算を用いたパターン認識処理、背景雑音を抑える識別フィルターの設計手法、光学的適応閾値処理が有効であることを示した。これらの手法は、図2の多重相関システムの中に組み入れることが可能である。ここでは、多重相関システムで達成できる処理速度や実現可能性について議論し、本稿のまとめとした。

3.1 処理速度

光ビジョン技術を用いる最大の狙いは処理の高速化にある。図2に示す多重相関システムの処理速度は、用いる空間光変調器の応答速度に大きく依存している。これまでに発表されている強誘電性液晶を用いた空間光変調器の応答速度は数十～数百 μs であり^{11,12)}、これを用いるならば多重相関システムによって1msの処理速度を達成することも可能である。

これと同じ処理(100個の多重相関処理)を、電子コンピュータで実行した場合の処理時間の推定値を図9に示す*1。512×512画素の入力像に対して、現在の最高速の並列スーパーコンピュータ(300 GFLOPS¹³⁾)

*1 演算速度の算定方法については付録にて述べる。

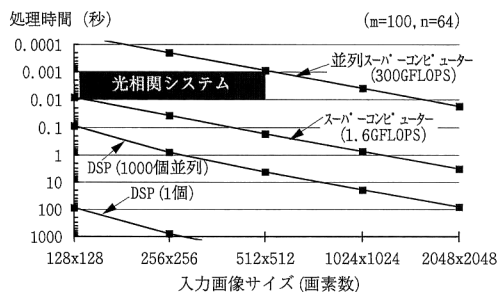


図9 光相関システムと電子コンピューターで実行できる相関演算速度の比較。識別フィルターのサイズを64×64画素、相関演算の多重度を100、DSPのサイクルタイムを50 nsとした場合。

を用いた場合、約1 msの処理時間が必要となる。また、50 nsの高速サイクルタイムをもつDSP (digital signal processor) を1000個完全並列で用いたとしても、識別フィルターの大きさが64×64画素の場合には、数秒の演算時間が必要になる。

このように、現在の最高速の並列スーパーコンピューターでようやく達成できる処理速度 (図9の網点領域) を、多重光相関システムで実現できる可能性があり、光ビジョンシステムを用いる意義は大きいと言える。

3.2 空間光変調器の感度

多重相関演算に必要な多重結像系で得られる像の照度は、近似的に結像にかかわるレンズのFナンバーの2乗に反比例する¹⁴⁾。Fナンバーが3~5程度のマイクロ・レンズアレイが既に作製されており¹⁵⁾、多重結像による像の明るさは大きな問題とはならない。

しかし、実世界像を直接撮影する初段の空間光変調器では、照明などによって光量を増やすことが困難であり、しかも1 msの高速動作で実世界像をとらえるためには感度の高い素子が必要となる。これまでに試作された液晶空間光変調素子の感度は数 $\mu\text{W}/\text{cm}^2$ ~数百 $\mu\text{W}/\text{cm}^2$ であり^{11,16)}、現在の一般的 CCD カメラの撮像面最低照度が約 $1 \text{ lx}^{17)}$ 程度 (波長 555 nm で約 $0.15 \mu\text{W}/\text{cm}^2$) であるのに比べると感度不足であり、今後の改良が期待される場所である。

3.3 空間光変調器の空間帯域積

システムで扱うことのできる画像の大きさは、用いる空間光変調器の空間分解能と有効サイズの積 (空間帯域積) で決まる。発表されている強誘電性液晶を用いた空間光変調器では、約2500以上の空間帯域積をもっており¹¹⁾、多重度100 (10×10) の場合、512×512画素までの入力像の多重相関を実行できることになる。この空間

帯域積を大きくできれば、より高分解能の入力画像を扱うことが可能となり、図9の光相関システムの扱える領域を広げ、最高速のスーパーコンピューターでも達成できない処理速度が実現できるようになる。今後のデバイス開発に期待したい。

実世界を対象にしたビジョン技術は、人とマシンが共存し協調するために欠かせない技術であり、人にやさしい技術開発が求められているなかで、今後ますます重要になってくると考えられる。今後、光ビジョン技術の研究開発が大いに進展して、小型で高速の知的光ビジョン・センサーなるものが開発され、このような要求に応えられるようになることを期待している。

