

マルチメディアに対する感性合成処理の コンセプトと実験システムの評価

長谷川 隆[†] 北原 義典[†]

本論文では、エンドユーザによる効果的でセンスの良いマルチメディア作品の作成を支援する目的で、雰囲気やイメージ等の形容語を入力することにより、作品に色彩・音楽等のマルチメディア情報を付与することを特長とする、感性合成処理の基本コンセプトを提案する。感性合成処理では、入力された形容語からマルチメディア情報生成に必要な様々な形態のデータを検索する感性一般モデルと、メディアごとの情報生成手段であるメディア依存モデルの2つのモデルを用いて、各メディアの情報が独立して生成される。本論文ではまた、提案した感性合成処理方式の一実現例であり、形容語の入力によって静止画像のテンプレートとバックグラウンド音楽を生成することを特長とする、マルチメディアプレゼンテーション作成支援実験システムについて、前記各モデルの具体的実現方法を中心に述べる。最後に、実験システムにより独立に生成されたメディア情報がマルチメディア情報として感性的にどのような相互作用を及ぼすかに関し、心理実験により評価を行い、本方式の有効性を示した。

Basic Concept of Multimedia Kansei Synthesis Method and Evaluation of Its Experimental System

TAKASHI HASEGAWA[†] and YOSHINORI KITAHARA[†]

In this paper, we propose the basic concept of the Kansei Synthesis Method, a technique that can be used to create multimedia information such as color and music from given descriptive words and phrases representing atmosphere or image. The purpose of this technique is to enable users to create impressive and tasteful multimedia products. The Kansei Synthesis Method consists of the Kansei General Model, which uses the descriptive words and phrases to retrieve various forms of data to synthesize multimedia information, and the Media Dependent Model, which is a means of separately creating audio and visual media information. We also describe an experimental system based on this method focusing on how to apply the two models. Finally, we conducted a psychological experiment to demonstrate feasibility of this method by evaluating the interaction in the sense of Kansei between the two types of media information created with this system.

1. はじめに

画像情報と音響情報をあわせて提示すること（マルチメディア提示）を特徴とするマルチメディア作品は古くから存在する。たとえば映画、テレビ番組、パソコン等のゲームがこれに当たる。一般にこれらの作品は、監督（ディレクター）・デザイナー・音楽家等の視覚・聴覚演出効果に関する専門家が、多くの資金と時間を費やし、共同作業をすることによって創られていた。

一方、近年のマルチメディア機器の普及により、演出効果に関する専門家でないエンドユーザが、マルチ

メディア作品を自分で創る機会が増加してきている。たとえばビジネスマンによる企画・報告等の様々なプレゼンテーション、研究者・学生による学会発表、一般店舗の店員が作成する電子カタログ、学校教師が作成する教育ソフト、家庭におけるビデオ作品等がマルチメディア作品として作成されるようになってきた。これらの作品は画像・音響といったマルチメディア情報を効果的に用いて相手に訴えかけることを目的とするので、それら演出効果の品質は作品の良し悪しを決める重要な鍵となる。ところが、エンドユーザにとって、効果的にマルチメディア情報を用いた高品質の作品を制作することは非常に困難である。そこでエンドユーザに対し、人間の感性にかかわるマルチメディア情報の演出を支援する手段が必要となってきた。

このような状況の下、従来の計算機科学では扱われ

[†] 株式会社日立製作所中央研究所
Central Research Lab., Hitachi Ltd.

なかった感性情報処理という分野が登場し、いくつかの機関で研究が行われている¹⁾。この中で雰囲気やイメージ等の形容語を入力しメディア情報を生成する感性合成処理に関し、今までに報告された研究例を以下にあげる。

まず、佐藤らによるイラスト画合成システム²⁾は、入力された形容語に適合する背景色の組合せとシンボル(描かれる物)を検索し、背景色をグラデーション等で混合した背景上に、属性として定められた位置にシンボルを配置することにより画像を合成するシステムである。同システムは、特定の作者の画風に依存した手法を用いているために、生成される画像に制限があるが、入力された形容語の雰囲気を持つイラストを生成できた。

また、才脇らによる作/編曲支援システム³⁾は、形容語と和音の断片の対応をデータベースとして記憶しておき、入力された形容語に対応する和音断片を組み合わせた和声を基に音楽規則に従ってメロディを生成するシステムである。一方、中村らによるBGM自動生成システム⁴⁾は、入力された形容語に対応するテンポ、ボリューム、音階、伴奏ボタンを用いて、音楽規則に従って生成された和声と与えられたモチーフを組み合わせて生成されるメロディからBGMを生成するシステムである。これらのシステムはいずれも、生成されるメロディに制限があるが、入力された形容語の雰囲気を持つ音楽を生成することができた。

以上の研究例はいずれも1種類のメディア情報を生成するシステムであった。一方、マルチメディア作品の演出効果付与の支援を行うためには、画像と音響のメディアをあわせて取り扱う処理が必要となる。そこで、我々は、エンドユーザがマルチメディア・プレゼンテーションを制作する際に、その作品の全体の雰囲気、イメージを表す形容語を指定するだけで、作品に使用する色彩やBGM等のマルチメディア情報から構成される演出効果を生成するマルチメディア・プレゼンテーション作成支援システムを開発している^{5),6)}。本システムは、入力された形容語をキーワードとして、あらかじめ保存されたマルチメディア・データや特徴量、規則等を検索し、それらの情報からマルチメディア情報を、画像と音響のメディアごとに独立に生成する。

本論文の目的は、マルチメディア作品に対する演出効果付与の支援を行うために必要となる、形容語を入力して画像と音響のマルチメディア情報を生成する処理を、マルチメディア感性合成処理として体系化し、その概念モデルを提案し、その一実現例であるマルチメディア・プレゼンテーション作成支援システムを評

価することにより同概念モデルを検証することである。そこで、本論文ではまず、上記概念モデルを提案し、その一実現例であるシステムについて詳細に述べ、最後に、同システムの評価を行った結果を述べる。

2. マルチメディア感性合成処理

2.1 用語の定義と処理対象の範囲

まず、マルチメディア感性合成処理に関連する用語を以下のように定義する。

- (1) 感性情報：雰囲気やイメージ、印象といった人間の感性に関する情報。
- (2) 感性物理量：人間の感性に影響する光や音等の物理現象を表す量の組。メディア情報は、感性物理量の組合せで合成されるので、感性物理量の一種である。
- (3) 感性情報処理：感性情報と感性物理量を相互変換する処理。
- (4) 感性合成処理：感性情報処理の中で、特に、エンドユーザが入力した感性情報からマルチメディア作品の制作にかかわる感性物理量を作成する処理。作成された感性物理量は、入力された感性情報に対応する感性を持つ。ここで「感性物理量が感性Xを持つ」とは、その視聴者(の多く)が、提示された物理量に対し、感性Xの印象を受けることを表す。
- (5) 感性合成処理モデル：感性合成処理を計算機で実現するためのモデル。

また、本論文では、処理対象をマルチメディア作品における画像と音響に限定し、その中で有用と考えられる特徴量を感性物理量とする。画像情報は、図形、文字、グラフのいずれかである部品が配置された静止画像、それら部品が動く動画像、ビデオ画像のいずれかとする。また、音響情報は、作品のほぼ全体に渡って流れているバックグラウンド音楽(BGM)、短い時間で発せられる音や音楽(ジングル)である効果音、解説のための音声のいずれかとする。したがって、取り扱う感性物理量は以下ようになる。

- (1) 色彩：画像情報に用いられる色の組合せ。
- (2) 形状：画像内の部品の形状。
- (3) レイアウト：画像内の部品の配置。
- (4) 動き：部品の速度ベクトル、ビデオ画像の変化率。
- (5) 音楽：音色情報、音符情報(発音時刻、発音時間)。
- (6) 効果音：波形。
- (7) 声質：音声の音色、声の高さ、しゃべる速さ、声の強さ。

