

# Neural Networkによるカテゴリカル色知覚モデルを用いたシーン中の色認識

矢田 紀子<sup>†1</sup> 長尾 智晴<sup>†1</sup> 内川 恵 二<sup>†2</sup>

本研究の目的は、人間が持つカテゴリカル色知覚と色の恒常性の機能を備えたモデルを獲得し、これを用いて未知の照明光下で撮影された画像中の物体色の色認識を行うことである。さまざまな照明条件下で撮影された画像中の色認識はコンピュータビジョンやマシンビジョンに有効であり非常に重要であるが、色恒常性を解決する手法はいまだに確立されておらず、特に、人間の色覚特性を調べた心理物理実験の結果を利用した手法は提案されていない。そこで我々は、さまざまな照明条件下で行われたカテゴリカルカラーネーミング実験の結果をニューラルネットワークで学習することで色恒常性も考慮したカテゴリカル色知覚モデルを獲得し、このモデルを画像中の各画素の色認識に適用してモデルが人間と同じように色認識をできていることを示す。

## Categorical Color Perception for Image Segmentation Using Neural Network

NORIKO YATA,<sup>†1</sup> TOMOHARU NAGAO<sup>†1</sup> and KEIJI UCHIKAWA<sup>†2</sup>

The purpose of this study is to get a model that can operate similarly to human categorical color perception and color constancy. It relates to a categorical color perception system for automatically judging a categorical color and is a technology for correctly judging under various environment. The color recognition is difficult in particular when color constancy is related. We decide to test a surprisingly simple hypothesis. Namely, that the relationship between the chromaticity of color chips under different illuminations and human categorical color perception for the color chips under the illumination can be learned by a structured neural network. The categorical color perception is the product of a categorical color naming experiment. In addition, we apply this categorical color perception model to color recognition. We classify all pixels in image in basic categorical colors.

### 1. はじめに

我々人間は、色の微妙な違いを見分けることができる一方で、その色を赤・青などのように大雑把にいくつかにまとめて表現することがある。この色の大雑把な分類を、カテゴリカル色知覚という<sup>1)</sup>。Berlin らは 100 種類以上の言語を調べて、白・赤・緑・黄・青・茶・オレンジ・紫・ピンク・灰・黒の 11 色が、言語や人によらず等しく用いられる基本カテゴリ色であることを示した<sup>2)</sup>。これら 11 色の色名が基本カテゴリ色であることは、Uchikawa らによる心理物理実験<sup>3)</sup>

や松沢によるチンパンジーの行動実験<sup>4)</sup>でも示されている。また、Franklin らは言語を習得する以前の幼児もカテゴリ色の知覚を行い、このカテゴリが基本カテゴリ色と一致することを示した<sup>5)</sup>。これらの事実から、基本カテゴリ色の認識は視覚システムの持つメカニズムによって決まっており、そのほかの色の認識の場合とは明らかに区別できるといえる。

また、人間は環境光のスペクトルが変化しても物体の色を安定して知覚することができる。これは一般的には恒常性 (constancy) として知られており、色知覚上では色恒常性と呼ばれている<sup>6)</sup>。我々が色を大雑把に表現するときを感じる赤色や青色といった感覚は、頭の中でどのように生じているのであろうか。視覚系は、目の内側にある網膜で光のエネルギーが電気信号に変換されることから始まる。その信号が動きと色や形の情報に分類され、動きの処理を行う部位や色や形の処理を行う部位へと伝達され、最終的に高次の知覚

<sup>†1</sup> 横浜国立大学大学院環境情報学府  
Graduate School of Environment and Information Sciences, Yokohama National University

<sup>†2</sup> 東京工業大学大学院総合理工学研究科  
Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology

へと発展してゆく。しかし、これらが全体でどのような処理を行うことで、色恒常性やカテゴリカル色知覚のような高度な処理を行っているかはいまだに不明である。

コンピュータビジョンにおいて色の認識を行いたい場合には、一般にカメラのセンサ値や画像の階調値を用いて色の境界を設定して特定の色を判断するが、すべての色を人間と同じように色名に分類して認識することは非常に困難である。

