Vol.2011-MUS-89 No.16
2011/2/12

音楽のムード分類結果を利用したホームビデオへの BGM 付与支援システム

小野佑大[†]　石先広海[††]　帆足啓一郎[††]
小野智弘[††]　甲藤二郎[†]

本研究では，ホームビデオへの BGM 付与支援システムを提案する．BGM 付与の作業工程は大まかに，i) BGM の選曲，ii) 選択楽曲から BGM としてもちいる区間の切り出し(トリミング)，の二つに分けられる．しかし，映像作品の制作に慣れないユーザにとってこれら作業は非常に労力を要する．そこで，本システムでは，音楽のムード分類結果を利用して BGM 選曲を行う．さらに，選択された楽曲から映像の動きと同期する区間を自動抽出する．これにより，ユーザは効率良く BGM の選曲ができ，かつ，満足度の高いコンテンツを容易に制作することができる．そして，被験者評価実験を実施し，本システムの有効性を検証する．

Supporting BGM Addition to Home Video Based on Music Mood Classification

Yudai Ono[†]　Hiromi Ishizaki[††]　Keiichiro Hoashi[††]
Chihiro Ono[††]　Jiro Katto[†]

We propose a system to support adding BGM to a home video, by utilizing the result of music mood classification. In this paper, first, we propose a method of music mood classification, and music visualization on a two-dimensional plane. Second, we propose a method to extract the synchronizing section with video's motion from the selected song. In these methods, users can select BGM efficiently which can give the appropriate psychological effect to a home video. Moreover, users can generate stylish video contents easily. Finally, we evaluate our system by subjective test.

1. はじめに

近年，デジタルビデオカメラなどを用いて容易に動画撮影を楽しめるようになり，撮影した動画を web 上で披露するなど，一般的なユーザにとって映像編集は身近になっている．映像編集とは，撮影した動画から冗長な映像のトリミングや，BGM の付与などによって新しい映像を制作するプロセスを指す．そして，一般的に，適切な“BGM の付与”は，印象深い映像作品を制作する上で重要であると言われている．たとえば，映画やドラマでは，別れのシーンに悲しい音楽を流すなど，シーンのムードに合った BGM を使用することで，シーンを効果的に演出している．

一方，ミュージックビデオでは，映像の動きなどを音楽のビートなどと一致させ，印象的な音楽体験を提供している．[1]によると BGM 付与の作業工程は，(i) 映像のムードに合った楽曲の選択, (ii) 選択した楽曲から BGM としてもちいる区間のトリミング，の二つに大きく分けられる．しかし，(i)の作業は，大規模楽曲データベースから映像に合った楽曲を探し出すために多大な労力を要す．さらに(ii)の作業は，時間だけでなく高度な技術も必要となり，映像編集は一般のユーザにとって困難なものと言える．

本稿では，映像編集作業の効率化のために，ホームビデオへの BGM 付与支援システムを提案する．(i)については，音楽のムード分類結果を利用した楽曲の絞り込みと，二次元平面を利用した選曲 GUI により，BGM 選曲の効率化を図る．(ii)については，楽曲から映像の動きと同期する区間を抽出し BGM として付与することで，BGM の編集作業の自動化を図る．最終的に，被験者による主観評価実験で，提案システムを被験者に使用してもらい，本システムの有効性を検証する．

2. 関連研究

映像編集において，BGM 付与を支援する研究が多く報告されている．Mulhem ら[3]は，映像編集のルールに基づき，映像と音楽の特徴量を，ダイナミクス，ピッチ，モーションの 3 つのカテゴリに分け，それぞれのカテゴリにおいて，映像の特徴と相関の高い特徴を持つ楽曲を BGM として推薦する手法を提案している．また，Foote ら[4]は，楽曲構造の境界に合わせて，映像の動きや輝度のちらつきが少ないショットを切り貼りすることでミュージックビデオを自動で生成する手法を提案している．これらの手法では，推薦された BGM の提示方法としてリスト型を想定しており，ジャンル情報を楽曲に付与することで効率的な BGM の選定ができると述べている．しかし，

[†] 早稲田大学理工学術院
Waseda University
[††] KDDI 研究所
KDDI R&D Laboratories Inc.

