

# Song Surfing: 類似フレーズで音楽ライブラリを散策する音楽再生システム

Song Surfing: Music playback system to explore a music library through similar-sounding phrases

堀内 直明, 藺田 俊行, 田中 浩司  
Naoaki Horiuchi, Toshiyuki Sonoda, Koji Tanaka

田中 淳一, 長沢 秀哉, 莪山 真一  
Junichi Tanaka, Hideya Nagasawa, Shinichi Gayama

**要旨** 近年個人でも大規模な音楽ライブラリを容易に構築し、楽しめる環境が整ってきた。しかし、楽曲名やアーティスト名などの書誌情報で楽曲を検索する従来のインターフェースでは、個人のライブラリであっても数千曲もの楽曲を把握するのは困難であった。そこで本研究では、楽曲の持つ時系列データに注目し、適度なユーザの介入を許すインターフェースと組み合わせて、大規模音楽ライブラリを容易に楽しめるシステム Song Surfing を開発した。Song Surfing を用いれば、音楽ライブラリの大きさに左右されず同じ労力でそのライブラリを楽しむことができるようになる。

**Summary** It has become relatively easy to construct a huge music library, but the existing interfaces for searching do not allow the user to fully enjoy the variety of songs in the library. Generally, the bigger the library is, the harder it becomes to select songs. In this paper, we describe Song Surfing, a new interface for a huge music library. First, we explain how our system analyzes songs. Second, we present the system to explore a music library through similar-sounding phrases. Finally, we discuss the usability of this system through an experiment with a library of 300 songs. Song Surfing allows exploration of the library regardless of its scale.

**キーワード** : 大規模音楽データベース, 音楽特徴量, 自動連続再生, インターフェース, 類似曲検索

## 1. まえがき

近年、ハードディスクや半導体メモリの大容量化・小型化に加え、定額制の音楽配信サービスが開始されるなど、個人でも大規模な音楽ライブラリを容易に構築し、楽しめる環境が整ってきた。しかし、楽曲名やアーティスト名などの書誌情報で楽曲を検索する従来のインターフェースでは、個人のライブラリであっても数千曲もの楽曲を把握するのは困難である。

例えば、ユーザが1時間音楽を楽しみたいと考えたとしよう。1曲5分であると仮定すれば、12曲分の再生リストを作成すれば良いことになる。このとき

ユーザのライブラリが100曲程度であればライブラリを約10%まで絞り込めば済み、こういった作業をこなせるユーザも少なくないかもしれない。しかしユーザのライブラリが5000曲からなるようなものであった場合、ユーザは自身のライブラリを0.24%まで絞り込まねばならない。この絞り込み作業は、多くのユーザに苦痛を強いることになるだろう。しかもその時々ユーザの気分や環境を考慮して作成する場合、かなりの困難を伴う作業となる。したがってライブラリがある程度の規模を超えた多くのユーザは、ランダム再生やシャッフル再生と呼ばれる機能を用いるか、

ユーザが把握できる小さな範囲で再生リストを作成してライブラリを楽しむことになる。

しかし、ランダム再生では曲間のつながりが考慮されることはないため、魅力的な再生リストができるかどうかは楽曲を聞いてみるまで分からないという問題点がある。一方、ユーザが把握できる範囲で再生リストを作成すれば、楽曲間のつながりについては心配ないものの、それでは巨大なライブラリのもつ楽しさを充分活用しているとは言いがたい。つまり現状のインタフェースを利用する限り、ユーザは大規模な音楽ライブラリを保有していたとしても、そのライブラリが持つ多様性や意外性を十分に活用することができなかった。

この問題を解決するため、後藤らは Musicream<sup>(1)</sup>を開発し、楽曲を効果的に視覚化した上で、自由かつ直感的な操作によって類似楽曲の集合化や過去の聴取楽曲の再現を容易にする、大規模な音楽ライブラリのための新たなインタフェースを提案した。一般に、楽曲は特徴の異なる多様なフレーズの時間的な結合と遷移によって音楽の曲調を表現している。それに対し、Musicreamの類似楽曲の集合化は楽曲単位で扱うものであり、楽曲の内部まで踏み込んだ類似性や多様性を扱っていない。

