



## ロボット視覚のための新しいカラービジョン システムの設計

山羽 和夫\*・三宅 洋一\*\*

\* 工業技術院機械技術研究所 〒305 つくば市並木 1-2

\*\* 千葉大学工学部 〒263 千葉市稲毛区弥生町 1-33

(1994年6月8日受付, 1994年11月14日受理)

### Design of a Newly Developed Color Vision System for Robot Vision

Kazuo YAMABA\* and Yoichi MIYAKE\*\*

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

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

(Received June 8, 1994; Accepted November 14, 1994)

Color is not only a physical quantity, but it is also a psychological phenomenon based on Purkinje effect and color constancy. This suggests that a human-like color vision system must process data using subjective parameters. Although this problem is of great importance, no one has proposed a solution using a human-like color instrumentation similar to a retinal processing unit until very recently. This paper describes a newly developed color vision model for robotic systems. This new model consists of a normal lens, a neutral density filter, an infrared filter, two different prisms, four different bands of optical pass filters, four monochromatic CCD cameras and a camera controller. This vision system is very effective by using an early processing technology to realize a higher level human-like color visual perception. A simple experiment was done to demonstrate human-like color vision by employing an opponent theory.

#### 1. はじめに

カラービジョンすなわち人間の色覚に関する二つの主な理論計算は Young-Helmholtz と Hering に帰する<sup>1)</sup>といえよう。このうちの一つは Helmholtz が Young の色覚理論を体系化したもの (Young-Helmholtz 説) であり、「人間の眼の色覚は赤と緑と青の3種の感覚の組合せでできている」というすべての色の基本を3種類の色と定めた基本仮定 (三色説; trichromatic theory) に基づいている。Young-Helmholtz の三色説が混色の原理に基づいているのに対し、Hering の提案は現象的な色の「見え」にその基礎を置いている。これは純粋な色としてあげられる四つの色 (主要4色相; primary color) に混色あるいは同時色対比において互いに反対の関係にある赤-緑、黄-青の組合せと、さらに、この二つ

の組合せに白-黒の感覚まで含めている。これが Hering の反対色説と呼ばれているもので、Young-Helmholtz の三色説とは黄を赤、緑、青とともに原色として扱っている点が異なっている。この反対色説では黄は赤と緑の混色によって生じるが、できあがった黄の感覚内容は赤からも緑からも性質が異なっている。

近年になって、上記二つの理論について、まず最初に、Young-Helmholtz の理論においてその生理的基礎が発見され、また、Hering の反対色説のほうは DeValois により、さるを使用した生理的事実が見つけだされている。Hering の反対色理論以来、Guth らによる Vector モデル<sup>2)</sup>とか Hurvich と Jameson の Cancellation 法<sup>3)</sup>のような反対色理論を補充するような多くの研究も紹介されてきている。

一方、これまでのメカトロニクスの代表格である産業

用ロボット等の視覚（色）システムにおいては赤，緑，青の3色の情報をそのまま処理していることから，通常の白黒のイメージセンサーと比べて約3倍の処理時間がかかるためこれまでは静止色画像を対象にその処理が行われてきた．ところが最近では，コンピュータをベースとした各種の並列画像処理システムが提案されてきている<sup>4)</sup>．このシステムの特徴は，明るさ，色，運動などの画像の特徴を独立した特徴抽出モジュールから空間的に並列に抽出した後，画像の認識のために簡単な重み付けを用いて特徴を実時間で統合するという点にある．しかしながら，ロボットビジョンとしての実用化を考えた場合システムが大がかりになることが欠点であった．

以上のことを背景にして，本研究では，ロボットの視覚システムなどでこれまで色情報処理といえば赤，緑，青の3色の色情報で通常の明暗の処理と比べて3倍の処理時間がかかり，それゆえ，それらの研究開発が半導体素子とかシステムそのものの処理速度に依存していたのを，新しく人間の色情報処理を模倣した処理方法を導入することで情報処理速度を短縮するカラービジョンシステムについて提案し，ニューラルネットワークを用いたパターン認識実験を介してそれらの有用性について確認した．