ここで、(5)~(7)の場合は単体で音響情報を構成

するが、(1)~(3)の感性物理量は通常、合成されて1つの画像情報になる。また、声質の場合は、音声合成、またはあらかじめ入力された音声情報の声質を交換することにより作成される。

2.2 感性合成処理モデル

次に、感性合成処理モデルに関して詳細に述べる。

感性合成処理では、エンドユーザが感性情報を入力するためのインタフェースとして**感性表現語**というシンボルを用いる。ここで、感性表現語は「明るい」、「落ち着いた」といった雰囲気を表す形容詞や形容動詞か、「海のような」、「ヒップホップ風の」といった比喩やイメージを表す名詞を用いた形容語である。したがって、本モデルは、感性表現語で表現できる感性のみを取り扱う。

また、マルチメディア作品に対する感性合成処理において、以下に示す仮定を設ける。

[メディア感性独立仮定] 画像と音響の各メディア情報が持つ感性が等しいならば、それらをあわせて提示したマルチメディア情報が持つ感性は個々のメディアが持つ感性に等しい。

これは入力表現語の感性を持つ画像情報と同じ感性を持つ音響情報を独立して作成し、それらを同時に提示すると、マルチメディア作品としてやはり同じ感性を持つという仮定である。この仮定に基づくと、感性合成処理は、メディア情報間の相互作用を無視して、メディア情報ごとに合成処理を独立に行えばよいことになる。

以上から、感性合成処理モデルの構造を図1のように定めることができる。本モデルでは、感性合成処理は、入力された感性表現語を、感性一般モデルとメディア依存モデルを順次用いて、感性物理量に変換する処理である。以下に、感性一般モデルとメディア依存モデルに関し、詳細に述べる。

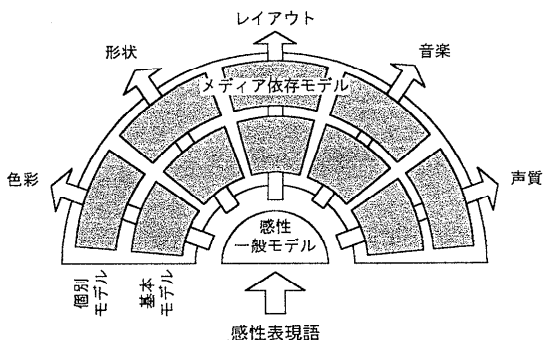


図1 感性合成処理モデル

Fig. 1 Kansei synthesis model.

2.2.1 感性一般モデル

本モデルは、感性合成処理の中でメディアに依存しない部分の処理モデルであり、感性表現語と感性物理量の作成に必要なデータ（感性データ）をフィールドとして持つデータベースと、そのマネジメントシステムから構成される。本モデルでは、入力された感性表現語をキーワードとして感性データが検索される。本モデルで用いられるデータベースを**感性データベース**という。以下に感性データの種類をあげる。

- (1) **第一種感性データ**（マルチメディア・データ）：感性物理量そのもの。専門家により創作されたデータを格納する。
- (2) **第二種感性データ**（感性特徴量）：感性物理量を生成するための特徴量の組であり、数値、または記号の組である。
- (3) **第三種感性データ**（感性規則）：感性物理量生成の際に用いられる制約条件、または生成規則であり、数式、またはルールである。
- (4) **第四種感性データ**：感性物理量生成アルゴリズムそのものであり、プログラムである。

また、感性表現語・感性データ間の関連付けは、以下にあげる2つの方法のいずれかを用いて行う。

- (1) **メディア専門家**（グラフィック・デザイナー、音楽家等）の知識を、データベース構築者が、アナウンスにより、または著書等から取得し、感性データに変換する。
- (2) 本モデルを用いて作成された感性物理量の持つ感性を、複数のエンドユーザを被験者とした心理実験により求め、元の感性データに対する感性表現語とする。具体的な方法は、後述（3.4.2項）する。

(1)は、メディア専門家がエンドユーザの一般的な感性を知っているという仮定に基づいた方法であり、同方法を採用した場合には、第一種感性データ以外のデータも直接作成することができ、少数の専門家の知識獲得のみでデータベースが構築できる。

一方、(2)の方法を採用した場合には、個人の個性に影響されない一般的なデータを得ることができるが、統計的手法を用いてデータを求めるので、比較的多数のエンドユーザに対する心理実験が必要となる。しかも、本来、感性表現語に対してエンドユーザが持っているイメージは様々であり、一般に実験結果の分散が非常に大きくなるので、通常の統計的手法による有意な結果は出にくい。この問題に対しては、有意水準を低く設定する代わりに、感性合成の際にメディア情報を複数作成し、エンドユーザがそれらの中から選択す

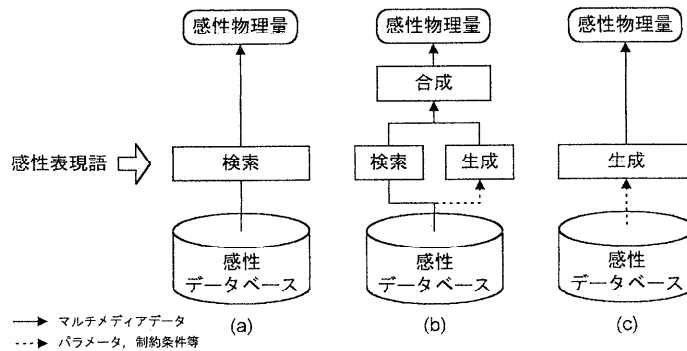


図2 感性合成処理のレベル
Fig. 2 Level of Kansei synthesis.

る方法をとることで解決する。また、本方法により作成されたデータベースは、被験者に共通の文化的・空間的・時間的背景に特有のものとなる。

2.2.2 メディア依存モデル

メディア依存モデルは感性一般モデルによって検索された感性データを用いて、メディアごとに感性物理量を生成するためのモデルで、メディア依存基本モデルとメディア依存個別モデルから構成される。

まず、メディア依存基本モデルに関して説明する。一般に、メディア情報を作成する場合、感性表現語にかかわらず一般的に満足すべき、または満足することが望ましい条件が存在する。たとえば、作成物理量が色彩の場合、センスの良い配色を作成することが望まれる。また、音楽の作成の場合には、作成された音符列が少なくとも音楽として聞こえることが必要である。これらの感性表現語に非依存の感性物理量に対する意味的条件を満たすことを「基本感性を満たす」と呼ぶことにすると、メディア依存基本モデルは、作成される感性物理量が基本感性を満たす場合に真となる制約条件の集合、または基本感性を満たす感性物理量を生成可能なプログラムである。

次に、メディア依存個別モデルは、感性表現語に応じた感性物理量を作成するモデルであり、たとえば「明るい」という感性表現語に対し、明るい配色や明るい音楽を生成するためのモデルである。感性一般モデルからの出力である感性データが第二種、または第三種である場合には、感性データで与えられたパラメータ、または制約条件から感性物理量を生成するプログラムであり、第四種である場合には、感性データに記述されたプログラムの実行エンジンである。いずれの場合も、メディア依存基本モデルの制約条件を満たすように感性物理量が作成される。

2.3 感性合成処理のレベル

また、感性合成処理には感性物理量の作成方法により3つレベルが考えられる。以下に各レベルの概要を述べる(図2)。

2.3.1 レベル0 (感性検索)

感性データベースに蓄積されている感性データが第一種のみである場合の感性合成処理を、レベル0の処理と定義する。この場合には、感性物理量の作成はマルチメディア・データベースの検索と等価になり(図2(a))、検索された物理量がそのまま出力されるので、メディア依存モデルは不要になる。また、第一種感性データは感性表現語と感性物理量を直接関連付けるので、出力されたメディア情報は入力された表現語の感性を精度良く持つことが期待できる。しかし、入力表現語に対してデータベースに蓄積されたデータのみしか出力しないので、各表現語に対して多数のマルチメディア・データを用意する必要がある。したがって、データ作成、および感性表現語のキーワード付けのコストを考慮すると、非常に小規模のシステムにしか適用できない。

さらに、音楽や動画像のように時間情報を持つ場合には、任意時間の感性物理量を作成する必要があるので、レベル0の処理では作成できない。また、声質の場合も、任意の発話内容を作成する必要があるので同処理では作成できない。したがって、レベル0の感性合成処理は静止画像情報の作成のみに対して適用可能である。

2.3.2 レベル1 (感性検索生成)

第一種とその他の感性データが感性データベースに蓄積されている場合の感性合成処理をレベル1の処理と定義する。この場合、第二種以上の感性データを用いて生成されたマルチメディアデータと検索されたマルチメディアデータを合成することにより感性物理量

を作成する(図2(b)).たとえば,音楽の場合には,短い音楽素片を複数用意しておき,パラメータや制約条件に基づいてそれを組み合わせることにより音楽を合成する場合がこれに相当する.