そこで我々は、楽曲の単位より細かいフレーズの特徴に注目し、似ているフレーズを手がかりに多くの異なる楽曲を連続的に楽しむことのできる新たな音楽

システム Song Surfing を開発した。また本システムは独自のユーザインタフェースを有しており、音楽再生の様子やシステム動作を効果的に視覚化する。このインタフェースを用いることにより、ユーザが大規模音楽ライブラリを楽しむ際の作業負担を大幅に軽減できる。本稿では最初に Song Surfing の動作原理である楽曲再生手法について述べ、次に本システムの機能とユーザインタフェースについて解説する。

## 2. Song Surfing の楽曲再生手法

従来、類似する楽曲の検出によく用いられる方法としては、楽曲の特徴を平均化して扱い、曲単位で類似度を計算する技術がある。つまり5分の曲であれば、5分間にわたって変化する楽曲内の特徴量を平均値で扱っていた。そのため、曲同士の類似度を判定することはできても、ある曲の途中部分と別の曲の途中部分が類似しているかどうかの判定はできなかった。また、楽曲の変化を平均化して扱っているため、楽曲から受ける実際の印象と、計算した特徴量が乖離することも珍しくなかった。例えば図1の曲Aに近い楽曲を楽曲B-D内で検索するとする。この場合、楽曲Dのもつ平均特徴量よりも楽曲Bや楽曲Cの平均値が楽曲Aの平均特徴量と近い。したがって曲Aと曲B・Cが、楽曲Aと楽曲Dよりも近いという判断になる。しかし特徴量の遷移に注目した場合、むしろ楽曲Dの方が、

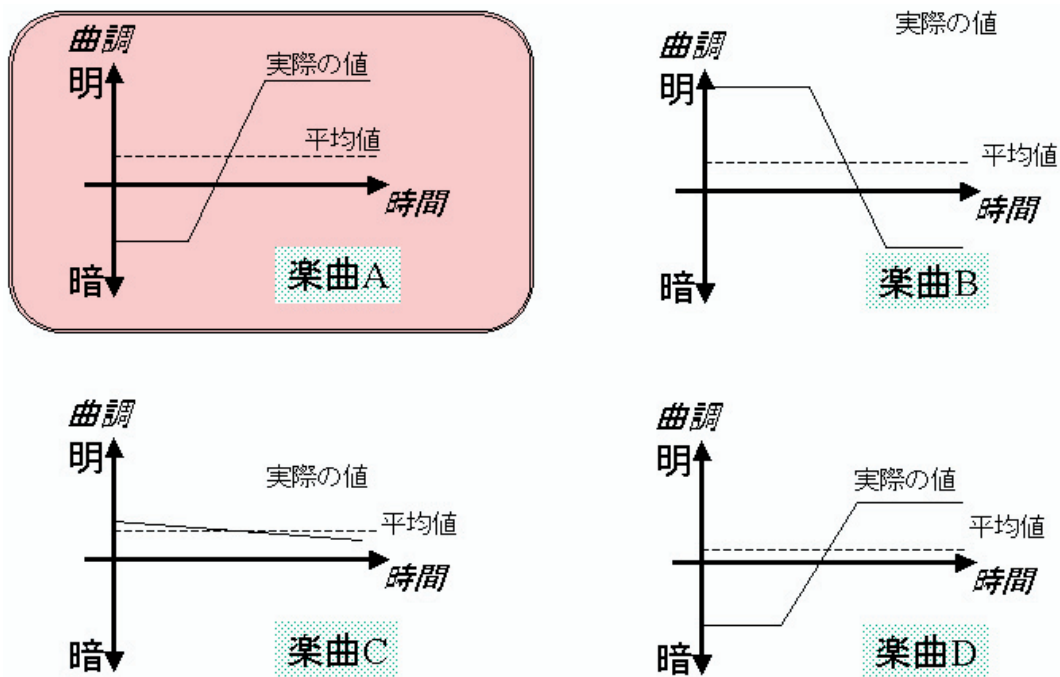


図1 従来の音楽解釈手法

楽曲 B・C よりも楽曲 A に近いと考えられる。これらの違和感は、本来楽曲が持っているダイナミックな特徴変化を数値的に平均化して扱ったため生じている。

そこで本システムでは、楽曲を時間的に結合されたフレーズの集合として扱い、類似度の計算には曲単位に平均化された特徴ではなく、各フレーズのもつ特徴量を用いた。これにより、楽曲のもつ時間的遷移を各フレーズの特徴量によって適切に表現するとともに、異なる楽曲中の互いに類似するフレーズ群の検出を可能とした。さらに我々は本技術を用いて、類似フレーズを介して異なる楽曲を違和感なく切り替えるという新たな音楽の連続再生手法を実現した。