ジャンル情報によって，ユーザが映像へ与えたい心理効果の楽曲を効率良く選択できるとは限られていないため,ユーザが期待した楽曲を付与できないと言う課題がある．また，映像に対して単純に，サビなどの楽曲構造の一区間を付与しているため，より効果的に映像を演出できる区間が存在していた場合に，最適な区間を BGM として付与できない可能性がある．

これらの問題に対して，小野ら[2]は，音楽のムード分類結果を利用して，映像へ適切な心理効果を付与できる楽曲を選定し，選択した楽曲から映像の動きと同期する区間を抽出する BGM 区間抽出法を提案している．BGM 区間抽出法では, 映像から動き，音楽から音量に関する特徴量を抽出し，時系列解析で利用される特異スペクトル変換[7](以下，SST と呼ぶ)で検出された変化度を一致させることで同期を図っている．SST とは，時系列のある時点の変化に対して過去と未来の部分系列から特異値分解によって特徴を抽出し, その非類似性で変化の度合い(変化度)を求める手法である．しかし，SST が特徴の“変化の大きさ”に着目しているため，特徴自体の大きさを反映できず音量が小さい区間でもその変化度が大きく計算されるために，視聴者とって映像との同期が分かりづらい区間が抽出されると言う課題や，楽曲特徴として音量のみを使用しているため，楽曲によって音量が小さい場合，この問題が顕著となると言う課題がある．

3. ホームビデオへの BGM 付与支援システム

そこで本稿では，第 2 章で記述した問題を解決するために，BGM 付与支援システムを提案する．具体的には，音楽のムード分類結果を利用し，ムードを表す二次元平面で楽曲を可視化することによって，BGM 選曲の効率化を図る．また，[2]の BGM 区間抽出法を，楽曲の音量の大きさを変化度へ反映させる処理と新たな特徴量を追加することによって改善する．

ここで，図 1 に本システムの概要を示す．本システムは大きく分けて，“BGM 選択部”と“BGM 区間抽出部”の二つの処理で構成される．はじめに，ユーザはホームビデオを入力し，“BGM 選択部”で，映像に付与したい心理効果を“動画のテーマ”から選択する．そして，選択した動画のテーマに適した楽曲が二次元平面上の点として表示される．二次元平面の各軸は楽曲のムードをより詳細に表した情報で，それに基づき BGM として使用する楽曲を選択する．次に，“BGM 区間抽出部”にて，選択された楽曲から映像を効果的に演出するような区間が自動で抽出される．最終的に，入力映像と BGM を同期再生する．以下の節にて，各処理についての詳細を述べる．

