

# 研究

## 中心窩を有するロボットカラービジョンシステム

山羽 和夫\*・富長 博\*\*・三宅 洋一\*\*\*・矢口 博久\*\*\*  
戸倉 暢史\*\*\*\*・塚本 満早\*\*\*\*

\* 工業技術院機械技術研究所 〒305 つくば市並木 1-2

\*\* 茨城県工業技術センター 〒311-31 茨城県東茨城郡茨城町長岡 3781-1

\*\*\* 千葉大学工学部 〒263 千葉市稲毛区弥生町 1-33

\*\*\*\* 九州松下電器(株)生産技術研究所 〒812 福岡市博多区美野島 4-1-62

(1994年8月1日受付, 1995年1月6日受理)

### Robot Color Vision System with Fovea

Kazuo YAMABA,\* Hiroshi TOMINAGA,\*\* Yoichi MIYAKE,\*\*\* Hirohisa YAGUCHI,\*\*\*  
Nobufumi TOKURA\*\*\*\* and Mitsuhaya TSUKAMOTO\*\*\*\*

\* Mechanical Engineering Laboratory, AIST, 1-2, Namiki, Tsukuba 305

\*\* Ibaraki Prefectural Industrial Technology Center,  
3781-1, Nagaoka, Ibaraki-machi, Higashi-Ibaraki-gun, Ibaraki 311-31

\*\*\* Faculty of Engineering, Chiba University, 1-33, Yayoi, Inage-ku, Chiba 263

\*\*\*\* Kyusyu Matsushita Electric Co., Ltd., 4-1-62, Minoshima, Hakata-ku, Fukuoka 812

(Received August 1, 1994; Accepted January 6, 1995)

Recent studies on the distributions of human and baboon retinal cones/rods have shown that their retinal cones/rods are mapped histochemically against visible lights. In order to realize new images corresponding to a retinal image for robotic vision, a novel color information model is proposed. This model is composed of a visual input unit with four CCD cameras, an image processor and a visual early-processing unit. In order to recognize colored objects, the visual early-processing unit, based on a neural network model, is used. To evaluate this model, a simple experiment was done to demonstrate human-like visual perception using the density profile around foveae of baboon eyes. Each profile of baboon eyes is solved by employing the multi-regression analysis from Marc's data. As the results, such colored shapes as triangles, squares and circles with five different colors could be recognized efficiently.

### 1. はじめに

近年, 半導体集積化技術の進歩に伴ってより高密度かつ小型の二次元半導体固体素子が登場してきている。この固体素子を次世代の高度ロボットの視覚として用いる場合, 最も問題になることは, ロボットの眼としての解像度に直接関連する素子の密度と視覚システムとしてカラーの高速処理が可能か否かであろう。

このうち視覚の解像度に関連する素子の密度について考えると, 1993年の現状では約600万画素の固体素子が製作されている。しかしながら, 1億以上ある人間の

網膜の光受容細胞と比べると実装密度, 機能(ダイナミックレンジおよび応答性)などの面でまだまだ劣っている。

ロボットビジョンシステムとしてのカラー情報処理手法については, 従来より赤, 緑, 青の3色の情報をそのままグレースケールに置きかえて処理されていた。この方法は通常の白黒のイメージセンサーと比べて約3倍の時間がかかるため, 色画像システムの演算の簡略化や色画像圧縮などの方法も検討されてきた。

一方, 人間や霊長類の場合, 中心窩付近で光受容細胞の密度が集中して, 視覚入力情報に対しより有意な情報

を選定し、かつ、網膜細胞層においては視覚入力情報を時間的、空間的の微積分を行うことで情報圧縮がなされていることが知られている。この情報圧縮などのメカニズムを解明することも生体内においてなぜ高速の視覚情報処理が可能であるかを探るキーポイントの一つになっている。

以上の観点からここでは、中心窩付近の密度分布に着目し、人間の分光分布特性に類似した視覚情報入力機器を使用して人間に近いといわれる霊長類の中心窩付近の光受容細胞の分布の近似式を求め、これにより得られた視覚情報を人間に類似した視覚プロセスによって処理した実験結果について述べる。