## 2. 人間の視覚構造のモデルの提案

本節では，人間の視覚構造をモデルにしたロボットビジョンのためのカラービジョンシステムの設計概念について述べる．

人間の視覚処理プロセスを模倣した視覚並列処理機構の総合モデルを Fig. 1 に示す．本モデルは初期視覚モデルとして視覚ニューラルモデル部 (Visual neural model) と大脳の判断機能を想定しているファジィ推論部 (Fuzzy reasoning) の二つのブロックから構成されている．本論文では，研究対象をこの二つのうち，ロボットビジョンということで初期視覚モデルに絞っている．Fig. 1 から，例えば，“A”，“B”などの文字情報，色情報，形状などの光情報入力が入力セルにある場合，人間の網膜構造と同じ前処理 (図中 R-G は赤信号-緑信号を示し，また，Y-BI は赤信号と緑信号で合成された黄色信号から青信号を減じた信号を示す．Bk/W は黒/白の明るさ信号を示す．本論文では，以降これらの表現を使用する) が行われた後，視覚ニューラルネットモデル内のニューラルネットワーク部でデータ変換が行われ，自然環境下での入力情報の区別 (図例では単一文字の種類，色の種類，形の選別) が行える．このデータ変換さ

れた情報をファジィ推論し，色，形，大きさなどの恒常性，組合せ，オクルージョンなどの画像の評価を行うことによりそれぞれの結果に対応した高次処理レベルの信号として取り出すことができる．

以上，述べてきた視覚並列処理機構の総合モデルを新しい視覚機能素子として実現していくためには視覚機能素子以後に接続されるシステムとの最適な接続ニズム (入力情報の変換，伝送，表現，記憶等の並列処理をいう) を決定しなければならない．そのための要素技術として，まず，視覚ニューラルモデル部すなわち初期視覚モデルのハードウェア化を含めた検討が最も必要になってくることが理解できよう．

## 3. 初期視覚モデルの試作

以上の視覚並列処理機構のモデルの提案に基づいてその初期レベルである初期視覚モデルを作成した．Fig. 2 に試作した初期視覚モデルの入力部の外観写真を示す．

Fig. 3 にロボットカラービジョンシステムのブロックダイヤグラムを示す．

試作した視覚初期モデルの入力部は主にレンズ系 (Lens, ND, IR)，2種類のプリズム (Prism 1 および Prism 2)，光学フィルタ (Sc 1~Sc 4)，4台の CCD カメラ (C CCD camera, R CCD camera, G CCD camera, B CCD camera) およびカメラ制御機器で構成されている．このなかで網膜は二次元視覚センサー (2種類のプリズム，光学フィルタ，4台のカメラ) と考えることができる．レンズ系は両凸レンズ (Lens)，ND フィルタ (ND)，赤外線カットフィルタ (IR) で構成してい

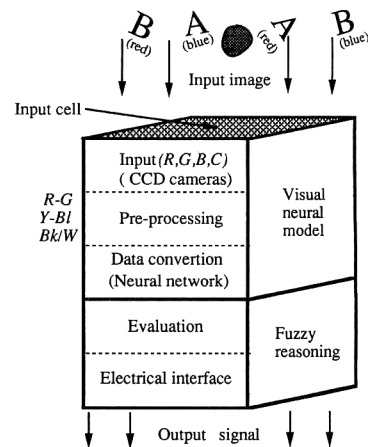


Fig. 1 Visual parallel processing model using a human-like performance similar to a retinal processing unit.