以下に、上述した楽曲特徴量の抽出手法、ならびに我々が実現した音楽の連続再生曲決定手法について概要を述べる。なお、これらの技術については文献(2)(3)(4)においてより詳しく述べられている。

## 2.1 楽曲特徴量抽出手法

本システムでは、以下に挙げる 5 つの要素を楽曲の特徴量として採用した。

1. 和音進行 (音楽の持つ曲調の流れ)
2. リズム (音楽の持つ大まかな速さ)
3. 音量 (聴感上の音量変化)
4. 周波数重心 (音楽の盛り上がり度合い)
5. 音程パワー分散 (音楽の持つ複雑さ)

まず和音進行の抽出手法について述べる。和音進行は、ある時間幅における楽曲の時間波形を解析して抽出されている。またここで用いる和音進行とは、自動採譜などで求められるような厳密なものではなく、主要な 3 音の時系列の変化を表現したものである。なぜなら、一般に和音は 4 音以上で構成されていても概ね 3 音の組み合わせで表現可能であり、また楽曲の時間的な遷移を数値的に表現する際に、複雑な和音表現は適さないためである。

これら 3 音の抽出は、指定時間窓における楽曲の平均律音階に相当する周波数パワーを計算して行う。具体的には、楽曲の先頭からこの時間窓を移動させて同様の計算を行い、結果として時系列の和音進行データが得られる。一般に和音進行が 1 秒以内に变化し続けることは極めてまれであることから、本技術では少なくとも 1 秒間は同じ和音進行が続くという仮定に基づいてデータ生成を行っている。このため、楽曲の時間波形全てを本解釈手法で処理しても、その結果のデータ量はあまり大きくなる。したがって、類似フレーズ検索の際に大きなシステム資源を必要としな

いという利点がある。

次に残りの特徴量抽出について述べる。リズムは低域信号の振幅ピークの数と間隔から推定、周波数重心は周波数パワーの周波数重み付き平均値として求める。また、音程パワー分散は平均律音階に相当する周波数パワーを 1 オクターブ (12 音階) に射影・加算した上で、それらの分散を演算した値である。次節では、これらの特徴量を用いた楽曲の連続再生曲決定手法について述べる。

## 2.2 連続再生曲決定手法

前節で述べた特徴量は、それぞれ楽曲の時系列に沿って変化している。本システムでは、和音進行抽出で得られた和音単位 (1 つの和音が維持される時間幅) で、その他 4 つの特徴量を平均化して扱っている。ここである楽曲とつながりやすい楽曲を検索する場合は、接続元の楽曲 A が接続点以降にもつ音楽特徴量と類似した部分を持つ楽曲集合を検索する。つまり、楽曲 A が持っている進行を崩さない部分を持つ楽曲を選び出している。

例えば、図 2 において楽曲 A の接続点 p1 で別の曲につなぎたい場合、接続点 p1 以降の音楽特徴列 (b) に近い音楽特徴列を持つ楽曲を探す。ここで楽曲 B が音楽特徴列 (b) に近い音楽特徴列 (b') を持っているとするれば、接続候補として楽曲 B が選ばれる。楽曲 C には音楽特徴列 (b) に近い音楽特徴列がないので選択されない。したがって、楽曲 A の持つ音楽特徴列 (a) から楽曲 B の持つ音楽特徴列 (b') に接続すれば、ユーザに違和感なく曲を切り替えられるということになる。