3.1 BGM 選択部

本処理では[5]に基づき，ムードを表す二次元平面で楽曲を可視化し，その空間上で楽曲をクラスタリングする．そして，各クラスタへ適切な印象語を付与する．

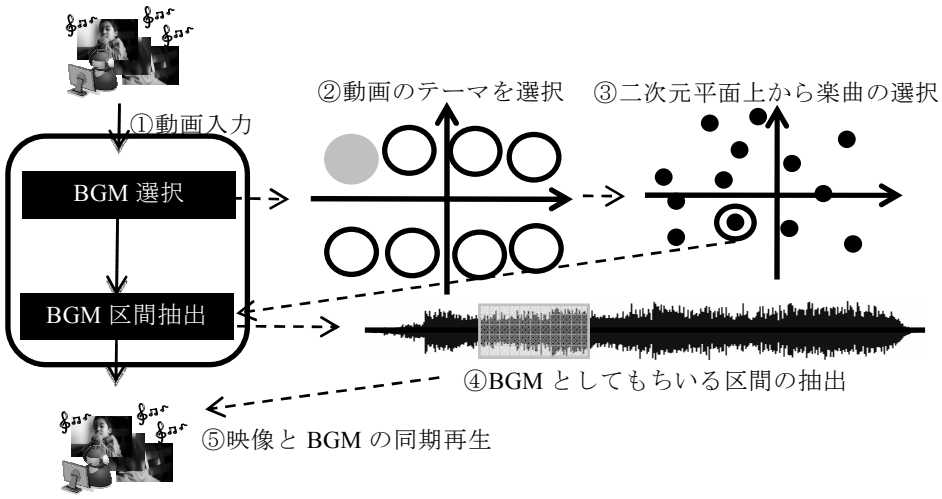

図 1 システム概要

3.1.1 ムードを表す二次元平面での楽曲の可視化

本システムでは，音楽のムードを表す二次元平面として心理学の分野で定義された AV 空間[6]をもちいる．これは，energetic-calm(Arousal)，positive-negative(Valence)の二軸(AV 軸)から成る二次元平面で，空間上の座標値(AV 値)で印象語を表現し，同一象限内の AV 値間の距離が近いムードは類似すると言う性質がある．

本システムにおける AV 値の計算方法について述べる．まず，楽曲のムードと関連する特徴量全 29 次元[5]フレームごとに抽出しその平均と標準偏差を使用)を抽出し平均 0，分散 1 に正規化後，主成分分析を行う．各主成分の因子負荷量を分析した結果を利用して，Arousal 値に第一主成分得点に-1 を乗じたもの，Valence 値に第二主成分得点をもちいる．

3.1.2 音楽のムード分類法

次に，AV 空間上で楽曲をクラスタリングする．二次元平面上の点に対してクラスタリングする際，k-means 法を適用することが考えられるが，AV 空間上で単純に k-means 法を適用した場合，3.1.1 で述べた AV 空間の性質を考慮できず，象限を跨いでクラスタが生成される．そこで本システムでは，Valence 値，Arousal 値それぞれに対して階層的に k-means 法を適用し, AV 空間の各象限に収まるようにクラスタを生成する．そして，生成したクラスタへ印象語を付与する．本稿では[6]と同様，予め楽曲ごとに音楽情報サイト all music guide の印象語を付与し，クラスタ内で重複したラベルをそのクラスタの印象語とする．

Vol.2011-MUS-89 No.16
2011/2/12

3.2 BGM 区間抽出部

本処理では，[2]の BGM 区間抽出法に対して，楽曲の音量の大きさを変化度へ反映させる処理と，新たな特徴量の追加の二つを適用する．

3.2.1 音量を変化度へ反映させる処理

まず楽曲から RMS を抽出する．そして，そのフレーム間差分を求め，SST を適用することで時系列の変化度を得る．最後に，その変化度と再度楽曲から抽出された RMS との績を求める．

3.2.2 新たな特徴量の追加

[2]にて実施した「映像と BGM を同期させるために重要な楽曲特徴は何ですか」というアンケートの回答として，楽曲構造の変化が最も多かった．そこで楽曲構造の変化を表す特徴として，Novelty Score[5]に着目する．これは，Foote[4]などによって提案された楽曲構造に関する特徴量で，各ピークは楽曲構造の変化の境界を表す．Novelty Score が大きければ大きいほど，楽曲の大きな変化を表している．

本稿では，抽出した Novelty Score のピークを検出し，それを楽曲構造が変化する境界を表す特徴としてもちいる．

3.2.3 BGM 区間抽出法への適用

まず，映像から動き特徴としてフレーム内の動きベクトルの総和で得られる Motion Activity[2]を抽出する．そして，Motion Activity に対して SST を施し，時系列の変化度を算出する．これに対し，3.2.1 節で得た特徴量との m フレーム目における相互相関係数を式(1)より求める．

$$R_m(MA, RMS) = \sum_{i=0}^{N-1} MA(i)RMS(i+m) \quad (m = 0,1,2...M-N) \qquad (1)$$

$MA(i)$ は i 番目のフレームにおける総フレーム数 N の Motion Activity の変化度，$RMS(i)$ は i 番目のフレームにおける総フレーム数 M の 3.2.1 節で得た楽曲の RMS フレーム間差分の変化度を示す．同ように 3.2.2 節で得た Novelty Score のピーク値との m フレーム目における相互相関係数を式(2)より得る．

$$R_m(MA, boundary) = \sum_{i=0}^{N-1} MA(i)boundary(i+m) \quad (m = 0,1,2...M-N) \qquad (2)$$

