

## 慶應義塾大学学術情報リポジトリ

## Keio Associated Repository of Academic resouces

Title	時間的な印象変化の分析によるマルチメディアデータ検索・可視化システム
Sub Title	Impression-based multimedia retrieval & visualization system with temporal aesthetic features analysis on media streams
Author	倉林, 修一(Kurabayashi, Shuichi)
Publisher	慶應義塾大学湘南藤沢学会
Publication year	2014
Jtitle	Keio SFC journal Vol.14, No.1 (2014. ), p.44- 61
Abstract	本論文では, Web上に拡大する動画像, および, 音楽データを対象とし, それらを多様なコンテキストを持つ時系列メディアデータとして捉え, 時系列的な分析による印象コンテキストの自動抽出, 検索, 可視化を行うマルチメディアデータベースシステムについて述べる。本システムは, これまで感性的な視点からの情報獲得が困難であった時系列メディアデータを対象として, 個人の感性的嗜好に合致する対象の自動配信環境を実現する, 新たなコンテンツ流通基盤として位置づけることができる。
Notes	特集 SFCが拓く知の方法論#招待論文 挿図
Genre	Journal Article
URL	<a href="http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=0402-1401-0003">http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=0402-1401-0003</a>

[招待論文]

# 時間的な印象変化の分析による マルチメディアデータ検索・可視化システム Impression-Based Multimedia Retrieval & Visualization System with Temporal Aesthetic Features Analysis on Media Streams

倉林 修一

慶應義塾大学環境情報学部専任講師

Shuichi Kurabayashi

Assistant Professor, Faculty of Environment and Information Studies, Keio University

**Abstract:** 本論文では、Web 上に拡大する動画像、および、音楽データを対象とし、それらを多様なコンテキストを持つ時系列メディアデータとして捉え、時系列的な分析による印象コンテキストの自動抽出、検索、可視化を行うマルチメディアデータベースシステムについて述べる。本システムは、これまで感性的な視点からの情報獲得が困難であった時系列メディアデータを対象として、個人の感性的嗜好に合致する対象の自動配信環境を実現する、新たなコンテンツ流通基盤として位置づけることができる。

This paper shows an impression-based video and music search and visualization system that realizes emotional semantic video retrieval by analyzing aesthetic features of a whole media stream. This system is effective to retrieve videos which are unfamiliar to users because users do not have to know detailed information. This system models multimedia data as time-series media data having various contexts, such as video and audio data distributed on the Web and live images acquired from a Web camera. This system finds media data from the viewpoint of impression concerning aesthetic presentation of a whole video stream, which conventional video and audio retrieval approaches have not support well because many of those approaches focus on fragments of media data.

**Keywords:** マルチメディア、時系列メディアデータ、マルチメディア検索インターフェイス、大規模データ処理

## 1 はじめに

近年、動画配信サイトや音楽配信サイトの発展に伴い、インターネット上の主要なコンテンツとして、動画像や音楽のように、時間軸に沿って内容が変化するメディアデータが飛躍的に普及してきている。これらの動画・音楽

---

配信を効率的に利用するためには、コンテンツのストーリー性や時間軸に沿った内容の変化を考慮した検索エンジンの実現が重要な研究課題となっている<sup>[23]</sup>。

Web上に大量に存在する動画像コンテンツ群の中から、利用者の嗜好に適合するコンテンツを探し出すための手段として、コンテンツに付与されたキーワードやテキストによるタグ情報（これらをアノテーションと呼ぶ）を対象とした文字列照合検索や、内容情報（色彩、構図、オブジェクトの動きなど）に応じた検索<sup>[8] [24] [25] [30] [31]</sup>が実現されている。これらの検索手段は、利用者が検索対象に関する詳細な情報を既知とする場合に有効である。しかしながら、文献<sup>[9]</sup>における調査では、動画共有サイトの利用者の多くは、動画を視聴する動機として視聴時の気分や雰囲気重要な要因として挙げており、利用者にとって未知のコンテンツであっても感性的に適合するものを、抽象的なキーワードを用いた検索や人手による探索を通じて獲得していることが指摘されている。

このような未知の動画像コンテンツを獲得するための検索、すなわち、発見的検索を実現するためには、印象に応じた検索方式<sup>[1]</sup>が、検索対象に関する詳細情報を必要とせず、利用者の感性的嗜好を表現する問い合わせを用いて対象を検索するため、有効である。これまでに、静止画像を対象とした印象検索方式<sup>[8]</sup>や、音楽データを対象とした感性検索方式<sup>[13]</sup>が提案されている。動画像を対象とした感性検索としては、ユーザが自発的に動画像に付与したキーワード文字列、タグ文字列、コメント文字列を対象とした感性検索方式<sup>[26]</sup>が提案されている。これらの文字列メタデータは、ソーシャルネットワークを通じて共有され、共同的に構築されるため、ソーシャルアノテーションと呼ばれる。視覚メディアとしての動画像が視聴者に与える直接的な印象に応じた検索を実現するためには、動画像中の内容変化を扱うことが必要であるが、動画像の内容を対象とした感性分析を行い、感性的な内容変化に応じた検索は、実現されていない。従来の動画検索システムは、動画像中から、特定のシーンや特定のオブジェクトを検索することを目的としており、動画像の内容全体が持つ感性的特徴から動画像を探し出すことは困難である。時間軸に沿った印象の変化に応じたメディア検索を実現するためには、次の3

