

画像の色彩と形状から受ける感性情報を用いたぬりえ支援システム

井上 尚[†] 岡田龍太郎^{††} 北川 高嗣^{†††}

[†] 筑波大学情報学群情報科学類 〒305-8573 茨城県つくば市天王台 1-1-1

^{††} 筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻 〒305-8573 茨城県つくば市天王台 1-1-1

^{†††} 筑波大学システム情報系 〒305-8573 茨城県つくば市天王台 1-1-1

E-mail: [†]inoue@mma.cs.tsukuba.ac.jp, ^{††}rytrkd@gmail.com, ^{†††}takashi@cs.tsukuba.ac.jp

あらまし 本稿では人間が画像から受ける印象を計算機上で取り扱う感性モデルを作成し、それを用いた人間の知的創造活動の支援について考える。今回は特に人間と計算機の相互のつながりを意識し、知的創造活動支援の実現を目指す。今回は知的創造作業の一つとしてぬりえ作成支援について実現する。予め閉領域に区切られている2値画像をシステム側が提示し、その閉領域に対し任意の色をぬっていく形式を取る。最初に、ユーザが選択したぬりえ画像の形状情報を用いて画像から受ける印象を提示する。次に、ユーザが表現したいイメージを複数の言葉として入力し、そのイメージを表現する色をいくつか提示する。さらにぬりえ作成時にユーザの色の塗りに応じて、画像から得られる印象をユーザに対しフィードバックする。画像から受ける印象のモデルは色彩や形状から受ける印象についての専門家の知識データベースを用い作成する。画像を構成する色彩情報、形状情報を抽出し、それらの情報から画像の印象を決定する。以上により各種専門家の知識をユーザにフィードバックする仕組みを構築する。また、感性のモデル化により、人間の作成物に対しての計算機による評価や、計算機による作品の自動作成などが考えられる。今回はその例としてぬりえ画像の形状情報を用いた自動着色を構成する。

キーワード 感性, 印象, 支援, 画像

Hisashi INOUE[†], Ryotaro OKADA^{††}, and Takashi KITAGAWA^{†††}

[†] College of Information Science, Tsukuba University. 305-8573 Japan

^{††} Department of Computer Science, Graduate school of SIE, University of Tsukuba. 305-8573 Japan

^{†††} Faculty of Engineering, Information and Systems. Division of Information Engineering.

E-mail: [†]inoue@mma.cs.tsukuba.ac.jp, ^{††}rytrkd@gmail.com, ^{†††}takashi@cs.tsukuba.ac.jp

1. はじめに

個々人が作成する視覚的なコンテンツにおいて、受け取り手がどのような印象を感じるかはコンテンツ作成において最も重要視される問題である。そこで、本稿では人間が画像から受ける印象を計算機上で取り扱う感性モデルを作成し、画像から受ける印象情報を提示することで知的創造活動の支援を行う。

今回は知的創造作業の一つとしてぬりえ作成支援について実現する。予め閉領域に区切られている2値画像をシステム側が提示し、その閉領域に対し任意の色をぬっていく形式を取る。

最初に、ユーザが選択したぬりえ画像の形状情報を用いて画像から受ける印象を提示する。次に、ユーザが表現したいイメージを複数の言葉として入力し、そのイメージを表現する色をいくつか提示する。さらにぬりえ作成時にユーザの色の塗りに応じて、画像から得られる印象をユーザに対しフィードバッ

クする。

画像から受ける印象のモデルは色彩や形状から受ける印象についての専門家の知識データベースを用い作成する。

また、ユーザの知的創造活動支援の他に印象モデルを活用した例として、ぬり絵画像の形状情報を用いた自動着色を構成する。これにより画像の形状が表現する印象に適した着色を確認することができる。

以上の手法で画像の印象に関連した情報をユーザに対しフィードバックすることにより、計算機が芸術作品に対して評価を行い、それをユーザに対し提示することができる。それによりユーザの知的創造活動に対して計算機が印象という側面から支援を行うことができる。

