

色彩の印象からの楽曲自動生成方式の実現

中西 崇文[†] 芳村 亮^{††} 北川 高嗣^{††}

本稿では、色彩の持つ印象から音楽を生成する楽曲自動生成の実現方式について述べる。本方式は、色彩、楽曲に関する心理学者による研究を用い、ユーザから与えられた色彩情報からその印象を言葉として抽出し、さらに楽曲の印象を表す言葉とを計量し、その印象と合致した楽曲を生成する方式である。本方式が実現されることで、画像と音楽との異種のメディア間の自由な連携が可能となり、新しい複合メディアのコミュニケーションの場を提供できる。さらにユーザの意図とする感情に合致した楽曲メディアデータが自動生成され、それによりコンピュータ環境の中で人間の感性を増幅し、コンピュータが人間と密着した感性を共有することが可能になると考えられる。

An Implementation Method of Automatic Composition of Music from Color Impression

TAKAFUMI NAKANISHI,[†] RYO YOSHIMURA^{††}
and TAKASHI KITAGAWA ^{††}

In this paper, we represent an automatic generation method for music data corresponding to impression of color information. This method can generate a music data by measuring relations between impression words extracted from color information and 8 categories of impression words given by a psychological research. This method can realize connections between heterogeneous media such as image and music. This method can realize communication places for compound media. When this method can realize, computers may be able to share and amplify Kansei with human each other.

1. はじめに

現在、ネットワークの高速化および広域化、コンピュータの高速化および廉価化により、コンピュータが我々の生活に密着し、接触することが多くなった。将来、コンピュータはさらに様々な環境に偏在し、人間の生活に欠かせないパートナーとなっていくだろうと予想される¹⁾。このような環境において、人間のマルチモーダル性を生かし、人間とコンピュータがその場に合った感性を共有し、増幅しあうシステムの実現が重要になると考えられる。

我々はこれまで、人間の心理的要素が反映されたメタデータを対象としたメディアデータ検索機能として、

楽曲メディアデータを対象とした人間の感性に合致した意味的連想検索²⁾を実現している。これは、楽曲メディアデータが人に与える印象を心理学の研究成果に基づいて、言葉によって表現されるメタデータとして自動抽出する方式^{3),4)}と言葉と言葉の関係の計量が可能な意味の数学モデル文献^{5)~7)}を適用し、さらに人間の物理的な刺激と感じ方の関係に基づく演算子を導入することにより実現されている。

また、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についても、色彩の印象についての統計データである、カラーイメージスケール⁸⁾を用いて印象語を抽出する方式として、文献^{9),10)}で述べている。

これら方式により、人間の感性に合致した、メディアデータの検索が可能となった。しかし、人間のマルチモーダル性を生かし、人間とコンピュータがその場に合った感性の共有や増幅を実現するためには、メディア間の自由な写像操作ができ、表現できるコミュニケーションの場の実現が重要となると考えられる。

本稿では、色彩情報と楽曲に着目し、色彩の持つ印象から音楽を生成する楽曲自動生成の実現方式について述べる。本方式は、ユーザから与えられた色彩情報

[†] 独立行政法人 情報通信研究機構
National Institute of Information and Communications
Technology
e-mail : takafumi@nict.go.jp

^{††} 筑波大学大学院システム情報工学研究科
Graduate School of Systems and Information Engineering,
University of Tsukuba
e-mail : yoshimura@mma.cs.tsukuba.ac.jp,
takashi@cs.tsukuba.ac.jp

から、文献^{9),10)}を用いて印象語を抽出し、その印象語と Hevner の研究^{11)~14)}による楽曲の印象を表す8つの印象語群とを意味の数学モデル^{5)~7)}を用いて関連を計量し、さらに、Hevner の研究^{11)~14)}を用いて、楽曲の特徴を抽出することにより、ユーザが発行した問い合わせに合致した楽曲メディアデータを生成するものである。

