

出演者の代表顔を利用した番組選択支援方式の有効性検証

The Evaluation of the Validity of a Method for TV-program Selection with Performer Representative Faces

石川真澄† 大網亮磨† 平田恭二十
Masumi Ishikawa Ryoma Oami Kyoji Hirata

1. はじめに

パーソナルビデオレコーダ(PVR)の大容量化に伴い、大量の番組が蓄積されつつある。その中から、利用者が視聴したい番組を効率的に選べるよう支援する番組選択支援方式が重要になってきた。

多くのPVRでは、番組選択支援のためにElectronic Program Guide(EPG)[1]に含まれる番組名や局、放映日時、出演者情報が提示される。中でも出演者情報は、映像内容を表現する重要な要素[2]であるが、EPGでは出演者名がテキストで記述されるため即時把握が困難だった。

我々はこれまでに、出演者の代表顔を提示することで出演者を即座に把握可能な番組一覧を提案し、その実現に向けて代表顔抽出方式について検討してきた[3]。

本稿では、さらに豊富な出演者情報の提示に向けて、代表顔による番組一覧に加え、代表顔をもとに番組映像から切り出されたビデオクリップを提示する番組選択支援方式を提案する。本方式により、蓄積された番組に対する一覧性を保持しつつ、個々の番組について実映像による詳細な出演者情報を提示できる。利用者は、番組一覧によって複数番組から出演者をもとに興味番組を選び、ビデオクリップによって、出演者の番組中での振舞いを確認した上で視聴番組を選択できる。

本方式による番組選択の有効性を評価するために、番組選択支援プロトタイプを構築した。本システムを用いて、番組の選択から視聴までの一連の流れの中で、テキストによる番組一覧、出演者の代表顔を利用した番組一覧、ビデオクリップの各提示機能を評価した。被験者8名に、本システムを利用して番組選択を行ってもらい、アンケート結果を分析することで有効性を評価した。

2. 代表顔による番組一覧に基づく方式

代表顔による番組一覧[3]では、出演者の即時把握を目的に、番組映像から抽出された各出演者の代表的な顔画像を出演頻度順に並べて提示する。

利用者は、代表顔の並び順から、主役あるいは脇役が誰かを即座に把握し、代表顔の表情・髪型・化粧等から、出演者の役柄を想像できる。また、利用者は名前を正確に知らなくても出演者を認知できる。

3. 番組一覧とビデオクリップに基づく方式

3.1 提案方式の概要

提案方式では、さらに豊富な出演者情報の提示を目的として、2章で述べた番組一覧に加え、代表顔をもとに番組映像から切り出したビデオクリップを提示する。本方式では、図1に示すように、出演者の代表顔を並べて提示するとともに、利用者が選択した番組について、各出演者のビ

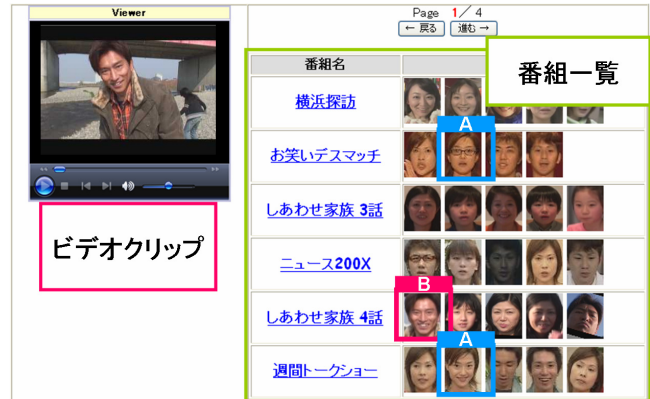


図1. 番組一覧とビデオクリップに基づく方式

デオクリップを再生する。

3.2 ビデオクリップの生成

ビデオクリップは、代表顔抽出方式[3]によって抽出された代表顔、あるいは代表顔抽出の過程で作成した出演者ごとの顔画像クラスタをもとに、顔画像の抽出されたフレームを含む前後数秒の映像区間を切り出すことによって生成される。

ビデオクリップの実例として、以下が挙げられる。

- (1) 代表顔の登場シーン
各出演者の代表顔を含むフレームをもとに切り出された映像区間。
- (2) 顔のクローズアップシーン
各出演者の顔画像クラスタ内で、最も顔のサイズが大きく顔の位置がフレーム中心に近い顔画像を含むフレームをもとに切り出した映像区間。
- (3) 初出シーン
各出演者の顔画像クラスタ内で、出現時刻が最も早い顔画像を含むフレームをもとに切り出した映像区間。
- (4) 発話シーン
各出演者の顔画像クラスタ内で、声と同時に検出された顔画像を含むフレームをもとに切り出した映像区間。

3.3 ビデオクリップの提示による効果

利用者は、ビデオクリップによって、出演者の実映像中の演技を確認し、他の出演者との関係や、場面設定、番組の雰囲気を感じとることができる。また、一覧から見つけた意外な風貌の出演者についてその演技を楽しんだり、何気なく選んだ番組の全出演者のビデオクリップを確認することで番組概要を把握する満足感を得たりと、番組選択自体を楽しむことができる。

ビデオクリップによる効果を、本方式の提示画面(図1)を用いて実例で説明する。利用者は、お気に入りの出演者“A”が、“お笑いデスマッチ”ではコンビ芸人として漫才をする様子を、“週間トークショー”では朗らかな表情で司会を務める様子を確認できる。また、一覧の顔の表情か

