

## Color constancy問題へのロボットビジョン向きアプローチ

塚田 正人 大田 友一  
筑波大学電子・情報工学系

ロボットの視覚を指向したcolor constancyの計算理論について述べる。ロボットが過去の画像情報を正確な数値として記憶できることに着目し、2枚の画像に基づいて情景の色情報を復元する方法を開発した。情景の色情報に関して特定の仮定をおく必要はなく、2枚の画像上で2個の物体が同定されていれば、その見かけの色から情景の照明や物体の色を復元できる。実験により、本手法の有効性を示した。

## An approach to color constancy problem for robot vision

Masato TSUKADA and Yuichi OHTA

Institute of Information Sciences and Electronics

University of Tsukuba

Ibaraki, 305, Japan

We propose a computational theory for color constancy suitable to robot vision. A robot can store image information observed in the past. Then it seems natural to use more than one image to achieve color constancy. In our algorithm, we can recover the illuminant color and the reflectance color only based on the RGB values of two objects identified on two images. It is not necessary to make a specific assumption on the scene. Experiments show the validity of the proposed algorithm.

## 1. はじめに

色情報は、ロボットが情景認識を行うための有用な手がかりとなる。しかし、屋外情景のように照明が制御できない環境では、画像を通して観測される色情報は、照明条件の変動に伴って大きく変動し、安定な特徴量として活用することは困難である。

一方、我々人間は、太陽光の下でも白熱灯の下でも、物体の見かけの色に関してあまり大きな変化を知覚せず、異なった照明環境の下でも色情報を手がかりに物体を同定することができる。すなわち、物体を見る場合、その物体からの反射光の色よりも、物体の反射特性そのものに対応する色に近いものを知覚できると言われている。この機能は、color constancy（色の恒常性）と呼ばれ、その計算理論的説明が重要な研究課題となっている<sup>(1)</sup>。

ロボットの視覚にも人間のcolor constancyと類似の機能を実装すれば、色情報を安定な特徴量として利用することができる。しかし、そこで用いる手法では、必ずしも、人間の持つ制約条件を基にした枠組みを踏襲する必要はない。すなわち、ロボットは、過去の画像情報を正確な数値として記憶することができる。従って、現在観測されている画像情報だけでなく、過去に観測されたり、モデルとして与えられている画像情報を併せて、複数の画像を同時に利用できる。これを活用すれば、1枚の画像のみを利用するという枠組みに従っている人間のcolor constancyの機能を模擬するよりも、さらに汎用的で高精度の色情報補正機能の実現が期待できる。

本稿では、以上の考え方に基づき我々が開発した、ロボットの視覚を指向したcolor constancyの計算理論と、その実験結果について述べる。

## 2. 問題点

Color constancyに関する計算理論は、画像として観測される見かけの色情報から、情景中の照明による効果と物体の反射率による効果を分離することを目標としており、対象とする情景や画像に関していくつかの仮定が成立していることを前提とする。