本処理では,少ないデータから多くの種類の感性物理量を生成できるため,レベル0に比べエンドユーザの選択範囲が大きく,また任意時間の感性物理量を作成できる.また,生成が困難な要素に対して専門家により作成されたマルチメディア・データを用いることにより,すべてを生成した場合(後述のレベル2)に比べ,より専門家が創作した感性物理量に類似した物理量(以下「上質の物理量」と記す)を生成できる.また,レベル0に比べ小規模のマルチメディア・データしか必要ないので,多くのデータを用意できる.

一方,感性データから様々な物理量が生成されるので,データと感性表現の対応付けが困難である.この問題は,メディア依存モデルを適切に構築し,その妥当性を心理実験により検証することで解決できる(4章「実験システムの評価」参照).

以上から,感性合成システムを構築する際に,本処理が最も実用的であると考えられる.

2.3.3 レベル2(感性生成)

各メディアに対し感性データベースに第二種以上の感性データのみが蓄積されている場合の感性合成処理を,レベル2の処理と定義する.この場合,感性物理量を直接生成する(図2(c)).本処理は,一般に大容量となるマルチメディア・データを蓄積する必要がなく,しかも多様な感性物理量を生成できる.一方,レベル1と同様に感性データと感性表現の対応付けが困難であり,上質の物理量を生成することが難しい.

したがって,本レベルの処理は,上質の物理量を生成可能なモデルが構築できる用途,もしくはレベル0,1では作成できない感性物理量に用いられる.

3. 感性合成実験システム

次に,感性合成処理の実験システムであるマルチメディア・プレゼンテーション作成支援システムについて述べる.本システムは,ビジネスマン等のエンドユーザが,カラーの背景上にカラーのテキスト・グラフ・表・イラスト・図形等を配置したOHPイメージの複数の静止画像と,それらの提示時間中に流れる音楽とを用いた,マルチメディア・プレゼンテーションを構築する際の感性的な演出支援を目的とする.プレゼンテーションの雰囲気・イメージといった感性表現語を入力することにより,画像テンプレートの作成,BGMの付与を行うことを特徴とする.ここで,画像テン

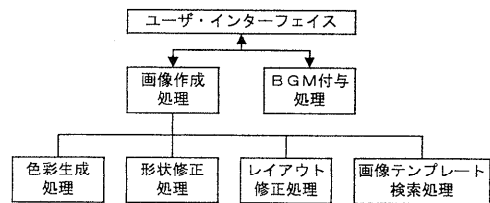


図3 マルチメディア・プレゼンテーション編集支援システム
Fig.3 Computer aided multimedia presentation product system.

プレートとは,表題等の文字やグラフ等の数値が未入力で背景や用いられるフォント,グラフの種類,それらの配置等が与えられている画像である.

なお,本システムは狭義のプレゼンテーションのほか,研究者による発表,一般店舗の店員による電子カタログ,学校教師が作成する教育ソフト,家庭における電子アルバム等の広義のプレゼンテーションの作成に対し感性合成が可能である.また本システムは,MS-WINDOWS上にVisual C++とVisual BASICを用いて開発された.

以下に,本システムの構成,各処理の詳細を順に述べる.

3.1 システム構成

本システムの構成を図3に示す.本システムは,入力された感性表現語から複数の静止画像作成の支援を行う画像作成処理,作成された画像列にBGMを付与するBGM付与処理,およびユーザー・インタフェースから構成される.この中で,画像作成処理には,形状,レイアウトをそれぞれ修正するための処理が含まれる.また,本システムでは,画像作成処理とBGM付与処理においてレベル1の感性合成処理を用いる.以下に感性合成にかかわる前記2つの処理と感性データベースの構築方法に関し,詳細に述べる.

3.2 画像作成処理

本処理では,画像テンプレート検索,色付け,およびエンドユーザによる画像の編集の支援を行う.第一種感性データである画像テンプレートは,背景の線画データと画像上に配置される属性付き部品の組を持つ.画像テンプレートの例を図4に示す.これは,部品を2つ持つ2色の背景と,縦棒グラフと色付きテキストから構成されるグラフ表示のための画像テンプレートである.入力された感性表現語をキーとして検索された画像テンプレートは,同じ表現語を用いた色彩生成処理により色付けされ,エンドユーザに提示される.エンドユーザはテンプレートに対し,テキスト,グラフの値等を入力,背景以外の部品の編集,イラスト等の図形の追加等の編集作業を行うことができる.なお,

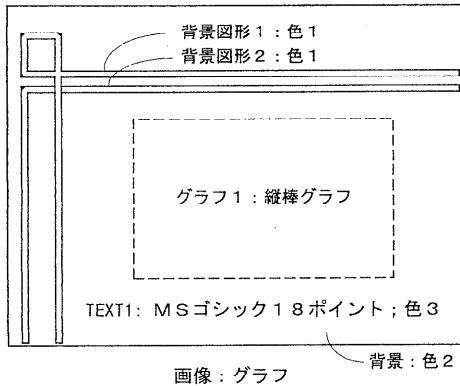


図4 画像情報の感性データの例

Fig. 4 Example of Kansei data for visual information.

イラストに関しては、レベル0の感性合成処理を行う。

3.2.1 色彩生成処理

一般に、システムにより生成された色彩は表示するディスプレイ、環境光等により変化する。一方、人間は、室内光・太陽光にかかわらず同じ物の色は同じに見える等の、色彩に関する順応性を持つ。したがって、色彩に関して厳密なモデルを構築することは無意味である。以上を考慮して本システムで採用しているモデルを述べる。

まず色彩は、人間の感覚を忠実に数値化した表現と考えられる修正マンセル記法により色相 (H; 色相目盛を用いる)、明度 (V), 彩度 (C) の3つの実数で表現する^{7),8)}。なお、Hは0~100, V, Cは0~10の実数値で表現される。

次に基本モデルは、視認性が良く、色彩が調和するための条件であり、任意の2つの色彩に関する条件をHVCに関する不等式の論理式で表したものである。これは第三種感性データに相当し、一般に以下のように表すことができる。

$$L_B(f_{k_b}(H_{i_1}, V_{i_1}, C_{i_1}, H_{i_2}, V_{i_2}, C_{i_2}) > 0) \\ (k_b = 1, \dots, n_b; i_n = 1, \dots, n_c, i_1 \neq i_2) \quad (1)$$

ここで f_{k_b} は数式、 L_B は n_b 個の不等式を述語とする論理式、 $H_{i_n}, V_{i_n}, C_{i_n}$ は色 i_n の値、 n_c は画像に用いられている色数である。本システムでは色の調和に関しムーン・スペンサーの調和理論を、視認性に関しては、背景色と文字、グラフ等の部品の色 (以下前景色と記す) との間に経験的に以下に示す条件を実装している⁵⁾。