一方、ニューラルネットワーク (Neural Network; NN) を用いて、視覚系が処理している課題を学習させ、その中間層に生体の神経細胞と同様な応答をする中間ユニットを獲得している先行研究が多くある。Zipser らは、網膜上に投影された刺激の網膜上の位置とその刺激を見たときの眼球の向きから、提示されている物体の位置を計算する課題をニューラルネットワークに誤差逆伝播法 (Back Propagation; BP 法) を用いて学習させた<sup>7)</sup>。その結果、ネットワークモデルはマカクザルの V7a 野における、同様の処理をしていると考えられている細胞と似た応答をする中間ユニットを獲得した。また Usui らは、ある単色光に対する L, M, S 錐体の応答から、V4 野に確認されている色に対する選択性を持つ細胞への写像を、Zipser らと同様に BP 法を用いてニューラルネットワークに学習させた<sup>8)</sup>。その結果、ネットワークの中間層には、マカクザルの外側膝状体内に見られる反対色応答を示す細胞と似た応答を示すユニットが現れた。

筆者らの研究グループでは、ニューラルネットワークを用いたカテゴリカル色知覚のモデルを提案している<sup>9),10)</sup>。このモデルでは入力値として環境の照明光が与えられたときに、対象の色を正確に基本カテゴリ色で認識することができた。しかし、このモデルは環境の照明光が未知の場合には適用することができず、撮影条件が未知な画像中の色認識への適用はできていなかった。また、白色に近い照明条件で撮影された画像中の色認識への適用実験では、特に照明条件の影響を受けやすい白色の物体の認識ができないという問題点があった。

そこで本論文では、画像中の白色点を照明光の手がかりとして用いることで照明条件が未知の画像中の色認識へ適用可能な新しいモデルを構築する。そして、このモデルの有効性を検証するために、10種類の照明光を用いたシミュレーション実験と、未知の照明光下で撮影された画像中の物体色認識実験を行い、モデルが人間と同じように色認識をできていることを示す。

## 2. ニューラルネットワークモデル

本論文では、これまでに提案されている特殊な構造のニューラルネットワーク<sup>9),10)</sup>を用いて、人間に対する視覚心理物理実験の結果を学習することで、人間の色覚と同様に色認識を行うことができるモデルを構築する。

### 2.1 ニューラルネットワークの構造

このモデルでは、多数回の試行実験による実験結果を基に、最終的に入力層 1 層、中間層 2 層、出力層 1 層からなる 4 層のフィードフォワード型ニューラルネットワークを採用し、中間層と出力層の入出力関数にはシグモイド関数を用いる。このモデルの構造を図 1 に示す。入力層 (Input layer) のユニットは、3種類の錐体 (L, M, S) に相当するユニットをサンプル色 (Test color), 照明光 (Illuminant) それぞれに対して 1 セットずつ、合計 6 ユニットの用いた。ここで、照明光の変化に頑健なモデルを獲得するために、ニューラルネットワークを 4 層にして 2 層ある中間層の 1 層目では照明光の補正は行わず、2 層目で照明光の補正を行うようなニューラルネットワークの構造を採用する。すなわち、入力層と中間層第 1 層は、ともにサンプル色に対する部分と照明光色に対する部分に分かれており、それぞれの部分内では、全結合であるが部分間では互いに結合のない構造である。

中間層 (Hidden layers) のユニット数は、試行実験を基にして最も適当なユニット数とする。中間層第 1 層は  $4 \times 2 = 8$  ユニットの、中間層第 2 層は 8 ユニットのからなる。また出力層 (Output layer) は、11 ユニットの

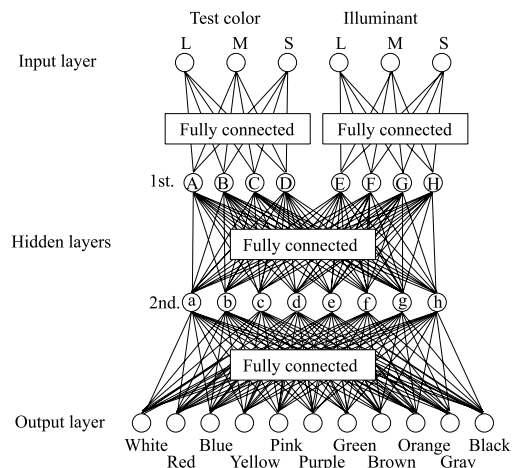


図 1 モデルの構築に用いたニューラルネットワークの構造  
Fig. 1 Structure of neural network for categorical color perception model.