機能が必要となる。

- メディアデータの印象は、時間軸に沿って推移していくため、メディアデータにおける印象変化の文脈を自動的に識別し、時系列印象メタデータを自動抽出する機能
- メディアデータ間の感性的な類似性、差異性は、静的な関係によって決定されるものではないため、多様な印象の文脈を識別し、利用者の感性的な情報の要求を記述した問い合わせを解釈し、検索ランキングを行う機能
- 利用者が検索結果として得たすべてのメディアデータを実際に再生し、問い合わせとの適合性を検証することは時間的に困難であるため、動画像中の感性的内容変化を直感的に表示し、利用者の理解を支援する可視化機能

本論文では、時間軸にそって内容が変化するデータ、“時系列メディアデータ”を対象として、そのコンテキストの自動抽出を伴う印象分析・検索方式、および、それを実現するシステム群について示す。人類が日々作成し続ける情報の多くは、時間情報を伴っており、その多くは時系列メディアデータとして位置づけることができる。本研究は、これらの時系列メディアデータに内在する文脈抽出の技術を中核としており、幅広い応用が可能である。具体的に、動画像を対象とする MediaMatrix システム<sup>[6][17]</sup>、電子書籍を対象とする Query-by-Apperance システム<sup>[18]</sup>、検索エンジンの検索履歴を対象とする QMAP システム<sup>[19]</sup>、音楽データを対象とする Chord-Cube システム<sup>[20]</sup>と VizMIR システム<sup>[21]</sup>について、本論文では述べる。具体的かつ学際的な応用として、SFC における外国語教育研究と情報システム研究の成果を統合し、学習者のユビキタス環境での学習を支援するシステムの構築を行っている<sup>[6]</sup>。この学際的共同研究成果の教育への応用として、文部科学省「質の高い大学教育推進プログラム（教育 GP）」事業“ユビキタス社会の問題発見解決型人材育成”([http://www.sfc.keio.ac.jp/research/research\\_projects/ubiquitous.html](http://www.sfc.keio.ac.jp/research/research_projects/ubiquitous.html))を実施し、スマートフォンを用いたユビキタス外国語学習環境を実現し、学

---

部学生へ提供してきた。さらに、本研究による時系列データ分析を産業界が有するビッグデータへ適用する実践として、JR 東日本との共同研究を、2008年以降継続して実施してきた。この共同研究を通じて開発したシステムは、日本の首都圏の鉄道ネットワークにおける約 400 駅から得られるセンシングデータ（1日約 670 万人の乗客）を取得し、乗客の移動状況を分析するために 2,420 万通りの移動経路とセンシングデータを対象とした大規模データマイニング処理を行うシステムである。本システムは、各運行車両、駅、時間ごとの乗降車状況を、1分単位で可視化し、運行状況に合致する広告情報などの実空間への配信を実現しており、世界的にも例を見ない大規模公共交通データマイニング・システムである。なお、この成果は、旅客流動に関するビッグデータ分析のコア技術の一つとして、共同で特許を出願している（特開 2012-108654）。

## 2 関連研究

メディアデータを対象とした感性分析・検索の手法は、印象メタデータの生成方法において、大きくは、人間がメタデータを記述する直接生成方式と、メディアデータの特徴量と印象メタデータの相関定義を用いてメタデータを自動生成する間接生成方式に分類できる。

直接生成方式は、人間がメディアデータを実際に鑑賞し、感性と関連する形容詞を用いてメディアデータのメタデータを記述するものである。動画像を対象とした感性メタデータの直接生成方式としては、動画像に付与されたソーシャルアノテーション（ソーシャルネットワークを通じて共有され、共同的に構築されるキーワードやタグ情報）を対象とした感性検索方式<sup>[26]</sup>がある。直接生成方式は、メディアデータを人間が鑑賞するため、人間が視聴時に受ける直観的な印象と適合するメタデータを高精度に生成することができる。しかしながら、人間がメディアデータを実際に鑑賞する必要があるため、大規模なメディアデータ群を対象として、印象メタデータを付与することは難しい。メタデータを付与する鑑賞者の人数が多い場合、付与する印象語の語彙の一貫性を保つことも困難となる。