2. 網膜モデルと設計概念

Fig. 1 に人間の場合の眼球の構造と網膜モデル<sup>1)</sup>および本研究で用いた視覚モデルを示す。Fig. 1 左に示すように、網膜は受容器細胞と呼ばれる視細胞（錐体細胞 C, 桿体細胞 R）、水平細胞 H, 双極細胞 B, アマクリン細胞 A, 神経節細胞 G の 5 種類の細胞で構成される三つの層からなっている。

人間の場合、網膜には、錐体細胞、桿体細胞を合わせて約 1 億 3000 万個程度の受容器細胞があるが、これらが眼球内に一様に分布しているのではなく、視角の中心に向かってそれらの密度が高くなっている。ただし、視

軸の中心線上にある中心窩と呼ばれる小さなくぼみの近辺では桿体細胞は存在せず、錐体細胞のみが存在し、視力も最も高くなっている。

また、中心窩、傍中心窩など網膜上の部位の違いによる系列処理、並列処理、ぼけ処理なども提案されている<sup>2,3)</sup>が、視覚心理物理学的に、まだ問題点も散在している。

錐体全体の分布について古くは Osterberg (1935)、最近では Curcio (1990) が研究発表している。Marc らは霊長類 (ヒヒ) の錐体の状況について、青色錐体が周辺に規則的に分布し、また、赤緑錐体はいたるところに不規則に分布、さらにこれらの錐体密度が M (緑) 錐体 > L (赤) 錐体 > S (青) 錐体の順であることを報告している<sup>4)</sup>。また、de Monasterio らはマカク猿の青色錐体の分布を計測し、青色錐体が中心窩ではゼロ、0.75~1.50° でピークに達し、その後、減少という人間の心理物理から求められたものに一致しているとの結果を得ている<sup>5)</sup>。

L 錐体と M 錐体の比は 4 対 3 ともいわれていた<sup>6)</sup>が、全体形状、先端のテーパが似ていることからこれまでそれらの比を正確に出すことは非常に困難であった。最近、Cicerone らの心理物理学的手法でその L/M 比が 1.3~2.6 で、L 錐体が M 錐体の約 2 倍の密度を有していることが報告されている<sup>6)</sup>。S 錐体の分布について、

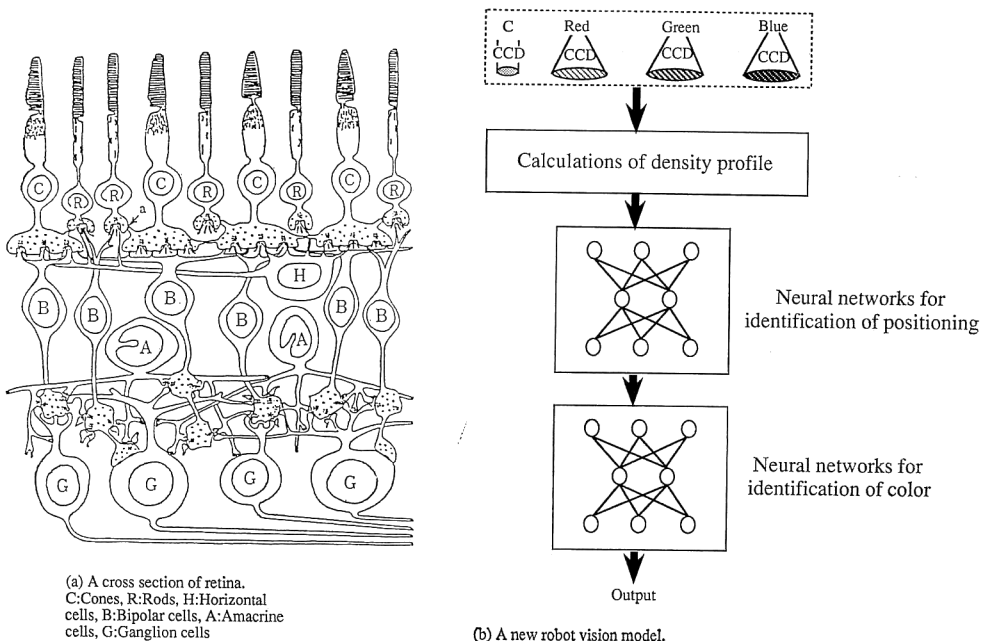


Fig. 1 A cross section of human retina which was proposed by Dowling and Boycott in 1966 and the new robot vision model.