トからなり、各ユニットが11色の基本カテゴリ色にそれぞれ対応している。

## 2.2 ニューラルネットワークの学習方法

ニューラルネットワークの学習には、BP法を改良した手法を用いる。この方法では、入力層-中間層第1層間はサンプル色に対する部分と照明光に対する部分で、それぞれまったく同じネットワークが形成される。たとえば、図1の入力層サンプル色部分のユニットLと中間層第1層のユニットA間の結合荷重を修正する際に、入力層照明光色部分のユニットLと中間層第1層のユニットE間の結合荷重にも同じ修正が加わる。また逆に、照明光色部分のある結合荷重を修正する際には、サンプル色部分の構造上対応する結合の結合荷重にも同じ修正が加わる。これは、互いに構造上対応する結合どうして共通の結合荷重を持ち、サンプル色成分に対する処理と照明光成分に対する処理で共用することに等しい。

このような構造と学習方法を採用することで、モデルの照明光色部分の学習にもサンプル色に対する学習結果が活かされるという利点があり、学習に用いる照明光の色数が少なくても、学習過程においてさまざまな信号に対する学習が行われることになる。この方法を用いることで一般的な照明光下での色恒常性を備えたモデルが獲得できることが示されている<sup>9),10)</sup>。

## 2.3 教師データセット

モデルの学習に用いた教師データセットについて示す。教師データセットは、3種類の異なる照明光を用いて行われたカテゴリカルカラーネーミング実験を基に作成する。

### 2.3.1 カテゴリカルカラーネーミング実験

人間のカテゴリカル知覚特性を調べるために行ったカテゴリカルカラーネーミング実験について述べる。カテゴリカルカラーネーミングとは、被験者が呈示された色票の見えを、最もよく表す色名を11色の基本カテゴリ色の中から1個答える方法である。この実験では、呈示刺激としてOSA色票424枚を用いる。刺激は暗室内で天井からLCDプロジェクタによって照明され、マンセルN5の明るさの灰色のボード上に1枚ずつ呈示される。照明光が変化したときのカテゴリカル知覚の変化を調べるために、照明光には3種類の異なる照明光(3,000 K, 6,500 K, 25,000 K)を用いる。これらを照明光o, w, bと呼ぶ。この3種類の照明光の相関色温度およびCIE(1931)<sub>xy</sub>色度を表1に示す。また、これらの分光分布を図2に示す。

被験者は、それぞれの照明光に順応した状態で提示される色票を見て、その色票の見えを最もよく表す色

表1 実験に用いた照明光の相関色温度と色度

Table 1 Color temperature and CIE(1931)<sub>xy</sub>-chromaticity of illuminants used in the experiments.

照明光	相関色温度	xy色度(x, y)
o	3,000 K	(0.439, 0.410)
w	6,500 K	(0.313, 0.332)
b	25,000 K	(0.255, 0.252)

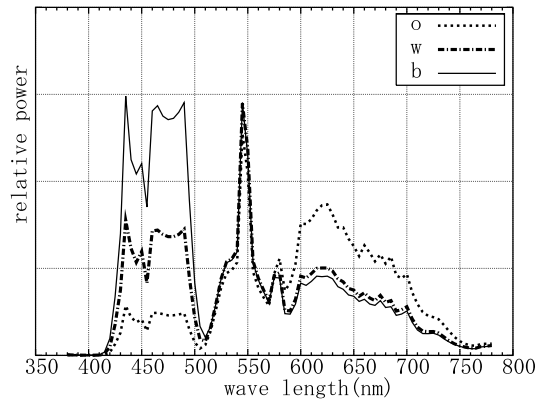


図2 実験に用いた照明光の分光分布

Fig. 2 Spectral distributions of illuminant used in the experiments.