2. 作成したシステムの概要

作成したシステムはブラウザ上で動作する Web アプリケー

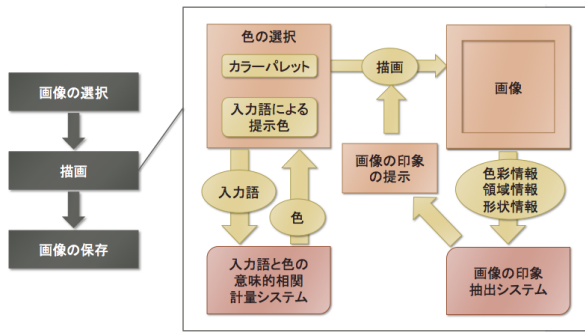


図 1 システムの全体図
Fig.1 Outline of the System

ションであり，線画の閉領域に対し色を対応させていくぬり絵の形式をとる．システムの全体図の概要を図 1 に示す．

システムの流れは大きく分けると三段階となり，それぞれ下地となる画像の選択，描画，完成した画像の保存となる．描画段階では入力語と色の意味的相関計量システム，画像の印象抽出システムにより画像作成支援を行う．具体的にはユーザが表現したい画像のイメージを任意の英単語の組み合わせとして入力させ，それらの単語群の持つ印象に対応する色をいくつか提示する．ユーザは提示された色，もしくはカラーパレットから任意の色を選択し，下地画像に対して着色していく．ユーザの着色により画像の状態が変化すると即座にその状態での画像から得られる印象を抽出し画面上にそれらを表示する．画像から印象の抽出には画像の色彩情報，形状情報を用いる．これにより実際に作成した画像がどのような印象をうけるかを確認しながら作業をすすめることができる．形状による印象については，描画中は変化しないため下地画像を選択した段階でその画像の形状から得られる印象をユーザに対して提示する．

3. 入力語と色の意味的相関計量システム

入力語と色の意味的相関計量システムとは入力語と色の意味的相関計量システムとは図 1 に示すとおり，ユーザからの入力語（複数 または単一の任意の英単語）を入力とし，それに対し意味的に相関の大きい基本色を出力する機構である．この機構を意味の数学モデル [1], [2]，及びカラーイメージスケール [3] を用いて構成する．

3.1 意味の数学モデル

本章では言葉と言葉の関係を計算することのできる意味の数学モデルの概要を示す．詳細は文献 [1], [2] に述べられている．

3.1.1 メタデータ空間 MDS の設定

対象となるメディアデータをベクトルで表現したデータにマッピングするための正規直交空間（以下，メタデータ空間 MDS ）を設定する．

初めに， m 個の基本データについて各々 n 個の特徴 $f_1, f_2, \dots, f_n$ を列挙した特徴付ベクトル $\mathbf{d}_i (i = 1, \dots, m)$ が与えられているものとし，そのベクトルを並べて構成する $m \times n$ 行列を M とおく（図 2）．このとき， M は，列ごとに

2 ノルムで正規化されている．

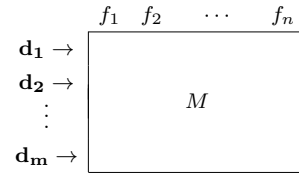


図 2 データ行列 M によるメタデータの表現.
Fig. 2 Data Matrix M

- (1) データ行列 M の相関行列 $M^T M$ を計算する．
- (2) $M^T M$ を固有値分解する．

$$M^T M = Q \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_\nu & \\ & & & 0 \dots 0 \end{pmatrix} Q^T, \quad 0 \leq \nu \leq n.$$

ここで行列 Q は，

$$Q = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n]$$

である．この $\mathbf{q}_i (i = 1, \dots, n)$ は，相関行列の正規化された固有ベクトルである．相関行列の対称性から，この固有値は全て実数であり，その固有ベクトルは互いに直交している．

- (3) メタデータ空間 MDS を以下で定義する．

非ゼロ固有値に対応する固有ベクトルによって形成される正規直交空間をメタデータ空間 MDS と定義する．この空間の次元 ν は，データ行列 M のランクに一致する．この空間は， ν 次元ユークリッド空間となる．

$$MDS := \text{span}(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_\nu).$$