間接生成方式は、印象語と、メディアデータから直接抽出可能な特徴量

との相関関係を、統計調査や機械学習により予め求めておき、メディアデータを対象とした自動索引付けを実現する方式である。本論文において示す MediaMatrix システムは、メディアデータの特徴量(色彩)と印象メタデータ(印象語)の相関定義を用いてメタデータを自動生成する間接生成方式に分類される。文献<sup>[12]</sup>では、アニメーション動画像を対象として、配色の傾向を抽出し、アニメーションの色彩演出に応じた自動索引付け手法を提案している。動画データと同様に、時間軸を伴うメディアである音楽データを対象とした感性検索技術も多数提案されている。文献<sup>[4]</sup>は、楽曲を視聴した際の印象の受け方が類似するユーザを予めグルーピングすることにより、音楽の聞き手であるユーザの嗜好に応じて、それぞれの楽曲へ検索のためのメタデータを付与する方式を示している。音楽データベースにおいて、時間軸に沿った印象メタデータ生成に取り組んだ研究として、文献<sup>[5]</sup>は、一つの楽曲内の時間的変化に応じて、楽曲の印象変化を分析し、その印象の変化を検索用メタデータ(印象メタデータ)として付与する音楽データベース・システムを提案している。文献<sup>[5]</sup>が示す印象メタデータの生成方式は、一楽曲中の隣り合う前後のセクション間の印象特徴量の差分を計量し、その差分を重みとしてメタデータに反映することにより、一楽曲内における印象の推移を反映したメタデータを生成する。

人間の感情の分析手法としては、心理学者のラッセル (Russell, J.)、メラビアン (Mehrabian, A.) らが、快 (Pleasure)、覚醒 (Arousal)、優越 (Dominance) で形成される 3次元の P-A-D 感情モデルを用いて、感情(情動反応)を符号化する手法を提案している<sup>[28]</sup>、この P-A-D 感情モデルを動画像検索に応用し、視聴者が感じる情動反応に応じた、動画像への自動的な索引付け手法が文献<sup>[7]</sup>で提案されている。動画像への索引付けとは、動画像中の特定のシーンに、そのシーンの意味や印象を表すキーワードや数値を付与することである。付与されたキーワードや数値を索引(インデクス)と呼び、動画像の内容検索では、この索引を利用して高速に目的のシーンや動画像を獲得する。文献<sup>[7]</sup>において提案されている手法は、P-A-D 感情モデルを用いて特徴付けした 6つの情動反応 Amusement、Violence、Neutral、Fear、Sadness、Happiness を用いて映像を、いくつかのシーンに分割し、映像を感性的な視点から構造化する

---

ことを可能にする。映像のシーン検出に関しては、様々な研究が行われている。例えば、文献<sup>[33]</sup>では、映像中の物理的な変化として、色の変化やカメラの動き、登場人物の動きなどのデータを用いて映像のシーンの境目を決定する手法が示されている。

### 3 時系列印象マルチメディアデータ検索・可視化機構

本節では、これまでに開発した、印象による情報検索を実現するシステム群について述べる。動画像・音楽を対象とした印象検索を実現する MediaMatrix システム<sup>[16][17]</sup>、電子書籍の表紙デザインを対象とした印象検索を実現する Query-by-Apperance システム<sup>[18]</sup>、検索エンジンへ入力される検索語と検索語の関係を「言葉の地図」としてダイナミックに可視化する QMAP システム<sup>[19]</sup>、楽曲同士の印象の類似性を 3次元空間における距離として可視化する Chord-Cube システム<sup>[20]</sup>、画像の印象を利用して楽曲を検索可能にする VizMIR システム<sup>[21]</sup>について述べる。

#### 3.1 動画像・音楽データを対象とした印象検索・可視化

MediaMatrix とは、これまでに筆者が開発してきた動画像・音楽データ対象アクティブ・感性マルチメディア配信を実現する統合システム・ソフトウェアである (図 1)。本システムの目的は、時間的な内容の変化を伴う動画像データ、および、音楽データを対象とし、利用者に感性的要求に適合するコンテンツを、印象語 (キーワード) を組み合わせた問い合わせを用いて獲得する機能を実現することである。MediaMatrix システムは、印象特徴量分析による感性的コンテキストの自動抽出、検索、配信を行うものであり、そのための感性自動分析処理機能と相関量演算機能、および、それらを組み合わせて実際のアプリケーションを記述するための動画像問い合わせ言語を特徴とする。ここで、動画像問い合わせ言語とは、ユーザが獲得したいと考える動画像の印象を記述するための簡易的なプログラミング言語である。この動画像問い合わせ言語により、ユーザは、マイクロソフト社の Access や、オラクル社の Oracle などに採用されている SQL と呼ばれる問い合わせ言語と類似する手法で検索することができる。本章では、実現した MediaMatrix の全体