名を11個の基本色の中から1個答える。被験者は4名で、それぞれ色票424枚に対してのネーミングを1セッションとして、同一照明光で2回のセッションを行い、3照明光×2回=6セッションを行う。ここで、内川らによる研究<sup>11)</sup>から、色空間を基本カテゴリ11色で領域分割する場合、いつ誰が見ても安定して分割される領域分割の安定性があることが分かっている。したがって、4名×2セッションに対する実験で、一般的なモデルを獲得するためのサンプルとして十分であると考えられる。

また、このカテゴリカルカラーネーミング実験では、色票全体の色度点の分布が照明光によって大きく変化しており、色味の強いoやbの照明光を用いた場合には色度が一定の領域に偏っている。そのため、色票の色度点から判断すると限られた色名しか得られないと考えられるが、被験者の応答には複数の色名が用いられていた<sup>9)</sup>。これは色名が色票の色度だけによって決まっているのではなく、照明光成分を差し引いて色票の分光特性の違いによって決まっていることを示している。つまり、色恒常性が成立しているといえる。

さらに、この実験結果から各照明光・色票に対しての回答について、4名×2セッション=8回中での色名の一致率を調べたところ、一致率100%が48.3%、一致率50%以上100%未満が50.1%であり一致率50%以

下は 1.7%と少なかった．このことから，今回の実験でも過去の研究と同様に領域分割の安定性が保たれていたといえる．

### 2.3.2 教師データの算出

ニューラルネットワークの学習に用いた教師データセットは 2.3.1 項のカテゴリカルカラーネーミング実験を基にして作成する．

まず，それぞれの照明条件下で測定した OSA 色票の色をサンプル色とする．モデルのサンプル色部分の入力データは，サンプル色の CIE (1931)  $xy$  色度 ( $x, y$ ) と輝度  $Lum$  から Smith-Pokorny の錐体分光感度関数<sup>12)</sup>を用いて求めた  $L, M, S$  3 刺激値とする．次に，照明条件が未知の画像中の色認識に適用可能なモデルを構築するために，すべての色票中の白色点を照明光の手がかりとして用いる．ここで，そのときの照明光下で最も輝度の高かった OSA 色票の色を白色点とする．そしてモデルの照明光部分の入力データは，サンプル色と同様に，白色点の色度 ( $x, y$ ) と輝度  $Lum$  から求めた  $L, M, S$  3 刺激値とする．ただし，ここで得られた ( $L, M, S$ ) を  $[0, 1]$  の間に正規化したものを入力データに用いる．

教師データには，実験で 4 名  $\times$  2 セッション中に得られた結果の，ある色票の見えに対してある基本色名が何度用いられたかを表す色名使用比率を， $[0, 1]$  に正規化した値を用いる．このような教師データを学習させることで，ニューラルネットワークは人間が行っている色の認識過程を  $L, M, S$  3 刺激値から基本カテゴリ色名への写像という計算課題として学習することが期待される．教師データセットには，3 照明光  $\times$  424 枚 = 1,272 セットを用意する．

### 2.4 学習結果

ここでは，ニューラルネットワークの学習結果について示す．学習が正しく行われたことを確認するために，学習データに対する最終的な誤差と色認識結果の正解率を求め，表 2 に示す．ここで誤差とは，ニューラルネットワークの出力値と教師データの値の二乗誤差を求め，出力ユニット数とデータ数で平均したものである．また，正解率は，出力値が最も大きかった出力ユニットに対応する色名と心理物理実験において

ずれかのセッションで回答された色名が一致した場合を正解とし，正解数をデータ数で割って算出した．表 2 より，モデルが学習データに対して高い認識率を示しており，学習に用いた照明条件での色恒常性を実現できたことが分かる．