これまで自動作曲方式として、様々な研究がされてきた。例えば、提示された限定された語、ジャンルをユーザに指定させ自動作曲する方式(例えば¹⁵⁾)、学習データとして既存の楽曲データを準備し、それらから得られたルールから楽曲を作曲する方式(例えば¹⁶⁾)などが提案されている。

本方式は、他のメディアデータの印象を楽曲にし、楽曲と異種メディアの連携を目的としており、従来方式とは目的が異なる。但し、従来方式と組み合わせ、より自然な楽曲生成が可能になると考えられる。

本方式が実現されることにより、利用者の意図とする感情に合致した楽曲メディアデータの自動生成が実現されることで、音楽を通じた情報の新しい価値を生み出すことが可能となる。また、本方式によりコンピュータ環境の中で人間の感性を増幅することできるシステムとなり、コンピュータが人間の生活を豊かにし、人間と密着した感性を共有するパートナーとして新しい関係を築く第一歩となると考えられる。

2. 意味の数学モデルの概要

本節では、言葉と言葉の関係の計量を実現する意味の数学モデルの概要を示す。詳細は、文献^{5)~7)}に述べられている。

(1) メタデータ空間 MDS の設定

検索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。

(2) メディアデータのメタデータをメタデータ空間 MDS へ写像

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。

(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。

(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. 色彩情報を対象としたメタデータ自動抽出方式

本節では、デジタル化された色彩情報を対象として、印象を抽出して、言葉によって表現されるメタデータを自動的に抽出する方式について述べる。詳細は、文献^{9),10)}に述べられている。

本方式は、色彩とその印象との関連を記述した統計データであるカラーイメージスケール⁶⁾を用いる。カラーイメージスケールは、130色とは10色相、12色調で表現される有彩色120色、および10階調で表される無彩色10色の計130色の基本色と180の印象語の関連性を5段階で表現している。

具体的な流れは以下の通りである。

(1) 色印象行列 C の作成

基本色130個の各々の色彩を、静止画像の定義に必要な十分な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。

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

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。ここで色印象ベクトル c_k を行ベクトルとし(c_1, c_2, \dots, c_{130})より構成される色印象行列 C とする。色印象行列 C は色と印象語との相関の強さを表す。

この色印象行列をカラーイメージスケールを用いて、有彩色120色、及び無彩色10色の基本

色 130 色とその印象を表す 180 語により色印象行列 C とする。

(2) 色彩情報の抽出

静止画像から色彩情報が抽出し、その色彩情報は静止画像全体における基本色 130 色の占める割合で構成される画像色彩ベクトル r によって表現される。画像色彩ベクトルを次に示す。

$$r = (r_1, r_2, \dots, r_{130})^T. \quad (2)$$

(3) 静止画像のメタデータ抽出

色印象行列 C 、及び画像色彩ベクトル r を用いて、静止画像メタデータベクトル z の抽出を行う。 z は、色印象ベクトル c_k に用いられる 180 個の印象語と同一の印象語で特徴付けられるベクトルである。

静止画像メタデータベクトル z を次のように表す。

$$z = Cr \quad (3)$$

静止画像メタデータベクトル z を以下のように定義する。

$$z = (z_1, z_2, \dots, z_{180})^T. \quad (4)$$

これが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により出力される重みつき印象語によるメタデータである。

4. 色彩の印象からの楽曲自動生成方式

本節では、色彩の印象からの楽曲自動生成の実現方式について述べる。本節では、先ず、4.1 節で、任意の印象語を対象とした楽曲自動生成実現方式について述べ、さらに 4.2 節で、本提案方式である色彩の印象からの楽曲自動生成の実現方式について述べる。