Land と McCann によるレティネクス理論では、画面内の照明が不均一であるときの見かけの明るさの変

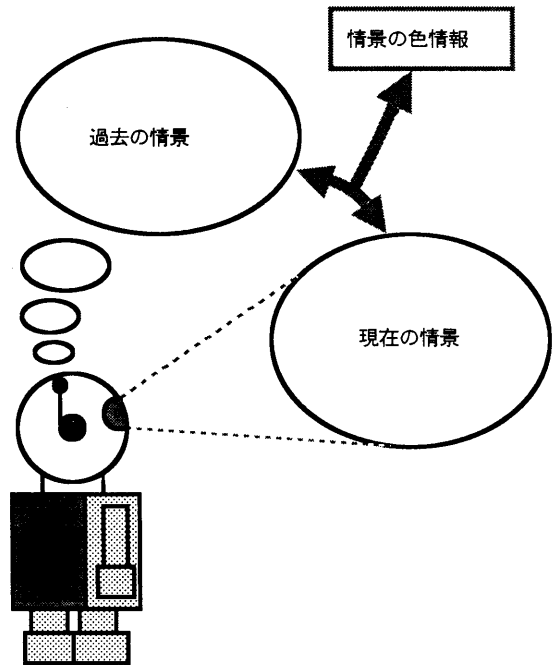


図1 ロボットによる color constancy の実現法

化を除去することにより、物体の反射率に対応する値を獲得する<sup>(2)</sup>。これは、画像上で見かけの明るさが急激に変化する部分のみが物体の反射率が変化する部分であるとする仮定に基づいている。一方、色彩画像に関しては、3原色の各チャネルに対して上記の処理を行い、最も反射率の大きな部分が白色であると仮定して、3原色間の調整を行なう。色情報の処理に当たって3原色の各チャネルを分離して扱うことは、照明光や反射率の分光特性がチャネル内で平坦であり、チャネルの境界で階段状に変化することを仮定している。

分光特性は、実際には全スペクトル域でなめらかに変化することが多く、階段状であることを仮定した取り扱いでは精度が悪い。Gershon は、分光特性を3個の基底ベクトルの加重和として表現し、この加重係数を情景の色情報として求める方法を提案した<sup>(3)</sup>。この方法では、情景中には十分に多種類の色の物体が混在し、しかも、その分光反射率の平均値は既知の値とみなせるという仮定を用いて、その値と画像中の見かけの色の平均値との比較から照明光の分光特性の加重係数を

得ている。画面内の照明は一樣であるとの仮定から、各物体の反射率が割り出せる。

情景中の照明光や反射率の色情報を導出するために、レティネクス理論の場合もGershon の場合も共通に用いている仮定は、情景中の色情報の「真の値」が少なくとも一個は既知であり、それに対応する見かけの色情報が画像から得られるということである。これは、1枚の画像のみを手がかりに情景中の色情報を復元しようとすることから必然的に生じる仮定であると考えられる。

これらの研究では、主として人間の視覚機能を模擬するという立場から、人間の場合と同様、現時点で観測されている1枚の画像のみが利用できるという制約の中で情景情報を復元する手法を模索している。これに対して、ロボットの視覚機能を実現するという観点にたてば、利用する画像を1枚に限る理由はない。ロボット、または、計算機は、過去の情報を正確に数値として記憶できる。従って、図1のように複数枚の画像を手がかりにして情景情報を復元するという枠組みを用いることが可能である。

本論文では、複数枚の画像を用いることにより、情景中の「真の値」に関しては全く仮定を置くことなく、情景中の照明光と物体反射率の色情報が復元可能なことを示す。具体的には、異なる照明の下で撮影された2枚の画像上で、情景中では同一であった物体の対が2組与えられれば、その見かけの色から情景の色情報をすべて復元できる。情景中の物体に特定の仮定をおく必要がないので、広範囲の情景に対して安定に適用可能であり、実際の情景と仮定との不一致に起因する誤差を避けることができる。1枚の画面中では照明は均一であるという仮定は必要であるが、Gershon の場合のようにできる限り広い範囲で色の平均を求める必要はなく、比較的狭い範囲に処理を限定できるので、均一照明という仮定も満たされ易い。