- (1) 色相: 類似した色相と補色は避ける。
- (2) 彩度, 明度: 互いの色の彩度, または明度の差が大きくなるようにする。

これらの条件を式 (1) の形式で表した具体例を図5

$$\text{not}(1 < D_H < 7 \text{ or } 12 < D_H < 28) \text{ and} \\ \text{not} \left[\{d_{vc}(0.15, 0.5) \geq 1 \text{ and } d_{vc}(0.5, 3) \leq 1\} \text{ or} \right. \\ \left. \{d_{vc}(1.5, 5) \geq 1 \text{ and } d_{vc}(2.5, 7) \leq 1\} \right]$$

where

$$D_H = \begin{cases} H_i - H_j & (0 \leq H_i - H_j \leq 50) \\ H_j - H_i & (-50 \leq H_i - H_j \leq 0) \\ H_i - H_j + 100 & (50 < H_i - H_j < 100) \\ H_j - H_i + 100 & (-100 < H_i - H_j < -50) \end{cases}$$

$$d_{vc}(a, b) = \sqrt{\frac{(V_i - V_j)^2}{a^2} + \frac{(C_i - C_j)^2}{b^2}}$$

i, j : any color

(a)

$$\overline{H}_i \cdot \overline{H}_j > 1 - \varepsilon_1 \text{ and } (|V_i - V_j| > \varepsilon_2 \text{ or } |C_i - C_j| > \varepsilon_3)$$

where

$$\overline{H}_i = \left(\cos \frac{2\pi H_i}{100}, \sin \frac{2\pi H_i}{100} \right)$$

i' : Background's Color; j' : Component's Color

ε_i : Constant

(b)

図5 色彩感性基本モデル

Fig. 5 Color Kansei basic model.

感性表現語	感性データ
重厚な	$\overline{H}_i \cdot e_{\frac{3\pi}{5}} < \zeta_1 \text{ and } V_i < \zeta_2 \text{ and } C_i - 3 < \zeta_3$
爽やかな	$\overline{H}_i \cdot e_{\frac{11\pi}{10}} > \zeta_4 \text{ and } V_i > \zeta_5 \text{ and } C_i - 6 < \zeta_6$
楽しい	$\overline{H}_i \cdot e_{\frac{\pi}{2}} > \zeta_7 \text{ and } V_i - 6 < \zeta_8 \text{ and } C_i > \zeta_9$

where

$$\overline{e}_\theta = (\cos \theta, \sin \theta); \zeta_i: \text{Constant}$$

図6 色彩感性個別モデル

Fig. 6 Color Kansei individual model.

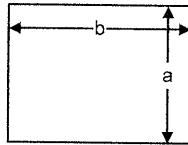
に示す。ここで図5(a)がムーン・スペンサーの調和理論に対応し、すべての色彩の組合せが満たさなければならぬ条件であり、背景色と前景色との間はさらに(b)も満たさなければならぬ。

次に、感性データベースに格納されている感性データは、感性表現語ごとに、色に関する条件を、HVCに関する不等式の論理式で表したものであり、以下のように表すことができる。

$$L_D(f_{k_d}(H_i, V_i, C_i) > 0) \\ (k_d = 1, \dots, n_d; i = 1, \dots, n_c) \quad (2)$$

ここで L_D は n_d 個の不等式を述語とする論理式である。

一般に、感性表現語に対応する色彩の特徴は古くから研究されており^{9),10)}、それらの特徴を上記不等式に表すことにより感性データが作成できる。図6に感性



$$\sqrt{\frac{1}{G}} \leq \frac{a}{b} \leq \sqrt{G} \rightarrow a' = b' = \sqrt{ab}$$

$$\frac{a}{b} > \sqrt{G} \rightarrow a' = \sqrt{Gab}; b' = \sqrt{\frac{ab}{G}}$$

$$\frac{a}{b} < \sqrt{\frac{1}{G}} \rightarrow a' = \sqrt{\frac{ab}{G}}; b' = \sqrt{Gab}$$

where $G = \frac{1 + \sqrt{5}}{2}$

図7 形状感性基本モデル

Fig. 7 Shape Kansei basic model.

データの例を示す。たとえば図6の中で、感性表現語「重厚な」に対応する特徴は、「紫系、および黄系以外の暗い色」であり、これをHVCで考えると、色相は（色環で考えた）黄、紫と直角に近い角度を成し、明度は低く、彩度はやや低い色になる。

感性個別モデルでは、まず式(2)の条件を満たす背景色をランダムに生成する。次に、複数の色彩を生成する場合には、すでに生成された色との間で式(1)を満たし、かつ式(2)を満たす色彩をランダムに順次生成する。

3.2.2 形状修正処理

本システムは、形状に関してメディア依存基本モデルのみを持つ。また、画像テンプレートとイラストの形状は基本感性を満たすように作られているので、ユーザによるグラフ・表・図のデータ入力・修正により各部品の形状が変化した場合のみ本処理が適用される。本モデルは第四種感性データであり、各部品の外接四角形が正方形、またはアスペクト比が黄金分割比になるように形状を変更する。これは前記形状が形態感情として「気持ち良い」と感じられるという研究結果に基づく¹¹⁾。具体的には、図7のようにグラフ、表、図の外接四角形の形状を変更する。ただし、テキストは通常、黄金分割比より横長になるのでテキスト外接四角形の変形は行わない。

3.2.3 レイアウト修正処理

本システムは、レイアウトに関してもメディア依存基本モデルのみを持つ。また、画像テンプレートは基本感性を満たすように作られているので、ユーザによるテキスト・グラフ・表・図の形状変更・追加/削除・移動により各部品のレイアウトが変化した場合のみ本処理が適用される。本モデルは第四種感性データであり、ゲシュタルト心理学をはじめとした認知心理学的見地^{12),13)}より以下に示す条件をできるだけ満たすよ

うに部品の配置を移動する。

- (1) 視認性要因：各部品は互いに重なり合わない。
- (2) 近接要因：同じ種類の部品は近くに配置する。
- (3) 類同要因：同じ種類の部品は位置を揃える。
- (4) シンメトリの法則：部品は左右対称に配置する。

ここで、部品の種類とは、各部品が持つ属性である。

以下に部品の種類の例を示す。

- (1) タイトル表示用画像テンプレート中の「表題テキスト」と「英表題テキスト」、および「氏名テキスト」と「所属テキスト」は同種である。
- (2) 図、表、グラフ表示用画像テンプレートではその本体（図/表/グラフ）部品と「本体タイトルテキスト」は同種である。

また、基本的に視認性要因である部品の重なりはユーザにより解消され、その他の要因もほぼ満たしているものとする。また、部品の数は数個であるとする。本仮定はプレゼンテーション作成支援において妥当と考えられる。

上記仮定の下で、以下の手順でレイアウト修正を行う。ただし、部品はすべてその外接四角形で近似する。

- (1) 部品に重なりが存在する場合、ユーザの要求によりSA (Simulated Annealing) 法を用いて重なりを除去する。この場合の評価関数はすべての2つの部品の組合せに対するの重なり長方形の面積の総和である。
- (2) 各部品に対し、他の部品とのX方向、およびY方向の距離のうち小さい方の値を求める。異種類間距離の最小値 D_1 が同種類間距離の最大値 D_2 より小さい場合、 D_2 の距離を持つ同種類部品どうしを D_1 まで近づける。前記処理は、近接要因を満たすまで、もしくはいずれの部品も移動できなくなるまで続ける。
- (3) 同じ種類の部品に対し、図8に示すように位置を修正する。ただし、図8には部品がY方向に並んでいる場合を示しているが、X方向に並んでいる場合もこれに準ずる移動を行う。
- (4) 同種類の部品を1つの部品と見なし、それらをX方向に画像の中央に移動する。