$boundary\ (i)$ は i 番目のフレームにおける総フレーム数 M の 3.2.2 節で得た楽曲構造の境界を表す．本システムでは楽曲の長さが入力動画の長さよりも長いことを想定し，$M > N$ としている．次に，式(1)，(2)から得た二つの相互相関係数に対して，式(3)より映像の動きと BGM との同期の強さを表す m フレーム目におけるスコアを算出する．

$$Score(m) = R_m(MA, RMS)R_m(MA, boundary) \qquad (3)$$

表 1. 各クラスタへ付与された印象語

クラスタ1	クラスタ2	クラスタ3	クラスタ4
Sweet	Happy	Melancholy	Angst-Ridden
Dramatic	Fun	Sad	Angry
Gentle	Party/Celebratory	Angst-Ridden	Nihilistic
クラスタ5	**クラスタ6**	**クラスタ7**	**クラスタ8**
Sentimental	Aggressive	Intimate	Energetic
Bittersweet	Intense	Sentimental	Confident
Laid-Back/Mellow	Angry	Atmospheric	Stylish

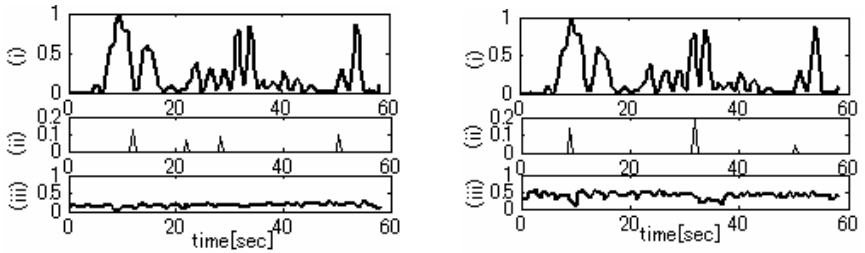

図 2.左:[2]の BGM 区間抽出法，右: 3.2.1 節と 3.2.2 節を追加後の BGM 区間抽出法
(i)映像の Motion Activity の変化度，(ii)楽曲構造の境界，(iii)RMS

最終的に，スコアが最大となるフレーム m を選択し，$N-1+m$ 番目のフレームまでの区間を 3.1 節で選択された楽曲から抽出し，入力の動画へ付与する．

ここで，[2]の BGM 区間抽出法と，3.2.1，3.2.2 節で得た二つの特徴量を追加した BGM 区間抽出法によって楽曲から抽出された区間における，映像の Motion Activity と BGM の楽曲構造の境界，RMS を図 2 に示す．まず(iii)の RMS を比べると，RMS の大きい区間が抽出されていることが分かる．よって 3.2.1 節の特徴量を追加することで，楽曲の音量の大きさを変化度へ反映できると言える．さらに(i), (ii)を比べると，[2]の BGM 区間抽出法では(i)と(ii)のピークが一致していないのに対し，3.2.1 節と 3.2.2 節の特徴量を追加後の BGM 区間抽出法では一致していることが分かる．したがって，3.2.2 節の特徴量を追加することによって，映像の動きの変化に対し，楽曲構造が変化する区間を楽曲から抽出できていることが分かる．

3.3 ホームビデオへの BGM 付与支援システムの実装

3.1，3.2 節の処理を実装したシステムを図 3 に示す．本システムでは，図 3 の①に AV 空間上で生成されたクラスタが異なる色で表示され，動画のテーマとして各クラスタの中央に印象語が表示される．尚，本稿ではクラスタ数を 8 に設定し，各クラスタへ付与された印象語を表 1 に示す．ユーザがクラスタを選択すると，システムは図 4(左)の①のように，クラスタ内の楽曲を AV 空間上の点として表示する．このように，

Vol.2011-MUS-89 No.16
2011/2/12

AV 空間上で楽曲を可視化することで，空間上のどの付近に動画のテーマに合う楽曲が存在するのか把握しやすくなり，BGM 選曲における操作性の向上が期待される．この点をクリックすると AV 値が表示され，ダブルクリックすると楽曲が再生される．ユーザは好みの BGM を発見したら，図 3 の“BGM 決定”ボタン(②)を押す．すると，システムは 3.2 節で述べた方式で楽曲から映像と同期する区間を抽出し，ホームビデオへ付与する．制作された作品は，図 3 の“作品視聴”ボタン(③)を押して視聴できる．