### 3. 分光特性のモデル化

#### 3.1 色情報の入力過程

画像に記録される色情報の一般的な入力過程を図2に示した。まず照明光が、物体にあたり反射する。この

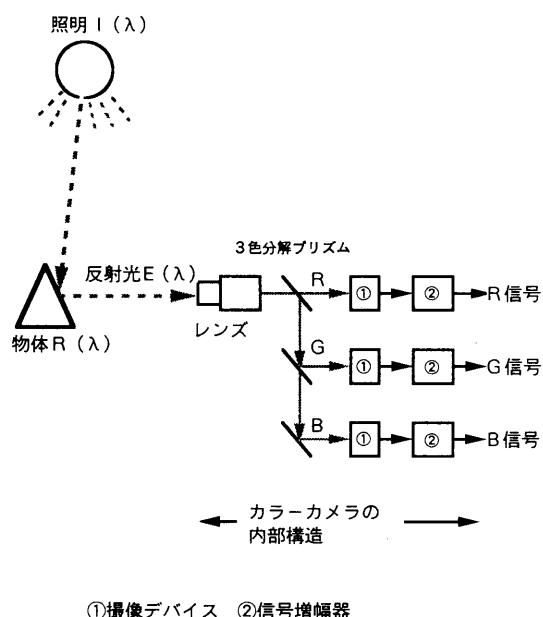


図2 色情報の入力過程

反射光はカメラレンズに入射し、3色分解プリズムにより、R, G, Bの3原色に分解される。そして、それぞれの3原色光は、撮像デバイスにより電気信号に変換され、増幅されてR, G, Bの出力信号となる。こうしてカメラにより撮影された物体のR, G, B値は、物体からの反射光の分光分布特性を $E(\lambda)$ 、撮像系の分光感度特性を $C_R(\lambda)$ 、 $C_G(\lambda)$ 、 $C_B(\lambda)$ とすると、式(3.1)のように表される。

$$\begin{aligned} R &= \int E(\lambda) \cdot C_R(\lambda) d\lambda \\ G &= \int E(\lambda) \cdot C_G(\lambda) d\lambda \\ B &= \int E(\lambda) \cdot C_B(\lambda) d\lambda \end{aligned} \quad (3.1)$$

( $\lambda$ は波長を表し、可視領域で積分)

ここで物体からの反射光 $E(\lambda)$ は、以下のように表すことができる。

$$E(\lambda) = I(\lambda) \cdot R(\lambda) \quad (3.2)$$

ただし、 $I(\lambda)$ は照明光の分光分布特性、 $R(\lambda)$ は物

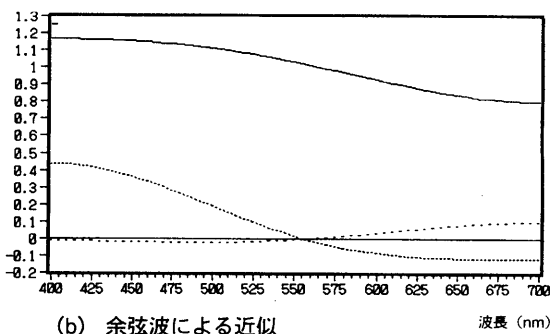
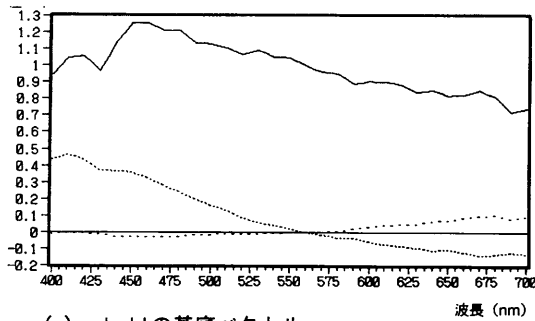


図3 照明光の分光特性のための基底ベクトル

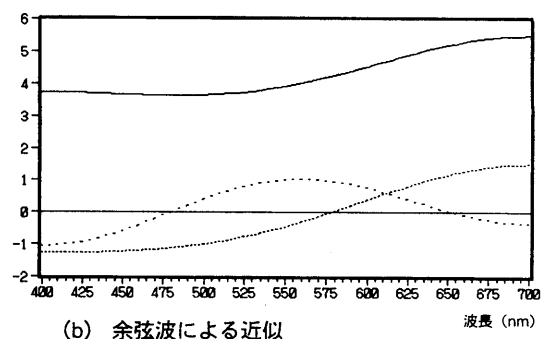
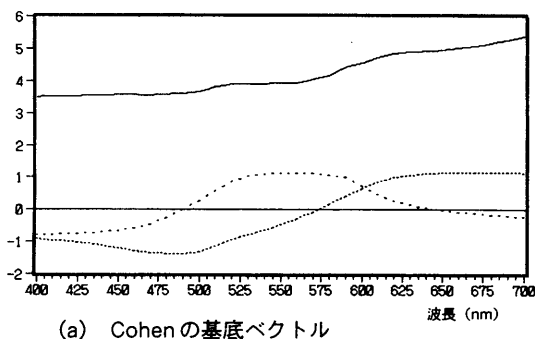