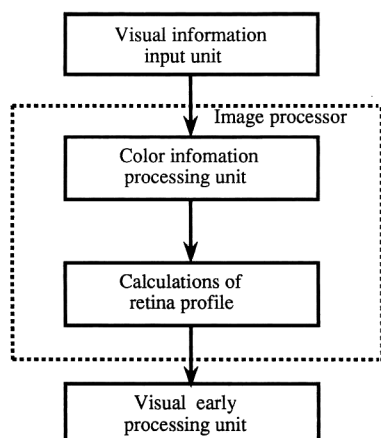


Fig. 2 Block diagram of a new type of color information system for robot vision.

このほか、Williams<sup>7)</sup>による研究も紹介されている。

Fig. 2に本視覚モデルシステムのブロックダイアグラムを示す。図に示すように本視覚モデルシステムは光受容細胞に相当する視覚情報入力部 (Fig. 1 右上の破線部に相当)、桿体、錐体に相当する四つの種類の視覚情報 (Fig. 1 右上の破線部内の四つの CCD カメラすなわち C CCD, Red CCD, Green CCD, Blue CCD から得られる情報を示す) を同時に処理可能な画像処理装置 (カラー情報処理部および分布演算処理部から構成されている)、マイクロコンピュータからなる視覚初期処理部で構成されている。本モデルは人間の目とは構造的に異なるが、視覚情報入力部の分光分布特性を、1台の固体カメラ出力は桿体細胞、3台の固体カメラ出力はそれぞれの種類の錐体細胞の分光吸収特性に類似させている<sup>1,8)</sup>。視覚処理部は二つの3層からなるニューラルネットワーク構造を持ち、ソフトウェアによってマイクロコンピュータ内で対象物認識のための最終処理がなされている。

### 3. 桿体分布および錐体分布の重回帰分析

前節で述べたように錐体の分布に関して、最近、サルを計測した事例などが数多く発表されているが、ここでは簡単のため、多くの文献に引用されている Woodson のデータ<sup>1)</sup>に基づいて人間の桿体および錐体の分布に関する重回帰分析を行った。その結果、各光受容細胞数 ( $y$ ) と中心窩からの隔たり ( $x^0$ ) の間の関係が導出することができた。桿体の受容細胞数  $y_r$  では式(1)のようになる。また、錐体の受容細胞数  $y_c$  は式(2)のようになった。

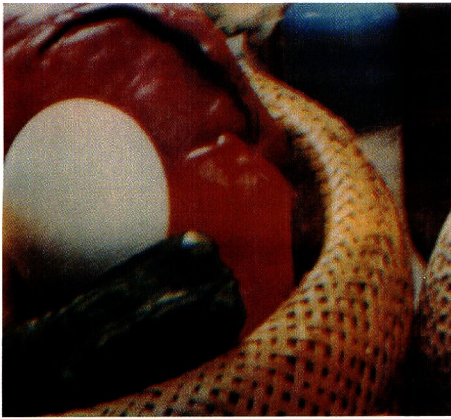
$$y_r = 0.411x^3 - 16.782x^2 + 283.761x + 0.714 \quad (1)$$

$$y_c = -0.005x^5 + 0.33x^4 - 8.24x^3 + 99.886x^2 - 619.583x + 1929.85 \quad (2)$$

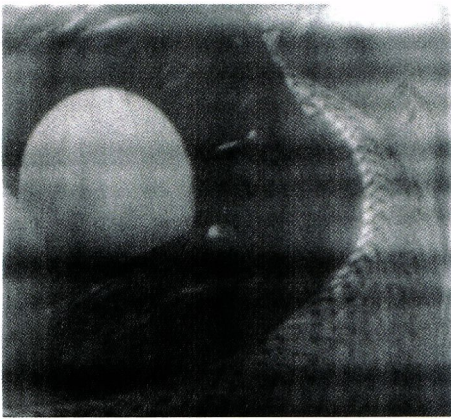
式(1)、式(2)とも  $F$  検定した結果、その有意差の確率  $P_t$  は  $P_t \leq 0.001$  であった。これらの式を各カメラの画像出力に適用した。本節では光受容細胞数の最大値を信号強度の最大値 255 とし、相対強度でディスプレイ画面上に表示することにした。また、ディスプレイ画面の縦方向、横方向の中心を中心窩と仮定し、この中心から同心円状に隔たり  $x$  をとった。  $x$  の値はディスプレイ画面の横方向の最大幅を  $10^\circ$  視野と仮定し、中心 (中心窩) から  $0.2^\circ$  ごとにドーナツ状に画像処理していった。この画面は視細胞モデルでは水平細胞前のモデルに相当する。今回はこの部位で、画像評価を行うことにした。評価の対象に画像処理でよく用いられるガウシアン分布した場合も求めてみた。