4. 評価実験

4.1 実験概要

本実験では，提案システムにおける，(i) 動画のテーマに合った BGM 選曲の効率性，(ii) BGM 選曲のしやすさ(操作性)，(iii) 制作された映像作品の満足度，を評価する．尚，本実験では[2，5]と同ように，all music guide の印象語を検索クエリとして収集した 4 種類の動画([2]とは異なる動画を使用)と，USPOP 楽曲 206 曲を使用した．

4.2 比較システム

ここで，本システムの比較対象として実装したシステムを図 4(右)に示す．比較システムでは，画面左側に楽曲をリスト形式で提示し，楽曲ごとに all music guide のジャンル情報が表示される．そして，リスト内のジャンル情報をダブルクリックすると楽曲が再生される．このように，比較システムでは，ジャンル情報を基にホームビデオへ付与したい心理効果を持つ楽曲を決定する．使用する BGM を発見したら，提案システムと同ように “BGM 決定”ボタンを押すと，入力動画へ BGM が付与される．

4.3 実験方法

本実験では，大学生 20 名に，比較，提案システムを通じて，指定されたホームビデオへ付与する BGM を実際に選曲してもらう．はじめに，被験者はシステム上で与えられたホームビデオを視聴する．そして，画面上に提示された動画のテーマ(動画収集時に検索クエリとして使用した印象語)と，視聴した動画に合った楽曲を選定するという条件で，BGM 選曲を行う．動画を再視聴する際は，“動画再視聴”ボタン(図 3 の④)を押し，好みの BGM を発見したら“BGM 決定”ボタンを押す．するとシステムは，選択された楽曲から，イントロ，[8]の手法で得たサビ区間，そして 3.2 節で述べた方式で抽出した映像の動きと同期する区間，をそれぞれ抽出しホームビデオへ付与する．最後に，被験者はシステムに制作された作品を視聴し，一番のお気に入りを選択する．

本実験では，これを比較，提案システムそれぞれ 4 種類の動画に対して繰り返し行う．また，客観評価項目として，(i)の評価のために，被験者が比較，提案システムを使用している間，BGM 選曲に要する時間(楽曲や動画の聴取時間を除く，システムの操作時間)と聴取曲数を計測する．また，提案システムの(ii)の評価をするために，被験者が聴取した楽曲の AV 値を計測する．