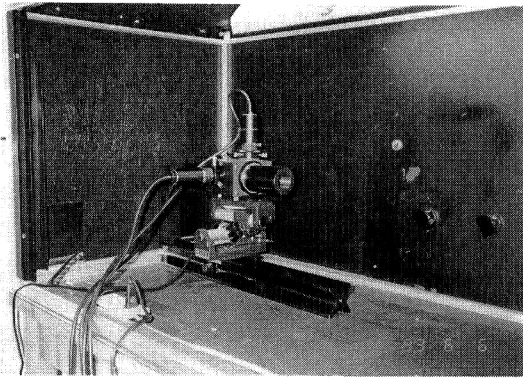


Fig. 2 A photograph of color input model based on human vision.

る。また、2種類のプリズムには光量を半分にするビームスピリット (Prism 1), 光学的に透明で屈折率の異なる誘電体を交互に多層真空蒸着したダイクロイック膜 (膜の厚さ, 屈折率を選択することである特定の波長領域だけ反射し, 残りの波長領域は透過する特性を持たせられる) で構成された色分解光学系プリズム (テレビカメラ用分解光学系プリズム<sup>5,6)</sup>: Prism 2) を使用した。この二つのプリズムの出口に, それぞれ光学フィルタ (Sc1~Sc4; ビームスピリットからの可視光領域出力と分解光学系プリズムの赤色出力に色補正フィルタを用いた, また, 緑色出力と青色出力にそれぞれ光学フィルタを用いた。色補正フィルタは厚さを 1mm にまで薄く加工仕上げた) と 4 台の単色固体撮像素子 (CCD カメラ; 画素数: 576 (水平) × 485 (垂直), 2/3 in × 2/3 in) を色分解プリズム (Prism 2) からそれぞれの素子面まで

の距離が同一になるように取り付けられた。試作した初期視覚モデルの画角は実測の結果, 最大 2.7° であった。これは C. I. E. が規程する 2° 視野を十分カバーできる値である。

この初期処理モデルの入力部は人間の目とは構造的に異なるが, 機能的には 1 台のカメラ出力は桿体細胞, 3 台のカメラ出力は錐体細胞に類似させている。すなわち Fig. 3 に示すように CCD カメラ制御部 (4 台の CCD カメラの画像入力, 制御を行う) の後のマイクロコンピュータ部では視覚並列処理モデルに示した前処理部の演算を行う。さらにその演算出力をニューラルネットワークに入力する構成を採用している。このニューラルネットワークには市販の階層構造可能なニューラルボードを取り付けた。

各カメラから出力される分光感度特性のうち, それぞれのピーク波長と半値幅を計測した結果を Fig. 4 に示す。Fig. 4 の各特性はレンズ, 色分解プリズム, フィルタ, CCD カメラまで含めた本システム全体の分光感度特性を示した。ここでは各カメラの分光感度を  $x(\lambda)$ ,  $y(\lambda)$ ,  $z(\lambda)$ ,  $c(\lambda)$  とし, それに対応する特性として, Red, Green, Blue, Brightness と表記した。Table 1 および Fig. 4 にそれらの結果を示す。Table 1 および Fig. 4 から同モデルのピーク波長が Dartnall らの計測した人間の光受容器の分光吸収特性<sup>7)</sup> に近いこともわかる。