Fig. 3 に Red CCD, Green CCD, Blue CCD の三つの CCD カメラおよび C CCD カメラからの原画像を示す。Fig. 3 (a) に示すように用いた原画像は黄色の竹編み籠の中に、上から赤いキャベツ、白い卵、その隣に赤いピーマン、下に緑のピーマン、さらに編み籠の外に青色のコップを置いた状態を撮映したものである。

Fig. 4, Fig. 5, Fig. 6 に式(1)、式(2)による結果およびガウシアン分布について示す。式(1)、式(2)から Fig. 4, および Fig. 5 への変換は以下のように行った。式(1)および式(2)で求めた光受容細胞数の最大数を  $Y_{max}$ , 変換する画素部分での細胞数を  $C_x$  とする。また、その画素部分での画素の明るさを  $B_x$  とすると、そのときそれらの間には  $Y_{max}/255 = C_x/B_x$  という関係があるのでこれを用いると  $B_x = 255 \cdot C_x/Y_{max}$  が導出できる。ここではこれに取込み画像のその位置における明るさのレベルを電圧に変換した値を乗じてその点の明るさとして表示した。また、Fig. 6 については  $B_x$  の代わりにガウス関数を用いて変換し、表示させた。これらの図から次のことがいえる。Fig. 4 に示すように桿体細胞に相当するカメラ出力から、その処理画像は、輪郭が元の画像より明確にとれるため空間微分しやすく、明暗覚センサー、形状認識するのに適していることがわかる。Fig. 5 から、式(2)を通した色画像はガウシアン分布を通した色画像 (Fig. 6) より表面性状を局部的 (約  $2^\circ$  の範囲で) に観察する必要があるといえよう。画像処理の面からみれば判定するターゲットを限定しやすいなどの効果があるともいえる。本実験では、光源には簡単のため、暗箱中で反射カバー付きインバータ蛍光灯 (18



(a) Output of RGB camera.



(b) Output of C CCD camera.

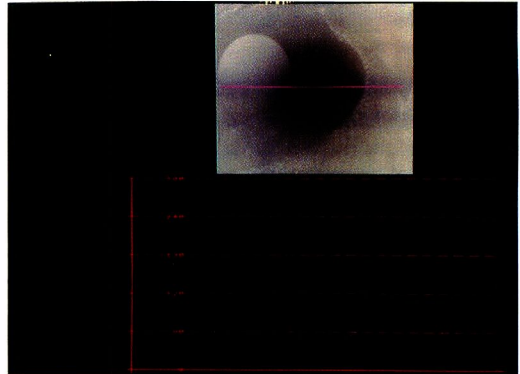
**Fig. 3** Input image before calculations of density profile.

W: FL 20 SSEX/18-G) を用い、これを斜め約  $45^\circ$  の上方約 15 cm 位置から被写体に照射させた。

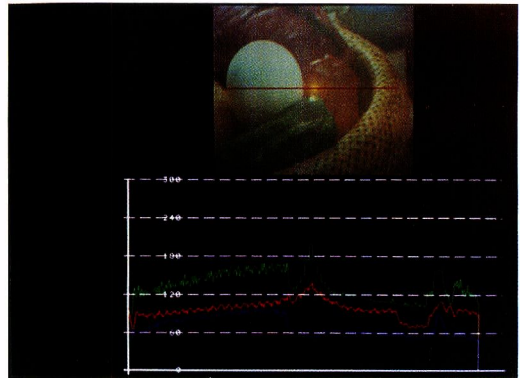
人間の場合、各錐体の分布がそれぞれ異なる<sup>5-7)</sup>ことはすでに前節で述べたが、それに基づいた画像評価を行うため、式(2)に代る各錐体別の特性を求めることにした。しかしながら、人間の錐体分布が実際に計測されはじめたのがごく最近であり、L/M 比についても発表論文によってまだばらつきがあるため、本論文では、V4以降の高次野まで解析が進んでいる霊長類を人間の代りに取りあげ、それをモデルにした各錐体分布の形成を行った。そのため、ここでは Marc ら<sup>4)</sup>が発表した霊長類(ヒヒ)の錐体の各錐体数を示すグラフから L, M, S 錐体の分布に関する重回帰分析を行った。Table 1 に Marc らによって導出された各錐体から読み取った細胞数を示す。これらの各細胞数を前記と同じ方法を使用し