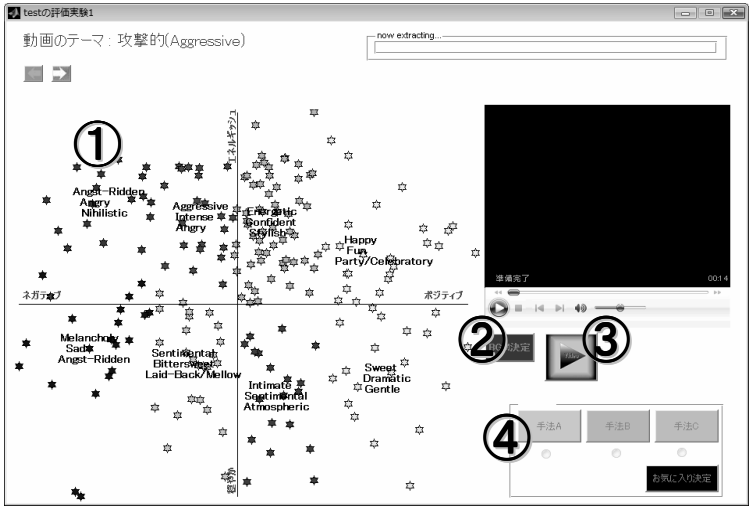

図 3. 提案システム

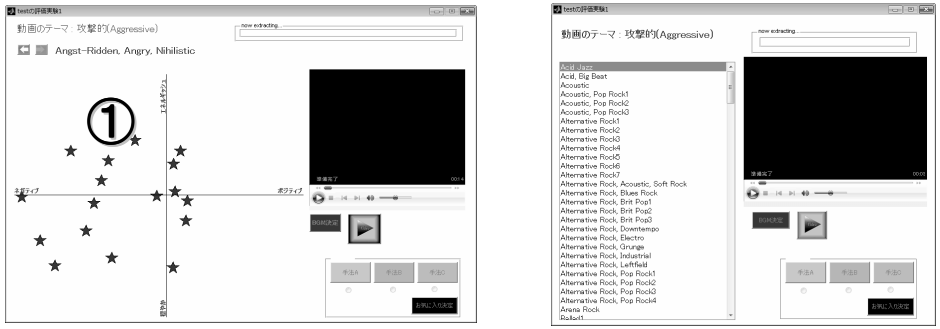

図 4. (左)提案システムのクラスタ拡大表示画面，(右)比較システム

さらに，各システムの評価が終了後，主観評価項目として，「BGM 選曲の効率性」，「今後比較，提案システムを使いたいと思うか」，「比較システムのジャンル情報，提案システムのクラスタ情報がどれくらい BGM の選曲に役立ったか」，という項目について 5 段階評価(5. 良い～1. 悪い)を行う．

(iii)は，被験者にお気に入りの作品として選択された回数で評価する．尚，“作品視聴”ボタンは，比較，提案システム共に 3 つ配置し，順序効果を考慮するためにボタンの配置順を使用するシステムごとに変える．

Vol.2011-MUS-89 No.16
2011/2/12

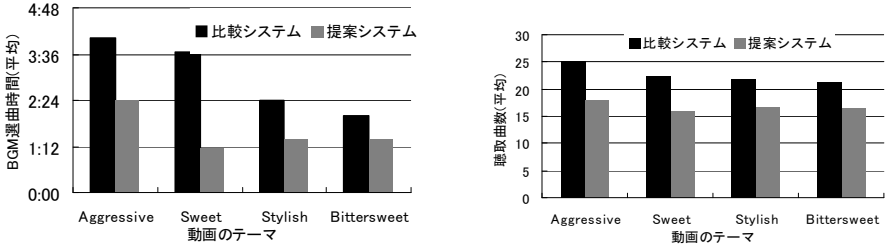

図 5. 各動画における(右)BGM の選曲時間，(左)の聴取曲数(α =0.05 で有意差あり)

4.4 実験結果

(i) BGM 選曲の効率性

図 5 に各動画における BGM の選曲時間と聴取曲数を示す．横軸は被験者に提示した動画のテーマを表し，図 5(左)の縦軸は被験者が BGM 選曲に要した時間の平均を，図 5(右)の縦軸は被験者が聴取した楽曲数の平均を示している．これらを見ると，いずれの動画においても提案システムは比較システムと比べ，選曲時間，聴取曲数が少ないことが分かる．また，BGM 選曲時間については，提案システムが平均して約 1 分 20 秒早めることができ，聴取曲数については，平均して約 6 曲減らすことができた．このことから，本システムをもちいることで効率良く動画のテーマに適した BGM を選曲できると言える．

また，図 6 の「BGM 選曲の効率性」と言う評価項目に着目する．図 6 の縦軸の評価値は，被験者による 5 段階評価の平均を表し，5 に近づくにつれて良い結果を表している．これを見ると，比較システムの評価値が平均して 2.5 と低い結果を示しているのに対し，提案システムの評価値は高く，平均して 4.5 であった．したがって，このことからも，提案システムを使用することで効率良く BGM 選曲を行えると言える．

ここで，被験者によるクラスタの選択回数を表 2 に示す．表中の太字になっている数字は動画のテーマに合ったクラスタの選択回数，括弧内の数字は使用する BGM として決定された楽曲数を表している．これらを見ると，動画のテーマに合ったクラスタが最も多く選択され，そのクラスタから映像作品の BGM として使用する楽曲が多く決定されていることが分かる．さらに，実験終了後に提案システムのメリットについて自由記述のアンケートを実施したところ，20 名中 14 名の被験者から，クラスタによる絞り込みが有効である旨の回答を得た．つまり，音楽のムード分類結果を利用することで，動画のテーマに合った楽曲を適切に絞り込むことができたと言える．したがって，これらの実験結果より，本システムは，ユーザが映像に付与したい心理効果を持つ BGM を効率良く選曲することができると言える．