図4 反射率の分光特性のための基底ベクトル

体の分光反射特性である。照明光  $I(\lambda)$  が変化すれば、反射特性が同じでも反射光  $E(\lambda)$  は変化し、異なる R, G, B 値として入力されることになる。

$$I(\lambda) = a_1 I_1(\lambda) + a_2 I_2(\lambda) + a_3 I_3(\lambda)$$

$$R(\lambda) = b_1 R_1(\lambda) + b_2 R_2(\lambda) + b_3 R_3(\lambda)$$

(3. 3)

### 3. 2 分光特性の基底ベクトルによる展開

照明光や物体反射率の分光特性は、波長  $\lambda$  の関数であるが、 $\lambda$  に関して適当な間隔で標準化することにより、標本値のベクトルとして表現できる。Color constancy 問題において、これらの分光特性ベクトルは未知であり、観測される画像からこれらを復元するためには、比較的少数のパラメータでモデル化することが必要となる。

物体の反射率や自然光の分光特性は、比較的なめらかな曲線を示すことが多い。したがって、照明の分光特性  $I(\lambda)$ 、物体の反射特性  $R(\lambda)$  の各々を、3 個の互いに線形独立な基底ベクトルの加重和で表現できるものと仮定し、以下のように表す。

このようなモデル化の妥当性に関してはいくつかの実験が報告されている。照明光に関しては、太陽光の典型的な 622 サンプルの分光分布特性を集め主成分分析したところ、第 3 主成分までをとればほぼ 100 % の近似が得られるという Judd の報告がある<sup>(4)</sup>。また、物体の分光反射特性についても、433 色からなるマンセル色表の中からランダムに選んだ 150 色について主成分分析し、やはり第 3 主成分までで 99. 18 % の近似が得られることが Cohen により報告されている<sup>(5)</sup>。このことから上記の仮定は十分に妥当性があるといえよう。式 (3. 3) の係数  $(a_1, a_2, a_3)$ 、 $(b_1, b_2, b_3)$  をそれぞれ照明の分光分布特性、物体の分光反射特性を表す特性パラメータと呼ぶ。

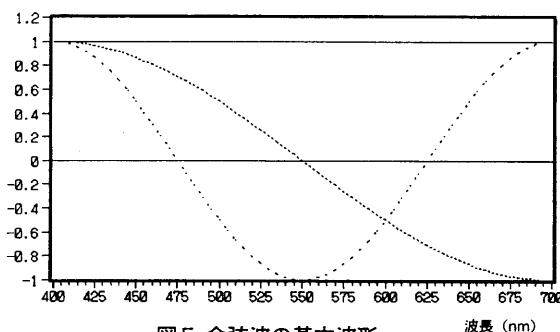


図5 余弦波の基本波形

### 3. 3 基底ベクトルの選択

照明の分光分布特性および物体の分光反射特性を3個の基底ベクトルの加重和で表現するとき、どのような基底ベクトルを選択するかは重要な問題である。図3(a)に照明光に関してJuddにより求められた第3主成分までに対応する基底ベクトルを、また図4(a)に反射率に関してCohenにより求められた第3主成分までに対応する基底ベクトルを示す。