て各光受容細胞数 ( $y$ ) と中心窩からの隔たり ( $x^\circ$ ) の間の関係を導出した。Fig. 7 に解析結果のグラフを示す。L 錐体の受容細胞数  $y_L$ , M 錐体の受容細胞数  $y_M$ , S 錐体の受容細胞数  $y_S$  とすると、それぞれ式(3)~(5)で表せる関係式が得られた。

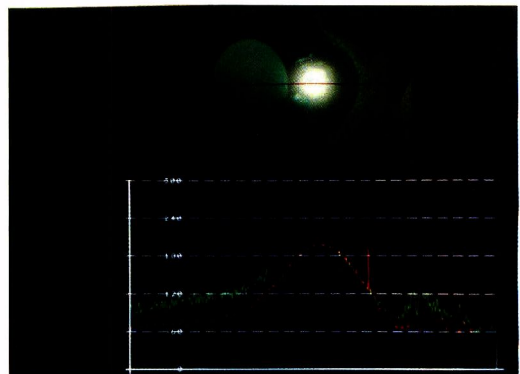
$$y_L = 8.915 \times 10^{-5} x^6 - 0.005 x^5 + 0.093 x^4$$



**Fig. 4** The results of image processing due to model of rods.



**Fig. 5** The results of image processing due to model of cones.



**Fig. 6** The results of Gaussian function.

**Table 1** Each cone profile data from Marc's experiment. Unit (1000 cell/mm<sup>2</sup>)

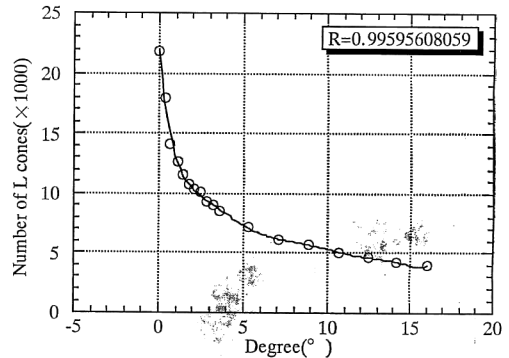
Degree (°)	S cones	L cones	M cones	Total
0	2.70	21.78	42.12	66.96
0.4	4.32	18.00	26.64	48.96
0.7	6.30	14.04	20.16	40.32
1.1	6.88	12.60	18.00	37.80
1.4	6.84	11.52	16.56	34.20
1.8	6.66	10.80	15.66	31.68
2.1	6.34	10.37	15.12	30.06
2.5	6.05	10.08	14.40	28.62
2.8	5.58	9.36	13.86	27.36
3.2	5.40	9.00	13.50	26.28
3.6	4.86	8.46	12.96	25.56
5.3	3.78	7.20	11.52	21.96
7.1	2.70	6.12	10.44	19.08
8.9	2.23	5.69	9.36	16.74
10.7	2.05	5.11	8.35	14.94
12.5	2.05	4.61	7.49	13.14
14.2	2.02	4.25	6.84	11.70
16.0	1.98	3.96	6.23	10.62
17.8	1.91	3.78	5.94	9.61
19.6	1.87	3.71	5.83	9.00
21.4	1.87	3.60	5.58	8.46
23.1	1.87	3.53	5.40	8.06
24.9	1.84	3.46	5.26	7.63
26.7	1.84	3.38	5.15	7.38
28.5	1.80	3.31	5.29	7.27
32.0	1.80	3.13	4.68	6.91
35.6	1.80	2.95	4.36	6.70
39.2	1.80	2.88	4.07	6.37

