Vol.2012-MUS-96 No.1
2012/8/9

Songrium: 多様な関係性に基づく音楽視聴支援サービス

濱崎 雅弘[1,a)]　後藤 真孝[1,b)]

概要：本稿では，楽曲やミュージックビデオ等の音楽コンテンツ間の多様な関係性を意識しながら，Web 上で新たな音楽コンテンツに出会うことができる音楽視聴支援サービス Songrium について述べる．従来の音楽情報検索・推薦では，関係性に基づく音楽コンテンツとの出会いは十分に支援されていなかった．本研究では，音楽コンテンツ間の明示的あるいは暗黙的な関係に名前を付ける「矢印タグ」という枠組みを提案し，Songrium 上で不特定多数のユーザが自由に矢印タグを定義して共有するソーシャルアノテーションを可能にする．ユーザが興味のある音楽コンテンツから出発して，矢印タグが指す先にあるコンテンツを次々と視聴していくことで，潜在的に関心のあるコンテンツに出会う機会が増える．さらに矢印タグが普及すれば，各音楽コンテンツの位置づけが視聴前に判明し，人々はより多くの関係性を意識でき，新たな音楽コンテンツを生み出す土壌となることが期待できる．

1. はじめに

音楽と音楽の間には関係がある．類似関係（歌詞のテーマや社会背景，曲調，雰囲気，演奏楽器などが同じか似ている）や，派生関係（アレンジ，リミックス，替え歌，カバー，映像付与），クリエータの人間関係（楽曲の作者・歌手・演奏者が同じ，あるいは，友人・師弟・ライバル・同所属）など，ある楽曲は他の楽曲と様々な関係を持っている．動画コミュニティサービス「ニコニコ動画」[1] 上に日々創作・投稿されている音楽コンテンツ間にも豊かな関係性があり，例えば，ある楽曲を別の人が歌ったり，演奏したり，踊ったりする派生した音楽コンテンツには，「歌ってみた」，「演奏してみた」，「踊ってみた」のようなジャンルを表す特殊なタグ（カテゴリタグ）が付与され，その関係性を人々は意識しながら音楽コンテンツを楽しんでいる．自分が好きな楽曲に関係がある他の楽曲にも関心を持つのは自然であり，関係性が，新たな楽曲に出会うときの手がかりになることも多い．「この曲を聴いていると，あの曲も聴きたくなる」という気持ちになるのも，「次にこちらを聴きたい」という関係があると見なすことができる．本研究では，こうした音楽の多様な関係性に着目し，人々と音楽コンテンツとの出会いをより豊かなものにすることを目的とする．

しかし従来の音楽視聴環境では，こうした音楽の多様な関係性は十分には活用されていなかった．近年，膨大な音楽コンテンツがいつでもどこでも視聴可能となり，音楽配信サイトや動画共有サイトの普及により，自分が知らない様々な音楽コンテンツを容易に視聴できるようになった．特に音楽に関連した CGM（Consumer Generated Media，消費者生成メディア）現象 [2] の広がりで，日々新たな音楽コンテンツが生まれ続けている．そのため視聴者にとって，膨大な音楽コンテンツの中から潜在的に関心のある音楽に出会うのは容易でない．音楽情報検索・推薦の進展により書誌情報や音響特徴量の類似度に基づく検索や [3], [4]，視聴履歴等のユーザモデリングや協調フィルタリングに基づく推薦等が可能になり [5], [6]，音楽と出会うためのインタフェースの研究も取り組まれてきた [7], [8], [9]．しかし，従来は音楽の関係性の一部を限定的に扱っていることが多く，関係性に基づく音楽コンテンツとの出会いは十分に支援されていなかった．

そこで本研究では，多様な関係性を明示的に扱うために，ある音楽コンテンツを起点として，別の音楽コンテンツがそれに対してどういう関係にあるかに名前を付ける**「矢印タグ」(Arrow Tag)** を提案する．矢印タグは，音楽コンテンツ間の明示的あるいは暗黙的な関係に対するソーシャルタギングを可能にする枠組みであり，人々が矢印タグを自由に名付けて定義し，共有していくことで，様々な関係性を扱えるようになっていく．これにより，矢印タグの先にある音楽コンテンツを選択して辿りながら視聴するだけで，様々な音楽コンテンツに出会うことが可能となる．さらに，この矢印タグが普及していくことで，(1) 音楽コンテンツの位置づけを視聴前に関係から捉えられる，(2) 気

1 産業技術総合研究所，〒 305-8568 茨城県つくば市梅園 1-1-1
National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan
a) masahiro.hamasaki(at)aist.go.jp
b) m.goto(at)aist.go.jp

© 2012 Information Processing Society of Japan

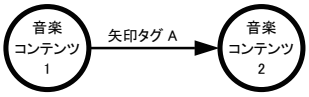

図 1 音楽コンテンツ間の関係を表す矢印タグ (Arrow Tag)

づいていなかった音楽コンテンツ間の関係を発見できる，(3) 同じ関係を持つような音楽コンテンツ群に注目できる，(4) 多様な関係が生まれて淘汰される，(5) 関係を意識して新たなコンテンツが生まれる，という様々なメリットを生じることが期待できる．