本分光感度特性の設計にあたっては人間の桿体細胞および各錐体細胞と同じピーク波長, 似た半値幅を持ち, また, その用途がロボットの視覚であるとの前提条件

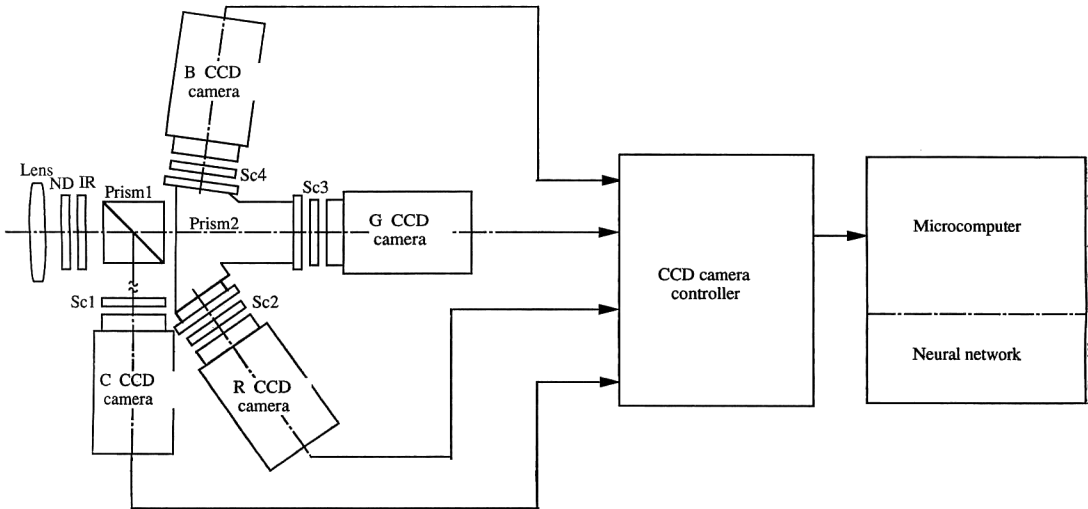


Fig. 3 A block diagram of a newly developed color vision system for robot vision.

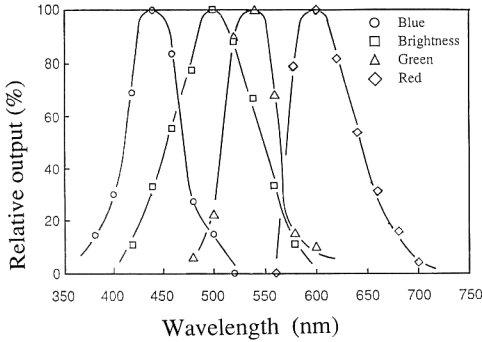


Fig. 4 The spectral responses of the newly developed color input model.

Table 1 Peak wavelength and half power spectrum of this system.

	Peak wavelength	Half power spectrum width
$x(\lambda)$	600 nm	70 nm
$y(\lambda)$	540 nm	55 nm
$z(\lambda)$	440 nm	58 nm
$c(\lambda)$	500 nm	92 nm

で、でき得るかぎり小形軽量にすることを念頭においた。これらの条件のうち、ロボットカラービジョンを実現するうえで特に重要なものはピーク波長を一致させることと小形軽量化であろう。実際に得られた分光感度特性は Fig. 4 で示すように人間の光受容器の分光吸収特性より重なりが少なかった。この重なりが少ないと単色光を色判定する場合に問題が生じる。いわゆる、色純度の高い青緑の 490 nm 付近と黄色の 570 nm 付近では感度が落ちることが予想される。ただロボットビジョンでは扱う対象が物体色であることで単色光でないため、色判定の際にはそれほど問題にはならない。もし、扱う対象が単色光に限られる場合は明るさ応答モデルからのフィードバック情報を使用してその部分の感度を上昇されるようにするなど別の手段も必要となろう。

Fig. 5 に本論文での初期視覚モデルの概念図を示す。この図では、本研究におけるカラー情報処理手法の第 1 ステップとして視覚心理からの反対色理論と輝度応答モデルを使用するための検討を進めた。明るさ応答モデルは Kaiser, 矢口, 中野らのモデルなど<sup>8-11)</sup>が報告されているが、ここでは薄暮時に対応できるシステムを考えてみた。この明るさ信号はニューラルネットワークにおける各セル間の結合の強さを制御する重み信号として機能できるものと思われる。