上記アルゴリズムはつねに最適解が得られるとは限らないが、上記仮定から準最適解が実用的な処理時間で得られる。

3.3 音楽生成処理

本処理では、入力された感性表現語から与えられた時間のBGMを生成する。

ここで、本システムがエンドユーザ向けのシステムであることを考慮し、本処理で生成するBGMは、調

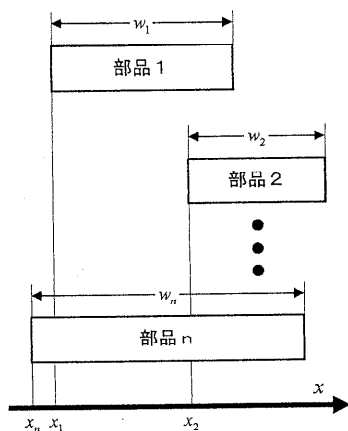


図8 レイアウト感性基本モデル

Fig. 8 Layout Kansei basic model.

$$d_i = \sum_{i=1}^n |x_i - x_i| \quad \text{where } x_i = \bar{x}_i$$

$$d_m = \sum_{i=1}^n \left| x_i + \frac{w_i}{2} - x_m \right| \quad \text{where } x_m = \bar{x}_i + \frac{w_i}{2}$$

$$d_r = \sum_{i=1}^n |x_i + w_i - x_r| \quad \text{where } x_r = \bar{x}_i + w_i$$



$$\min(d_i, d_m, d_r) = d_i \Rightarrow x'_i = x_i$$

$$\min(d_i, d_m, d_r) = d_m \Rightarrow x'_i = x_m - \frac{w_i}{2}$$

$$\min(d_i, d_m, d_r) = d_r \Rightarrow x'_i = x_r - w_i$$

性音楽であり、かつ単旋律音楽であるとする。さらに音楽を音符の記号列として取り扱う。したがって、演奏タイミングや波形データは直接は取り扱わず、演奏は音符情報に忠実に行われ、波形はあらかじめ用意された限定された数の音色情報から生成するものとする。また、各音の基本周波数は音符情報に含まれる音高情報から平均律に準拠して決定されるとする。以上の限定によりエンドユーザが使用するPCでも処理を実用的な計算時間で行うことができる。また、コンピュータ音楽における標準規格であるMIDI規格との親和性も増す。

以下に、感性データベースに格納される、音楽生成に関係する感性データ、および実際の処理を順に述べる。なおここで、感性データベースから検索された感性データ自身がメディア依存個別モデル、処理アルゴリズムがメディア依存基本モデルに相当する。

3.3.1 音楽感性データ

ラルーらのスタイル・アナリシス¹⁴⁾では音楽作品の様式を組織的に取り扱うために以下に示す5つの要素を考慮する必要があると述べている。

- (1) サウンド：音色，その組合せ，および音の強度。
- (2) ハーモニー：和声のような，音の垂直的結合の連続に関連した現象。
- (3) メロディ：音高の結合によって形成された線的輪郭。
- (4) リズム：音の持続時間の列であり，テンポ，拍子を含む。(1)～(3)と直交する概念であり，それらとの組合せで，それぞれサウンド・リズム，ハーモニック・リズム，コントゥア・リズム(音価列)という要素になる。
- (5) グロウス：上記(1)～(4)を統合する要素であ

り，音楽の階層的構造を表現する。モーション(動的過程)とシェイプ(静的形態)から構成される。

本システムでは音楽のマクロ構造を考慮しないので，グロウスは省略する。また，BGMは単旋律音楽であることから1つのメロディと伴奏から構成されると考える。一方，ハーモニーの中の和声は，メロディと伴奏双方に関係する。また，音価列はメロディに，ハーモニック・リズムは和声にそれぞれ関係する。さらに，サウンド，和声以外のハーモニー，サウンド・リズムは伴奏に関係する。

以上の考察から本システムにおける音楽の要素は以下ようになる。

- (1) メロディ：音符の列。したがって音高の列と音価の列から構成される。
- (2) 和声：時間構造(ハーモニック・リズム)を持った和音の列。
- (3) 伴奏：使用される音色の組と音の強弱情報(サウンド)，リズム楽器の演奏情報，和声楽器(和音を演奏する楽器)のリズム情報と和音中の演奏音情報から構成される。

次に，感性データの内容を述べる。本データは上記3要素を作成するために必要となる第一種，および第二種のデータであり，各感性表現語に対し，以下に示すデータを各々複数持つ。

- (1) 音価列情報：演奏時間，テンポ，拍子，音価列。
- (2) 和声素片：小節数，拍子，音階，和音列。
- (3) 伴奏パターン：拍子，メロディ音色，コード・ベース・リズム楽器演奏情報

ここで，音価列は，メロディのリズムであり，音高がC3のみのMIDI形式のデータとして，和音列は音

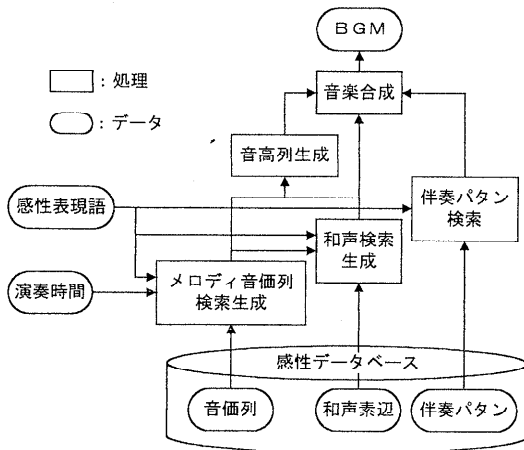


図9 BGM 生成処理
Fig.9 BGM creating process.

楽基本感性を満たす和音進行とその持続時間の組の短い列として、コード楽器の演奏情報は音色番号と1小節分を C_7 のコードで演奏する場合の MIDI データ、ベース楽器の演奏情報は音色番号とすべての音を音高 $C1$ で演奏する場合の1小節分の MIDI データ、リズム楽器の演奏情報は楽器セット（ドラムセット）番号と1小節分の MIDI データとして表現される。また、コード楽器情報は1伴奏パターンデータ中に任意数個、ベース、リズム楽器情報は0個か1個ずつ持つことができる。

3.3.2 BGM 生成処理概要

処理の概要を図9に示す。まず、与えられた感性表現語から伴奏パターンが検索される。同時に、感性表現語と演奏時間より音価列データが検索生成される（音価列検索生成）。次に、感性表現語と生成された音価列の小節数から感性データベース中の和声素片を用いて和声検索生成される（和声検索生成）。次に、和声の制約を考慮して音価列の各音符に対する音高を生成することによってメロディが合成される（音高列生成）。最後に生成されたメロディと和声、伴奏パターンを合成することにより BGM が作成される。

この中で、音高列は経験的に感性表現語への影響が他の要素に比べ少ないこと、および後述するように統計的確率分布を用いるので1つのモデルを構築するために多くの既存曲データを必要とすることから、音高列生成処理は感性基本モデルのみから構成される。また、音楽合成処理も感性表現語とは独立の処理であることから、感性基本モデルのみから構成される。その他の処理は感性依存モデルである。

以下に各処理の詳細を順に述べる。

3.3.3 メロディ音高列生成処理

メロディ音高自動生成に関する研究は数多く成されているが、生成されるメロディが調性音楽であり、あらかじめユーザがモチーフや楽曲の構造等の音楽的な情報を入力しない音高列生成手法は、ほとんどの場合マルコフ過程等の確率過程に基づく手法である¹⁵⁾。本システムでも、生成するメロディの所要時間の柔軟性、生成可能なメロディの種類多様性等を考慮して、確率過程に基づく生成を採用した。また、一般には確率分布として n 次のマルコフ過程を用いることが考えられるが、本システムの目的が BGM として聴くに耐える（基本感性を満たす）メロディを生成することであることから、以下に示す2つの経験的規則から2つの確率分布を用いる。