この矢印タグを誰でも Web ブラウザ上で利用することができるように，我々は，音楽コンテンツ間の多様な関係を矢印タグで辿りながら，様々な未知の音楽コンテンツに出会うことができる音楽視聴支援サービス**「Songrium」（ソングリウム）**（http://songrium.jp）を実現して提供する．Songrium は，二つの音楽コンテンツ間に矢印タグを定義して共有するソーシャルアノテーションを可能にする Web 上のサービスである．本研究での「音楽コンテンツ」は，楽曲だけでなく，音楽に密接に関連した音楽連動動画（ミュージックビデオやダンス動画等）を意味する．Songrium は本来は対象コンテンツを限定するサービスではないが，初期段階では動画コミュニティサービス「ニコニコ動画」上の歌声合成 (VOCALOID [10]) 技術に関連した動画を対象とする．これは，クリエータと視聴者がソーシャルアノテーション（タグやコメント）を積極的に活用して，日々新たなコンテンツを生み出しているコミュニティ [11], [12] であるからであり，ニコニコ動画の外部プレーヤを活用して，Songrium のサービス上で実際に動画を視聴可能なメリットもある．

2. 矢印タグ

矢印タグ (Arrow Tag) とは，ある音楽コンテンツから別の音楽コンテンツへの方向性のあるリンク（矢印）に，その関係を表現する名称を付与したものである．図 1 のように，音楽コンテンツ 1 から音楽コンテンツ 2 へ向けた矢印タグ A は，「音楽コンテンツ 1 を基準に考えたときに，音楽コンテンツ 2 は音楽コンテンツ 1 に対して矢印タグ A の関係にある」，つまり，「音楽コンテンツ 2 と矢印タグ A の関係にあるのが音楽コンテンツ 1 である」ということを意味する．矢印タグが表現する関係は，明示的でも暗黙的でもよく，客観的でも主観的でもよい．例えば「こちらを踊ってみた」という矢印タグが，音楽コンテンツ 1 から音楽コンテンツ 2 に向けて付与されている場合，音楽コンテンツ 2 の楽曲を誰かが踊るような内容を，音楽コンテンツ 1 が含んでいるという関係があることがわかる．それが「P の次回作」[*1] という矢印タグの場合には，このクリエータが，音楽コンテンツ 1 の次に音楽コンテンツ 2 を創作したということがわかる．

様々な矢印タグが付与された音楽コンテンツ群は，各ノードが音楽コンテンツである有向グラフ (directed graph) になり，ノード間の辺が矢印タグとなる．矢印タグは，二つの音楽コンテンツ間のみに付与できる．方向性のない関係は他の既存手段でも表現しやすいため，本研究が命名して導入する「矢印タグ」では，必ず方向性のある関係を扱うこととする．したがってデータ構造としては「始点となる音楽コンテンツ，関係の名称，終点となる音楽コンテンツ」の三つ組となる．同じ二つの音楽コンテンツ間に複数種類の矢印タグを付与してもよいし，同じ名称の矢印タグを音楽コンテンツ間に逆方向に二つ付与してもよい．ただし，重複する矢印タグは付与できない．

2.1 矢印タグが表現できる音楽コンテンツ間の関係性

矢印タグは，類似関係や派生関係等の，音楽コンテンツ間の様々な関係を表現できる潜在的な能力を持っている．まず類似関係は，「こちらと歌詞のテーマが同じ」，「こちらと雰囲気が似ている」のように，「こちらと ... が同じ」や「こちらと ... が似ている」といった矢印タグで表現できる（「こちらと」は省略してもよい）．次に派生関係は「こちらを歌ってみた」，「こちらを演奏してみた」，「こちらを踊ってみた」，「こちらに PV[*2]をつけてみた」のような，「こちらを ... してみた」や「こちらに ... してみた」といった矢印タグで表現できる．最初の三つは，ニコニコ動画上の音楽コンテンツの代表的なジャンルを表すカテゴリタグの「歌ってみた」，「演奏してみた」，「踊ってみた」を表現した例である．「こちらを」は方向性をわかりやすくするために頭に追記した．音楽コンテンツのクリエータ自身が，影響を受けた相手に対して「こちらにインスパイアされました」や「こんな曲を目指しました」のような矢印タグを入力することも考えられる．他にも楽曲のアーティスト（音楽コンテンツのクリエータ）同士の人間関係は，例えば，「こちらと作者が同じ」，「こちらとギタリストが同じ」，「P（クリエータ）同士が親しい」のような矢印タグで表現できる．

視聴者が勝手に感じるような関係を，矢印タグで付与してもよい．例えば，ある楽曲が他の楽曲のアンサーソング（返答として創作された歌）のように感じたら，「こちらの曲へのアンサーソングと勝手に思ってみた」という矢印タグを付与してもよいし，楽曲の雰囲気が似ていて続けて聴くと心地よい曲を見つけたら，その曲へ向けて「次にこち

*1 ここでの「P」はプロデューサーの頭文字を意味する．ニコニコ動画では動画のクリエータ・投稿者が「（名前）P」というニックネーム（これを「P 名」という）で呼ばれることが多いことから，P がクリエータを意味するようになった．

*2 PV とは，プロモーションビデオ (promotion video) の略で，音楽連動動画（ミュージックビデオやダンス動画等）の一種である．元々は主にポピュラー音楽の販売促進（プロモーション）を目的として制作されることが多く，その目的に向いた内容・形式を持つ動画が PV と呼ばれる．

© 2012 Information Processing Society of Japan

らを聴きたい」という矢印タグを付与してもよい．自分が好きなクリエータが新曲を作ったら，「次回作はこちら」や「P の次回作」という矢印タグを付与し，他の人たちも気づけるようにして応援することも考えられる．一般にソーシャルタギングでは，主観的なタグも客観的なタグも両方書ける特長を持つが [13]，このように矢印タグでも，主観的な関係と客観的な関係が混在することが重要である．それによりユーザは気軽に矢印タグを付与することが可能となり，その中で便利な矢印タグがあれば多くのユーザに真似されて広まっていくことが期待できる．