初期視覚モデルの R, G, B の各 CCD カメラからは人間の場合と同様に周囲が明るいときには物体の表面の状態 (surface properties) いわゆる色の情報と明るさ情報  $L$  が入力される設計にしている。また、桿体細胞に相当する C CCD camera からはエッジとか形状、運動などの情報  $L'$  を入力するよう構成させた。図中、 $Bk/W$  の入力部には明るさ知覚として次のような信号が入る。

$$L = r \log R + g \log G + b \log B \quad (1)$$

式 (1) は中野らが提案した錐体出力のレベルで対数的非線形性を導入して得られた明るさ知覚一定という式<sup>11)</sup>を変形したもので  $R, G, B$  は各錐体に相当するカメラの出力で刺激値に相当するよう設定し、また、 $r, g, b$  はある任意の明るさのもとの各カメラの出力を相対値に変換するために用いた係数であり、実験で決まる値である。実際には、これらの値として  $r=1.20, g=0.980, b=0.860$  を採用した。

Fig. 5 の桿体細胞に相当するカメラ (C CCD camera) の下に位置する加え合わせ要素では、周囲の環境状態が暗くなってきたときいわゆる明所視 (photopic vision) から薄明視 (mesopic vision), さらには暗所視 (scotopic vision) の状態になったとき錐体から桿体へ主たる受容器が移行するために引き起こされる比視感度  $V(\lambda)$  の変化に対応する動きをさせている。したがって、ここでの出力  $Bm$  は

$$Bm = a \log L' + (1-a) \log L \quad (2)$$

で表される。ここで  $a$  は調節係数で桿体細胞が輝度レベル (luminance level) に及ぼす量に相当し、 $0 > a > 1$  の間で照度から実験的に決まる値である。本論文では  $a$  について人間の視覚を計測した値<sup>12)</sup>を参考に実験的に設定

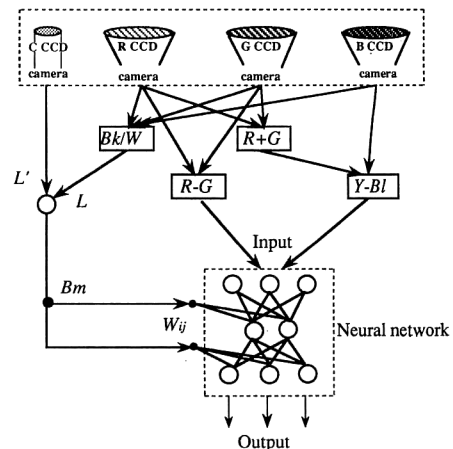


Fig. 5 The new concept of visual neural model robot vision.

した。式(2)は明るさに応じた2種類の細胞からの信号の入力量によって操作される状況を表しており、プルキンエ効果 (Purkinje effect) のような感覚量に相当していることが理解できる。

そこで式(2)で求められた値を階層形のニューラルネットワークの学習時の重み信号 ( $W_{ij}$ ) として利用することで初期視覚モデルが完成できる。

#### 4. 実験および評価

**Fig. 6** に R, G, B カメラからの原画像とその被写体の大まかな説明図を示す。本実験では被写体が 2.7° の角角に入るようカメラと被写体との距離を設定した。

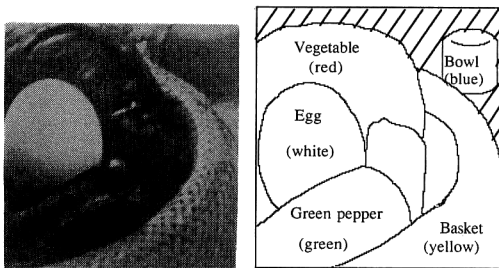
**Fig. 7** に C CCD camera からの画像出力を示す。また、**Fig. 8** に反対色信号のニューロボードへの取り込み画像を示す。実験は暗箱中で行い (Fig. 2 参照), 光源にはインバータの蛍光灯 (FL 20 W) を用いた。被写体の表面での照度は計測の結果, 900 lx であった。Fig. 8 より, 明るさ知覚, R-G, Y-BI 信号のなかでは明るさ知覚の信号よりも反対色信号のほうが, 特に, 色物の形態を判定するのに有利であることがわかる。また, Fig. 6 の被写体の説明図の中心部にある白色物体の画像入力情