4.1 任意の印象語を対象とした楽曲自動生成方式

本節では、入力された言葉から印象のメタデータを抽出し、それに合致した楽曲を自動的に生成する、楽曲自動生成方式について述べる。

4.1.1 任意の印象語を対象とした楽曲自動生成方式全体の概要

この方式の全体の流れは、次のような手順により実現される。

- Step1: 楽曲印象語群ベクトルの出力
意味の数学モデルにより、入力された言葉(コンテキスト)と Hevner の 8 つの印象語群それぞれの相関を示す楽曲印象語群ベクトルが出力される。詳細は 4.1.3.2 で述べる。
- Step2: 楽曲構造要素ベクトル生成機能
まず楽曲印象語群ベクトルを正規化する。正規化された楽曲印象語群ベクトルと、8 つの印象語群

		c6	bright cheerful gay happy joyous merry		c5	delicate fanciful graceful humorous light playful quaint sprightly whimsical		c4	calm leisurely lyrical quiet satisfying serene soothing tranquil
	c7	agitated dramatic exciting exhilarated impetuous passionate restless sensational soaring triumphant							
	c8	emphatic exalting majestic marial ponderous robust vigorous							
	c1	awe-inspiring dignified lofty sacred serious sober solemn spiritual							
			c2	dark depressing doleful frustrated gloomy heavy melancholy mournful pathetic sad tragic					
						c3	dreamy longing plaintive sentimental tender yearning yielding		

図 1 Hevner による 8 つの印象語群

表 1 各印象語群に対する楽曲構造要素の相対重要性の表

印象語群名	c1	c2	c3	c4	c5	c6	c7	c8
構造要素名								
key	長調 4	短調 12	短調 20	長調 3	長調 21	長調 24	—	—
tempo	遅い 14	遅い 12	遅い 16	速い 20	速い 6	速い 20	速い 21	速い 6
pitch	低い 10	低い 19	高い 6	高い 8	高い 16	高い 6	低い 9	低い 13
rhythm	固定 18	固定 3	流動 9	流動 2	固定 8	流動 10	固定 2	固定 10
hamony	単純 3	複雑 7	単純 4	単純 10	単純 12	単純 16	複雑 14	複雑 8
melody	上昇 4	—	—	上昇 3	下降 3	—	下降 7	下降 8

と 6 つの楽曲構造要素との相関を示す変換行列により、6 つの楽曲構造要素によって特徴付けられた楽曲構造要素ベクトルが出力される。詳細は 4.1.3.3 で述べる。

- Step3: 楽曲特徴値生成機能
楽曲構造要素ベクトルの値をもとにして、生成する楽曲の特徴値を決定する。詳細は 4.1.3.4 で述べる。
- Step4: 楽曲生成・出力機能
Step3 で決定した特徴値をもとに楽曲を生成する。詳細は 4.1.3.5 で述べる。

4.1.2 Hevner の研究

Hevner の研究^{11)~13)}では、楽曲構造要素として調性 (key) ・テンポ (tempo) ・音高 (pitch) ・リズム (rhythm) ・和声 (harmony) ・旋律 (melody) の 6 つを挙げている。Hevner は、この 6 つの楽曲構造要素と 8 つの印象語群 (図 1) によって表現される印象との相関関係を調べた (表 1)。8 つの印象語群は、印象語間で類似性があるものをまとめて一つの印象語群を作り、さらに印象語群間で類似性があるものを隣接するよう、円形に配置している。

4.1.3 楽曲自動生成の実現方式

本節では、提案方式の各機能の詳細について示す。

4.1.3.1 楽曲の仕様

ここでは、本方式において生成する楽曲の仕様について示す。

楽曲の拍子・長さ 本方式では本格的な曲を生成する必要はなく、入力された言葉の持つ意味や感情が短いフレーズに反映されればよいので、生成する楽曲は4/4拍子、長さは4小節とする。

和音 一つの楽曲に使用しうる和音は非常に多いので制限を与える。使用する和音はI度からVII度までの7つで、和音の構成音として使用する音はその調の音階に沿ったものとする。長調ではI, II_m, III, IV, V7, VI_m, VII_m7⁽⁻⁵⁾、短調ではIm, II_m7⁽⁻⁵⁾, III, IV_m, V7, VI, VIIである。また、和音の長さは四分音符と八分音符の2種類のみとする。