## 3. カテゴリカル色知覚モデルの検証実験

獲得したモデルの有効性を検証するために，10 種類の照明光を用いたシミュレーション実験と，未知の照明光下で撮影された画像中の物体色認識実験を行った．

### 3.1 シミュレーション実験

まず，獲得したカテゴリカル色知覚モデルの学習に用いた 3 種類の照明光以外の照明条件下でもモデルが同様に色恒常性を示すかどうかを検証するために，色認識シミュレーション実験を行った．この実験では未知の照明光下での OSA 色票の色を計算してモデルへの入力とした．未知の照明光には色温度が 5,000 K ~ 20,000 K の 10 種類の Daylight データ (Color Science, p.8 ~ p.10<sup>13)</sup>)を用いた．このとき用いた Daylight データの分光分布を図 3 に示す．

未知の照明光に対する出力結果の正解率を求めするために，2.3.1 項のカテゴリカルカラーネーミング実験の結果を各色票の色名の正解として用いる．ここで，このカテゴリカルカラーネーミング実験では，用いる照明光によって色票の色度が大きく変化しているにもかかわらず同一の色票に対して同じ色名が答えられており，照明光 o (3,000 K) と照明光 w (6,500 K)，照明光 w (6,500 K) と照明光 b (25,000 K) 間で色恒常性が見られる．したがって，この間の色度となる照明条件でも，同様の色名が知覚されるであろうと推定できる．そこで，Daylight の 5,000 K ~ 6,000 K の出力結果は照明光 o か照明光 w の実験結果のいずれかと一

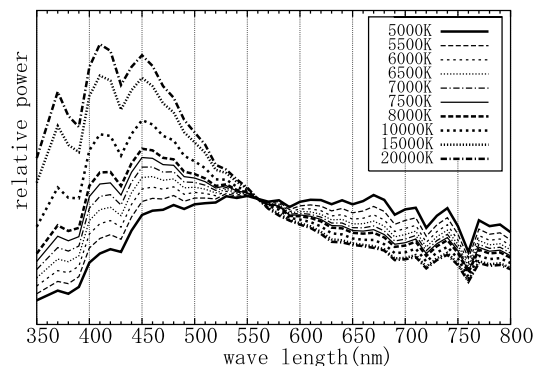


図 3 シミュレーション実験に用いた照明光の分光分布  
Fig. 3 Spectral distributions of DayLight used in the simulation test .

表 2 教師データとニューラルネットワークの出力結果の誤差と正解率

Table 2 Accuracy rates of NN with training data.

Illuminant	Mean square error	Accuracy rate (%)
o	0.031	87.1
w	0.023	89.8
b	0.015	93.7

表 3 シミュレーション実験の結果  
Table 3 Results of the simulation test.

Illuminant	Accuracy rate (%)
DL5000K	89.8
DL5500K	88.8
DL6000K	89.8
DL6500K	93.4
DL7000K	93.9
DL7500K	93.4
DL8000K	93.4
DL10000K	94.4
DL15000K	93.4
DL20000K	88.3
Average	92.8

致したものを正解とした。同様に、7,000 K ~ 20,000 K の出力結果は照明光  $w$  か照明光  $b$  の場合、6,500 K の出力結果は照明光  $w$  の場合と一致したものを正解とした。この結果、正解数をデータ数で割って算出した正解率を表 3 に示す。

表 3 から、獲得したカテゴリカル色知覚モデルが、学習に用いた 3 つの照明光とは異なる色の照明条件下でも高い色認識率を示していることが分かる。このことから、モデルは学習に用いた以外の一般的な照明光でも色恒常性を実現していると考えられる。

### 3.2 実画像への適用実験

画像中の色認識は、コンピュータビジョンにおいて最も重要な問題の 1 つである。特に、人間が知覚する見え方と同じように色を認識できるモデルは、人間と共存するロボットや人物を自動認識するセキュリティカメラに欠かせない重要な技術である。しかし、照明条件やその他の条件によって人間の色の知覚は変化するので、カメラやその他のセンサ値から人間の知覚と同様に色認識を行うことは非常に困難である。そこで筆者らは、ニューラルネットワークを用いたカテゴリカル色知覚モデルを用いて画像中の色認識を行った。