このように矢印タグは幅広く活用できるが，以上はあくまで例に過ぎず，実際に矢印タグでどう表現するかはユーザの自由である．矢印タグで表現しやすい関係と，表現しにくい関係があることも考えらし，ユーザの創意工夫で思いもしない矢印タグが登場するかもしれない．音楽視聴支援サービス Songrium を多くのユーザに使ってもらうことで，そうした矢印タグの様々な可能性が広がっていくことを我々は期待している．

2.2 矢印タグの普及が切り拓く世界

自分が好きな音楽コンテンツから出て行く矢印タグや，逆に向かってくる矢印タグがある場合，その先の音楽コンテンツに関心を持つのは自然であり，矢印タグを次々と辿りながら音楽コンテンツを視聴したくなる．それが未知の音楽コンテンツとの出会いにつながり，様々な関係性に気付くこともできるようになる．

そして，矢印タグを誰もが自由に付与し，様々な音楽コンテンツ間に多様な矢印タグが増えて普及していくことで，さらに図 2 に示す 5 つのメリットが生じることを我々は期待している．以下では，これらを順に説明する．

2.2.1 音楽コンテンツの位置づけの把握

音楽コンテンツに出入りする矢印タグを見ることで，他の音楽コンテンツに対してどういう位置づけにあるのかが，視聴する前に把握できるようになる．特に矢印タグが増えると効果的である．例えば，図 2 (1) では，音楽コンテンツ 1 は，音楽コンテンツ 2〜4 のすべてから矢印タグ A の関係で指されている特別な位置づけにあることがわかる．もし矢印タグ A が，「こちらを歌ってみた」や「こちらを踊ってみた」だった場合には，音楽コンテンツ 1 は歌いたくなったり踊りたくなるような魅力を持つ楽曲であるとわかる．「こちらを伴奏としてお借りしました」だった場合には，音楽コンテンツ 1 は伴奏として使えるような楽曲である可能性が高い．

ある音楽コンテンツに出入りする矢印タグの数が多ければ，人気のある音楽コンテンツである可能性も高くなる．ニコニコ動画の場合，再生回数やマイリスト数（お気に入りリストへの登録数）なども人気度を測る指標として有用だが，矢印タグの数もそれらとは違う側面で人気度の参考

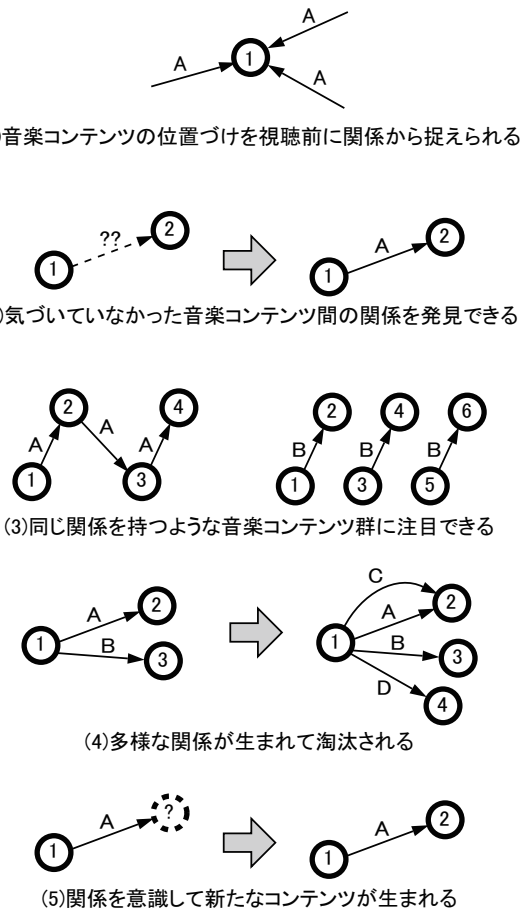

図 2 矢印タグの普及によってもたらされる 5 つのメリット

になる．

ソーシャルタグでもある程度は位置づけが把握できるが，あくまでもその音楽コンテンツ一つに関する記述である．一方，矢印タグは他の音楽コンテンツとの関係に関する記述であるため，把握できる位置づけの質が異なるのが特長である．

2.2.2 気付いていなかった関係の発見

様々な矢印タグを見ているうちに，これまで関係があることに気付いていなかった音楽コンテンツ間に，ある矢印タグの関係があることに気付けるようになる．図 2 (2) では，矢印タグ A を既に知っていることで，音楽コンテンツ 1 と 2 の間にも同じ関係の矢印タグ A を付与すればよいことに気付ける例を示している．

このように関係に名前を付けて矢印タグとして共有することで，人々がその関係に容易に気付くことができ，そこで付与した矢印タグがさらに広まることにつながっていく．

2.2.3 同じ関係を持つ音楽コンテンツ群に注目

ある矢印タグの関係に興味をもったら，それと同じ矢印タグが付与されている音楽コンテンツ群に注目して視聴することが可能になる．図 2 (3) の左側では，音楽コンテンツ 1〜4 がすべて矢印タグ A（例えば「次にこちらを聴きたい」）で関連づけられており，矢印タグ A があるからこそ，これら一連の音楽コンテンツ群に注目できる．このよ

うに矢印タグの共通性は，一種のグルーピング手段として使える．

さらに図 2 (3) の右側では，矢印タグ B の関係にある，音楽コンテンツ 1 と 2 のペア，3 と 4 のペア，5 と 6 のペアに注目し，それだけを抜き出して視聴できる．例えば，矢印タグ B が「こちらをアレンジしてみた」の場合，アレンジ前後を聞き比べることが，様々なペアに対して可能になる．