旋律 旋律に使用する音はその調の音階に沿ったものとする。また、Hevnerの研究では旋律音の長さやリズムには言及していないので、旋律音の長さは四分音符のみとする。

4.1.3.2 楽曲印象語群ベクトルの出力 (Step1の詳細)

2章に示した意味の数学モデルを用いて、入力に与えられた任意の言葉と、Hevnerによる8つの各印象語群との相関を求める。さらに、その各相関値を、1) 相関値のうち最大のものを1に写像、2) 相関値のうち最小のものを-1に写像することにより正規化し、楽曲印象語群ベクトル $c = (v_{c_1}, v_{c_2}, \dots, v_{c_8})$ (v_{c_i} は語群 c_i への重み) を生成する。なお、この正規化の方式は、いくつかの予備実験の結果をもとに定めたものである。また、ここでの定式化においては上記に限定しない。

4.1.3.3 楽曲構造要素ベクトルの出力 (Step2の詳細)

表1で示した、Hevnerによる各印象語群に対する楽曲構造要素の相対重要性の表に正負の符号をつけ、行ごとに正規化したものを変換行列 T とする(図2)。楽曲印象語群ベクトルを入力として、変換行列 T を作用させることにより、6つの楽曲構造要素への重みを表す楽曲構造要素ベクトル $f = (key, tempo, pitch, rhythm, harmony, melody)^T$ を生成する。

$$f = Tc \quad (5)$$

なお、変換行列 T の規定方式は予備実験によるもので、ここでの定式化においては上記に限定しない。

	印象語群							
	c1	c2	c3	c4	c5	c6	c7	c8
key	0.0476	-0.1429	-0.2381	0.0357	0.2500	0.2857	0.0000	0.0000
tempo	-0.1217	-0.1043	-0.1391	-0.1739	0.0522	0.1739	0.1826	0.0522
pitch	-0.1149	-0.2184	0.0690	0.0920	0.1839	0.0690	-0.1034	-0.1494
rhythm	0.2903	0.0484	-0.1452	-0.0323	0.1290	-0.1613	0.0323	0.1613
harmony	0.0405	-0.0946	0.0541	0.1351	0.1622	0.2162	-0.1892	-0.1081
melody	0.1600	0.0000	0.0000	0.1200	-0.1200	0.0000	-0.2800	-0.3200

図2 変換行列 T

表2 楽曲の特徴値

kn	楽曲の調性
tem	テンポ
$mnap$	旋律の音高の平均
$unac$	四分音符の和音の総演奏時間
ac	八分音符の和音の総演奏時間
tc	三和音の総演奏時間
oc	三和音以外の和音の総演奏時間
um	上昇する旋律音の総演奏時間
dm	下降する旋律音の総演奏時間
lm	水平な旋律音の総演奏時間

表3 各調の調性度を示す表

長調の調性度									
C	G	D	A	E	B	F#	F	Bb	Eb
0.9375	1.0000	0.7500	0.7500	0.4375	0.6250	0.5000	0.6875	0.1250	0.4375
A _b	D _b								
0.3125	0.0000								
短調の調性度									
a	e	b	f#	c#	d	g	c	f	bb
0.0000	-0.6250	-0.6250	-0.6250	-0.8750	-0.7500	-0.5000	-1.0000	-0.7500	-0.8750

4.1.3.4 楽曲の特徴値の出力 (Step3の詳細)

出力されたMFVの値をもとに、生成する楽曲の特徴値 $kn, tem, mnap, unac, ac, tc, oc, um, dm, lm$ を決定する。調性とテンポはここで定める値を楽曲に使用し、音高・リズム・和声・旋律については、ここで定める特徴値をもとに次の行程でより具体的に決定する。 kn の決定方法については文献¹⁴⁾に、それ以外の各特徴値の決定方法は文献³⁾中の式にもとづいて定めている。表2に楽曲の特徴値の意味を示す。

調性 kn の決定

調性の値から、生成する楽曲の調 kn を定める。梅本の研究¹⁴⁾による表(表3)から特徴値に近いものを探し、それを生成する楽曲の調性とする。候補となる調性が複数ある場合は、その中からひとつをランダムに選択する。