- (1) メロディに極端な高音や低音が含まれない。
→ 音高分布
- (2) メロディは極端に音が飛ぶことはない。
→ 音程（音高差）分布

なお、確率分布データは、エンドユーザに馴染みが深いと考えられる童謡（100曲）と邦楽ポップス（100曲）を用いて求めた。以下に具体的手順を示す。

- (1) 検索生成された音価列の各音符について(2)~(6)の処理を行う。
- (2) 音程の確率分布を1つ前の音符の音高分平行移動した確率分布 D を求める（求める音符が最初の音符の場合には、すべての音高の確率が1である分布を D とする）。
- (3) D と音高確率分布との積の分布 E を求める。
- (4) 音階と現在の音符に対応する和音を和声より求め、非音階音と非和声音の音高の確率を非常に小さな定数 ϵ とする。
- (5) E を正規化した確率分布を用いてランダムに音高を求める。
- (6) 求めた音高を現在の音符の音高とする。

3.3.4 メロディ音価列検索生成処理

音価列（リズム）の自動生成に関する研究は音高列生成に比べ少ないが、音高列と同様に確率による生成¹⁶⁾が知られている。しかし、音価列は音高列に比べ繰返しが多く、繰返しが生じない前記手法では不十分である。一方、短いリズムパターンを繰り返す方法ではメロディが単調になる。そこで、本システムでは、あらかじめ人手で創られた複数の音価列情報からランダムに検索された情報を用いてメロディ音価列を生成する手法を採用した。具体的な手順を以下に示す。

- (1) 与えられた感性表現語のキーワードを持つ音価列情報の中で与えられた所要時間 T に最も近

いデータ D を求める。もし T がデータベース上の対応するすべての音価列所要時間より大きいならば (2) を実行する。次に (3) を実行する。

- (2) 音価列繰返数 $C = T/\text{最大所要時間} + 1$ (“/”は商を求める演算子) を求め、 $T' = T \div C$ として T' に対して (1) を実行する。 D の音価列 S_D を C 回繰り返した音価列を新たに S_D とし、 D の所要時間 T_D も C 倍する。
- (3) D のテンポ $M_D = M_D \cdot T_D \div T$ とする。

求められた音価列は元の音価列を最少回繰り返したもので、かつ元のテンポに近いテンポとなる。

3.3.5 和声検索生成処理

音価列が検索生成されるとメロディの拍子と小節数 N_M が決まる。本処理ではまず、入力された感性表現語と音価列の拍子に対応する和声素片データをすべて検索する。次に検索された和声素片を、以下に示す手順で並べることにより和声を生成する。

- (1) N_M 以下の小節数を持つ和声素片を小節数で分類し、小節数 N_i の和声素片集合を B_i とする。
- (2) $N_M = \sum C_i \cdot N_i$ を満たす整数係数 C_i を動的計画法により求める。
- (3) 各 B_i の中から C_i 個の和声素片の繰返しを許してランダムに選択する。
- (4) 選択されたすべての和声素片をランダムに並べる。

3.3.6 音楽合成処理

メロディ、和声、伴奏パターンが検索または生成された後、以下に示す手順でメロディ、コード、ベース、リズムの各トラックから構成されるBGMを合成する。

- (1) メロディに伴奏パターンに含まれるメロディ音色を付与し、メロディ・トラックを合成する。
- (2) 伴奏パターンに含まれるコード楽器演奏情報を、和声に含まれる対応する和音に変更しながら小節数分繰り返すことによりコード・トラックを合成する。
- (3) 伴奏パターンに含まれるベース楽器演奏情報を、和声に含まれる対応する和音のベース音に変更しながら小節数分繰り返すことによりベース・トラックを合成する。
- (4) 伴奏パターンに含まれるリズム楽器演奏情報を小節数分繰り返すことによりリズム・トラックを合成する。

3.4 感性データベース

3.4.1 データベース構造

本システムの感性一般モデルでは、第一種～第三種

の感性データを取り扱う。また、感性表現語と感性データは多対多の関係になっている。したがって、感性データベースは、リレーショナルデータベース (RDB)、マルチメディアデータファイル (MM ファイル) 集合、データベースアクセスプログラムで構成することにより実現される。RDBは、感性データごとにテーブルを持ち、各テーブルは1つの感性表現語文字列のフィールドと1つ以上の感性データ用フィールドを持つ。感性データは、第一種の場合、MMファイルとして格納され、RDB上にはファイル名文字列が格納される。第二種の場合には、数値や記号文字列が直接RDB上に格納される。第三種の場合には、不等式やルールが文字列として格納される。MMファイル名文字列や第三種データ文字列はデータベースアクセスプログラムにより解釈され感性データに変換される。

3.4.2 キーワード付け

本システムでは、2.2.1項「感性一般モデル」で述べたキーワード付けの方法の中でエンドユーザを被験者とした心理実験による方法を採用した。心理実験は以下のステップで行う。

- (1) 20～30代の、画像と音楽に関する非専門家の被験者に対し複数の感性表現語対と感性データから生成された複数の感性物理量を提示する。
- (2) 被験者は各感性物理量についてそれぞれ、すべての表現語対について評価を行う。

評価は、キーワードとする感性表現語 X を、その反対の意味を持つ対表現語 Y とともに提示し、「非常に X である」「やや X である」「どちらでもない」「やや Y である」「非常に Y である」の5段階で行う。

本システムにおいてキーワード付けされる感性データをまとめると以下ようになる。

- (1) 第一種感性データ：画像テンプレートの背景、音価列、コード・ベース・リズム楽器演奏情報
- (2) 第二種感性データ：テンポ、拍子、音階、和音列
- (3) 第三種感性データ：色彩制約不等式

ここで、第一種感性データは、感性物理量として直接提示する。第二種感性データの場合は、できるだけ他の要素が感性的に中立であると考えられる感性物理量を、該当するデータを用いて生成する。具体的には、テンポと拍子の場合には対応するクリック音を、音階と和音列の場合には中音域である $C3 \sim B4$ をベースとするスケールと和音をそれぞれ用いる。第三種感性データの場合は、該当するデータを用いて感性物理量を複数生成し、各々に対して心理実験を行い、得られた結果すべてをその感性データの結果と見なす。

得られた評価は $-2 \sim 2$ の整数で数値化し、各感性

表1 感性表現語対
Table 1 Descriptive words.

1	アダルトな	ヤングな
2	エレガントな	ラフな
3	おだやかな	激しい
4	シンプルな	複雑な
5	ソフトな	ハードな
6	フォーマルな	カジュアルな
7	高級な	庶民的な
8	知的な	野性的な
9	落ちついた	活発な

データと感性表現語に対し、実験結果の被験者に関する平均と標準偏差を求め、 $|\text{平均値} - C \cdot \text{標準偏差}| > 0.5$ である場合に平均値の符号の表現語をキーワードとして付与する。なお、ある感性表現語 X が前記式を満たす場合を「感性 X を持つ」ということにする。ここで $C \approx 0.84$ であり、基準系の累積正規分布関数 $F(z) = 0.8$ となる z の値に相当する。

4. 実験システムの評価

最後に、本実験システムが与えられた感性表現語を満たすマルチメディア作品を合成できるかどうかを評価することを主な目的として心理実験を行ったので、その方法と結果を示す。