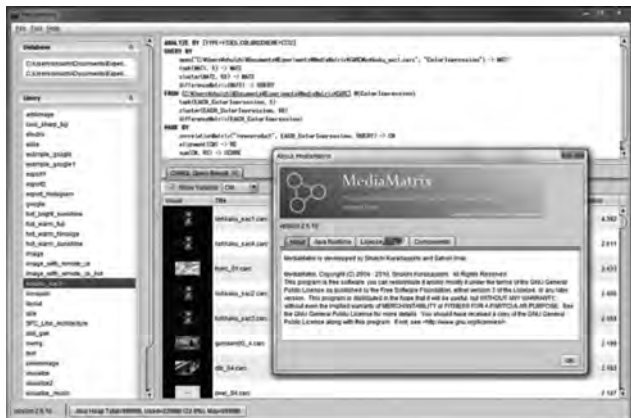


図1 動画像・音楽メディアを対象とした印象分析・可視化・配信を実現する MediaMatrix

設計、および、主要機能について述べる。

MediaMatrix システムは、動画像データから、その感性的特徴を表すメタデータを自動抽出し、得られたメタデータを用いた関連性の計量機能による情報獲得を実現する。本システムは、色彩心理学において定義された130色のマンセル基本色の組み合わせである182配色セットを用いて、182次元の色彩印象空間を構築する。この182配色セットにおいて、1つの配色は1つの固有の印象語に対応しており、構築した182次元の色彩印象空間は、2の182乗の種類の印象のコンテクストを識別することができる(図2)。ここで、印象のコンテクストとは、基本的な182種類の印象キーワードを複数の組み合わせ、かつ、それぞれのキーワードに重みを付けて表現されたものであり、ユーザーが獲得したいと考えるメディアデータの印象を示すデータである。さらに、西洋音楽における24調性に対応する24次元の音楽感性空間を構築し、2の24乗の種類の音楽印象コンテクストを識別可能な音楽感性分析機構(図3)、および、2の24乗の種類の音楽印象コンテクストの可視化機構(図4)を開発した。この感性分析・可視化機構は、動画像における時間軸に沿った色彩印象の推移を反映したメタデータ生成のために、動画像を任意秒数ごとの静止画像として抽出し、それら静止画像群を対象として、映像中の色彩情報を用いて感性特徴量を分析する機能を実現している。また、音楽データ分析にお





図 2 MediaMatrix は 182 次元の色彩印象空間による  $2^{182}$  種類の印象コンテキストを識別可能な動画感性分析機能を実現

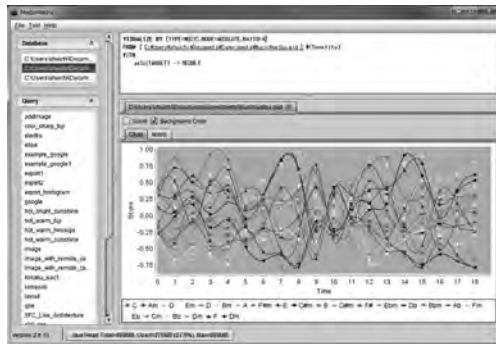


図 3 MediaMatrix は、24 次元の音楽感性空間による  $2^{24}$  種類の音楽印象コンテキストを識別可能な音楽感性分析機構を実現

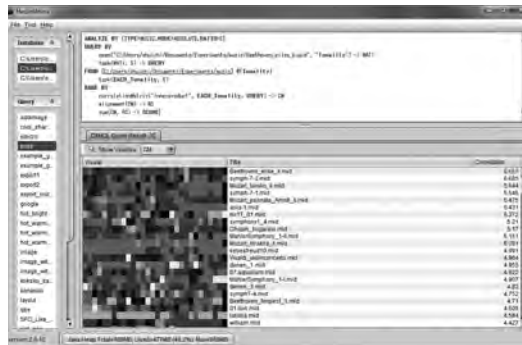


図 4 MediaMatrix の感性時系列メディア分析・可視化機構として、楽曲の時系列内容分析・相関量計量・可視化機構を実現

いては、音楽データを任意の小節数毎の音高（ピッチ）の発音時間のヒストグラムとして抽出し、それら音高ヒストグラムの列を対象として、調性を自動的に判定し、24の調性に応じた感性特徴量を分析する機能を実現した。

本システムが実現する動画像印象検索方式の特色は、動画像全体において支配的に出現する印象、および、動画像の特定箇所において局所的に強く出現する印象を自動的に抽出する特徴選択の演算を適用し、動画像全体をメタデータとして要約することにより生じる印象の曖昧性を排除する点にある。具体的には、本システムは、動画像全体において支配的に出現する印象 (Dominant Impression)、および、動画像の特定箇所において局所的に強く出現する印象 (Salient Impression) を自動的に抽出する特徴選択の演算を適用し、動画像の全体的なイメージと局所的なイメージを表現する2種類の印象メタデータを抽出する。この特徴選択の演算を、本研究では、“時系列印象変化分析による特徴選択演算 (Feature Selection Operation by Time-Series Impression Transition Analysis)” と呼ぶ (図5)。この演算は、動画像を静止画の集合としてとらえ、すべての静止画中の色彩の傾向から、多くのシーンで頻繁に出現する色彩が与える印象を Dominant Impression として抽出する。さらに、この支配的印象と逆の振る舞いをする色彩の傾向、すなわち、局所的には頻繁に表れるが、全体的にはあまり出現しない色彩が与える印象を、Salient Impression として抽出する。本システムでは、色彩印象空間において設定されたストーリー特徴選択演算により、それぞれの印象語がストーリー中において位置づけられる文脈を、動画像の内容変化に応じて動的に計算することが可能である。また、ストーリーの変化に応じた特徴選択の機能を導入することにより、動画像全体を一つのメタデータとして要約することにより生じる印象の曖昧性を排除することを可能としている (図6、図7)。なお、図6と図7は、同一の動画像からの抽出結果である。

MediaMatrix による動画像印象分析において、色彩と、視聴者の感性的感覚との関係を定義するデータ構造が、色彩印象空間である。色彩印象空間  $K$  は、色彩心理学において定義された  $h$  個の色彩印象語  $cs$  について各々  $g$  個の基本色  $(c_1, c_2, \dots, c_g)$  との関連  $(f_1, f_2, \dots, f_g)$  を列挙した特徴付ベクトル  $cs_i$  ( $i=1, \dots, g$ ) が与えられているとし、そのベクトルを並べて構成する  $g \times h$  行列で