$$-0.915x^3 + 4.605x^2 - 11.888x + 21.514 \quad (3)$$

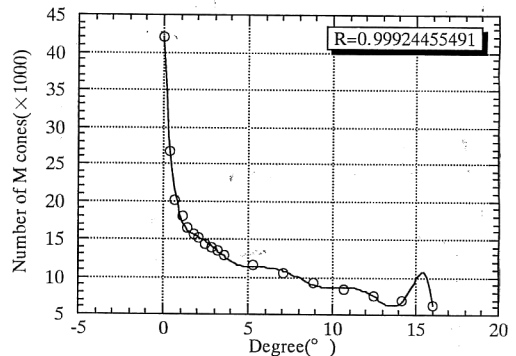
$$y_M = -2.346 \times 10^{-6}x^9 + 1.714 \times 10^{-4}x^8 - 0.005x^7 + 0.089x^6 - 0.897x^5 + 5.573x^4 - 20.968x^3 + 45.648x^2 - 53.485x + 41.98 \quad (4)$$

$$y_S = -7.013 \times 10^{-7}x^8 + 5.199 \times 10^{-5}x^7 - 0.002x^6 + 0.027x^5 - 0.267x^4 + 1.57x^3 - 5.245x^2 + 7.976x + 2.497 \quad (5)$$

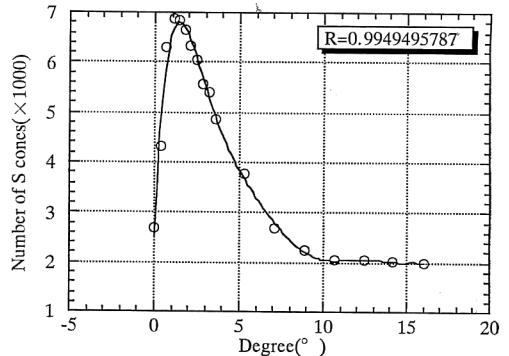
Fig. 4 を得たときと同じ方法で上式から得られた色画像の状況を Fig. 8 に示す。実際には中心視および近中心視に相当する画像の中心位置から 4° までを上式によって処理し、それから外側は誤差が発生する恐れがあることから適宜、重回帰分析を行ったグラフから錐体数を読み取って、これをその点における受容細胞数とした。また、上式を重回帰分析で求める際、中心点を計算上の理由から便宜的に 0.01° とし、これより内側は Table 1 の 0° の各値から読み取って用いた。Fig. 8 の S 錐体に相当する画像から中心窩にあたる付近での信号変化率が大きいことがわかる。また、L、M 錐体に相当する画像からは、その波長領域の入力値に対してのみ感度が上昇



(a) The results of the multi-regression analysis from L cones.



(b) The results of the multi-regression analysis from M cones.



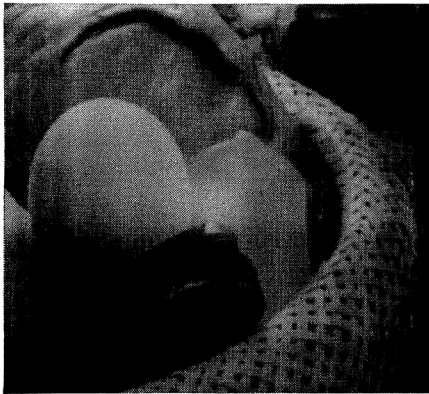
(c) The results of the multi-regression analysis from S cones.

**Fig. 7** The results of solution by employing the multi-analysis from Marc's data. Each table shows coefficients of the multi-regressing analysis.

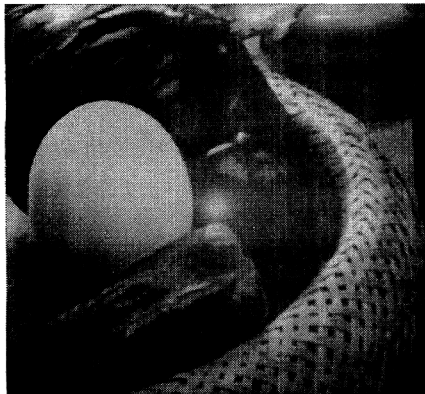
していることもわかる。この図から中心窩付近での各錐体の違いの状況が把握できよう。

#### 4. 実験および評価

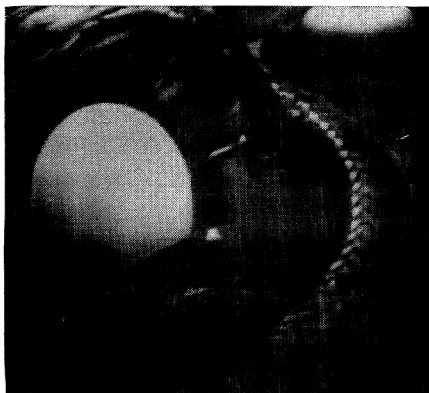
以上の結果に基づいて簡単な認識実験を行った。実験では中心窩付近での中心視 (ここでは中心視を 0°~2° の円内と定義する)、近中心視 (同様に、近中心視を 2°~4° のドーナツ状の部分と定義する)、および周中心視 (同



(a) The results of the image of L cones.