これらの基底ベクトルは、十分に多種類のサンプルに基づいて求められており、他に適当な基底ベクトルが存在しない限り、そのまま採用するのが妥当と考えられる。しかし、照明光や反射率の分光特性について、これらの基底ベクトルによりほぼ100%の近似が得られることの本質的な理由は、これらの基底ベクトル自体に大きな意味があるのではなく、自然界に存在する分光特性が一般にはなめらかであることによると仮定することもできる。この場合、図3、4の(a)に示したそれぞれの分布を波形と見なすことにより、低周波の基本波形に展開可能なはずである。この基本波形として、図5に示した直流成分、半周期成分、1周期成分の余弦波を用い、JuddとCohenの基底ベクトル波形を近似した結果をそれぞれ図3(b)、図4(b)に示す。図から判るように、比較的良好な近似が得られている。したがって、照明光や反射率の分光特性を表現する基底ベクトルとして、余弦波形を採用することも、充分検討の余地があると考えられる。

## 4. 情景中の色情報の復元

### 4. 1 観測方程式

情景中の照明光と物体反射率の分光特性を表す特性パラメータと画像情報との関係式として、前節の式(3.3)を式(3.1)に代入し、 $R$ 、 $G$ 、 $B$ 値と特性パラメータに関する非線形連立方程式を作成する。ただしカメラの分光感度特性 $C_R(\lambda)$ 、 $C_G(\lambda)$ 、 $C_B(\lambda)$ は既知であるとする。 $R$ については、下式のようなになる。

$$R = \int \sum_{i=1}^3 a_i I_i(\lambda) \sum_{j=1}^3 b_j R_j(\lambda) C_R(\lambda) d\lambda$$

$$= \sum_{i=1}^3 \sum_{j=1}^3 a_i b_j \left\{ \int I_i(\lambda) R_j(\lambda) C_R(\lambda) d\lambda \right\} \quad (4.1)$$

ここで、 $I_i(\lambda)$ 、 $R_j(\lambda)$ 、 $C_R(\lambda)$ は既知であるので、各々の積の項について先に積分を行うことができる。つまり、式(4.1)は

$$R = \mathbf{b}^T \mathbf{C}_R \mathbf{a} \quad (4.2)$$

という特性パラメータに関する2次の非線形方程式の形に書くことができる。ただし、 $\mathbf{C}_R$ は $3 \times 3$ 定数行列であり、その $ij$ 要素は $\int I_i(\lambda) R_j(\lambda) C_R(\lambda) d\lambda$ により求められる。また $\mathbf{a}$ は、照明分光分布の特性パラメータベクトル、 $\mathbf{b}$ は物体反射率の特性パラメータベクトルを表し、 $\mathbf{a} = (a_1, a_2, a_3)^T$ 、 $\mathbf{b} = (b_1, b_2, b_3)^T$ である。 $G$ 、 $B$ についても同様に非線形方程式を作成できる。これらの3つの方程式は、情景情報に関する6個の特性パラメータと、画像を通して観測できる3個の値 $R$ 、 $G$ 、 $B$ の間の関係を示す観測方程式と考えられ、以下のように略記する。

$$\mathbf{C} = \mathbf{F}(\mathbf{a}, \mathbf{b}) \quad (4.3)$$

ただし、 $\mathbf{C} = (R, G, B)^T$ であり、 $\mathbf{F}(\cdot)$ は $R$ 、 $G$ 、 $B$ それぞれの $\mathbf{a}$ 、 $\mathbf{b}$ との関係を表す。

Color constancyの問題は、(4.3)式に基づいて未知量 $\mathbf{a}$ 、 $\mathbf{b}$ を求める問題と考えられる。ある画像において、1個の物体の $R$ 、 $G$ 、 $B$ 値を得た場合、(4.3)式は、未知数6個、方程式3個の非線形連立方程式となる。一般に、 $n$ 枚の画像上で $m$ 個の物体の $R$ 、 $G$ 、 $B$ 値を得た場合、未知数は、照明に関して $3n$ 個、物体に関して $3m$