2.2.4 関係の種類が増加

様々な考えのユーザがボトムアップに関係を発見・考案して，矢印タグとして自由に追加していくことで，新たな矢印タグが生まれ，その種類が増え続けていくことが期待できる．音楽コンテンツ間の関係を，事前に網羅的に定義することは困難である．人によって音楽コンテンツの捉え方は違い，その違いも新たな矢印タグを生む要因となる．そうして図 2 (4) のように，矢印タグの種類は A, B, C, D と増えていく．

ただし矢印タグを付与したからといって，それを他の人たちが利用するとは限らない．音楽コンテンツ間の関係を的確に表現した矢印タグは，幅広く受け入れられて，不特定多数の人々によって様々な箇所に付与してもらえる可能性がある．一方，誤っていると思われれば，削除されるかも知れない．こうして淘汰されながら矢印タグの種類は増え続け，便利な矢印タグは繰り返し利用されるようになり，音楽コンテンツ間の関係をより豊かに表現していくと考えられる．

2.2.5 新たな音楽コンテンツ創作を触発

これまでの説明では音楽コンテンツが先に生み出され，そこに関係を矢印タグとして付与することを前提に述べてきたが，矢印タグが今後普及すれば，ある矢印タグに触発されて新たなコンテンツが生まれる可能性がある．例えば図 2 (5) では，矢印タグ A の関係を意識することで，既存の音楽コンテンツ 1 に対してちょうど矢印タグ A の関係になるように，新たな音楽コンテンツ 2 を創作する様子を示している．

これまでもニコニコ動画では，「踊ってみた」のようなタグが普及した結果，そのタグを意識してある楽曲を自分で踊る音楽コンテンツを作る，という現象が既に起きている [14]．矢印タグによって，様々な関係が人々に意識されることで，それが新たな音楽コンテンツの創作にも貢献できればと考えている．

3. 多様な関係性に基づく音楽視聴支援サービス Songrium の機能

Songrium (http://songrium.jp) は，矢印タグのソーシャルアノテーションを可能にする音楽視聴支援サービスである．Web 上で視聴できる音楽コンテンツ間の関係を，ユーザが矢印タグとして自由に付与し，それを辿りながら音楽コンテンツを視聴できる．Songrium 公開初期の段階では，ニコニコ動画上の音楽コンテンツのうち，歌声合成技術 VOCALOID を用いた音楽コンテンツを中心に扱う．それらのコンテンツでは特に，あるオリジナル楽曲の音楽コンテンツが一次創作として起点となって，その楽曲を歌ったり，踊ったり，映像化したりした音楽コンテンツが二次創作，N 次創作 [15] として，さらに投稿されることが多い．そこで Songrium 上では，オリジナル楽曲を含む音楽コンテンツを**一次コンテンツ**，それを元に N 次創作として作られた多様な音楽コンテンツを**派生コンテンツ**と呼び，大きく二つに分けて扱う．

Songrium では，様々な関係をユーザが矢印タグで辿りながら次々に音楽コンテンツに出会うことで，さらにはそれらの関係を矢印タグとして自由に追加することで，従来の音楽情報検索・推薦とは異なる音楽視聴支援を実現することを狙っている．矢印タグの一部は自動抽出できるが，基本的にはユーザに付与してもらえるのを待つ必要がある．派生コンテンツはその定義上，一次コンテンツへ矢印タグを持ち，ニコニコ動画上の書誌情報等から自動抽出しやすい．しかし，一次コンテンツ間の矢印タグは最初の段階では乏しく，矢印タグに基づいて視聴するだけでは辿り着けない一次コンテンツも多い．

そこで Songrium では，各一次コンテンツを信号処理して分析し，その曲調（雰囲気）を表す音響特徴量を抽出した後に，その類似度から一次コンテンツ間の類似度を求める．その結果に基づき，曲調が似た一次コンテンツ同士が近くになるように，それらのノードを 2 次元平面上に配置する．これにより 2 次元平面上のノードをブラウジングすることで，矢印タグがない状態でも任意の一次コンテンツを視聴しやすくなる．Songrium のこの 2 次元平面は，一次コンテンツの各ノードが星に見え，矢印タグで繋がれた様子が星座に見えることがあるため，**音楽星図**と名付けた．

以下では，一次コンテンツの音楽星図を可視化する機能，一次コンテンツの詳細情報を可視化する機能，そして矢印タグを付与する機能について述べる．

3.1 音楽星図の表示機能

音楽星図は，音楽コンテンツに出会うために多様な関係性を辿る上で基本となるインタフェースである（図 3）．音楽星図には，Songrium に登録されている全ての一次コンテンツが音響特徴量の類似度に基づいて自動配置されている．ただし，一度に全てを表示するとユーザが理解しずらいため，再生回数が多いコンテンツだけが大きくアイコン付きで表示され，再生回数が少ないと段階的に小さく表示される．

ドラッグによるスクロールとズーミングにより，音楽星図の様々な場所を見ながら一次コンテンツを俯瞰できる．

図 3 Songrium の音楽星図．5 万件超の一次コンテンツが曲調に基づいて 2 次元平面上に配置されており，ドラッグによるスクロールとズーミングにより俯瞰できる．図中の軌跡および画面下部のアイコンリストは直近の視聴履歴を示している．