4.1 評価実験方法

- (1) 被験者：20～30代の画像と音楽に関する非専門家の男女14名（キーワード付けの際の心理実験における被験者とは別人）。
- (2) 刺激：本システムにより、感性表現語9語から生成された静止画像、および同じく約10秒のBGM各6種類。
- (3) 手続き：各被験者に対し、静止画像のみとBGMのみの提示、および静止画像とBGMのすべての組合せのマルチメディア提示がランダムな順序で行われる。画像はすべてPCの画面上に表示され、BGMはPCとシンセサイザーにより演奏され、ヘッドホンを通して流される。被験者は、表1に示す感性表現語対に対し、「非常にX、ややX、どちらでもない、ややY、非常にY」の5段階評価を、提示されたすべての画像、BGM、およびそれらの組合せについて行う。また、BGMに関しては、さらに「音楽らしい」、「音楽らしくない」という語対の評価も行う。被験者は評価の間、任意回刺激を視聴することができる。また、全表現語対の評価が終了すると、被験者の合図で次の刺激が提示される。

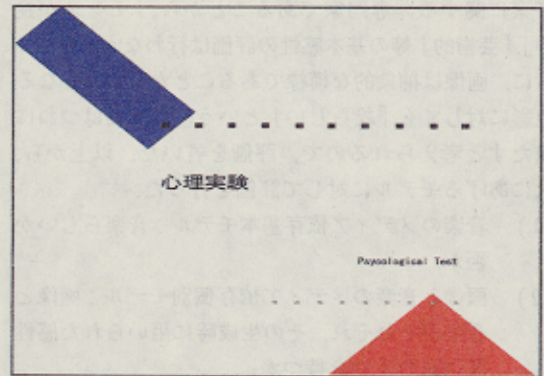


図10 提示画像の例

Fig. 10 Example of given picture.

なお、画像は、抽象的な模様が描かれている背景とテキストのみのものとした。これは具体的な物（海、自動車等）が描かれていたり、具体的な意味を持つイラストが含まれていることにより、その内容でイメージが固定されることを避けるためである。提示された画像の例を図10に示す。

また、感性表現語対は、視覚に関する様々な感性情報処理の心理実験例^{17),18)}で用いられた形容語対から、以下の点を考慮して表1に示したものに絞った。

- (1) 実験した感性物理量特有と考えられる表現語対は削除する（ファッションに関する実験における「ドレスリーな ⇔ ドレスリーでない」等）。
- (2) 主観的な嗜好を表す表現語対は削除する（「好きな ⇔ 嫌いな」等）。
- (3) 同義、もしくは類似の意味を持つと考えられる表現語対は1つにまとめる（「ソフトな ⇔ ハードな」と「柔らかい ⇔ 堅い」等）。
- (4) 主成分分析により1つの主成分にまとめられた表現語対は1つにまとめる。

ここで(2)の制約を考慮したのは、嗜好にかかわる表現語に対する感じ方は個人差が大きく、本実験システムのように多くの視聴者に対するプレゼンテーションを作成する場合には感性表現語として採用すべきではないと考えられるからである。

なお、本評価システムでは、感性物理量の提示から評価データ収集まではMS-WINDOWS上にVisual BASICを用いて開発されたプログラムですべて自動的に行った。また、データの解析は、MS-Accessを用いて行った。

4.2 評価結果と考察

4.2.1 単メディアの評価

まず、提示した画像とBGM各々のメディア依存モデルの評価結果を以下に示す。ここで、被験者が画像、

音楽に関する非専門家であることから、「センスが良い」、「芸術的」等の基本感性の評価は行わなかった。さらに、画像は抽象的な模様であることから、いかなる色彩に対しても「絵らしい」という基本感性はつねに満たすと考えられるので、評価を省いた。以上から、次にあげるモデルに対して評価を行った。

- (1) 音楽のメディア依存基本モデル：音楽らしいか否か。
- (2) 画像と音楽のメディア依存個別モデル：画像と音楽がそれぞれ、その生成時に用いられた感性表現語の感性を持つか。

まず(1)に関しては、「非常に音楽らしい」を5、「非常に音楽らしくない」を1とすると、被験者全員が提示されたすべてのBGMに対し3以上の値を回答した。この結果より、音楽基本モデル(BGM生成アルゴリズム)が、おおむね音楽らしい音楽を生成できるという意味で、妥当であると考えられる。

次に(2)に関しては、感性表現語Xの入力に対し生成された画像とBGMはすべて、感性Xを持っていた。この結果より、画像と音楽に関するメディア個別モデルが、入力された感性表現語に対して、おおむね多くの人がその表現語の感性を持つような画像もしくは音楽を生成できると考えられる。

4.2.2 マルチメディアの評価

次に、画像とBGMのマルチメディア提示により、マルチメディア作品としての感性はどのようになるかを評価した結果を以下に示す。

専門家が画像とBGMの調和を考慮して制作したマルチメディア作品の場合には、各メディアに共通ないくつかの形容語に対して、双方のメディアの感性に強い相関があり、2つのメディアが作品全体の感性を強調することが知られている¹⁹⁾。一方本論文では、メディア感性独立仮定に基づき感性合成処理をメディアごとの部分問題に分けることによって単純化している。ところが、本システムのように、それぞれ独立に生成されたメディアのマルチメディア提示の場合におけるメディア間の相互作用は、一般には知られていない。

そこで、本システムで生成された画像とBGMのすべての組合せの評価結果と、それぞれの単メディアでの評価結果との関係性を評価した。なお、同評価で用いたメディア情報は画像、BGM各々6種類で、すべて互いに異なる感性データから合成されたものである。

単メディア提示とマルチメディア提示の評価結果の関係を表2に、また表2で用いられている記号の意味を表3に示す。表3においてA、Bはそれぞれ画像(“P”)またはBGM(“M”)であり、(X, Y, 0)

表2 評価結果
Table 2 Evaluation results.

メディア		感性表現語								
画像	BGM	1	2	3	4	5	6	7	8	9
1	1	M△	M△	M△	□	□	M△	□	M△	M△
1	2	□	□	M◎	□	□	M◎	M△	□	M◎
1	3	M△	M△	M△	□	□	M△	M△	□	M◎
1	4	□	□	M◎	□	M◎	□	□	M△	M△
1	5	□	M△	M◎	□	M◎	M△	□	□	M△
1	6	M△	M△	□	□	□	M△	□	□	M◎
2	1	M△	M△	M◎	□	□	M△	□	M△	M△
2	2	□	?	M◎	□	□	M△	M△	□	M◎
2	3	M△	M△	M△	□	□	M△	M△	□	M◎
2	4	□	□	M◎	?	M◎	□	□	M△	M◎
2	5	□	M△	M◎	?	M△	M△	□	□	M△
2	6	M△	M△	□	□	□	M△	□	□	M◎
3	1	M◎	M△	M◎	P△	□	M◎	□	M◎	M◎
3	2	□	□	M△	P△	□	M△	M△	□	M◎
3	3	M△	M△	M△	P△	□	M△	M△	□	M△
3	4	□	□	M◎	P◎	M◎	□	□	M△	M◎
3	5	□	M◎	M◎	P△	M◎	M△	□	□	M◎
3	6	M△	M◎	?	P△	□	M△	□	□	M△
4	1	M△	M◎	□	P◎	□	M△	□	M△	○
4	2	□	□	■	P△	□	M△	M△	□	■
4	3	M△	M△	■	P△	□	M△	M△	□	■
4	4	□	□	○	P◎	M◎	□	□	M△	○
4	5	□	M△	○	P◎	M◎	M△	□	□	○
4	6	M△	M△	P△	P△	□	M△	□	□	■
5	1	M△	M△	○	P△	□	M△	□	M◎	M◎
5	2	□	□	■	P△	□	M◎	M△	□	M◎
5	3	M△	M△	■	P△	□	M△	M△	□	M△
5	4	□	□	○	P◎	M△	□	□	M△	M◎
5	5	□	M△	○	P◎	M△	M△	□	□	M◎
5	6	M△	M△	P△	P△	□	M△	□	□	M△
6	1	M△	M△	M◎	P△	□	M△	□	M△	M◎
6	2	□	□	M△	P△	□	M◎	M△	□	M△
6	3	M△	M△	M△	P△	□	M△	M△	□	M△
6	4	□	□	M◎	P◎	M△	□	□	M△	M◎
6	5	□	M△	M◎	P△	M△	M△	□	□	M◎
6	6	M△	M△	□	P△	□	M△	□	□	M△

表3 評価結果における記号
Table 3 Notation for evaluation results.