テンポ tem の決定

以下の式により、生成する楽曲のテンポ tem を定める。

$$tem = \begin{cases} tstan + (tmax - tstan)tempo, & tempo \geq 0 \\ tstan + (tstan - tmin)tempo, & tempo < 0 \end{cases} \quad (6)$$

但し、 $tempo$ はMFVにおけるテンポの値、 $tmax$ は

テンポの最大値(ここでは 184), $tmin$ はテンポの最小値(ここでは 44), $tstan$ はテンポの標準値(ここでは 88)である。また, tem , $tmax$, $tmin$, $tstan$ の単位は拍/分である。

音高 $mnap$ の決定

以下の式により, 生成する楽曲の旋律音の高さの平均 $mnap$ を定める。

$$mnap = \begin{cases} pstand + (pmax - pstand)pitch, & pitch \geq 0 \\ pstand + (pstand - pmin)pitch, & pitch < 0 \end{cases} \quad (7)$$

但し, $pitch$ は MFV における音高の値, $pmax$ は音高の最大値(ここでは 88 鍵ピアノの最高音にあたる MIDI ノートナンバー) $pmin$ は音高の最小値(88 鍵ピアノの最低音にあたる MIDI ノートナンバー), $pstand$ は音高の標準値(88 鍵ピアノの中央 A 音にあたる MIDI ノートナンバー)である。

リズムの特徴 $unac, ac$ の決定

以下の式により, 生成する楽曲のリズムの特徴 $unac, ac$ を定める。

$$rhythm = \frac{unac - ac}{unac + ac} \quad (8)$$

但し, $rhythm$ は MFV におけるリズムの値である。 $unac + ac$ は曲全体の演奏時間に相当するため, これを 1 とすると $unac, ac$ を一意に定めることができる。

和声の特徴 tc, oc の決定

和声の値から, 生成する楽曲に使用する和声の配分を定める。

$$harmony = \frac{tc - oc}{tc + oc + nc} \quad (9)$$

但し, $harmony$ は MFV における和声の値, nc は和声のない時間の合計(今回は 0 に規定)である。このとき, リズムの場合と同様に tc, oc は一意に定まる。

旋律の特徴 um, dm, lm の決定

以下の式により, 生成する楽曲の旋律の特徴を定める。

$$melody = \frac{um - dm}{um + dm + lm} \quad (10)$$

但し, $melody$ は MFV における旋律の値, である。文献¹⁴⁾より, 次の旋律音が同じ音となる割合は 26% なので, lm が 0.26 に近い値となるように乱数を用いて定める。但し, $lm > 1 - |melody|$ となったときは, $lm \leq |melody|$ となるように乱数を用いて定める。 lm が定まると, リズムや和声の場合と同様に um, dm は一意に定まる。

4.1.3.5 楽曲の生成・出力 (Step4 の詳細)

楽曲の特徴値に沿って, 新たな楽曲を生成する。以

下に楽曲データの生成手順を示す。

和音による伴奏の生成

決定した特徴値のうち, $kn, mnap, unac, ac, tc, oc$ を用いて, 和音を構成する音を以下の手順により生成する。

- (1) 使用する和音を kn によって決定する。
- (2) $unac$ と ac の値から四分音符の和音と八分音符の和音の個数・配列を決定する。 $unac$ の値を四分音符の和音の個数, ac を八分音符の個数に対応させる。
- (3) tc, oc の値をもとに各和音を「三和音」と「それ以外の和音」に振り分ける。但し, 最後の 2 拍または 2.5 拍は必ず「三和音」とする。
- (4) 各和音の具体的な和音名を決定する。「三和音」のものは I 度, IV 度, V 度から, 「それ以外の和音」は II 度, III 度, VI 度, VII 度からランダムに選択する。但し, 最後の 2 拍または 2.5 拍 (3. において「三和音」に固定した部分) は必ず I 度の和音とする。