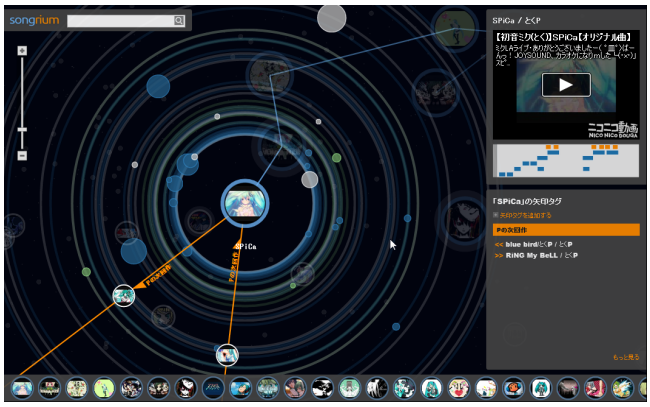

図 4 一次コンテンツの詳細表示画面．選択した一次コンテンツが中心に表示され，派生コンテンツがその周りを回る衛星のように表示される．画面右には再生プレーヤと音楽地図，矢印タグが表示される．

各一次コンテンツの近くには曲調（音響特徴量）が似た一次コンテンツ群が配置されているため，矢印タグだけではなく音楽星図上に表現される音響特徴量の類似関係を辿って視聴することができる．探したい一次コンテンツがあれば検索もできる．検索語を入力すると検索結果がリストアップされ，選んでクリックをすれば，その一次コンテンツが中央にくるように自動スクロールする．

音楽星図の広大な空間で矢印タグを辿りながら視聴していくと，全体のどのあたりを見ているのかわからなくなることがある．そこで Songrium では過去数回の視聴履歴が，音楽星図上にノード間のパスとして表示される．図 3 上の水色の軌跡はユーザが辿った一次コンテンツを示している．また，画面下部には常に視聴した履歴が表示されているので，いつでも過去に見た音楽コンテンツへ移動することができる．

3.2 一次コンテンツの詳細表示機能

音楽星図上の一次コンテンツのノードをクリックして選択すると，その詳細情報が表示される（図 4）．一次コンテンツは画面中央に表示され，その周囲には派生コンテンツの丸アイコンが衛星のように等速で回転しながら表示される．画面右には動画視聴用プレーヤ（ニコニコ動画の外部プレーヤ），コンテンツの内部構造（サビ区間等）を示した音楽地図 [16], [17]，そして他の一次コンテンツとの関係を示す矢印タグが表示される．

他の一次コンテンツとの矢印タグには，タグの名称と向きが示されている．矢印タグにつながった一次コンテンツをクリックすると，その一次コンテンツへフォーカスが移る．音楽星図は曲調の類似度に基づいて作られているため，近傍へ移動するだけでは，似た一次コンテンツばかりを聞くことになってしまう．矢印タグはいわば道路交通網における高速道路のような役割を果たし，音楽星図上の離れた場所へ瞬時に行き来するのを助ける．

この矢印タグを活用してユーザは探索的にいろいろな音楽コンテンツを視聴する．そのような視聴スタイルにおいては，じっくり視聴するだけでなく，手軽に試し聴きできることも重要である．そこで Songrium に，サビ出し機能付き音楽試聴インタフェース SmartMusicKIOSK[16], [17] の機能を搭載し，サビ区間と繰り返し区間を示した音楽地図を見ながら，サビ区間をクリックして容易に再生できるようにした．これによりユーザは音楽コンテンツ内の様々な箇所を迅速に試聴可能になった．

各派生コンテンツの丸アイコンは，その属性に応じて回転半径やサイズ，回転速度，色が異なる．回転半径（中心からの距離）は，一次コンテンツが投稿されてから派生コンテンツが投稿されるまでの期間である．古い派生コンテンツほど中央寄りに表示される．丸アイコンのサイズ（半径）は人気度を示しており，現在の実装ではニコニコ動画上での再生回数とした．丸アイコンの回転速度は注目度を示しており，現在の実装では動画のマイリスト数÷再生回数の値を用いた．丸アイコンの色は矢印タグによって示される派生関係の種類を表している．図 4 では青色の派生コンテンツが目立つが，青色は「こちらを歌ってみた」という矢印タグを持つ動画である．一方で図 4 では赤色の派生コンテンツが目立つ．赤色は「こちらを踊ってみた」という矢印タグを持つ動画である．こうした派生コンテンツの数や種類は，CGM コンテンツにおいて視聴回数とは違う価値を示す指標として有用である．Songrium では上記のように，その様子をユーザが直感的に把握できるように工夫している．

3.3 矢印タグ機能

一次コンテンツの詳細表示をしているときに，「矢印タグを追加する」と書かれたリンクをクリックすると，図 5 のような矢印タグの入力インタフェースが表示される．そこでユーザは検索語を入れて，矢印タグを付与したい音楽

© 2012 Information Processing Society of Japan

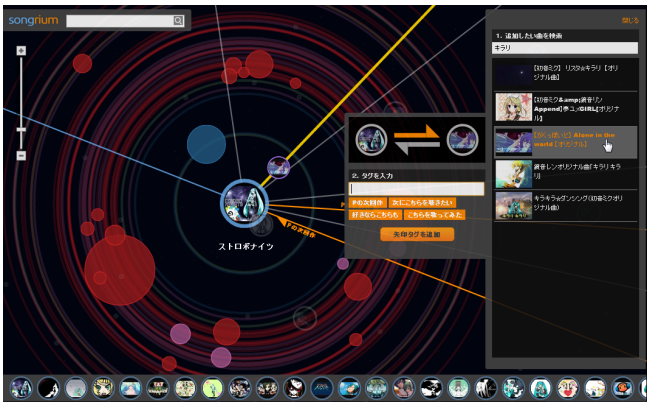

図 5 矢印タグの入力インタフェース．閲覧中の音楽コンテンツから矢印タグをつける音楽コンテンツを選択し，タグの名称および向きを入力する．入力フォーム下部にはすでに入力されている矢印タグが例示を兼ねてサジェストされる．