$\{\mathbf{q}_1, \dots, \mathbf{q}_\nu\}$ は MDS の正規直交基底である．

3.1.2 メディアデータのメタデータ空間 MDS への写像

設定されたメタデータ空間 MDS へ，メディアデータのメタデータをベクトル化し写像する．これにより，検索対象データのメタデータが同じメタデータ空間上に配置されることになり，検索対象データ間の意味的な関係を空間上での距離として計算することが可能となる．

具体的には，メディアデータのメタデータを以下のようにベクトル化する．

- (1) メディアデータの特徴づけ
メディアデータ P を t 個の印象語（あるいは， t 個のオブジェクト） $\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_t$ を用いて，次のように特徴づける．

$$P = \{\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_t\}.$$

ここで，各印象語 $\mathbf{o}_i$ は，データ行列の特徴と同一の特徴を用いて表現される特徴付ベクトルである．

$$\mathbf{o}_i = (\mathbf{o}_{i1}, \mathbf{o}_{i2}, \dots, \mathbf{o}_{in}).$$

- (2) メディアデータ P のベクトル表現

メディアデータ P を構成する t 個の印象語 $o_1, o_2, \dots, o_t$ が、それぞれ n 次元のベクトルで定義されている。印象語 $o_1, o_2, \dots, o_t$ は、合成することでひとつの n 次元ベクトルとして表現され、メディアデータベクトル p を形成する、さらにこのメディアデータベクトル p をメタデータ空間 MDS に写像する。これにより、同じ空間上に言葉とメディアデータが配置されることになり、言葉とメディアデータの関係を空間上の距離として動的に計算することが可能となる。

3.1.3 メタデータ空間 MDS の部分空間 (意味空間) の選択

検索者は与える文脈を複数の単語を用いて表現する。検索者が与える単語の集合をコンテキストと呼ぶ。このコンテキストを用いてメタデータ空間 MDS に各コンテキストに対応するベクトルを写像する。これらのベクトルは、メタデータ空間 MDS において合成され、意味重心を表すベクトルが生成される。意味重心から各軸への射影値を相関とし、閾値を超えた相関値 (以下、重み) を持つ軸からなる部分空間 (以下、意味空間) が選択される。

3.1.4 メタデータ空間 MDS の部分空間 (意味空間) における相関の定量化

選択されたメタデータ空間 MDS の部分空間 (意味空間) において、メディアデータベクトルのノルムを検索語列との相関として計量する。これにより、与えられたコンテキストと各メディアデータとの相関の強さを定量化している。この意味空間における検索結果は、各メディアデータを相関の強さについてソートしたリストとして与えられる。

3.2 カラーイメージスケール

本研究では色彩情報と印象語の相関量を計算するためのデータベースとして 130 の単色と 180 の印象語を関連付けているカラーイメージスケール [3] を用いる。カラーイメージスケールとは 130 の基本色と 180 の印象語を関連付けている研究成果である。130 の基本色はそれぞれ RGB 成分で指定されている単色であり、180 の印象語は日本語及び英語で表記されている。基本色と印象語の関連の大きさについてそれぞれ 6 段階で評価してある。

3.2.1 色印象行列 C の作成

基本色 130 色それぞれについて、180 個の印象語を特徴としたベクトルとして表現する。このベクトルを、色印象ベクトルとして次のように表す。

$$c_k = (w_{k1}, w_{k2}, \dots, w_{k180}) \quad k = 1, 2, \dots, 130$$

c_k の各要素は、色と印象語の関連の強さを示す数値データである。カラーイメージスケールは各々の色について、それから受ける印象を 6 段階 (値なしを含む) で評価している。それらに基づいて c_k の各要素を 0 から 1 の 0.2 飛びの値で設定する。ここで色印象ベクトル $c_k(c_1, c_2, \dots, c_{130})$ により構成される色印象行列を C (図 3) とする。 C は各色と各印象語との相関の強さを表す。これにより、 C はカラーイメージスケールにおける基本色 130 色と印象語 180 語からなる 180 行 130 列の行列