(b) The results of the image of M cones.



(c) The results of the image of S cones.

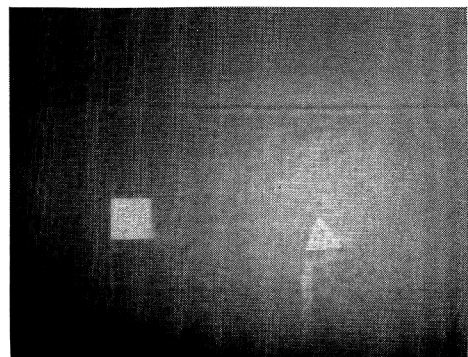
**Fig. 8** The results of the output images after the multi-regressing analysis.

様に、周中心視を近中心視の外周の外側  $4^\circ \sim 9^\circ$  の部位と定義する)に被写体をおき、それらの認識プロセスの比較を行った。最初に学習すべき3種類の形状の被写体を呈示し、学習させた。この被写体には、赤(修正マン

セル値 4R 4.5/14), 黄(同 5Y 8/13.5), 緑(同 4G 5.5/10.5), 青(同 10B 4/12), 紫(同 2P 3.5/12.5), 白(同 N9.5)の色票の上に黒紙(同 N1の色票)を円形, 三角形, 四角形にくり抜いて用いた。**Fig. 9**に用いた被写体の一例, ここでは, 三角形(黄)と四角形(黄)を示す。

**Fig. 2**で示す視覚情報入力部に入力された画像はカラー情報処理部(**Fig. 1 (b)**の破線部内に相当し, ここでは単に各カメラの制御などを行っている)を経由して分布演算処理部で各成分ごとに処理された後, 視覚初期処理部(**Fig. 1 (b)**の二つのニューラルネットワークに相当している)に入る。この視覚初期処理部は対象物同定用ニューラルネットワークモデルと色処理用ニューラルネットワークモデルの2段構造で構成されている。各ニューラルネットワークモデルは入力層, 中間層, 出力層の3層構造をなしており, 学習はバックプロパゲーションに基づいてなされる。本研究ではニューロンの入出力特性を定める非線形関数(シグモイド関数)の傾き係数を0.75, 荷重ゲイン0.2, 閾値ゲイン  $-0.2$ と設定した。また, 今回の学習では対象物同定用ニューラルネットワークモデルの入力層を1,024セル, 中間層を5セル, 出力層を9セル(対象物の置かれている三つの位置, すなわち中心視, 近中心視, 周中心視と用いた図形のカテゴリー数に対応)とし, 色処理用ニューラルネットワークモデルでは入力層を3セル, 中間層を5セル, 出力層を7セル(色のカテゴリー数に対応)とした。

入力層のセル数による演算時間などの制限から, ここでは簡単のため対象物同定用ニューラルネットワークモデルの入力パターン画素数を  $32 \times 32$ (入力層の1,024セルに対応している)に設定した。そのため  $512 \times 512$ の画像の中の  $20 \times 20$ の1ブロックをニューラルネットワークモデルの入力パターンの1画素分とした。各カメ



**Fig. 9** The example of the input image.

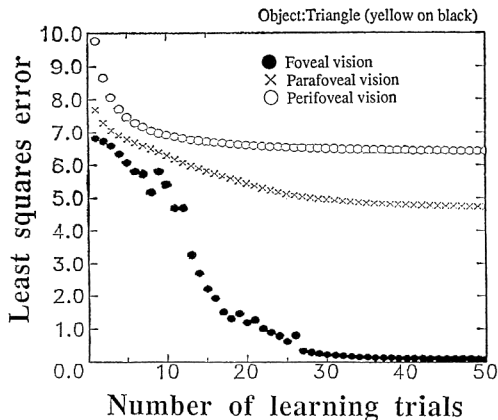


Fig. 10 The results of learning for recognizing.