コンテンツを検索し，選択する．すると，図 5 に示されるように矢印タグの名称と向きを入力するインタフェースが表示される．デフォルトでは現在見ている一次コンテンツから，検索した音楽コンテンツへ向かう矢印タグが追加されるが，ここで逆方向や両方を選択することも可能である．名称は入力フォームに直接タイプ入力してもよいが，入力フォーム下部には，既に入力済みの矢印タグの名称候補がサジェストされる．これによりユーザの負荷を軽減しつつ，意図しない名称の揺れを防ぐ．名称の付け方に制約はないが，Songrium 側で事前に定義した一次コンテンツと派生コンテンツとの関係を示す矢印タグ（例えば「こちらを歌ってみた」「こちらを演奏してみた」など．詳しくは 4.1 節にて述べる．）だけは，Songrium 上での表示方法に影響を与える特殊な矢印タグとして扱われる．

矢印タグの数が増えてきた場合にどれを優先的に表示するかという問題がある．これはソーシャルタギングにおけるタグの自然淘汰をどうデザインするかという問題ともいえる．多くの場合，2 つのアプローチのいずれかがとられている．一つはユーザはそれぞれ別々にタグ付けを行い，多くのユーザが付けたタグが優先されるアプローチである．主にソーシャルブックマークや写真共有サイトで用いられている方法である．もう一つはユーザは誰でもタグを編集することができ，最後まで残っていたタグが優先されるアプローチである．Wiki の共同編集に似ており，ニコニコ動画や facebook の写真アルバムなどで用いられている．前者は特定の個人に依存せず，賛成の集積（多くの人がそのタグをつける）によって全体にとって重要なタグが決定される．一方で後者は特定の個人の発想に対して，反対（タグの書き変え）がないことによって決定される．Songrium では関係が作られていくこと，更新されていくことに期待しているため，更新が容易な後者のアプローチをとる．具体的には一番最後に作られたまたは利用された（ユーザがその矢印タグを辿って移動した）矢印タグを優先的に表示する．

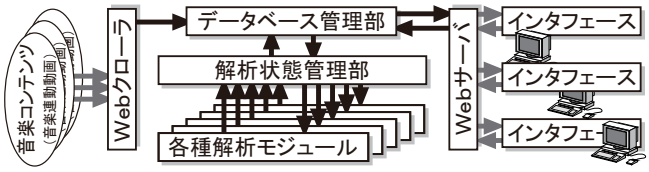

図 6 Songrium のシステム構成図

4. Songrium の実装

Songrium のシステム構成図を図 6 に示す．Web クローラはニコニコ動画から音楽コンテンツを収集し，データベースに登録する．次に各種解析モジュールが，各音楽コンテンツを処理する．例えば，楽曲構造（サビ区間抽出）と音響特徴量は別々のモジュールが推定する．処理が終わった解析モジュールから解析状態管理部へリクエストがあると，次に処理すべき音楽コンテンツが引き渡される．解析モジュールがその推定処理を終えると，推定結果は解析状態管理部を経てデータベース管理部に渡される．データベース管理部では，その推定結果を保存し，処理状態を管理する．最後に，Web サーバは，Songrium のインタフェースを提供する Web サイトとして動作する．なおユーザが音楽コンテンツを視聴する際には，元のニコニコ動画の Web サイトから配信された動画が，Songrium を経由せずに直接ユーザのブラウザ上で再生される（ニコニコ動画の外部プレーヤを埋め込んで実装した）．

一連の機能のサーバ側動作は主に，Web アプリケーションフレームワーク Ruby on Rails，プログラミング言語 Ruby，Web サーバ Apache，データベース MySQL を用いて実装した．Web クローラのみプログラミング言語 PHP を用いて実装している．クライアント側のユーザインタフェース機能は，スクリプト言語 JavaScript を用いて実装した．可視化には M. Bostock らが開発した可視化ライブラリ d3.js [18] を用いている．

4.1 コンテンツの分類と関係の抽出

Songrium ではニコニコ動画のタグ検索機能を用いて VOCALOID およびそれに関連する音楽コンテンツを収集している．具体的には「VOCALOID」「ミクオリジナル曲」「歌ってみた」「MikuMikuDance」などのタグを用いてタグ検索をし，Songrium で扱う音楽コンテンツ候補を収集する．収集した音楽コンテンツ候補はまず一次コンテンツと派生コンテンツとそれ以外とに自動分類する．それ以外とは音楽コンテンツではないものであり，Songrium では扱わない．分類に用いる特徴量には動画に付与されたタグや説明文，説明文中のハイパーリンクなどを利用する．他にも自動分類に参考になりそうな補足情報は，様々な Web サイトから収集している．

© 2012 Information Processing Society of Japan

次に，一次コンテンツと派生コンテンツとの関係を推定する．Songrium が対象とする派生コンテンツには，一次コンテンツを利用した様々なパフォーマンス（歌う，踊るなど）や編集・加工したコンテンツ（アレンジや PV 化など）がある．その関係は多岐にわたり，正確な関係を自動推定することは困難であった．そこで関係の種類を，予備調査に基づいて以下の 7 種類の矢印タグに限定し，分類した．

- 「こちらを歌ってみた」楽曲を歌った音を収録した音楽コンテンツ．
- 「こちらを踊ってみた」楽曲に合わせて踊った映像を収録した音楽コンテンツ．
- 「こちらを演奏してみた」楽曲を演奏した音を収録した音楽コンテンツ．
- 「こちらをアレンジしてみた」楽曲をアレンジした曲を収録した音楽コンテンツ．
- 「こちらに PV をつけてみた」映像を追加した音楽コンテンツ．
- 「こちらに MMD をつけてみた」MMD[*3]で作成した映像を追加した音楽コンテンツ．
- 「こちらの曲を使ってみた」上記のいずれにも当てはまらない音楽コンテンツ．