旋律の生成

特徴値のうち $kn, mnap, um, dm, lm$ を用いて旋律音を以下の手順により生成する。

- (1) kn より使用する音階を決定する。長調ならば長音階, 短調ならば短音階(和声的短音階)とする。
- (2) um, dm, lm の割合に応じて上昇音, 下降音, 水平音の個数を決定する。
- (3) 旋律の大まかな流れを決定するため, 各旋律音を上昇音, 下降音, 水平音のいずれかに振り分ける。但し, 最後の音については, 後ろの音が存在しないので水平音と定める。
- (4) 旋律を具体的に決定する。最後の音をその調の主音に固定して後ろから順に決定していく。ここで, 文献¹⁴⁾には, 隣接する二つの旋律音の音程とその出現する割合についての研究が示されている。この研究では, 歌曲の旋律について, 1 度から 8 度までの音程の出現頻度を調べている(表 4)。上昇音と下降音ではその音程の変化を表 4 に基づいた確率でランダムに決定する。
- (5) 生成された旋律が $mnap \pm 1$ オクターブ以内に収まらない場合は, (3) からやり直す。

4.2 色彩の印象からの楽曲自動生成の実現方式

本節では, 色彩情報からその印象に合致した楽曲を自動生成する方式について示す。具体的な方式の流れは図 3 のようになる。

Step1: 色彩情報から印象語を抽出

表 4 歌曲の旋律音における各音程の出現頻度

音程	1 度		2 度		3 度		4 度	
	短 2 度	長 2 度	短 3 度	長 3 度	完全 4 度	増 4 度		
出現頻度 (%)	26	17.5	24	10	6	9	0.5	
音程	5 度		6 度		7 度		8 度	
	減 5 度	完全 5 度	短 6 度	長 6 度	短 7 度	長 7 度		
出現頻度 (%)	0.5	3	1	1	0.5	0	1	

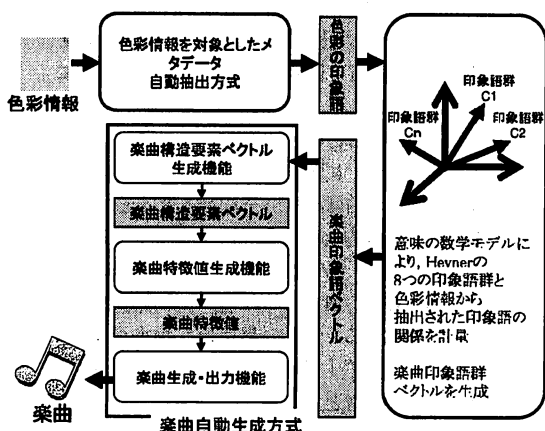


図 3 色彩の印象からの楽曲自動生成方式全体像

3 節で述べた色彩情報を対象としたメタデータ自動抽出方式により、色彩情報からその印象に合致した印象語を抽出する。

Step2: 色彩情報の印象語と楽曲ための印象語群の関係の計量

Step1 で抽出された色彩情報の印象語と Hevner^{(11)~(14)} によって定められた楽曲の印象を表す 8 つの印象語群の関連を 2 節で示した意味の数学モデルを用いてそれぞれ計量する。それにより楽曲印象語群ベクトルを生成する。これは、色彩情報の印象を楽曲に表現できて、かつ色彩の印象に近い印象を抽出することを意味する。

Step3: 楽曲生成・出力

4.1 節で示した、楽曲自動生成実現方式によって、Step2 で生成された楽曲印象語群ベクトルから楽曲を生成、出力する。

これにより、任意の色彩情報から、その色彩に合致した楽曲が生成できる。

5. 実験例

本節は、本方式の実現性を検証するため、本方式に基づく実験システムを構築した。

5.1 実験環境

意味の数学モデルにおける、メタデータ空間 MDS

作成については、“Longman Dictionary of Contemporary English” という英英辞書を使用した。同辞書は、約 2,000 の基本語だけをを用いて約 56,000 の見出し語を説明している。ここで基本語を特徴と見なし、各見出し語を説明する基本語が肯定の意味に用いられていた場合 “1”，否定の場合 “-1”，使用されていない場合 “0”，見出し語自身が基本語である場合その基本語の要素を “1” として、2 章データ行列 M を作成した。これより、約 2000 次元の正規直交空間であるメタデータ空間 MDS を生成した。約 2000 次元のメタデータ空間 MDS では、約 2^{2000} 通りの意味の様相が表現可能である。