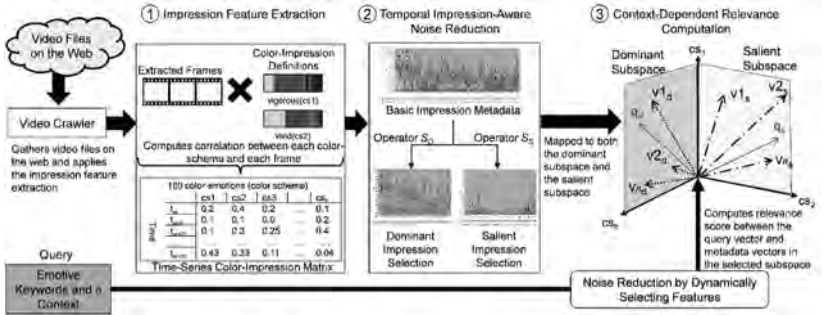


図5 時系列印象変化分析による特徴選択演算の概要



図6 動画画の特定箇所において局所的に強く出現する印象 (Salient Impression) を自動的に抽出する特徴選択の演算の適用例

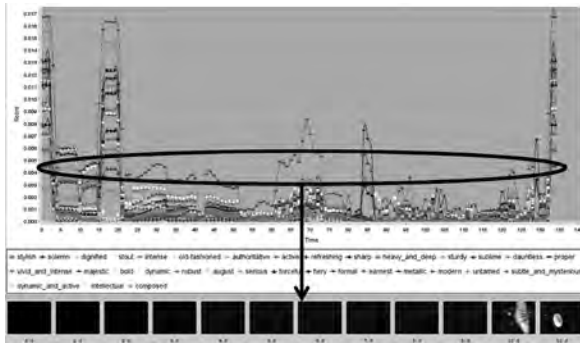


図7 動画画像全体において支配的に出現する印象 (Dominant Impression) を自動的に抽出する特徴選択の演算の適用例

ある(図8)。ここで、特徴は0から1までの実数として表現される。0は無相関を示し、1は最大相関を示す。このとき、色彩印象空間 $K$ は列毎、すなわち、ベクトル $c_s$ 毎に2ノルムで正規化される。色彩印象空間 $K$ は、色彩心理学の分野において数多く定義されている配色イメージ辞書<sup>[14][2][3]</sup>を用いて、作成することができる。図8に示されている“vigorous”と“vivid”は、文献<sup>[2]</sup>に掲載の色彩データを用いて作成した、182次元の色彩印象空間の一部である。配色中の各色の面積は当該色の重みに対応しており、例えば、赤色は、感性語“vigorous”に強い関連性を有し、黄色は、感性語“vivid”に強い関連性を有する。

### 3.2 電子書籍の表紙デザインを対象とした印象検索・可視化

印象による情報検索を実現するシステムの、電子書籍分野への応用例として、人間の感性に強い影響を及ぼす、書籍の表紙デザインを対象とした新たな検索システム<sup>[18]</sup>を開発した(図9)。本システムは、利用者が視覚メディアデータのラフなスケッチと色合いを選択すると、システムが“エディトリアル・デザイン”という知識を用いて利用者の意図した問い合わせを解釈・拡張し、高度なデザインによる問い合わせを容易に実行可能とする。これにより、メディアデータの「見た目」に応じた検索を実現し、利用者が直感的に視覚メディアデータを検索することができる。本システムの特徴は、視覚メディアデータのレイアウトや配色をデザインする知識であるエディトリアル・デザインを用いて、利用者が入力した線画のみのラフな問い合わせデータを拡張することにより、類似するデザインを検索可能にする点にある。図10に、開発した電子書籍検索・可視化システムの実働状況を示す。

### 3.3 Web 検索エンジンへ入力される検索語と検索語の関係の可視化・分析システム

直観的な情報検索システムの、モバイル環境への応用として、タブレット端末上における、利用者による問い合わせ入力を支援するシステム QMAP システム<sup>[19]</sup>について示す。QMAP システムは、近年普及が進んでいるスマートデバイス(スマートフォン、タブレット端末)を対象として、動画像・音楽デ

