

シーン中の領域に対するカテゴリカル色知覚

Categorical Color Perception for Image Segmentation

矢田 紀子†

長尾 智晴†

内川 恵二‡

Noriko Yata

Tomoharu Nagao

Keiji Uchikawa

1. まえがき

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、その色を他者に伝えようとするときには、赤・青などのように色をいくつかのカテゴリにまとめて表現することが多い。これを色のカテゴリカル知覚という。この色のカテゴリの中には、誰もがよく用いる色があり、これを基本カテゴリ色という。基本カテゴリ色には白・赤・緑・黄・青・茶・オレンジ・紫・ピンク・灰・黒の11色があることが知られており^[1]、これは言語によらず等しく用いられることが示されている。

また、我々人間は、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。これを色恒常性という。ある物体の色の見えがどのカテゴリ色になるかは、その物体の反射光スペクトルだけではなく、周囲の環境にも影響され、色恒常性を伴って決まる^[2]。

現在行なわれているほとんどの画像処理においては、色を階調値などの数値で扱っており、しきい値を設けて特定の色を判断しているが、照明光の変化に伴って変化する多くの物体色に対して、その色名を求めることは非常に困難である。

そこで我々は、複数の照明光の下で人間のカテゴリカル色知覚を測定した心理物理実験結果をもとに特殊な構造をもつ階層型ニューラルネットワークを用いて学習することで、人間と同様な色恒常性を備えたカテゴリカル色知覚モデルを獲得し^[3]、このカテゴリカル色知覚モデルを用いて画像中の全ての画素を基本カテゴリ色に分類する。

2. ニューラルネットワークを用いたモデル

学習に用いたニューラルネットワークの教師データセットは、複数の照明光下でのカテゴリカル色知覚を測定した内川らの心理物理実験から用意した。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もっとも良く表す色名を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、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o, w, bの3種類を用いた。この3種類の照明光の相関色温度およびCIE (1931) xy色度を表1. に示す。また、呈示刺激としてOSA色票424枚を用いた。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た灰色 (equivalent of N5)

†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

‡東京工業大学 大学院総合理工学研究所,
Interdisciplinary Graduate School of Science and
Engineering, Tokyo Institute of Technology

表1. 心理物理実験に用いた照明光

照明光	相関色温度	xy色度 (x, y)
o	3000 K	(0.439, 0.410)
w	6500 K	(0.313, 0.332)
b	25000 K	(0.255, 0.252)

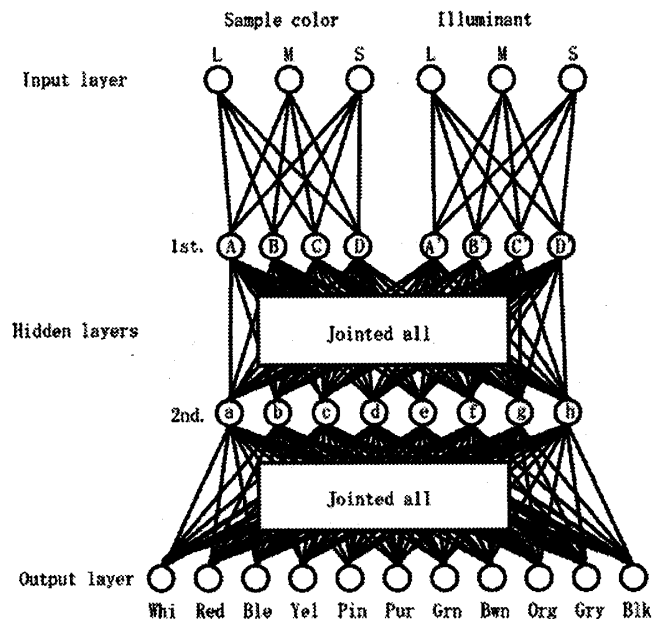


図1. ニューラルネットワークの構造

のボード上に1枚ずつ提示した。被験者はそれぞれの照明光に完全に順応した状態で刺激の見えを判断し、色票を11色の基本カテゴリ色に分類した。

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、合計で3照明光×2回=6セッション行った。

また、ニューラルネットワークの構造は多数回の試行実験による実験結果を基に、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た。

入力ユニットは3種類の錐体 (L, M, S) に相当するユニットをサンプル色、照明光色それぞれに対して1セットずつ、合計6ユニットを用いた。ここで、照明光の変化に頑健なモデルを獲得するために、図1のような、入力層と中間層第1層は共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無い構造を採用した。