報を消去または減少できていることから R-G, Y-BI 信号とも色検出のみ関与した信号が得られていることがわかる。このことは色情報のみを目的に画像処理をするときには二つの信号処理のみで良いことを示しており, また, 本反対色説の方式が従来の三原色法より処理時間が短縮できる点明るさ情報に無関係の色情報のみが取り扱える点などで優れた手法であるといえよう。

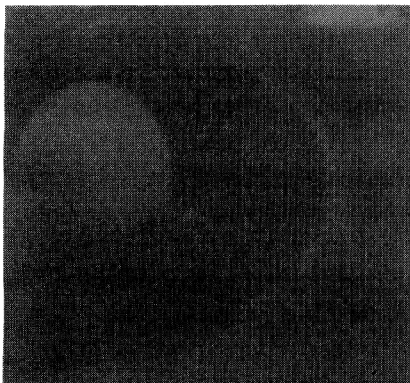
本実験においてさらに光源に覆いを使用し, 被写体の照度を 62.7% (564 lx), 31.2% (281 lx) に落として画像を採取した。

ニューラルネットワークのカテゴリーとしてここでは野菜 (赤色部分), 茶碗 (青色部分), ピーマン (緑色の部分), 箸 (黄色の部分), 卵 (白色部分) を設定し (Fig. 6 参照), 照度 900 lx の条件下で学習を行った。また, 比較のために従来の赤, 緑, 青の三色方式による学習も行った。それらの結果, および照度を落とした場合の2段階の反対色による結果を **Fig. 9** および **Fig. 10** に示す。

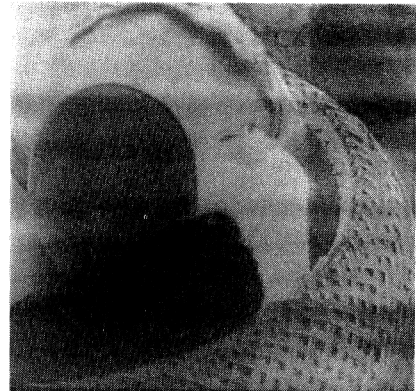
**Fig. 8, Fig. 9** から提案した反対色方式では照度が約 2/3 に低下しても画像入力データはほとんど変化してい



**Fig. 6** Objects and their brief explanatory drawing.



**Fig. 7** Output image from C CCD camera.



(a)



(b)

**Fig. 8** Input image under the fluorescent lamp (900 lx). (a) R-G image, (b) Y-BI image.

ないことがわかる。したがって、このままのデータでも、明るさによる影響は少なく、色、形状の判定を行うような場合に支障をきたさないことが予測される。ところが Fig. 10 のように明るさが約 1/3 になると画像入力データはかなり変化する。

このような画像入力データの変化に対処するため、ここでは式(2)の  $a$  の値にこの照度の大きさ、すなわち覆いによって減じられた照度の比を代入した。この操作は人間の場合の明度が低くなると桿体細胞が受容器となるような視覚機能に類似しており、この結果をニューラルネットワークの重み係数 ( $W_{ij}$ ) とすることで、この重み係数により、照度が低下することによる被写体の状況を補充するものと考えることができる。

本研究ではデータフロープロセッサ 4 個と 256 k (1ワード=8ビット) のイメージメモリを搭載したボードをニューラルネットワークモデルのバックプロパゲーション演算用ボードとして用い、このソフトウェア処理を C 言語によって行った。このニューラルネットワークモデルはフィードフォワード型階層構造 (入力層, 中間層, 出力層) を有しており、それらの初期値としてニューロ

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。この 3 層のバックプロパゲーションによるニューラルネットワークシステムでは各層のニューロンは隣の層のすべてのニューロンと結合し、この結合の強さすなわち荷重, しきい値を変化させることが学習につながる。式(2)の演算結果はこの荷重情報として利用しており、実際には式(2)を用いてシグモイド関数の傾きを変えようとした。