† 日本電気株式会社, NEC Corporation

ら興味引かれた出演者”B”が、”しあわせ家族 4 話”の中で、家族と仲良く遊ぶ姿を確認し、同番組の他の出演者について演技や会話を確認することで、番組概要を把握する満足感を得られる。

4. 主観評価実験

4.1 評価実験の概要

提案方式の有効性を評価するため、番組選択支援プロトタイプシステムを構築し、主観評価実験を行った。実験では、番組選択機能として本システムに実装した①テキストによる番組一覧、②代表顔による番組一覧、③ビデオクリップの各提示機能を比較する。各機能による提示内容を以下に示す。

① テキストによる番組一覧の提示

番組名、局、放映日時、概要、出演者名をテキストで一覧提示する。

② 代表顔による番組一覧の提示

番組名、局、放映日時をテキストで提示するとともに、出演頻度の高い上位 4 名の代表顔を出演頻度順に一覧提示する。4 名は“人が一度に処理できる要素数の限界[4]”に基づき妥当として用いた。

③ ビデオクリップの提示

番組一覧との関連性を考慮して、ビデオクリップとして 3.2 の(1)代表顔の登場シーンをを用いる。利用者の選択した番組について、一覧に提示した 4 名の代表顔を含むフレームをもとに切り出した各 10 秒の映像区間を出現時刻順に連続再生する。

4.2 評価方法の詳細

有効性の検証には、アンケート評価を用いた。アンケートの回答方式は、各々「非常に有効だと思う・有効だと思う・ふつう・あまり有効でないと思う・全く有効でないと思う」までの 5 段階評価と、評価の理由および改善点の記述形式とした。

評価には、ドラマ・ニュース・スポーツ・バラエティ等の様々なジャンルを含む番組 90 本を用いた。予め利用者が選択的に録画した蓄積番組中から見たい映像を選択する場面を想定し、8 名の被験者が番組選択から視聴までの一連の流れを実際に行う中で、各機能の有効性を評価した。

4.3 結果と考察

4.3.1 5段階評価の分析結果

被験者 8 名による 5 段階評価の平均スコアを図 2 に示す。テキストによる番組一覧は、一番低く 2.4 点だった。代表顔による番組一覧には、全被験者がテキストに基づく方式と同等または高い得点をつけ、平均スコアはテキストに基づく方式より 1.4 点高い 3.8 点だった。この平均スコアの差が有意である確率は、t 分布を仮定した一対比較において 99%であった。ビデオクリップは、テキストによる一覧よりも高い 2.9 点で、被験者の半数から 4 点以上で採点された。

さらにスコアを分析すると、全被験者が、代表顔による番組一覧とビデオクリップの一方または両方に、4 点以上を与えていることがわかった。これにより、提案方式が様々な利用者の番組選択支援に有効であるといえる。

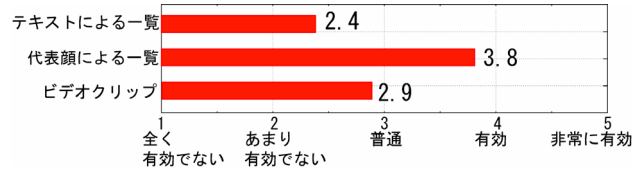


図 2. 有効性の平均スコア

4.3.2 評価理由の分析結果

テキストによる番組一覧は、見たい番組が定まっている場合に有効だが、文字で記述されるために読まねばならず視認性に欠ける点が問題とされた。

代表顔による番組一覧は、複数番組中の出演者を即座に認知できる点や、毎回出演者が変わるドラマシリーズにおいて特定の回の番組を判別しやすい点、番組映像から切り出された画像であることから一度視聴した番組をすぐに見つけられる点が有効とされた。

ビデオクリップについては、バラエティやドラマなど、出演者の振る舞い自体が番組選択に有用なジャンルについて有効とされた。また、お気に入りのコメディアン出演番組を一覧から見つけだし、ビデオクリップでその演技を確認するなど、番組選択自体を楽しめる点も評価された。より多くの出演者の提示、あるいは各出演者について複数の登場シーンを提示することで、さらに効果的であるとの指摘もあった。これは、ビデオクリップの視聴をきっかけに、番組の他の場面に利用者の興味を向かせる効果があったことを示している。低く採点した被験者は、複数出演者のビデオクリップの連続再生による音声の不連続を不快に感じたことを理由とした。今後、ビデオクリップの抽出・提示方法を検討する必要がある。

5. まとめ

出演者の代表顔による番組一覧とビデオクリップによる番組選択方式を提案した。本方式では、複数番組の各出演者の代表顔を一覧にし、各出演者に関するビデオクリップを提示する。提案方式の有効性を評価するために番組選択支援プロトタイプを構築し、アンケートに基づく有効性検証を行った。5 段階評価において、代表顔による番組一覧はテキストによる番組一覧よりも 1.4 点高い平均スコア 3.8 点を得た。また全被験者が、代表顔による番組一覧とビデオクリップの一方または両方を 4 点以上で評価したことから、提案方式が様々な利用者の番組選択を支援できることが示された。

参考文献

- [1]八木伸行,吉村俊郎,加井謙二郎,“データ放送技術入門,”オーム社,2005.10.
- [2]Zhu Liu and Yao Wang, “Major Cast Detection in Video Using Both Audio and Visual Information,” ICASSP’01, vol.3 7-11, pp.1413-1416, Apr.2001.
- [3]石川真澄,大網亮磨,平田恭二, “出現時刻による距離補正を用いた顔画像クラスタリング”, FIT2007 D-12-87, 2007.03.
- [4]George A. Miller, “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information”, The Psychological Review, 1956, vol.63, pp.81-97.