またネットワークの学習は BP 法をもとに改良した学習法で行ったが、サンプル色に対する部分と照明光色に対する部分の構造上同じ位置にある結合荷重は互いに全く同じになるように学習した。中間層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を決定し、中間層第 1 層は $4 \times 2 = 8$ ユニット、中間層第 2 層は 8 ユニットからなる。また出力層は 11 ユニットからなり、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。

3. シーン中の色識別

2. に示すモデルを用いて入力画像中の全ての画素を基本カテゴリ色に分類した。

ニューラルネットワークへの入力値を求めるために、まず入力画像中の画素の階調値 $sRGB$ をガンマ変換して RGB を求め、式 (1) を用いて XYZ 刺激値へ変換した。

$$\begin{aligned} X &= 39.33R + 36.51G + 19.03B \\ Y &= 21.23R + 70.10G + 8.58B \\ Z &= 1.82R + 11.17G + 95.70B \end{aligned} \quad (1)$$

次に、得られた XYZ を式 (2) に示すに Smith-Pokorny の錐体分光感度関数^[4]を用いて LMS 錐体応答値へ変換した。

$$\begin{pmatrix} L \\ M \\ S \end{pmatrix} = \begin{pmatrix} 0.15514 & 0.54312 & -0.03286 \\ -0.15514 & 0.45684 & 0.03286 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (2)$$

このようにして得られた LMS 錐体応答値をサンプル色としてニューラルネットワークに入力した。

また、今回は画像中のシーンにおける照明光は白色であると仮定し、色度 $(x, y) = (0.313, 0.332)$ に対応する LMS 錐体応答値を照明光としてニューラルネットワークに入力した。

ニューラルネットワークの出力結果から、出力ユニット 11 ユニットのうち最も大きい値を出力したユニットに対応するカテゴリ色をその画素のカテゴリ色とした。

4. 結果

図 2. に結果を示す。左の画像は入力画像で、右の画像はニューラルネットワークの出力結果から求めたカテゴリ色で塗り分けた画像である。

図 2 から、入力画像中の様々な色に対して人間が感じるのとはほぼ同様なカテゴリ分類がされていることが分かる。

ただし、入力画像で白色に見える部分が白に分類されず灰や紫に分類されてしまっている。これは、カテゴリカル色知覚モデルを作成する際に用いた実験データが実験ブースの中で色票 1 枚毎に対する評価を行うものであったために、少しでも色味のある色に対しては無彩色が回答されていなかったことが原因であると考えられる。また、今回は照明光を白色と仮定して分類を行ったことも原因のひとつであると考えられる。

また、赤い物体や緑の物体の陰の部分茶色と分類されるなど、陰影の影響で物体そのものの色以外に分類されている部分が見られた。



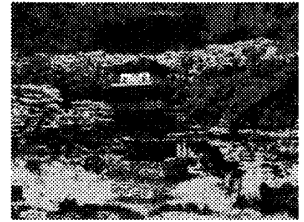
(a) 入力画像 1



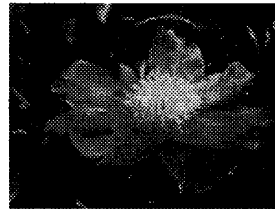
(b) 出力画像 1



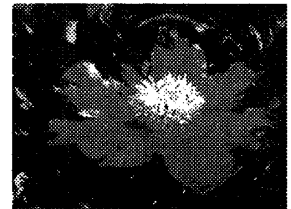
(c) 入力画像 2



(d) 出力画像 2



(e) 入力画像 3



(f) 出力画像 3

図 2. 識別結果

5. まとめ

階層型ニューラルネットワークを用いてシーン中の領域に対するカテゴリカル色知覚を行った。今後、画像中のシーンにおける照明光色を画像から推定する技術と組み合わせることや、画像中の無彩色を正確に分類することが出来るモデルの獲得を課題とする。

参考文献

- [1] Berlin, B. and Kay, P. : "Basic Color Terms: Their Universality and Evolution", University of California Press, Berkeley (1969)
- [2] 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)
- [3] 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照明光の変化を考慮したカテゴリカル色知覚モデル", 映情学誌, vol.59, no.12, pp.1809-1815 (2005)
- [4] 池田 光男 : "色彩工学の基礎", 朝倉書店(1980)