この 7 種類の分類では，一次コンテンツと派生コンテンツの双方のニコニコ動画上でのタグや説明文中のキーワードから，それらの関係がどの種類に分類されるかを判定した．この 7 種類は暫定的な分類であり，Songrium 運用中の矢印タグの利用状況から，より適切な分類に改良していければと考えている．

4.2 音楽星図の作成

一次コンテンツの音響特徴量ベクトルに対して主成分分析を用い，寄与率の高い成分を音楽星図上の座標として用いた．曲調の音響特徴量を抽出するためには，MARSYAS [19] を利用した．まず，音楽コンテンツ中の音響信号から 35 次元の音響特徴量ベクトルを求めた．その内訳はスペクトル特徴量（セントロイド，ロールオフ，フラックス）およびゼロクロス率の曲全体にわたる平均と分散で 8 次元，13 次元メル周波数ケプストラム係数の平均と分散で 26 次元，サビ区間のテンポで 1 次元である．次に，主成分分析を用いて得られた第 1 主成分と第 2 主成分をそれぞれ横軸と縦軸の座標に用いた．

4.3 楽曲構造の推定

サビ区間は，ポピュラー音楽に対するサビ区間検出手法 RefraiD [16][17] を用いて推定した．RefraiD は，様々な繰り返し区間の相互関係を調べることで，転調の有無に関わらず，楽曲中で繰り返されるすべてのサビ区間を網羅的に検出しようとする特長を持つ．

[*3] MikuMikuDance（MMD）は，キャラクタを中心とした動画制作を容易にする 3 次元 CG（コンピュータグラフィクス）制作用ソフトウェアで，VOCALOID に関連した CGM 現象の発展に大きな貢献を果たしている．

5. 関連研究

Songrium は多様な機能を持つサービスであり，様々な関連研究に対して位置づけられるが，以下では音楽のアノテーションと可視化の観点から，関連研究に対する位置づけを議論する．

5.1 音楽のアノテーション

音楽コンテンツ間の関係性や，音楽コンテンツ自身が持つ特徴を記述するためのオントロジーが提案されている [20]．近年では，このような関係性や特徴を記述するための語彙を整備し，それを用いて記述されたデータを公開する Linked Open Data (LOD) という動きが注目されている．LOD は音楽に限らず，政府，学術，地理など様々なデータを対象とした取り組みであるが，音楽は MusicBrainz や Last.fm，BBC Music などが数億トリプル（主語，述語，目的語の三つ組からなるデータ）を公開しており，LOD においても主要なデータカテゴリの一つとなっている．矢印タグはラベル付きの関係を記述し共有しようとする点で SW と似ているが，SW がデータの整理や再利用性を高めることを目的とし，整備された関係ラベルを作ろうとしているのに対し [21]，矢印タグはユーザが自由に関係ラベルを作成し，ユーザのコンテンツ視聴をより豊かにすることを目指している点で異なる．Web のハイパーリンクもラベル付き有向グラフを作るという点では同じだが，作者がコンテンツクリエータに限定される点，ラベル（アンカーテキスト）の共有を促す仕組みを持たない点で異なる．

アイテムに対してユーザに自由にタグをつけさせるソーシャルタギングは，数多くのサービスで採用され，また様々な研究がなされている．ラベル単体ではなく，属性名と属性値のペア情報をタグ付けするソーシャルタギングも提案されている [22][23][24]．しかしいずれもアイテムに対してタグ付けを行い，アイテム間の関係にタグをつける矢印タグとは異なる．

5.2 音楽の可視化

複数の音楽コンテンツやアーティスト，ジャンルなどを 2 次元平面上にマッピングして可視化する方法は数多く提案されている．我々のように類似した音楽内容を持つ音楽コンテンツやアーティストが互いに近くなるように 2 次元平面上で配置するものが多い．例えば従来研究では配置アルゴリズムとして，MDS [25] や主成分分析 [9]，SOM [26]，バネモデル [27] などが用いられてきた．

Songrium では 2 次元平面による可視化に加え，矢印タ

© 2012 Information Processing Society of Japan

グによる視点移動が可能となっている．大規模な音楽コンテンツを対象とすると，2 次元平面を連続的に移動するだけでは見られる範囲が局所的になってしまう可能性がある．Songrium のズーミングインタフェースと矢印タグの組み合わせは，閲覧範囲が局所的に陥らないような視点移動を可能にする．また，2 次元平面による可視化では，扱える関係性が距離に落とし込みやすいものに制限されてしまうが，Songrium では矢印タグによって多様な関係性を扱うことができる．

6. おわりに

本稿では，音楽コンテンツ間の方向性のある関係に対してタグを付与する「矢印タグ」を提案し，それをニコニコ動画上の音楽コンテンツ（VOCALOID 関連動画）の視聴で活用できる音楽視聴支援サービス Songrium を提案した．今後は試験公開を開始して，実際に矢印タグのソーシャルアノテーションが可能な状況を継続しながら，より有用なサービスを目指して改善を重ねていく予定である．そうして Songrium を多くのユーザに使ってもらうことで，矢印タグが持つ可能性を明らかにしていきたい．また，従来のタグが CGM 現象の分析に大きな役割を果たしたように，矢印タグも CGM 現象のさらなる分析に貢献できればと考えている．

謝辞 Songrium の Web サービスの実装を担当して頂いた石田 啓介 氏に感謝する．本研究の一部は，科学技術振興機構 OngaCREST プロジェクトによる支援を受けた．