表 2. 提案システムにおけるクラスタの選択回数と BGM に使用した楽曲数

	クラスタ1	クラスタ2	クラスタ3	クラスタ4	クラスタ5	クラスタ6	クラスタ7	クラスタ8
Aggressive	0.20	2.00(4)	0.20	2.05(4)	0.25	**8.80(11)**	0.55	4.00(1)
Sweet	**11.0(17)**	2.35(1)	0.20	0.00	0.25	0.00	1.60(2)	0.40
Stylish	0.00	1.10(2)	0.00	0.90(1)	0.20	0.95(1)	0.75(1)	**12.8(15)**
Bittersweet	0.50(2)	1.35(1)	1.20	0.20	**11.7(16)**	0.25	0.80(1)	0.60

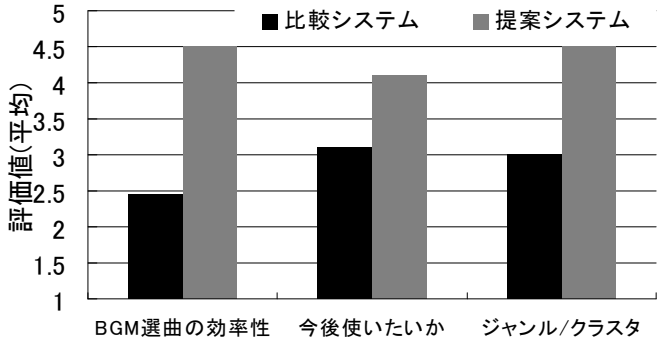

図 6. 各システムについてのアンケート結果(α =0.01 で有意差あり)

(ii) BGM 選曲のしやすさ(操作性)

まず，図 6 の，「ジャンル/クラスタ」という評価項目のアンケート結果に着目する．尚，「ジャンル/クラスタ」という評価項目は，比較システムならジャンル情報が，提案システムならクラスタ情報が BGM 選曲にどれくらい役立つか，という質問への 5 段階評価の結果である．これを見ると，比較システムに比べ，提案システムの方が評価値が高く，平均して 4 を超えている．このことから，クラスタ情報はジャンル情報よりも BGM 選曲の役に立つ情報であることが分かる．

さらに，図 6 の「今後使いたいか」と言う評価項目に着目する．これは，被験者が比較，及び，提案システムを今後も使ってみたいと思うか，と言う質問に対する 5 段階評価の結果である．これを見ると，多くの被験者が本システムを今後も使ってみたいと感じたことが分かる．これは，クラスタ情報がジャンル情報よりも BGM 選曲に役立つという結果を踏まえると，本システムが BGM 選曲を行いやすいインタフェースであることが理由であると考えられる．ここで，被験者が実際にどのような本システムを操作して楽曲を選択・聴取していたのか確かめるために，提案システムにおける，ある被験者の AV 空間上での楽曲の聴取の遷移を図 7 に示す．