個の合計  $3m + 3n$  個、方程式の数は  $3m$  個となり、未知数の方が方程式より常に多い。方程式の数を減らさず未知数の数を減らす方法は、複数枚の画像上で同一物体（すなわち、同一の反射特性パラメータを持つ物体）が観測可能とすることである。いま最も基本的な  $n = 2$  の場合を考えることにし、2枚の画像上で  $k$  組の物体が、何らかの手がかりにより同定されているとする。この場合、未知数は  $3m + 6 - 3k$  個、方程式は  $3m$  個となり、 $3m + 6 - 3k \leq 3m$ 、すなわち  $k \geq 2$  の場合には解を求めることが可能となる。このうち、最も単純なのは、 $m = 4$ 、 $k = 2$ 、すなわち2枚の画像上で同定されている2対の物体のみを観測する場合であり、 $m > 4$  の場合でもまず  $m = 4$  として方程式を設定することが一般的に可能なので、以後、この場合に関する解法を考察する。

## 4. 2 照明光と反射率の特性パラメータの導出

照明の異なる2枚の画像上で、2種類の物体について のべ4組の  $R$ 、 $G$ 、 $B$  値を観測することにより得られる観測方程式を以下に示す。

$$\begin{aligned} C_{11} &= F(a_1, b_1) \\ C_{12} &= F(a_1, b_2) \\ C_{21} &= F(a_2, b_1) \\ C_{22} &= F(a_2, b_2) \end{aligned} \quad (4.4)$$

ただし、 $C_{ij}$  は照明  $i$  の画像中における物体  $j$  の  $R$ 、 $G$ 、 $B$  値であり、 $a_i$  は照明  $i$  の特性パラメータ、 $b_j$  は物体  $j$  の特性パラメータである。(4.4) 式は、未知数、方程式とも12個の非線形連立方程式であり、一般的に解くのは困難である。

今、(4.4) 式のうち  $C_{11}$  の方程式において、もし  $a_1$  が既知と仮定すれば、これは  $b_1$  に関する3元1次連立方程式に帰着し、容易に  $b_1$  を求めることができる。同様に  $C_{12}$  の方程式から  $b_2$  も求められる。 $b_1$ 、 $b_2$  が求められれば、 $C_{21}$  または  $C_{22}$  の方程式から  $a_2$  が求められる。このように、12個の未知数のうち、3個を適当に固定することにより、他の9個は1次方程式の解として求めることが可能である。一方、(4.1) 式からも判るように、(4.4) 式の右辺は全て  $a_i b_j$  の項で構成されて

おり、12個のうち任意の1個を固定しないと解は定まらない。(すなわち、(4.4) 式のうち1個の式は冗長である。) したがって、一般性を失うことなく  $a_1$  の第1パラメータ  $a_{11}$  を定数に固定することができる。以上の考察に基づき、以下の手順で (4.4) 式の解を求める。

- (step1)  $a_1$  の第2,第3パラメータ  $a_{12}$ 、 $a_{13}$  が形成する2次元パラメータ空間を適当な刻みで分割し、全ての  $(a_{12}, a_{13})$  の組について step2, 3を実行する。
- (step2)  $a_1$  を (4.4) 式の上2組の方程式に代入し、 $b_1$ 、 $b_2$  を解く。
- (step3) step2で求められた、 $b_1$ 、 $b_2$  を (4.4) 式の下2組の方程式に代入し、 $a_2$  に関する6個の方程式を得て、最小2乗法で解く。
- (step4) step3における2乗誤差が最小であった  $(a_{12}, a_{13})$  の組と、それによって求められる  $a_2$ 、 $b_1$ 、 $b_2$  を解として選択する。

実際には、step1における刻みを段階的に細かくし、粗密探索法により計算量を削減する。

step1で固定するパラメータとしては、反射特性パラメータ  $b$  でもよいが、一般には照明の特性パラメータの方が変動が小さいと考えられ、探索範囲を限定できるので都合がよい。

また (4.4) 式は、非線形方程式なので複数個の解を持ち得るが、照明や反射率の特性を部分的にでも負にするような解を棄却することにより、ほとんどの場合、正解のみを選択することが可能である。