このように本技術を用いた場合、楽曲の持つ時系列の変化を特徴量として持っているため、従来技術と比較してより人の感覚に近い類似楽曲検索が可能となる。また、従来技術のほとんどは曲単位での接続を目的としており、曲の途中での接続は不可能であった。しかし、本技術を用いれば、楽曲の再生位置を考慮した上で、楽曲の途中同士をスムーズにつなげられる。さらに曲単位を中心とする従来の手法では、ある大規模音楽ライブラリにおいて検索された類似楽曲のつながりやすさと、そのライブラリの大きさには強い相関がなかった。これは従来手法が楽曲全体で平均化された特徴量で類似性を判断しており、ライブラリに含まれる曲が増えれば平均化によって類似性が高くなる曲も増えるためである。しかし、本手法では楽曲を時系列データとして扱っているため、ライブラリに含まれる曲が増えれば、よりつながりやすい楽曲が検索され

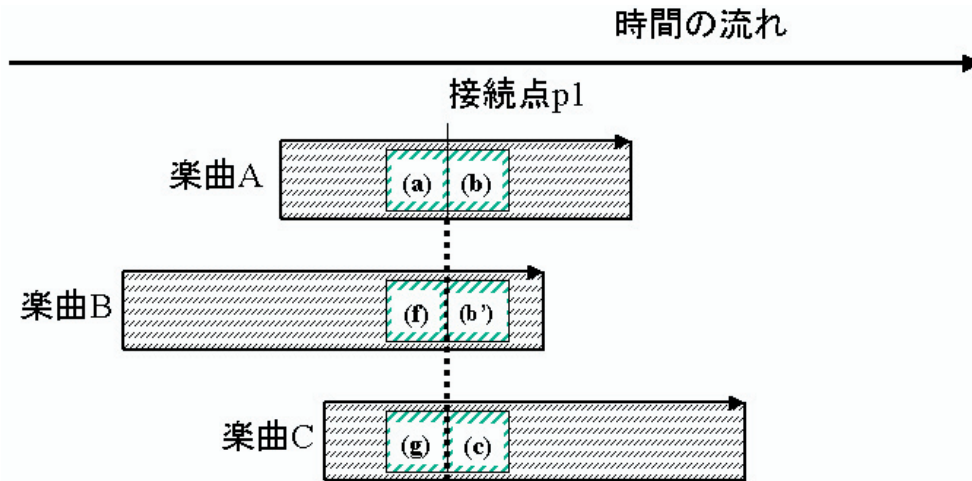


図2 楽曲接続概念

るようになる。つまり本システムでは、音楽ライブラリの大きさと検索された類似楽曲のつながりやすさに強い相関があり、大規模音楽ライブラリの持つ多様性を従来技術より引き出しやすいといえる。

次章以降では、この技術に適切なユーザインタフェースを組み合わせた Song Surfing の機能について述べる。

### 3. Song Surfing の機能とユーザインタフェース

Song Surfing は、類似フレーズを含む複数の楽曲を自動で見つけ出せるので、異なる楽曲を違和感なくつなげて連続的に再生できる。したがって Song Surfing を用いれば、大規模な音楽ライブラリを活用して楽

曲の持つ多くの多様性を引き出すとともに、あたかも DJ が演出しているかのような新鮮さと心地よさをユーザに提供できる。本章ではこれらの機能とユーザインタフェースについて解説する。ここで、図3は Song Surfing の実際の動作画面であり、図4はその画面構成図である。

#### 3.1 Auto Surfing 機能

Song Surfing は、現在再生中の楽曲における各フレーズの特徴量から次曲を自動的に選択する。この特徴量抽出については、前節で述べた。Song Surfing では、画面左端が過去、右端が未来、中央が現在を表している。中央の白い帯は音楽視聴の流れを表しており、白い帯に中央にある楽曲が現在聴いている曲 (i) であ