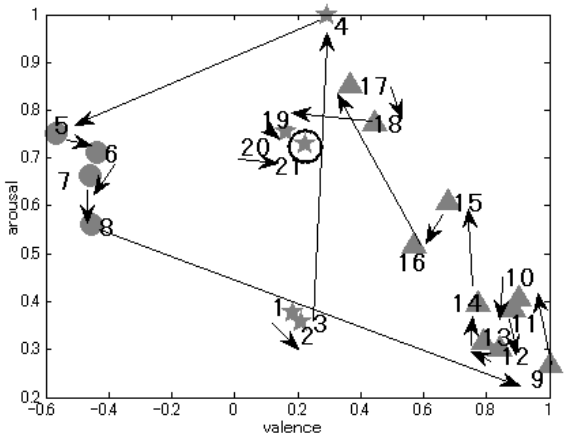

図 7. ある被験者の AV 空間上での楽曲の聴取の遷移

同じクラスタの楽曲は点の形状が同一で，星型は提示した動画のテーマと同一のクラスタを表している．また，各点に添えられている数字は，被験者の聴取した楽曲の順番を表し，矢印で示される順番で楽曲を聴取していることを示している．丸で囲まれた点は被験者が最終的に決定した BGM を表す．これを見ると，この被験者は，動画のテーマと同一のクラスタ，それに隣接する丸型と三角型の点のクラスタ，再度動画のテーマと同一のクラスタ，とクラスタを行き来している．そして，三角型から星型の点へ移る際，空間上の楽曲を部分的に聴取しながら，動画のテーマと同一のクラスタへ近づいていることが分かる．その他 14 名の被験者も同様の手順で楽曲を聴取していた．このことから，多くの被験者は楽曲のムードを把握するために，AV 値を頼りに動画のテーマに合った楽曲が，どのクラスタの，空間上のどこに存在するか判別していることが分かる．

したがって，AV 空間による楽曲の可視化は，動画のテーマに合わないクラスタを選択しても，次に参照するクラスタを決定する手がかりとして利用できる．以上より，本システムは，楽曲の可視化によって，BGM 選曲に役立つクラスタ情報を効果的に利用でき，選曲を行いやすいインタフェースであると言える．

(iii)　制作された映像作品の満足度

被験者によってお気に入りとして選ばれた手法ごとの映像作品数を図 8 に示す．これを見ると，3.2 節で述べた手法によって制作された作品は，その他の手法と比べ，最も多くお気に入りとして選択されていることが分かる．したがって，本システムをもちいることで，3.2 節で述べた方式により，ユーザにとって満足度の高い作品を制作できると言える．

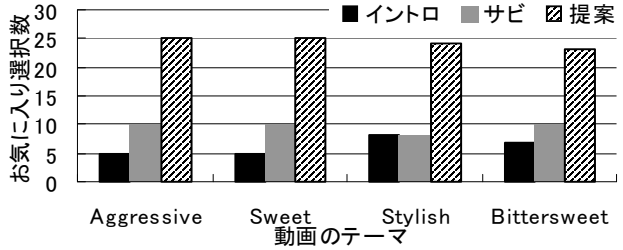

図 8. 手法ごとのお気に入りとして選択された映像作品数(α =0.01 で有意差あり)

5. おわりに

本稿では，音楽のムード分類結果をユーザに提示し，BGM 選曲の効率化を図る方式と，選択された楽曲から映像の動きと同期する区間を抽出し，映像へ自動付加する方式を実装した，ホームビデオへの BGM 付与支援システムを提案した．そして，主観評価実験により本システムの有効性を示した．評価実験時に提案アプリケーションのデメリットについてのアンケートを実施したところ，多くの被験者がクラスタ内から楽曲を選択してしまう故に，未知の楽曲と出会う機会が減るのではないかと言う指摘をしていた．また，動画の音声と BGM との音量のバランスを調整する必要性についての指摘も多かった．そこで今後は，入力映像の特徴を加味した楽曲推薦や，動画の音声と BGM との音量のバランス調整について検討する予定である．

参考文献

1) Herbert Zettl, “Sight, Sound, Motion: Applied Media Aesthetics,” Wadsworth Pub Co, 1998
2) 小野佑大, et al “音楽のムード分類結果を利用したホームビデオへの自動 BGM 付与・同期手法,” 第 9 回情報科学技術フォーラム, E-033 , 2010.
3) Philippe Mulhem, et al., “Pivot Vector Space Approach for Audio-Video Mixing,” IEEE Multimedia 2003, Vol.10, No.2, pp.28‐40, 2003.
4) Foote J et al., “Creating music videos using automatic media analysis,” Proceedings of ACM multimedia, New York, pp.553–560, 2002.
5) 小野佑大 et al., “ホームビデオへの自動 BGM 付与のための心理学に基づく音楽分類手法,” 第 72 回情報処理学会全国大会, 1T-2, 2010.
6) J. A. Russell, “A circumplex model of affect,” J. Personality Social Psychology, 1980.
7) T. Id´e et al., “Knowledge discovery from heterogeneous dynamic systems using changepoint correlations,” In Proc. SIAM Intl. Conf. Data Mining, pp.571–575, 2005.
8) Ewald Peiszer et al., “Automatic Audio Segmentation: Segment Boundary and Structure Detection in Popular Music,” Proceedings of the 2nd International Workshop on LSAS , 2008