として定義される。

以上の手法を用いて入力語と色の意味的相関を計量する。ここで、意味の数学モデルにおけるメタデータ空間 MDS の設定に “Longman Dictionary of Contemporary English” [4] という英英辞典を利用した。同辞書は約 2,000 語の基本語のみを用いて約 56,000 語の見出し語を説明している。この基本語を特徴とみなし、見出し語の説明で肯定的に使われている語を “1”、否定的に使われている語を “-1”、説明に使われていない語を “0” としてデータ行列 M を作成した。つまり、このメタデータ空間 MDS では、約 2^{2000} 通りの単語の組み合わせで意味空間を作成することができ、その文脈での意味が解釈可能となる。

これにより、任意の単語の組み合わせとカラーイメージスケールに掲載されている色の意味空間内での距離を計算することができるようになり、任意の単語の組み合わせに対応する印象を持つ色を提示することができるようになる。

4. 画像の印象抽出システム

本節では静止画像を入力とし、それからうける印象を出力とするシステムについて説明する。具体的には、まず静止画像から感性を考慮した色彩情報、形状情報を抽出し、それを用いて画像から受ける印象を抽出する。色彩情報はカラーイメージスケール、形状情報は単純図形から受ける印象についての先行研究 [7] を用い印象を抽出する。

また、文献 [5] に基づき、人間の感性についての研究である Fechner の法則を適用した計算式を印象の抽出に用いることで人間の感性に近い印象を導出できると考えた。

4.1 Fechner の法則

物理的な刺激と人間の感覚の関係を調べた研究に Fechner の法則 [6] がある。

E.H.Weber は人間が物の間の差異ではなく、物の大きさに対する差異の比率を知覚している、と述べた。この事実を Weber の法則という。Weber の法則が成立すると、刺激の大きさ (以下刺激強度) に対し我々が感じる感覚の大きさについて、

$$d\gamma = k \frac{d\beta}{\beta} \quad (1)$$

の式が成り立つ。

但し、 k : 比例定数、 β : 刺激強度、 γ : 感覚の大きさ、 $d\beta, d\gamma$: 刺激強度と感覚の大きさの微小増分とする。上の式を積分して整理すると、

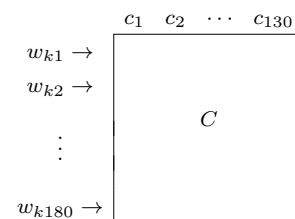


図 3 色印象行列 C .

Fig.3 Color Matrix C

$$\gamma = k \log \frac{\beta}{b} \quad (2)$$

となる．ここで b は積分定数である．この数式が表す関係を Fechner の法則と呼ぶ．

以上により刺激の大きさとそれから人間が受ける感覚の大きさは指数関数によって表現することができる．

4.2 色彩情報ベクトル r の抽出

静止画像からカラーイメージスケールに示されている 130 色に対応する 130 次元の色彩情報ベクトル $r = (r_1, r_2, \dots, r_{130})^T$ を抽出する．また、同時に 4.1 節で示した Fechner の法則に基づき、感覚の大きさと刺激強度の関係を定義する．以下に具体的な手順を示す．

4.2.1 色差による補正

ユーザの単語列の入力により提示された色はすべてカラーイメージスケールで指定されているが、ユーザが独自にカラーパレットから色を選択した場合、色彩情報ベクトルを作成する際に、カラーイメージスケールに掲載されている 130 色のうち適切な色に近似する必要がある．そこで閉領域ごとに $L^*a^*b^*$ 表色系に変換し、カラーイメージスケールに掲載されている 130 色それぞれについて色差を求める．色彩情報ベクトルを求めるときは、色差が一番小さい色はその領域に塗られていると考える．色差が 0、つまりカラーイメージスケール内の色と全く同じ場合に 1 を重みがけし、色差が大きくなるほど重み付けの値を小さくするために以下の式を適用した．