図3 SongSurfing 動作画面

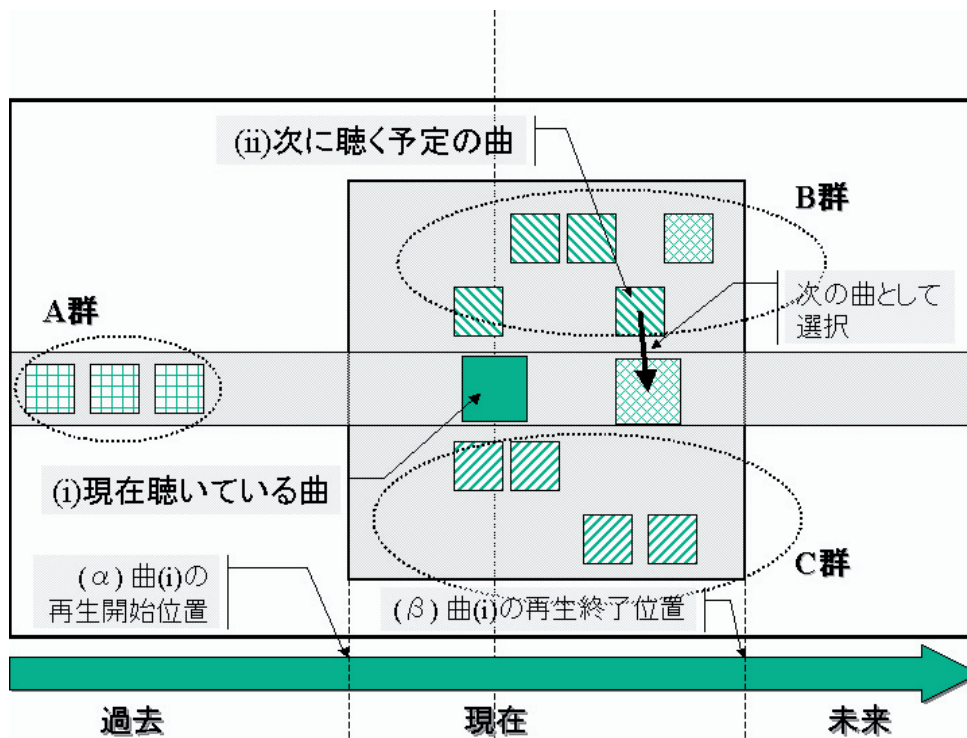


図4 SongSurfing 画面構成

る。したがって、曲 (i) のジャケット画像は再生終了するまで中央から動かない。ここで、曲 A 群は今まで聴いた曲の履歴である。そして現在再生中の曲 (i) の上部にある曲 B 群は、曲 (i) とスムーズにつながる箇所を持つ曲であり、現在再生中の曲 (i) の下部にある曲 C 群は、現在聴いている曲 (i) と特徴度が大きく違う箇所を持つ曲である。曲 (ii) は未来にある楽曲であり、曲 (i) とつながる予定の楽曲である。Song Surfing は曲 (ii) として、曲 B 群の中で最も曲 (i) に近いフレーズをもつ楽曲を選ぶ。曲 B、C 群のジャケット画像を囲む、大きなジャケット画像は曲 (i) のジャケット映像を拡大表示したものであり、このジャケット画像の左右両端が曲 (i) の開始点と終了点を表している。つまり、左端の位置 (α) が開始点を、右端の位置 (β) が終了点を表しており、曲 (i) の再生が進むに従って位置 (α)、(β) と曲 (ii) と B、C 群が未来 (右) から過去 (左) へと移動する。また、曲 (ii) の位置は現在聴いている曲 (i) とスムーズにつながる地点を表しており、曲 (ii) が曲 (i) と重なる位置まで来れば、自動的につなげて再生される。

したがって Song Surfing を用いれば、ユーザはシステムに対して明示的な指示を与えなくても音楽を楽しむことができる。さらに Song Surfing は類

似フレーズでスムーズに曲をつなぐので、ランダム再生のようにそれまでの視聴雰囲気を壊すことがない。

### 3.2 類似曲提示 / 選択機能

ユーザはシステムの自動再生に任せるだけではなく、他の候補曲を次に聴く曲として選べる。実際の操作としては、曲 B 群から次に聴きたい曲を選べばよい。ここで、曲 B 群にあるそれぞれのジャケットの表示位置は、曲 (i) のどの部分でスムーズにつながるかを表している。したがってユーザが現在視聴中の曲を知っていれば、つながる箇所から対象曲をある程度類推できる。曲 B 群から曲を選べば、その曲が音楽再生ラインを表す中央の帯上に配置される。図 1、2 では、曲 (ii) が次の曲として選ばれているが、別の曲を選べば曲 (ii) の代わりにその曲が表示される。