---

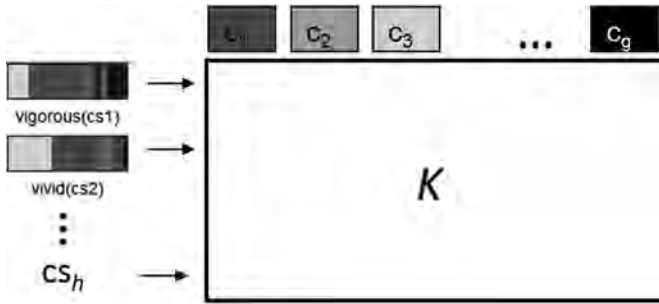


図8 データ行列  $K$  による色彩印象空間の表現

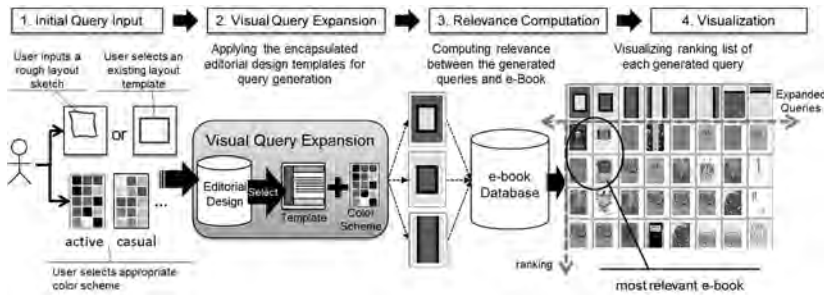


図9 MediaMatrix システムを応用し、メディアデータの「見た目」に応じた検索を実現し、利用者が直感的に視覚メディアデータを検索できるシステムを構築

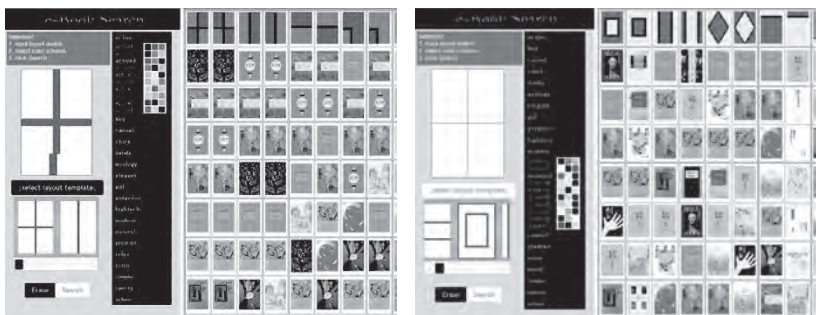


図10 利用者が視覚メディアデータのラフなスケッチと色合いを選択すると、システムが<sup>44</sup>「エディトリアル・デザイン」という知識を用いて利用者の意図した問い合わせを解釈・拡張し、高度なデザインによる問い合わせを容易に実行可能

ータの検索に用いられる検索語について、検索語と検索語の関係を「言葉の地図」としてダイナミックに可視化することにより、スマートフォンやタブレット端末などのタッチパネルから、検索エンジンへの入力を容易にするシステムである。図 11 に、開発した検索語自動推薦・可視化システムの実働状況を示す。本方式の特徴は、検索エンジンに蓄積された検索履歴のストーリー性を分析することにより、検索をより詳細化する単語、および、検索の視点を切り替える単語、という二つの観点からの検索語推薦を網羅的に実行する点にある。

### 3.4 楽曲同士の印象の類似性の可視化

Chord-Cube システム<sup>[20]</sup>は、楽曲同士の印象の類似性を 3 次元空間における距離として可視化する音楽検索・可視化システムである。本システムは、楽曲の A メロディー、B メロディー、サビにおける特徴量 (調性) をそれぞれ 3 次元空間上の軸にマッピングすることにより、楽曲と楽曲の関係を 3 次元空間上の点としてダイナミックに可視化する。図 12 に、開発した音楽検索・可視化システムの実働状況を示す。本システムは、Web 上における 3 次元画像表示の標準仕様である WebGL を用いて実装され、特別なプラグインなどを必要とせずに、Web 上で自由に 3 次元空間を回転・移動させながら楽曲を探すことが可能となっている。

### 3.5 画像の印象を利用して楽曲を検索可能にするクロスメディア検索

VizMIR システム<sup>[21]</sup>は、画像の印象を利用して楽曲を検索可能にするクロスメディア音楽検索システムである。このシステムは、利用者にとって未知の楽曲や、曖昧な印象のみ記憶している楽曲を検索する際、その印象を示す画像データを入力することにより、楽曲の検索を行うことができる。図 13 に、開発した Web 検索語自動推薦・可視化システムの実働状況を示す。本方式の特徴は、画像の列を与えることにより、問い合わせに印象変化の「ストーリー性」を付与し、そのストーリーに応じた楽曲の特徴量の変化を検出する点にある。図 13 に示す右側のスクリーンショットは、画面下部の画像群の中から、ユーザが、探したい楽曲の印象に適合する画像を選択する様子を示している。