参考文献

[1] 戀塚昭彦：ニコニコ動画の創造性: 動画コミュニティサービス「ニコニコ動画」の 5 年間，情報処理 (情報処理学会誌)，Vol. 53, No. 5, pp. 438–488 (2012).
[2] 後藤真孝：初音ミク，ニコニコ動画，ピアプロが切り拓いた CGM 現象，情報処理 (情報処理学会誌)，Vol. 53, No. 5, pp. 466–471 (2012).
[3] 後藤真孝：音楽情報検索，電子情報通信学会「知識ベース」2 群 9 編 1-6，電子情報通信学会 (2012).
[4] Grosche, P., Müller, M. and Serrà, J.: Audio Content-Based Music Retrieval, *Multimodal Music Processing* (Müller, M., Goto, M. and Schedl, M., eds.), Dagstuhl Follow-Ups, Vol. 3, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, pp. 157–174 (online), DOI: http://dx.doi.org/10.4230/DFU.Vol3.11041.157 (2012).
[5] 吉井和佳，後藤真孝：音楽推薦システム，情報処理 (情報処理学会誌)， Vol. 50, No. 8, pp. 751–755 (2009).
[6] Celma, O.: *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space*, Springer (2010).
[7] Goto, M. and Goto, T.: Musicream: Integrated Music-Listening Interface for Active, Flexible, and Unexpected Encounters with Musical Pieces, *IPSJ Journal*, Vol. 50, No. 12, pp. 2923–2936 (2009).
[8] 吉谷幹人，宇佐美敦志，浜中雅俊：BandNavi: バンドメンバーの変遷情報を辿るアーティスト発見システム，WISS 2010 論文集，pp. 29–34 (2010).
[9] Saito, Y. and Itoh, T.: MusiCube: a visual music recommendation system featuring interactive evolutionary computing, *Proc. of VINCI '11*, pp. 5:1–5:6 (2011).
[10] 剣持秀紀：歌声合成の過去・現在・未来: 「使える」歌声合成のためには，情報処理 (情報処理学会誌)， Vol. 53, No. 5, pp. 472–476 (2012).
[11] Hamasaki, M., Takeda, H. and Nishimura, T.: Network Analysis of Massively Collaborative Creation of Multimedia Contents - Case Study of Hatsune Miku videos on Nico Nico Douga -, *Proc. of uxTV2008*, pp. 165–168 (2008).
[12] 濱崎雅弘，武田英明，西村拓一：動画共有サイトにおける大規模な協調的創造活動の創発のネットワーク分析-ニコニコ動画における初音ミク動画コミュニティを対象として-，人工知能学会論文誌，Vol. 25, No. 1, pp. 157–167 (2010).
[13] Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F. M. and Riedl, J.: tagging, communities, vocabulary, evolution, *Proc. of CSCW '06*, pp. 181–190 (2006).
[14] 濱野智史：ニコ二コ動画の生成力，思想地図 vol.2：ジェネレーション，NHK 出版 (2008).
[15] 濱野智史：ニコニコ動画はいかなる点で特異なのか: 「擬似同期」「N次創作」「Fluxonomy（フラクソノミー）」，情報処理 (情報処理学会誌)，Vol. 53, No. 5, pp. 489–494 (2012).
[16] 後藤真孝：SmartMusicKIOSK: サビ出し機能付き音楽試聴機，情報処理学会インタラクション 2003 論文集，pp. 9–16 (2003).
[17] Goto, M.: A Chorus-Section Detection Method for Musical Audio Signals and Its Application to a Music Listening Station, *IEEE Transaction on ASLP*, Vol. 14, No. 5, pp. 1783–1794 (2006).
[18] Bostock, M., Ogievetsky, V. and Heer, J.: D3: Data-Driven Documents, *Proceedings of InfoVis 2011* (2011).
[19] Tzanetakis, G. and Cook, P.: MARSYAS: A Framework for Audio Analysis, *Organized Sound*, Vol. 4, No. 3, pp. 169–175 (2000).
[20] Raimond, Y., Abdallah, S., Sandler, M. and Giasson, F.: The Music Ontology, *Proc. of ISMIR 2007*, pp. 417–422 (2007).
[21] Heath, T. and Bizer, C.: *Linked Data: Evolving the Web into a Global Data Space*, Morgan & Claypool (2011).
[22] 濱崎雅弘：サジェスト機能によるゆるやかなオントロジー構築を可能にするシステムの提案，第 22 回セマンティックウェブ＆オントロジー研究会 (2010).
[23] Hamasaki, M., Goto, M. and Takeda, H.: Social Infobox: collaborative knowledge construction by social property tagging, *Proc. of CSCW 2011*, pp. 641–644 (2011).
[24] Chae, G. and KimCan, J.: Social Tagging Be a Tool to Reduce the Semantic Gap between Curators and Audiences? Making a Semantic Structure of Tags by Implementing Facetted Tagging System for Online Art Museums, *Proc. of Museums and the Web 2011* (2011).
[25] Cano, P., Kaltenbrunner, M., Gouyou, F. and Batlle, E.: On the use of FastMap for Audio Retrieval and Browsing, *Proc. of ISMIR 2002* (2002).
[26] Pampalk, E. and Dixon, S.: Exploring Music Collections by Browsing Different Views, *Computer Music Journal*, Vol. 28, No. 2, pp. 49–62 (2004).
[27] van Gulik, R., Vignoli, F. and van de Wetering, H.: Mapping Music In The Palm Of Your Hand, Explore And Discover Your Collection, *Proc. of ISMIR 2004* (2004).

© 2012 Information Processing Society of Japan