生成した楽曲データは、いったんフリーソフト MF2T/T2MF⁽¹⁷⁾ で規定された形式のテキストファイルを出力し、それを T2MF によりスタンダード MIDI ファイルに変換する。

5.2 出力例

入力対象の色彩情報として 1 色のみからなる画像メディアデータを入力することで、楽曲を出力させた。画像メディアデータとして、カラーイメージ上で比較的印象がはっきりしていると思われる、橙色 (オレンジ)、うすはなだ色 (シャドープルー) について行った。

まず、橙色 (オレンジ) を入力とした場合の、楽曲印象語群ベクトル、楽曲構造要素ベクトル、楽曲特徴値の各値の結果を表 5 に示す。楽曲印象語群ベクトルでは、印象語群 c6、つまり “merry”, “happy” などを表す印象が高くなっている。これはカラーイメージスケールの印象とほぼ一致している。これにより、楽曲特徴値では key が “A ♭” というメジャーキー (長調) の明るい曲が抽出されていることを確かめられる。実際に出力された楽曲の楽譜を図 4 に示す。

次に、うすはなだ色 (シャドープルー) を入力とした場合の、楽曲印象語群ベクトル、楽曲構造要素ベクトル、楽曲特徴値の各値の結果を表 6 に示す。楽曲印象語群ベクトルでは、印象語群 c1、つまり “serious”, “sober” などを表す印象が高くなっている。これはカラーイメージスケールの印象とほぼ一致している。これにより、楽曲特徴値では key が “Am” というマイナーキーで、テンポ (tem) が “70” という暗く遅い曲が抽出されていることを確かめられる。実際に出力さ

表5 橙色(オレンジ)の各値

楽曲印象辞ベクトル	楽曲構造要素ベクトル	楽曲特徴値
C1: -1.0000000000	Key: 0.8120021004	kn Ab
C2: 0.0175016580	Tempo: 0.2149308775	tem 118
C3: 0.0655665839	Pitch: 0.2637280047	mnap 79.2853921833
C4: 0.4952288534	Rhythm: -0.4115832135	unac 0.2942088938
C5: 0.2973258546	Harmony: 0.3059455935	ac 0.7057916067
C6: 1.0000000000	Melody: -0.1788307499	tc 0.6529727967
C7: -0.0445466093		oc 0.3470272083
C8: 0.1720978135		um 0.2574026823
		dm 0.4362333822
		lm 0.3063639856

表6 うすはなだ色(シャドーブルー)の各値

楽曲印象辞ベクトル	楽曲構造要素ベクトル	楽曲特徴値
C1: 1.0000000000	Key: -0.0491524761	kn Am
C2: -0.7077000098	Tempo: -0.8941746793	tem 70
C3: -0.3977501571	Pitch: 0.0889668235	mnap 72.4697061165
C4: 0.4412077834	Rhythm: 0.3739388652	unac 0.6869669826
C5: -0.0904689111	Harmony: 0.1252251212	ac 0.3130390674
C6: -1.0000000000	Melody: 0.5280727800	tc 0.5626125606
C7: -0.7080591578		oc 0.4373874394
C8: -0.3250471890		um 0.6401914442
		dm 0.1141186642
		lm 0.2456898917