$$\alpha = b \cdot e^{-\frac{d}{k}} \quad (3)$$

ここで e は自然対数の底、 d は色差、 b, k は定数である．今回は定数 $b = 1, k = 1$ とした．

4.2.2 領域の面積による補正

さらに領域の面積についても Fechner の法則に基づいた式を適用する．直感的に面積が大きい場合と小さい場合のそれぞれに対し、一定の面積が変化する場合、元の面積が小さい方が変化に対する影響が大きくなると考えられる．これらを表現するために以下の式を適用した．

$$\beta = k' \cdot \log\left(\frac{S}{b'} + 1\right) \quad (4)$$

ここで S は領域の画素数、 b', k' は定数である．今回は定数 $b' = 1, k' = 1$ とした．

上記の α, β の積を色彩情報ベクトル r の対応する要素の値とする．

4.3 印象ベクトルの抽出

3.2.1 章で作成した色印象行列 C に上で求めた色彩情報ベクトル r をかける事で静止画像から印象語 180 語に対応する 180 次元の印象語ベクトルを求める事ができる (図 4)．

4.4 形状情報から印象の抽出

静止画像を構成する形状情報から画像の印象を抽出する．ぬり絵画像は形状がユーザの操作によって変化しないので、最初に下地画像を選択した段階でその画像の形状から得られる印象をユーザに対して提示する．

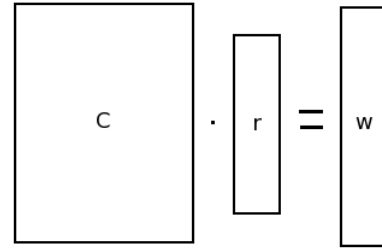


図 4 印象ベクトルの抽出

Fig. 4 Impression Vector

4.4.1 単純図形から受ける印象

単純図形から受ける印象についての先行研究 [7] に基づき、静止画像の形状から受ける印象の抽出を行う．先行研究では図形の心理的特徴として曲線性、複雑性、規則性、開閉性の 4 つの属性と図形の意味的特徴として柔和性、安定性、活動性との関係性について述べている．それぞれの性質について以下に示す．

- 心理的特徴

図形の生成特徴のこと．

- 曲線性
- 複雑性
- 規則性
- 開閉性

- 意味的特徴

図形から受ける印象のこと．

- 柔和性

「やわらかい」、「かわいらしい」、「美しい」がポジティブ

- 安定性

「美しい」、「どっしりした」、「静かな」がポジティブ

「不安な」がネガティブ

- 活動性

「活気のある」がポジティブ、「静かな」がネガティブ

上記のそれぞれの性質の組み合わせのうち、曲線性と柔和性、規則性と安定性、複雑性と活動性が密接に関連していることが分かっている．また、図形中に複数の心理的特徴が含まれている場合、図形の印象に及ぼす影響の強さは「曲線性」が一番大きく、「複雑性」、「規則性」、「開閉性」の順で小さくなる．図 5 は閉図形の単純図形を曲線性、複雑性、規則性によって分類したものである．

4.4.2 意味的特徴に対応する英語

上記の先行研究の成果を意味の数学モデルで用いるために、各種意味的特徴（柔和性、安定性、活動性）に対応する言葉を以下のように定義した．

- 柔和性

soft, pretty, beautiful

- 安定性

beautiful, balance, quiet, worry

- 活動性

life, storm

以上の言葉を意味の数学モデルに与えることで対応する印象を

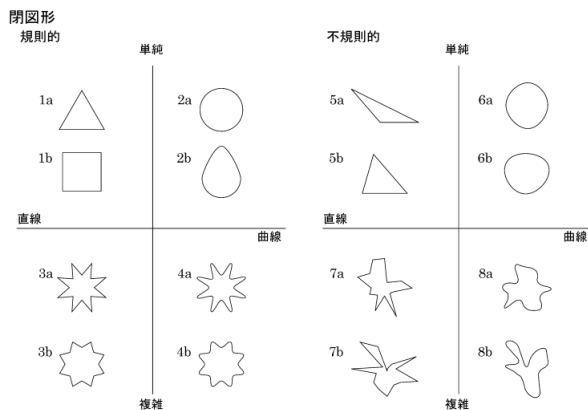


図 5 心理的特徴による単純図形の分類
Fig. 5 Psychophysical of shapes

抽出できる。

4.4.3 閉領域ごとの属性の付与

今回は予めシステム側が用意する画像の閉領域に対して、図形の生成特徴を指定する。それぞれの閉領域に対して、曲線性・複雑性・規則性をそれぞれ 11 段階で評価し、指定する。