また、曲 B 群にある候補曲と中心の白い帯との距離にも意味がある。中心の帯と各ジャケット画像の距離が現在聴いている曲との特徴量の違いの大きさを表現している。ジャケット映像が中心の帯に近ければ、現在聴いている曲により近い特徴をもっている。したがって、現在聴いている曲が既知の曲であれば、それぞれの候補曲がどのような特徴をもつのかおおよそ想像できる。また、それぞれの候補曲が既知であれば、現在聴いている曲がどのように展開していくのかお

よそ想像できる。一般に曲をそのまま特徴度で表現した場合、ユーザにとってその曲のイメージをつかむのは難しい。しかし音楽に代表される、言葉や数値による表現が難しいコンテンツであっても、具体的なコンテンツとの類似/比較を用いればユーザにとって理解しやすい。

### 3.3 雰囲気チェンジ機能

現在聴いている曲の雰囲気を大きく変更したい場合、聞いている曲から特徴度が異なる曲を選択・再生する機能である。実際の操作としては、曲C群から次に聴きたい曲を選べばよい。この機能がランダム再生と異なるのは、ランダム再生では必ずしも曲調の変化が保証されないのに対して、本機能は抽出済みの音楽特徴列を参照利用して曲調の異なる曲を再生する点である。

本システムは従来のインターフェースと比較して、Auto Surfing 機能と類似曲選択 / 雰囲気チェンジ機能とを用いてユーザの音楽視聴に必要な労力を軽くしている。つまり、ユーザはシステムに干渉したいときにだけ干渉すればよく、現在聴いている曲に不満がなければシステムに再生を任せればよい。図5にこの視聴スタイルの概念を示す。

## 4. システムを使用感と考察

300曲からなる音楽ライブラリを本システムにて視聴し、その操作性を確かめた。起点となる曲を一度

決定すればその後はシステム任せで音楽ライブラリを楽しめるという点で、本システムは特にBGMとして音楽を楽しみたいときに有効であると感じた。ランダム再生と違い、現在感じている曲の雰囲気があまり変わらないことも、BGMとして音楽を楽しむ際に有効であった。また、常に次再生曲候補が複数表示されているため、音楽を聴いていて生じる「もう少し違う曲を聴きたい」、「全く違う曲で雰囲気を変えたい」という欲求をすぐにシステムへ伝達できた。つまり、あらかじめ予測し難い選曲衝動に柔軟に対応できるシステムであることを確認できた。さらに本システムを用いれば、次再生曲候補の数によって「現在聴いている曲と類似した部分を持つ曲」が多いのか少ないのかを確認できる。この表示によって自分のライブラリにおける楽曲の偏りを視覚的に確認できただけでなく、未所持曲への興味をかきたてられることもあった。

従来、大規模音楽ライブラリを楽しむためには「各楽曲へ属性付けを行う」、「ユーザ自身が選んで、再生リストを作成する」などといった大きな労力を必要とした。つまり、従来システムではライブラリが大きくなるにしたがって、ユーザの選曲に必要な労力が増加する傾向にあった。しかしSong Surfingでは、提示する候補曲数とライブラリの大きさとは比例せず、ライブラリが大きくなればよりつながりやすい曲が出現しやすくなるだけである。そのため大きなライブラリ