(4.4) 式の右辺が全て  $a_i b_j$  の項からなるということは、 $p$  を任意の定数とし、上で求められた特性パラメータの  $a_1$ 、 $a_2$  を  $p$  倍し、 $b_1$ 、 $b_2$  を  $1/p$  倍した特性パラメータの組も (4.4) 式を満たすことを意味する。すなわち、照明光をすべてのスペクトル帯において  $p$  倍の強さにし、物体反射率をすべてのスペクトル帯において  $1/p$  倍にしても、観測される画像に変化を生じない。逆に画像から得られる情報のみでは、照明光強度と反射率の相対的大きさを示す1個のスケールファクターが決定できないことになる。しかし、コンピュータビジョンの目的には物理的な絶対量は、必ずしも重要で

はなく,種々の情景パラメータ間の相対関係が把握できれば充分であり,全パラメータを通じて1つの自由度が残ることはほとんど問題にならないと考えられる。

### 4. 3 見かけの色情報

画像における照明の特性パラメータ  $\mathbf{a}$  が与えられれば,その画像中の任意の物体について,反射率の特性パラメータ  $\mathbf{b}$  を導出することができる。すなわち,その物体の見かけの色を  $\mathbf{C} = (R, G, B)^t$  とし,  $\mathbf{a}$ ,  $\mathbf{C}$  を (4. 3) 式に代入すれば (4. 3) 式は  $\mathbf{b}$  に関する3元1次連立方程式になり,容易に解ける。また,反射率  $\mathbf{b}$  が未知の場合に,任意の照明  $\mathbf{a}$  の下での見かけの色  $\mathbf{C}$  を導出することも (4. 3) 式により可能である。従って, 4. 2において2枚の画像における照明の特性パラメータ  $\mathbf{a}_1, \mathbf{a}_2$  が得られれば,一方の画像における見かけの色  $\mathbf{C}$  から,もう一方の画像における見かけの色の予測値  $\mathbf{C}'$  を導出できる。これを全画面について行えば,同一情景に対して照明のみを切り換えた場合の画像を生成することが可能である。

### 5. 実験

本論文で述べた手法の有効性を示すために, 図7に示したような12色の色表を含む情景を用いて実験を行った。色表の色は,左上段を①とし,左から右,上から下の順に以下の通りである。

①緑 ②濃桃 ③青 ④赤 ⑤黄 ⑥橙  
⑦淡緑 ⑧水色 ⑨黄緑 ⑩山吹 ⑪青紫 ⑫茶

撮像系としては,3原色 R, G, B ごとに別個の CCD を用いた3板式カラーTVカメラを用い, 512×480画素, 8ビット/画素でデジタル化する。撮影は,晴天時の日陰の太陽光の下で行い, (画像1) 3200K→5600Kへの色温度変換フィルタをかけた状態, (画像2) フィルタなしの状態, の2枚の画像を入力した。すなわち, 画像1はフィルタを通過させた太陽光で照明した場合と等価である。

撮像系の分光感度特性は, メーカーより提供されたデータを用いている。また, 画像より得られる R, G, B 値は撮像系の  $\gamma$  特性を考慮した変換を施した後, 特性パラメータの導出に使用した。