	低い				高い	
直線	-5	...	0	...	5	曲線
単純	-5	...	0	...	5	複雑
不規則的	-5	...	0	...	5	規則的

4.4.4 属性による色の決定

上で求めた各性質に対する値（-5 ～ 5）を（-1 ～ 1）に正規化し、その単語の重みとする。ここで”beautiful”は曲線性・複雑性の値の平均値を重みとする。求めた単語と重みを 3.1 節で述べた意味の数学モデルに入力し、閉領域に対応する色を決定する。

また決定された色で描画を行うことにより、形状情報による自動着色が実現できる。

4.4.5 形状による印象の決定

形状情報により決定された色彩情報を用いて、上に示した色彩情報による印象を決定することで、画像の形状による印象の抽出ができる。

4.5 印象語リストの表示

以上の各手法により導出される印象語ベクトルから、画像から受ける印象をリストとして画面に表示する。

これら一連の処理はほとんど待ち時間もなく、ユーザが領域に着色すると印象語リストが直ちに再計算され、表示されるリストや背景の棒グラフが更新される。

5. 動作画面

5.1 入力語による色の提示

図 6 に例として “shock” を入力とした時の色の提示の様子を示す。入力の単語列と色の関連度は計算により数値で表されるので、それらの値をもとに各色との関連度を円の中心からの距離として表示する。最も関連度の高い色と入力語の関連度を 1 とし、他の 9 つの値を正規化している。

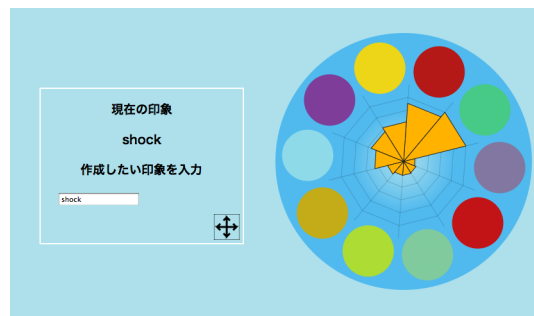


図 6 色の提示
Fig. 6 Color Suggestion

5.2 印象語の提示

静止画像から印象語を提示している様子を図 7 に示す。左部に表示した画像から受ける印象が右部の上から順に表示されている。画像の印象抽出システムにより抽出された、画像から得られる印象語ベクトルの要素のうち値の大きな 10 個の要素を選択し、それらに対応する印象語を表示している。最も値の大きい要素の値を 1 とし、残りの 9 つの値を正規化する。その値を基に割合を決定し、印象語の背景に棒グラフのように提示してある。ここで、ユーザの任意語の入力と提示される色を用いて着色していった場合の出力される印象語のリストが必ずしも一致しないという事がある。これはカラーイメージスケールで示されている特徴語以外を入力した場合におきる。

5.3 形状情報による自動着色

画像の形状情報を基に自動着色した画像を図 8 に示す。それぞれ左側の画像がもととなった画像であり、右側が自動着色した画像である。自動着色による画像は柔らかな印象の部分は暖色系、固い印象の部分は寒色系となっていると考えられる。

6. おわりに

本研究では任意の印象語からその印象を持つ色を提示し、さらにユーザの色の塗りに応じて、画像から印象のメタデータを作成し、フィードバックを行うことで、画像から受ける印象を用いた画像作成支援を行うことができた。