Single Media		Multimedia		
A	B	X	0	Y
X	X	○	×	×
	0	A◎	A△	×
	Y	A	■	B
0	0	?	□	?

はそれぞれ、該当するメディアが感性Xを持つ場合、感性表現語Xの対となる表現語の感性Yを持つ場合、どちらも持たない場合を表す。次に、表2で得られた組合せの数をまとめたものを表4に示す。ここで、単メディアにおけるA、Bは表2と同様に画像またはBGMである。

表4の第1行は、画像、BGM両方が単独提示の際に同じ感性を持つ場合で、9通りの組合せすべてにお

表4 評価結果要約
Table 4 Evaluation summary.

Single Media		Multimedia		
A	B	X	0	Y
X	X	9	0	0
X	0	55	129	0
X	Y	0	7	
0	0	4	120	

(A, B): (画像, BGM) or (BGM, 画像)
(X, Y): (表現語 X を持つ, 表現語 Y を持つ) or
(表現語 Y を持つ, 表現語 X を持つ)

いて、マルチメディア提示でも同じ感性を持った。このことから、本システムの基本的考え方であるメディア感性独立仮定は成立すると考えられる。

次に、表4の第2行は、一方のメディアのみ感性Xを持つ場合で、いずれの組合せでもマルチメディア提示の場合において同じ感性Xを持つか、X, Yのいずれの感性も持たないかのいずれかであり、対感性Yを持つことはなかった。このことは、感性を持つ度合を適当に数値化すると、メディア間にある種の感性加法性が成立することを示唆する。すなわち、Aの感性Xに対する度合が大きい場合にマルチメディア提示でXを持つと考えることができる。

次に、表4の第3行は、双方のメディアが逆の感性を持つ場合で、同時提示においてそのどちらの感性も持たなかった。これは、互いのメディアが感性を打ち消し合っていると考えられる。

最後に、表4の第4行は、双方のメディアとも、感性Xも対感性Yも持たない場合で、わずかの例外を除き(124通り中4通り)、マルチメディア提示の場合、どちらの感性も持たなかった。これもメディア間の感性加法性を示唆する。すなわち、双方のメディアが弱い感性Xを持っていて、通常は、マルチメディア提示でそれらを加えても感性Xを持つに至らないが、双方とも比較的強い感性Xを持っている場合にのみ、マルチメディア提示により感性Xを持つと考えることができる。

5. まとめ

本論文では、エンドユーザが入力した感性(雰囲気, イメージ, 印象)に基づいてマルチメディア作品に色彩・音楽等を生成・付与するための感性合成処理の基本コンセプトを提案した。また、そのモデルに基づいた実験システムである、複数のカラー静止画像から構成されるBGMをともなったプレゼンテーションの作成支援を行うマルチメディア・プレゼンテーション作成支援システムの具体的な処理方式を示した。最後に

システムが合成した画像やBGMが入力された感性を持つかどうか、マルチメディアプレゼンテーションとして、それらをあわせて提示した場合に相互作用があるかどうかに関して評価を行い、本システムの有効性を示した。

本実験システムでは、感性合成処理モデルが完備している感性物理量は色彩とBGMのみであった。今後は合成する感性物理量の種類の増加(特に動画像)と、各メディアごとの感性合成処理の実現手法の精緻化を行いたいと考えている。

また、4.2.2項「マルチメディアの評価」において、感性が加法性の成立する量として表現される可能性が示唆された。この場合の数量は、実数とは限らずFuzzy数、区間値、分布自身等が考えられるが、もし感性の数量化が可能となれば、たとえば画像のみで表現できない感性も、BGMを付与することにより全体としてその感性を持たせることが可能となる。今後、感性の数量化、メディア間の感性加法性に関し、検討したいと考えている。

参考文献

- 1) 井口征士: 感性情報処理が目指すもの, 情報処理, Vol.35, No.9, pp.792-798 (1994).
- 2) 佐藤宏介, 加藤博一: 計算機による絵画の描画技法とイメージ語との自動対応付け, 感性情報処理の情報学・心理学的研究(文部省科学研究費補助金重点領域研究平成6年度成果報告書), pp.101-104 (1995).
- 3) 才脇直樹, 片寄晴弘, 井口征士: 感性情報を利用した作・編曲支援システム, 電子情報通信学会全国大会, p.6-213 (1990).
- 4) Nakamura, J., Kaku T., Hyun, K., Noma, T. and Yoshida, S.: Automatic Background Music Generation Based on Actor's Mood and Motions, *The Journal of Visualization and Computer Animation*, Vol.5, pp.247-264 (1994).
- 5) 北原義典, 長谷川隆: ユーザのマルチメディア感性演出支援方式の提案, 信学技報, HC94-95, pp.63-70 (1995).
- 6) 長谷川隆, 北原義典: マルチメディア感性演出支援システムにおけるBGM自動生成方法の提案, 第11回ヒューマン・インターフェイス・シンポジウム論文集, pp.611-619 (1995).
- 7) 日本色彩学会: 色彩科学ハンドブック, p.1494, 東大出版会, 東京 (1980).
- 8) 近藤恒夫: 色彩学, p.154, 理工図書, 東京 (1992).
- 9) 岡部慶三: 色のイメージ事典, p.205, 同朋社, 京都 (1991).
- 10) 小林重順: 新・カラーイメージ事典, p.175, 講談社, 東京 (1993).

- 11) 中村 秀:心理学, p.366, 朝倉書店, 東京 (1957).
- 12) 和田陽平, 大山 正, 今井省吾:感覚+知覚ハンドブック, p.1030, 誠信書房 (1969).
- 13) 仲谷洋平, 藤本浩一:美と造形の心理学, p.229, 北大路書房, 京都 (1993).
- 14) ヤン・ラルー, 大宮眞琴:スタイル・アナリシス, 1, p.500, 音楽乃友社, 東京 (1988).
- 15) 西村恕彦:コンピュータと音楽, p.272, 共立出版, 東京 (1987).
- 16) 田中康裕, 渡邊哲史, 高田正之, 小谷善行:音程・リズムの遷移確率と主題・対位の構造に基づく作曲, 第38回情報処理学会全国大会論文集, pp.1867-1868 (1989).
- 17) 井口征士:感性情報処理, p.176, オーム社, 東京 (1994).
- 18) 長町三生:感性工学, p.138, 海文堂出版, 東京 (1989).
- 19) 若宮眞一郎:オーディオ・ヴィジュアル・メディアを通しての情報伝達における視覚と聴覚の相互作用に及ぼす音と映像の調和の影響, 日本音響学会誌, Vol.48, No.9, pp.31-39 (1992).

(平成8年9月20日受付)

(平成9年6月3日採録)



長谷川 隆 (正会員)

1962年生. 1986年東京大学工学部機械工学科卒業. 1988年同大学院情報工学専攻修士課程修了. 同年(株)日立製作所入社. 現在,(株)日立製作所中央研究所マルチメディアシステム研究部研究員. 主たる研究テーマは感性情報処理, 音楽情報処理, マルチメディア情報検索, 人工知能学会会員.



北原 義典 (正会員)

1955年生. 1979年広島大学総合科学部情報行動科学コース卒業. 1981年同大学院修士課程修了. (株)日立製作所入社. 1986~1989年(株)ATR視聴覚機構研究所勤務. 現在,(株)日立製作所中央研究所マルチメディアシステム研究部主任研究員. 工学博士. 主たる研究テーマは音声合成, 音声認識, ユーザインタフェース, 感性情報処理の研究. 電子情報通信学会, 日本音響学会, 計測自動制御学会HI部会各会員.