それぞれの明るさで 50 回の学習の後、認識実験を行ったが、これらのカテゴリーの範囲内では、誤認識はゼロであった。この学習回数の 50 回は一般のニューラルネットワークで行われる学習に比べ、極めて少ない学習回数である。これはロボットビジョンではできる限り少ない教示回数で認識を行うことが理想であることから経験的<sup>13)</sup>に選定した回数である。この回数が増加すれば、誤認識数は減少するが、反面、汎化性が低下し、また、教示時間も長くなり、これはロボットビジョンを設計する立場からいって好ましくないことである。

さらに、上記と同じ条件で重み情報をなくして学習し

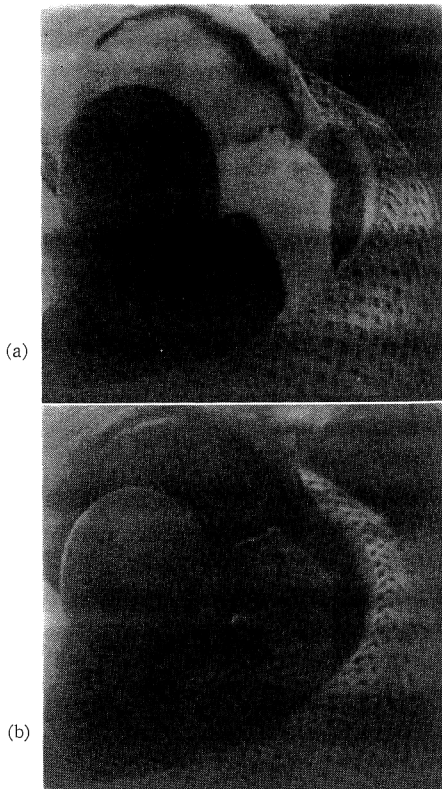


Fig. 9 Input image under the fluorescent lamp (564 lx). (a) R-G image, (b) Y-BI image.

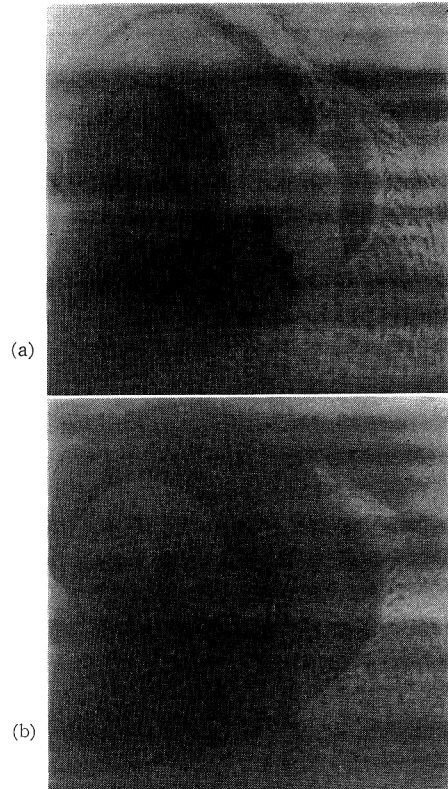


Fig. 10 Input image under the fluorescent lamp (281 lx). (a) R-G image, (b) Y-BI image.