また、画像の形状情報を用いたぬり絵の自動着色を行うことで、形状の印象に合った色をユーザに対し提示することができた。

これらにより、画像の印象に関する情報を提示することによ

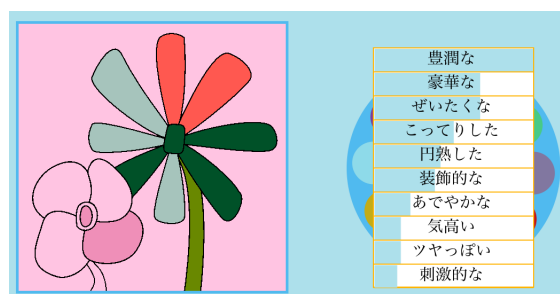


図 7 画像による印象語の提示
Fig. 7 Impression Word Suggestion

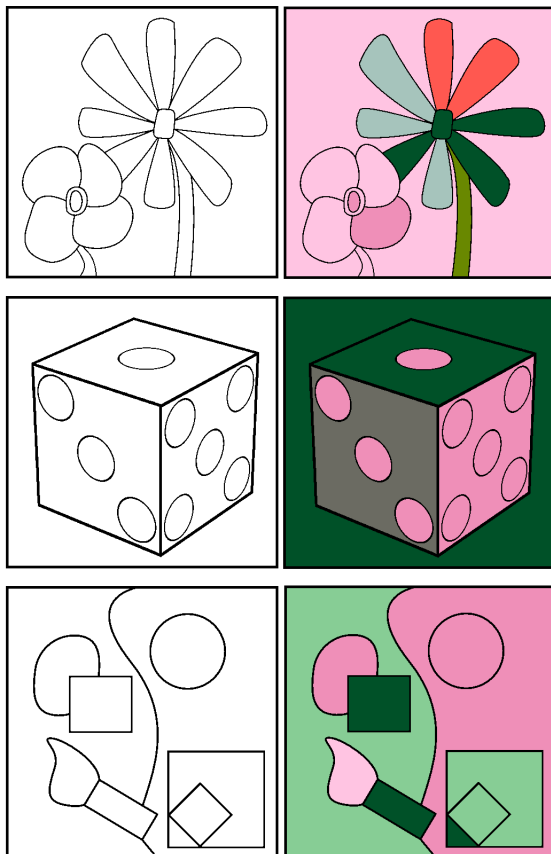


図 8 形状情報を用いた自動着色

Fig. 8 Auto-coloring List

りぬり絵という人間の知的創造活動に対して印象モデルを用いた支援活動ができたと考えられる。

今後の課題としては、今回は形状の閉領域を持つ図形特徴を人間が評価したが、任意のイラストについて形状情報を用いて印象を抽出するために任意の形状に対して各特徴量を自動で計測する手法について検討する。

また、形状が予め固定されているぬり絵ではなく、ユーザの自由描画について動的に印象情報を提示するための手法について検討する。

文 献

- [1] Kitagawa, T. and Kiyoki, Y.:The mathematical model of meaning and its application to multidatabase systems,Proceedings of 3rd IEEE International Workshop on Research Issues on Data Engineering: Interoperability in Multidatabase Systems(1993).
- [2] Kiyoki, Y., Kitagawa, T. and Hayama, T.:A Metadatabase System for Semantic Image Search by a Mathematical Model of Meaning,Multimedia Data Management ? using metadata to integrate and apply digital media ?,McGrawHill,,A. Sheth and W. Klas(editors) (1998).
- [3] 小林重順 , カラーイメージスケール改訂版, 講談社, 2001 .

- [4] Longman Dictionary of contemporary English.(1987).
- [5] 北川高嗣, 中西崇文, 清木康, 静止画像メディアデータを対象としたメタデータ自動抽出方式の実現とその意味的画像検索への応用
- [6] 大山 正, 今井 省吾, 和氣 典二, 新編 感覚・知覚心理学ハンドブック, 誠信書房 (1994) .
- [7] 山口由衣, 王晋民, 椎名健, 図形の心理的特徴と意味的特徴の対応関係, 認知心理学研究 第一巻第一号, 2004,45-54