今回の実験では画像中の画素ごとに色認識を行い、その結果を基に基本カテゴリ色 11 色で色分けした画像を結果画像として示す。

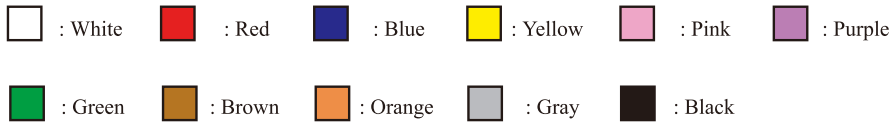
まず、画像中の色認識の手順を示す。ある画素の色名を求める場合、注目画素の値から求めた  $L, M, S$  3 刺激値をカテゴリカル色知覚モデルのサンプル色部分に入力する。また、画像中で最も高輝度な点を白色点と仮定し、白色点の階調値から求めた  $L, M, S$  3 刺激値をモデルの照明光色部分に入力する。そのときのモデルの出力値を見て、最も高い値を出力している出力ユニットに対応する色を注目画素の色名とする。これらの過程を繰り返してすべての画素に対して行うことで、色認識結果画像を得ることができる。

今回用いた 3 種類の原画像 image1 ~ image3 とその色認識の結果画像を図 4 に示す。図 4-(a), (c), (e) が原画像で、図 4-(b), (d), (f) が結果画像である。図 4 から、照明光が未知の条件でも太陽光・人工照明光にかかわらず、また自然風景・人工物などの対象にかかわらずモデルが人間と同じように色認識ができていることが分かる。たとえば、image 1 (図 4-(a), (b)) の木や緑の部分は多くの色が含まれており一般的なセグメンテーション手法では統一することが難しいが、すべて緑として認識できている。次に、image 2 (図 4-(c), (d)) の背景部分では、従来モデル<sup>9),10)</sup>で認識の難しかった白色の物体に対して、正しく白と認識できている。また、image 3 (図 4-(e), (f)) では濃い色の植物が大きく写されているために画像のヒストグラムが偏っており画像が撮影された照明条件を推定することは困難であるが、本手法では葉・花卉・めしべをそれぞれ緑・ピンク・黄色と認識できている。

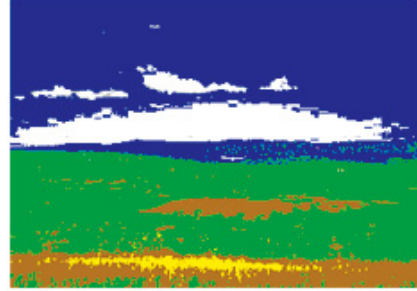
しかし、いくつかの問題点もある。たとえば、image 3 (図 4-(e), (f)) ではピンク色の花の陰影の部分が紫色と認識されている。これは、注目画素自体の色としては紫色という認識で正しいものの、本来は全体を見てピンク色の花卉の一部が影の影響で暗くなっているものであると認識する必要があったと考えられる。また、image 2 (図 4-(c), (d)) では白色のキャンディの一部や背景部分で黄色のキャンディに近い部分が黄色と認識されている。これは、黄色のキャンディからの反射光が白色の部分に反射しており、注目画素だけから判断したために全体として見ると白色と認識できる物体の一部でも黄色と認識してしまったことが原因であると考えられる。これらの問題は、画素ごとに色認識を行ったことが主な原因であり、画像のセグメンテーションを用いてある程度の領域ごとに色認識を行うことで改善されると考えられる。

## 4. ま と め

本論文では、人間のカテゴリカル色知覚と同様の色認識を行うことができるモデルを獲得するために、さまざまな照明光の下で行われたカテゴリカルカラーネーミング実験の結果をニューラルネットワークを用いて学習した。そして、獲得したモデルの性能を評価するために未知の照明条件下での OSA 色票の色認識シミュレーションを行い、モデルが色恒常性を実現できたことを示した。また、獲得したモデルを用いて未知の照明条件下で撮影された実画像中の物体色認識を行い、モデルが人間と同様に物体色を認識できたことを示した。



(a) Original image 1



(b) Output image 1



(c) Original image 2



(d) Output image 2



(e) Original image 3



(f) Output image 3

図 4 実画像への適用実験の結果  
Fig. 4 Results of the real image test.