---

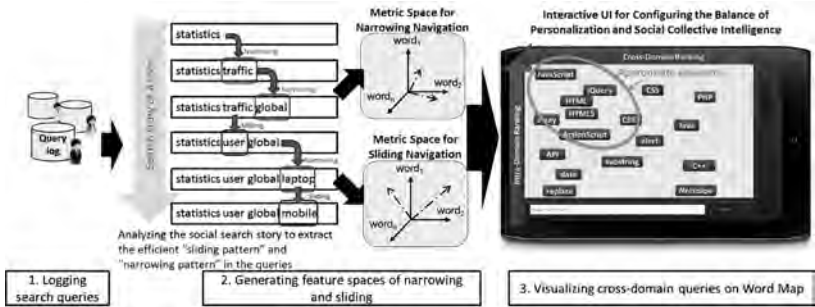


図 11 動画像・音楽メディアを対象としたクロス・メディアナビゲーションシステム QMAP

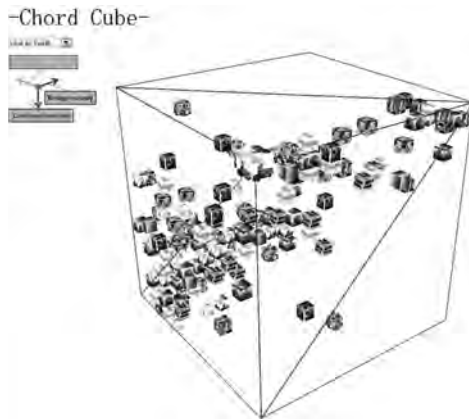


図 12 楽曲同士の印象の類似性を 3 次元空間における距離として可視化する Chord-Cube



図 13 楽曲の時間軸に沿った印象変化を画像の並びとして表現する音楽検索システム VizMIR

ユーザが画像を選択すると、画面上部に、左から右へ印象の変化のストーリーが、画像として表示される。この例では、4枚の画像を用いて、印象の変化を指定している。図13右側は、4枚の画像を用いて、実際に音楽検索を行った結果を示している。画面下部左側に、検索結果として、楽曲のランキングが示されている。このランキングでは、上位の楽曲ほど、ユーザが指定した印象に近いものとして判定されていることを示している。画面下部右側は、当該楽曲を再生するプレーヤである。この場合、YouTube上の楽曲を再生している。

## 4 おわりに

本論文では、動画・音楽データを対象とした感性マルチメディア配信を実現する統合システム・ソフトウェア MediaMatrix を中心として、メディアデータの感性的な特徴量に応じた可視化・検索システムについて述べた。本稿で紹介したシステム以外にも、分散ディスプレイ環境における、自動的な内容分析による情報配信システムの開発<sup>[15]</sup>、分散並列型感性メディア分析フレームワークの設計、分散並列型感性メディア分析フレームワークの実装<sup>[10]</sup>、Service-Oriented Architecture(SOA)と呼ばれる大規模システム構築技術と、筆者が開発した MediaMatrix との統合を行うシステムの実現可能性の検証<sup>[29]</sup>、および、MediaMatrix システムによるメタデータ生成、および、動画画像ダウンロードに関する高速化技術の開発<sup>[32]</sup>を進めている。また、近年普及しているスマートフォン上で動作するマルチメディアデータベースシステムの開発、および、実験を進めており、モバイル環境における自動的なマルチメディア情報分析・検索・配信を行う、モバイル・マルチメディア視聴環境を実現していく予定である。

## 参考文献

- [1] 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-D-2、No.4、1996年、pp.509-519。
- [2] 小林 重順『カラーイメージスケール改定版』講談社、2001年。



- [3] 南雲 治嘉『配色イメージチャート』グラフィック社、2000年。
- [4] 熊本 忠彦「印象に基づく楽曲検索のためのユーザモデリング手法」『情報処理学会論文誌』データベース、Vol.47、No.8、2006年、pp.157-164。
- [5] 伊地 智麻子・清水 康「音楽データの印象の時間的推移を扱う印象メタデータ自動生成方式」『情報処理学会論文誌』データベース、Vol.44、No.18、2003年、pp.1-16。
- [6] 藁谷 郁美・太田 達也・マルコ・ラインデル・倉林 修一「インフォーマル・ラーニングを支援する体験連動型外国語学習環境の構築」『日本教育工学会論文誌』36巻2号、2012年、pp.91-102。
- [7] Arifin, S. and Cheung, P. Y. K., "A computation method for video segmentation utilizing the pleasure-arousal-dominance emotional information," In *Proceedings of the 15th ACM International Conference on Multimedia*, 2007, pp. 68-77.
- [8] Cheung, S. S. and Zakhor, A., "Fast similarity search and clustering of video sequences on the world-wide-web," *IEEE Transactions on Multimedia*, 7(3), 2005, pp.524-537.
- [9] Cunningham, S. J. and Nichols, D. M., "How people find videos," In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries*, 2008, pp.201-210.
- [10] Fukuda, M., Kurabayashi, S., Hall, J. Kiyoki, Y., "Morphing Parallelization Strategy to Support On-the-Fly Video Analysis," In *Proceedings of the 2010 International Conference on Parallel and Distributed Processing Techniques and Applications*, 2010, pp.541-547.
- [11] Godlove, I. H., "Improved Color-Difference Formula, with Applications to the Perceptibility and Acceptability of Fadings," *Journal of the Optical Society of America*, 41(11), 1951, pp.760-770.
- [12] Ionescu, B., Coquin, D., Lambert, P., Buzuloiu, V., "A Fuzzy Color-Based Approach for Understanding Animated Movies Content in the Indexing Task," *EURASIP Journal on Image and Video Processing*, Article ID 849625, SpringerOpen, 2008, 17 pages.
- [13] Kitagawa, T., Kiyoki, Y., "Fundamental Framework for Media Data Retrieval Systems using Media-lexico Transformation Operator - in the Case of Musical MIDI Data," *Information Modelling and Knowledge Bases XII*, IOS Press, 2001, pp.316-326.
- [14] Kobayashi, S., "The aim and method of the color image scale," *Color Research & Application*, 6(2), John Wiley & Sons, 1981, pp.93-107.
- [15] Kurabayashi, S. Dubnov, S., and Kiyoki, Y., "JoinSee: A Real-Time and Collaborative Hyper-Media System for Participatory Performances in the Opera of Meaning," *Information Modelling and Knowledge Bases XXI*, IOS Press, 2010, pp.258-262.
- [16] Kurabayashi, S. and Kiyoki, Y., "MediaMatrix: A Video Stream Retrieval System with Mechanisms for Mining Contexts of Query Examples," In *Proceedings of the 15th International Conference on Database Systems for Advanced Applications (DASFAA2010)*, Springer, 2010, pp.452-455.
- [17] Kurabayashi, S. and Kiyoki, Y., "Impression-Aware Video Stream Retrieval System with Temporal Color-Sentiment Analysis and Visualization," In *Proceedings of the 23rd International Conference on Database and Expert Systems Applications (DEXA 2012)*, Springer, 2012, pp.168-182.
- [18] Kurabayashi, S. and Koike, Y., "Query-by-Appearance: Visual Query Expansion