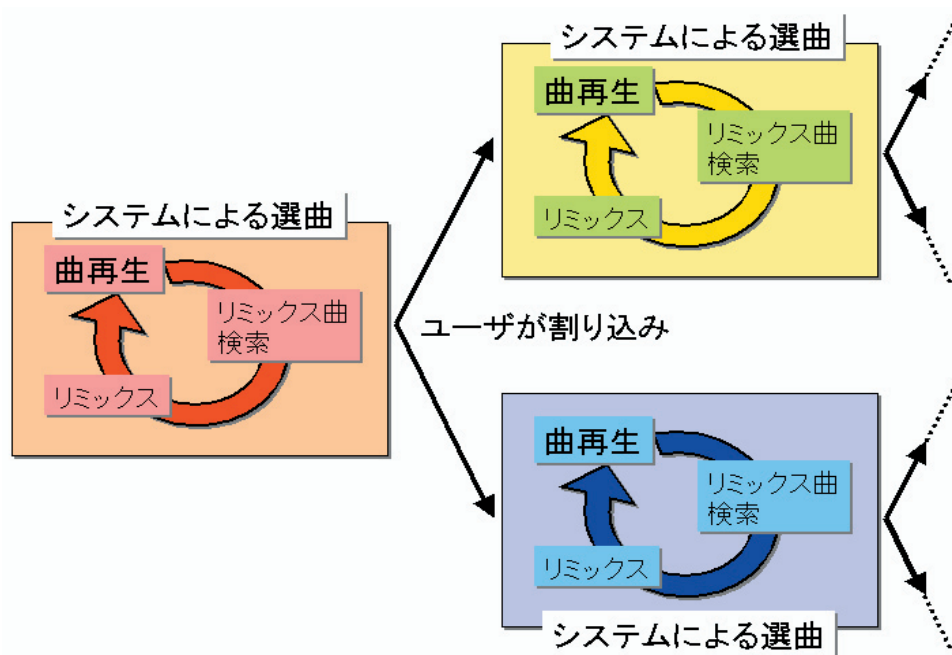


図5 視聴概念

のもつ利点を引き出しているにもかかわらず、ライブラリが大きくなった場合でも同じ操作感が保たれている。なおかつ、従来システムでは実現困難であった「自動再生であっても、現在の再生中の曲が持つ雰囲気を大きく壊さない」、「ユーザの選曲衝動に柔軟に対応する」ことを実現している。したがって Song Surfing は、従来システムと比較して大規模音楽ライブラリに極めて有効なシステムであるといえる。

個人の音楽ライブラリの大規模化が予想される今後、具体的に決まっている目的の1曲を聴く従来システムだけでは不十分である。Song Surfing のような、音楽ライブラリを簡便に取り扱うシステム/インタフェースの必要性は高まっていくだろう。

## 5. まとめ

似ているフレーズを手がかりに多くの異なる楽曲を連続的に楽しむことのできる音楽再生システム Song Surfing を開発した。さらに、本システム特有の音楽再生の様子やシステム動作を効果的に視覚化する独自のインタフェースによって、ユーザの作業負荷を大きく軽減した。今後は候補曲の選択に注目して研究を進め、「より明るい曲につなぐ」、「より遅い曲につなぐ」といった操作を組み込んだインタフェースを実現する予定である。また本研究は、RWC 研究用音楽データベース(クラシック音楽、ジャズ音楽、ポピュラー音楽、音楽ジャンル)<sup>(5)</sup>を利用した。

## 参考文献

- (1) 後藤 孝行, 後藤 真孝, "Musicream: 楽曲を流してくっつけて並べることのできる音楽再生インタフェース," WISS2004, pp.53-58, 2004.

- (2) 莪山 真一, "類似楽曲検索を目的とした楽音からの和音進行抽出手法," FIT2003, pp.245-246, 2003.
- (3) 莪山 真一, "和音進行の連続性を用いた楽音信号の連続再生手法," FIT2005, pp.277-278, 2005.
- (4) Shinichi Gayama, "Continuous Song Playback Based on Its Chord-Progression Continuity," ICCE2006, pp.1-2, 2006.
- (5) 後藤 真孝, "RWC 研究用音楽データベース: ポピュラー音楽データベースと著作権切れ音楽データベース," 情処研報音楽情報科学 2001-MUS-42-6, p.35-42 (2001).

## 筆者紹介

堀内 直明 (ほりうち なおあき)

技術開発本部 総合研究所 システム研究センター 情報メディア技術研究部

藺田 俊行 (そのだ としゆき)

パイオニアデザイン株式会社 プロダクトデザイン部 3スタジオ

田中 浩司 (たなか こうじ)

パイオニアデザイン株式会社 コミュニケーションデザイン部

田中 淳一 (たなか じゅんいち)

パイオニアデザイン株式会社 プロダクトデザイン部 2スタジオ

長沢 秀哉 (ながさわ ひでや)

パイオニアデザイン株式会社 プロダクトデザイン部 3スタジオ

莪山 真一 (がやま しんいち)

技術開発本部 総合研究所 システム研究センター 情報メディア技術研究部