今後の課題としては、陰影の部分の色認識の精度向上や、環境光の反射の影響を考慮することなどがあげられる。

#### 参 考 文 献

- 1) 池田光男：色彩工学の基礎，朝倉書店（1980）.
- 2) Berlin, B. and Kay, P.: *Basic Color Terms: Their Universality and Evolution*, University of California Press, Berkley (1969).
- 3) Uchikawa, K. and Boynton, R.M.: Categorical color perception of Japanese observers: comparison with that of Americans, *Vision Re-*

- search, Vol.27, No.10, pp.1825–1833 (1987).
- 4) 松沢哲郎：認知科学選書 23 チンパンジーから見た世界，東京大学出版会 (1991).
  - 5) Franklin, A. and Davies, I.: New evidence for infant color categories, *British Journal of Developmental Psychology*, Vol.22, No.3, pp.349–377 (2004).
  - 6) 内川恵二：色彩科学選書 4 色覚のメカニズム 色を見る仕組み，朝倉書店 (1998).
  - 7) Zipser, D. and Anderson, R.A.: A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons, *Nature*, Vol.331, No.25, pp.679–684 (1988).
  - 8) Usui, S., Nakauchi, S. and Miyake, S.: Acquisition of color opponent representation by a three-layered neural network model, *Biological Cybernetics*, Vol.72, No.1, pp.35–41 (1994).
  - 9) 矢田紀子，長尾智晴，内川恵二：ニューラルネットワークによるカテゴリカル色知覚モデル，映像情報メディア学会誌，Vol.59, No.12, pp.1809–1815 (2005).
  - 10) Yata, N., Nagao, T. and Uchikawa, K.: A categorical color perception model using structured neural network, *Proc. International Workshop on Advanced Image Technology (IWAIT-2006)*, Okinawa, Japan, pp.737–742 (2006).
  - 11) 内川恵二，栗木一郎，篠田博之：開口色と表面色モードにおける色空間のカテゴリカル色名領域，照明学会誌，Vol.77, No.6, pp.346–354 (1993).
  - 12) Smith, V.C. and Pokorný, J.: Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm, *Vision Research*, Vol.15, No.2, pp.161–171 (1975).
  - 13) Wiszecki, G. and Stiles, S.R.: *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd Edition, John Wiley and Sons, New York (2000).

(平成 19 年 8 月 7 日受付)

(平成 19 年 9 月 4 日採録)



矢田 紀子 (学生会員)

1981 年生。2005 年横浜国立大学大学院環境情報学府博士課程前期修了。現在，同博士課程後期在学中。視覚情報処理，コンピュータビジョン等に関する研究に従事。



長尾 智晴 (正会員)

1959 年生。1985 年東京工業大学大学院博士後期課程中退。同年同大学工学部附属像情報工学研究院施設助手。1995 年同大学工学部附属像情報工学研究院施設助教授。2001 年横浜国立大学大学院環境情報研究院教授，現在に至る。画像処理，進化計算法，神経回路網，マルチエージェント，進化経済学等に関する研究に従事。工学博士。



内川 恵二

1950 年生。1980 年東京工業大学大学院博士課程修了。1980 年カナダ・ヨーク大学博士研究員。1982 年東京工業大学大学院総合理工学研究科助手。1986 年から 1987 年にかけて米国カリフォルニア大学サンディエゴ校心理学科客員研究員。1989 年東京工業大学大学院総合理工学研究科助教授。1994 年同大学工学部附属像情報工学研究院施設教授。2001 年同大学大学院総合理工学研究科教授，現在に至る。視覚情報処理，色覚認識，心理物理学等に関する研究に従事。工学博士。