情景中の中央の2枚の色表 (⑥橙⑦淡緑) が2画像間

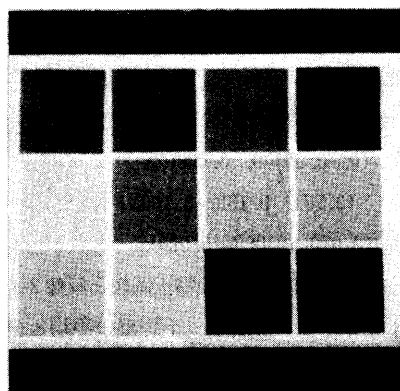


図7 12色の色表

で同定されているものとして導出した画像1, 2の照明の特性パラメータは, 以下の通りである。

a) 照明に Judd, 反射率に Cohen の基底ベクトルを用いた場合。

$$\mathbf{a}_1 = (0.5, 9.0, 14.0)^t$$

$$\mathbf{a}_2 = (1.67, 3.15, 8.50)^t$$

b) 照明・反射率とも余弦波形の基底ベクトルを用いた場合。

$$\mathbf{a}_1 = (0.5, 0.581, 0.168)^t$$

$$\mathbf{a}_2 = (0.683, 0.354, -0.159)^t$$

いずれの場合も,  $a_{11} = 0.5$  に固定してある。

次に, この特性パラメータに基づいて, 画像1の照明  $\mathbf{a}_1$  を照明  $\mathbf{a}_2$  に切り換えた画像を生成した。これを補正画像と呼ぶ。12枚の色表の補正画像上における R, G, B 値と画像2における R, G, B 値の間のユークリッド距離の平均は, a) の場合 9.36, b) の場合 8.98 であった。基底ベクトルに余弦波形を用いた場合でも, 良好な結果が得られている。

2枚の画像間で同定されているとする2色の選び方による影響を調べるため, 12色から2色選ぶ66通りのすべての組合せについて, 上記と同様の実験を行った。66回の試行により求められた照明の特性パラメータの平均と分散を以下に示す。

a) 照明に Judd, 反射率に Cohen の基底ベクトルを用いた場合。

$$\mathbf{a}_1 = (0.5, 8.17, 14.8)^t$$

$$\text{分散} : (0, 0.835, 0.915)$$

$$\mathbf{a}_2 = (1.70, 3.09, 20.3)^t$$

$$\text{分散} : (0.039, 2.105, 95.95)$$

b) 照明・反射率とも余弦波形の基底ベクトルを用いた場合.

$$a_1 = (0.5, 0.512, 0.161)^t$$

分散: (0, 0.00779, 0.00302)

$$a_2 = (0.771, 0.139, -0.109)^t$$

分散: (0.010, 0.068, 0.028)

また, 66回の全試行を通しての補正画像と画像2上での色表のR,G,B値のユークリッド距離の平均と分散は,

a) の場合 平均 = 12.76, 分散 = 18.44

b) の場合 平均 = 13.47, 分散 = 26.31

であり, 平均的には著しい悪化は見られず, 特性パラメータ導出に用いる色表の選択に関しても比較的安定であるといえる. ただし, (⑤黄, ⑩山吹) や (⑧水色, ⑨黄緑) などの同系色を選んだ場合に, 誤差が大きくなる傾向がある.

## 6. むすび

本稿では, ロボットにcolor constancyの機能を実装するという観点にたち, 複数枚の画像を用いることにより, 情景中の照明光と物体反射率の色情報を復元する方法について述べた.

今後は, 2枚の画像間で同定されている物体が2個未満の場合や, 3個以上の場合に本手法を拡張することを検討する.

## 参考文献

- 1) Marr, d.: VISION, W. H. Freeman and Company, 1982;  
乾, 安藤訳: ビジョン—視覚の計算理論と脳内表現—, 産業図書, 1987.
- 2) Land, E. H., MacCann, J. J.: Lightness and retinex theory, J. of the Optical Society of America, 61, pp. 1-11, 1971.
- 3) Gershon, R., Jepson, A. D., Tsotsos, J. K.: From R, G, B to Surface Reflectance: Computing Color Constancy Descriptors in Images, Proc. of 10th Int. Joint Conf. on Artificial Intelligence, pp. 755-758, 1987.
- 4) Judd, D. B., MacAdam, D. L., Wyszecki, G.: Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature, J. of the Optical Society of America, Vol. 54, No. 8, pp1031-1040, 1964.
- 5) Cohen, J.: Dependency of the spectral reflectance curves of the Munsell color chips, Psychonomical Science, Vol. 1, pp369-370, 1964.