文 献

- 1) A. Vanderlugt: "Signal detection by complex spatial filtering," IEEE Trans. Inf. Theory, **IT-10** (1964) 139-145.
- 2) D. Casasent: "Unified synthetic discriminant function computational formulation," Appl. Opt., **23** (1984) 1620-1627.
- 3) G. Ravichandran and D. Casasent: "Minimum noise and correlation energy optical correlation filter," Appl. Opt., **31** (1992) 1823-1833.
- 4) C. F. Hester and D. Casasent: "Multivariate technique for multiclass pattern recognition," Appl. Opt., **19** (1980) 1758-1761.
- 5) H.-Y. S. Li, Y. Qiao and D. Psaltis: "Optical network for real-time face recognition," Appl. Opt., **32** (1993) 5026-5035.
- 6) Y. Osugi, Q. Zhan and T. Minemoto: "Hybrid binary subtracted joint transform correlator for a large number of reference patterns using a $\text{Bi}_{12}\text{SiO}_{20}$ (BSO) spatial light modulator," Opt. Rev., **1** (1994) 159-162.
- 7) 伊藤義之, 亀丸俊一: "特徴抽出に基づくマッチングとニューラルネットワークを用いたパターン認識", 光学, **24** (1995) 32-39.
- 8) K. Matsuoka, M. Taniguchi and Y. Mokuno: "Application of optical multiple-correlation to recognition of road signs: the ability of multiple-correlation," *Proceedings of Optical Computing* (Inst. Phys. Publ., London, 1995) pp. 305-308.
- 9) 松岡克典, 谷口正樹, 野野由明, 大谷和男: "因子分析を用いたパターン識別フィルタの設計", 光学連合シンポジウム旭川'93講演予稿集 (1993) p. 35.
- 10) K. Matsuoka, M. Taniguchi and Y. Mokuno: "Optical adaptive processing for an intensity invariant pattern recognition," *CLEO/Pacific Rim '95* (1995) p. 290.
- 11) 岩城忠雄: "強誘電性液晶ライトバルブとその応用", 光学, **19** (1990) 295-301.
- 12) 黒川隆志, 福島誠治, 平林克彦: "液晶を用いた光処理デバイス", 光学, **21** (1992) 142-146.
- 13) "並列スーパーコンが出そろう", 日経コンピュータ, 7月号 (1994) 22.
- 14) 鶴田匡夫: 光の鉛筆, (新技術コミュニケーションズ, 1988) p. 8; または辻内順平: "光学系の特性", 光応用技術 (日本オプトメカトロニクス協会, 1987) p. 13.
- 15) K. Hamanaka and T. Kishimoto: "Multiple imaging and

multiple Fourier transformation using microlens arrays," Jpn. J. Appl. Phys., **29** (1990) L1277-L1280.

- 16) 平林克彦, 福島誠治, 大野正善, 黒川隆志: "STN 液晶を用いた空間光変調器", 秋季応用物理学会学術講演会 (1990) p. 753.
- 17) 黒田隆男: "最近の CCD における高感度低雑音技術", Optics E, No. 115 (1989) 81-86.

(付録) 図 9 の演算速度の算定

入力画像を $N \times N$ 画素, 識別フィルターの大きさを $n \times n$ 画素, 相関演算の多重度を m とすると, すべての相関演算に必要な掛け算回数は, $mn^2(N-n+1)^2$ 回

となる. この掛け算回数は, FFT (fast Fourier transform) を用いて空間周波数面で処理することにより, $mN^2(1+\log_2 N)$ 回に減らすことができる. この場合, 空間周波数面での広いダイナミックレンジを確保するために, 一般的に浮動小数点演算が必要となる.

図 9 では, $m=100, n=64$ の場合について示した. DSP (50 ns) を用いた処理速度の算定には前者の掛け算回数から, スーパーコンピュータを用いた処理速度の算定には後者の掛け算回数 (FFT を用いた回数) から求めた.