た場合には野菜(赤色部分)が認識不可能な結果が得られた。これは人間の場合の暗くなってきた場合に生ずるプルキンエ効果に相当しているのではと考えられよう。このような結果からさらに約 1/10 に照度を落としてその効果を確認したところピーマン(緑色部分)の判定が不可能になった。しかしながら、ニューラルネットワークに重み情報を付加することで 1/10 くらいの照度の減衰では認識には支障がない結果が得られた。なお、本実験システムの有用性を確認するために市販のカラー CCD カメラを使用して反対色ではなく三原色の画像情報を本ニューラルネットワークに入力して同様の実験を行ったが、照度を約 10% 減少させた段階ですべての認識が不可能になった。このとき使用した CCD カメラは、汎用のカラーカメラで、そのピーク波長がそれぞれ 620 nm, 530 nm, 420 nm であり、かつ、三つの波形間の重なりはほとんどない。この実験から、ピーク波長も各波形も異なるときは明るさになどの影響を大きく受けることがわかる。逆に、これらの分光特性を人間の分光感度曲線に近づけることにより、特に色彩識別や明るさ特性において人間の感覚にさらに近づけることができることも予想されよう。

以上の実験を通じて本ロボットカラービジョンの概念の妥当性が実証できたと思われる。

## 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って、カラー情報の他に明暗情報を取り入れた画像に対する検討を進めた。この概念はこれまでの市販のカラー入力機器にはなかった概念であり、今後のカラービジョンシステムに大きなインパクトを与えるのではと思われる。さらに、この概念と反対色説を導入したニューラルネットワークによる実験を通じて明るさ環境によらない人間に類似した高機能なロボットビジョンシステムの実験の可能性が確信できた。

本研究を進めるにあたり装置の製造に協力いただいた

九州松下電器(株)生産技術研究所戸倉暢史氏、塚本満早氏および富士写真光機(株)村上隆氏にお礼申し上げます。また、研究の過程で、貴重な意見を賜りました千葉大学工学部画像工学科矢口博久先生に謝意を表します。

## 文 献

- 1) E. Hering: *Grenzüge der Lehre von Lichtsinn* (Springer-Verlag, Berlin, 1920).
- 2) S.L. Guth, R.W. Massof and T. Benzschawel: "Vector model for normal and dichromatic color vision," *J. Opt. Soc. Am.*, **70** (1980) 197-212.
- 3) L.M. Hurvich and D. Jameson: "An opponent-process theory of color vision," *Psychol. Rev.*, **64** (1957) 384-402.
- 4) 長谷川修, 横澤一彦, 石塚 満: "生体の視覚系をモデルとした画像認識システムの試作", 第7回生体・生理工学シンポジウム論文集 (1992) pp. 59-64.
- 5) 樋口 武: "第6章光学系の設計, ホワイトシリーズ 88 固体撮像素子とカメラへの応用技術", トリケップス (1988) pp. 121-136.
- 6) 伊藤清次: "テレビカメラ色分解光学系のプリズム (その2), (その3), (その4)", 放送技術 **4**, **5**, **6** (1982) 339-342, 435-438, 531-534.
- 7) 川上元郎, 児玉 晃, 富塚 直, 大田 登編: 色彩の事典 (朝倉書店, 1987) p. 163.
- 8) S.L. Guth and H.R. Lodge: "Heterochromatic additivity, foveal spectral sensitivity, and a new color model," *J. Opt. Soc. Am.*, **63** (1973) 450-462.
- 9) 矢口博久, 池田光男: "明るさ感覚における反対色チャンネルの非線形介入", 光学, **9** (1980) 44-51.
- 10) Y. Nakano, M. Ikeda and P.K. Kaiser: "Contributions of the opponent mechanisms to brightness and nonlinear models," *Vision Res.*, **28** (1988) 799-810.
- 11) K. Sagawa and K. Takeichi: "Data on luminous efficiency for a 10-degree field in photopic, mesopic and scotopic vision obtained by heterochromatic brightness matching," *Bull. Indust. Prod. Res. Inst.*, **122** (1992) 1-39.
- 12) K. Sagawa and K. Takeichi: "System of mesopic photometry for evaluating lights in terms of comparative brightness relationships," *J. Opt. Soc. Am.*, **A/9-8** (1992) 1240-1246.
- 13) K. Yamaba and Y. Miyake: "Color character recognition method based on human perception," *Opt. Eng.*, **32-1** (1993) 33-40.