図4 橙色(オレンジ)の楽曲



図5 うすはなだ色(シャドーブルー)の楽曲

れた楽曲の楽譜を図5に示す。

これらにより、色彩情報からその印象に合致した楽曲を自動生成が可能であることが確認できた。

6. おわりに

本稿では、ユーザから与えられた色彩情報から、その印象に合致した楽曲を生成する方式について示した。

本方式により、利用者の意図とする感情に合致した楽曲メディアデータの自動生成が実現されることで、音楽を通じた情報の新しい価値を生み出すことが可能となる。また、本方式によりコンピュータ環境の中で人間の感性を増幅することできるシステムとなり、コンピュータが人間の生活を豊かにし、人間と密着した感性を共有するパートナーとして新しい関係を築く第一歩となると考えられる。

今後の課題は、他の作曲方式との連携によるより自然な楽曲生成機構の実現他メディアを対象とした連携・統合方式の実現、および、本方式の一般化によるメディア間の自由な計量の実現が挙げられる。

参考文献

- 1) 原島博, 井口征士(監修): “感性的ヒューマンインタフェース最前線 感じる・楽しむ・創り出す感性情報学,” 工作舎, 2004.
- 2) T.Kitagawa, T.Nakanishi, Y.Kiyoki: “An Implementation Method of Automatic Metadata Extraction Method for Music Data and Its Application to Semantic Associative Search,” Systems and Computers in Japan, Vol.35, No.6, pp59-78, 2004.
- 3) T. Kitagawa, Y. Kiyoki, “Fundamental

- Framework for Media Data Retrieval System Using Media Lexico Transformation Operator”, Information Modeling and Knowledge Bases, IOS Press, 2000.
- 4) 吉野太智, 高木秀行, 清木康, 北川高嗣, “楽曲データを対象としたメタデータの自動生成とその意味的連想検索への適用,” 情報処理学会研究報告, 1998-DBS-116(2), pp.109-116, 1998.
 - 5) T.Kitagawa, Y.Kiyoki, “The Mathematical Model of Meaning and its Application to Multidatabase Systems,” Proc. 3rd IEEE International Workshop on Research Issues on Data Engineering, Interoperability in Multidatabase Systems, pp.130-135, April, 1993.
 - 6) Y.Kiyoki, T.Kitagawa, T.Hayama: “A Metadata System for Semantic Image Search by a Mathematical Model of Meaning,” ACM SIGMOD Record, vol.23, no.4, pp.34-41, 1994. (also published as: Multimedia Data Management - using metadata to integrate and apply digital media -, McGrawHill, Amit Sheth and Wolfgang Klas(editors), Chapter 7, 1998).
 - 7) 清木康, 金子昌史, 北川高嗣: “意味の数学モデルによる画像データベース探索方式とその学習機構,” 電子情報通信学会論文誌, Vol.J79-D2, No.4, pp.509-519, 1996.
 - 8) 小林重順: “カラーイメージスケール,” 講談社, 1984.
 - 9) 小谷拓矢, 清木康, 北川高嗣: “色彩情報による静止画像メタデータ生成方式と意味的画像検索への適用,” 第9回データ工学ワークショップ(DEWS'98) 論文集, 電子情報通信学会, 1998.
 - 10) 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適用, 情報処理学会論文誌:データベース, Vol.44,No.SIG8(TOD18), pp21-36, 2003.
 - 11) K. Hevner, “Expression in music: A discussion of experimental studies and theories,” Psychological Review, vol.42, pp.186-204, 1935.
 - 12) K. Hevner, “Experimental studies of the elements of expression in music,” American J.Psychology, vol.48, pp.246-268, 1936.
 - 13) K. Hevner, “The affective value of pitch and tempo in music,” American J. Psychology, vol.49, pp.621-630, 1937.
 - 14) 梅本堯夫(編), 音楽心理学, 誠信書房, 1966.
 - 15) (有) ミューテック, “自動作曲システム,” <http://hp.vector.co.jp/authors/VA014815/music/autocomp.html>
 - 16) 吉川厚, 京野勉, 大場勇治郎: “ファジー演算を使った作曲支援システム,” 電子情報通信学会論文誌, Vol.J74-A, No.5, pp.735-742, 1991.
 - 17) P. V. Oostrum, MF2T/T2MF,