- to Support Domain-Specific Retrieval of e-Books,” *International Journal on Advances in Internet Technology*, Vol.5, No.3&4, IARIA, 2012, pp.162-172.
- [19] Kurabayashi, S. and Shimaoka, R., “A Cross-Domain Query Navigation and Visualization System for Touchscreens that Exploits Social Search History,” *International Journal on Advances in Internet Technology*, Vol.5, No.3&4, IARIA, 2012, pp.186-195.
- [20] Kurabayashi, S. and Imai, T., “Chord-Cube: Music Visualization and Navigation System with an Emotion-Aware Metric Space for Temporal Chord Progression,” *International Journal on Advances in Internet Technology*, Vol.7 No.1&2, IARIA, 2014, pp.52-62.
- [21] Kurabayashi, S. and Kato, Y., “VizMIR: A Cross-Media Music Retrieval System Supporting Mutual Transformation Between Emotive Color Changes and Tonal Changes,” *International Journal on Advances in Software*, vol.7 no.1&2, IARIA, 2014, pp.161-170.
- [22] Luo, M.R., G. Cui, and B. Rigg, “The development of the CIE 2000 colour-difference formula: CIEDE2000,” *Color Research & Application*, 26(5), John Wiley & Sons, 2001, pp.340-350.
- [23] Lehane, B., O’Connor, N. E., Lee, H. and Smeaton, A. F., “Indexing of Fictional Video Content for Event Detection and Summarisation,” *EURASIP Journal on Image and Video Processing*, Article ID 14615, SpringerOpen, 2007, 15 pages.
- [24] Lew, M. S., Sebe, N., Djeraba, C. and Jain, R., “Content-based multimedia information retrieval: State of the art and challenges,” *ACM Transactions on Multimedia Computing, Communications and Applications*, 2(1), 2006, pp.1-19.
- [25] Muneesawang, P. and Guan, L., “Adaptive Video Indexing and Retrieval,” *Multimedia Database Retrieval: A Human-Centered Approach*, Springer, 2006, pp. 127-156.
- [26] Nakamura, S. and Tanaka, K., “Video Search by Impression Extracted from Social Annotation,” In *Proceedings of the 10th International Conference on Web Information Systems Engineering (WISE2009)*, 2009, pp.401-414.
- [27] Newhall, S. M., Nickerson, D. and Judd, D. B., “Final Report of the O.S.A. Subcommittee on the Spacing of the Munsell Colors,” *Journal of the Optical Society of America*, 33(7), 1943, pp.385-411.
- [28] Russell, J. A. and Mehrabian, A., “Evidence for a three-factor theory of emotions,” *Journal of Research in Personality*, Vol. 11, 1977, pp. 273-294.
- [29] Šaša, A., Kiyoki, Y., Kurabayashi, S., Chen, X., and Krisper, M., “A Service-Oriented Framework for Personalized Recommender Systems Using a Colour-Impression-Based Image Retrieval and Ranking Method,” *Information Modelling and Knowledge Bases XXIII*, IOS Press, 2012, pp. 258-262.
- [30] Smeaton, A. F., “Techniques used and open challenges to the analysis, indexing and retrieval of digital video,” *Information Systems*, 32(4), Elsevier Science, 2007, pp.545-559.
- [31] Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A. and Jain, R., “Content-based image retrieval at the end of the early years,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 2000, pp.1349-1380.
- [32] Sillberg, P. Kurabayashi, S., Rantanen, P. Yoshida, N., “A Model of Evaluation: Computational Performance and Usability Benchmarks on Video Stream Context Analysis,” *Information Modelling and Knowledge Bases XXIV*, IOS Press, 2013,

- pp.188-200.
- [33] Yi, H., Rajan, D., and Chia, L.T., “A motion based scene tree for browsing and retrieval of compressed videos,” In *Proceedings of the 2nd ACM international workshop on Multimedia databases (MMDB '04)*, 2004, pp.10-18.

[受付日 2014. 5. 7]