ラの各画素が64階調で構成されており、これを1ブロックごと(400画素分)に集計(0~25, 200階調)した。さらにその集計したものを0.0~1.0に正規化し、ニューラルネットワークモデルの各画素の入力信号とした。対象物同定用ニューラルネットワークモデルで対象物の形状および位置を同定した後、逆投影図法に近傍法を加えた手法<sup>9)</sup>を使用して同定した対象物の位置の色に関する情報処理を色処理用ニューラルネットワークモデルで行った。このニューラルネットワークモデルでの入力手法は変形反対色理論<sup>10)</sup>に基づいて処理した。

本実験はまず円形、三角形、四角形の被写体を用いた学習回数を50回行った後、色の学習を行った。また、データは、学習回数による誤認識数の結果と出力層誤差の最小自乗和の値を採取したが、評価は最小2乗和を用いて行うことにし、学習回数に比べて誤差の最小自乗和が早くゼロに近くなる特性をその学習効果が良いと定義した。学習の結果をFig. 10に示す。各グラフとも縦軸は出力層誤差自乗和を示し、その最大値がグラフ上端にくるように正規化した。

本実験結果から、色、形状の双方とも評価では学習効果は中心視で最も良く、続いて、近中心視、周中心視の順であった。色の判定については、準備したすべての色で中心視が良好な結果が得られたが、青色については中心視で若干ではあるが他の色での結果に比べて学習効果が下がる傾向が認められた。また、形状については、すべての図形で中心視が最も良い学習効果を持ち、以下、近中心視、周中心視の順であった。

## 5. おわりに

本研究では、これまでのカラー情報処理機器と違う中

心視、近中心視、周中心視に相当する色画像について検討を進め、光受容細胞の分布の統計的解析を進めるとともにこの結果を画像処理機器に導入することで視覚探索が可能となり得るロボットカラービジョンシステムを提案した。この提案に基づいて、中心窩に相当する入力画像の中心付近が周辺より高解像度で、かつ、周辺領域が粗い画像となることで注視物のより正確な視覚情報処理を高速で処理し、対象物を認識可能なシステムが完成できる確証を得た。

また、最近、人間のL/M錐体比などに対して心理物理学的にも詳細に解明されつつあり、今後、これらの研究結果を利用したヒューマンカラービジョンの構築も図ってゆく予定である。

本研究の概念はこれまで市販のカラー入力機器にはなかった概念であり、それらの実験結果は屋外など広い視野内で動き回ることのできる自律ロボットなどのロボットカラービジョンシステムの設計の際に多くのインパクトを与えるのではと考えられる。

本研究を進めるに当たり貴重なご意見を賜りました松下電器産業(株)照明研究所大竹史郎氏に感謝いたします。

## 文 献

- 川上元郎, 児玉 晃, 富塚 直, 大田 登編: 色彩の事典(朝倉書店, 1987) pp. 157-158, p. 163.
- A. Fiorentini: "Differences between fovea and parafovea in visual search processes," *Vision Res.*, **29** (1989) 1153-1164.
- 熊田孝恒: "中心視と近中心視における視覚探索特性の検討", 日本心理学会第57大会(1993) p. 790.
- R.E. Marc and H.G. Sperling: "Chromatic organization of primate cones," *Science*, **196** (1977) 454-456.
- F.M. de Monasterio, E.P. McCrane, J.K. Newlander and S.J. Schein: "Density profile of blue-sensitive cones along the horizontal meridian of macaque retina," *Invest. Ophthalmol. Visual Sci.*, **26** (1985) 289-302.
- 大竹史郎, C.M. Cicerone: "L/M錐体比の評価に関する最近の心理物理学的研究", *VISION*, 日本視覚学会, **5** (1993) 125-128.
- D.R. Williams: "Topography of the foveal cone mosaic in the living human eye," *Vision Res.*, **28** (1988) 433-454.
- 山羽和夫, 三宅洋一: "ロボットビジョンのための新しいカラービジョンシステムの設計", *光学*, **24** (1995) 110-116.
- K. Yamaba and Y. Miyake: "Color character recognition method based on human perception," *Opt. Eng.*, **32** (1993) 33-40.
- K. Yamaba, H. Tominaga and Y. Miyake: "Learning effect in color character recognition," *IMEKO Proc. 1st Int. Symp. Measurement and Control in Robotics, ISMCR '90* (1990) pp. G.1.3.1-6.