

WA-TV: 番組コンテンツのウェブ化と情報補強～オンライン化とユーザ評価実験～

WA-TV: Webifying and Augmenting TV Programs – Online Processing and User Study Experiment

宮森 恒[†] 馬 強[†] 田中 克己^{†‡}

Hisashi MIYAMORI Qiang MA
Katsumi TANAKA

本稿では、テレビ番組をウェブコンテンツにメディア変換し、その結果を、インターネット上での補完情報検索の結果と統合することで、番組の一覧性と理解性を向上させる方式を提案する。TV番組からウェブコンテンツへのメディア変換によって、テレビ番組の一覧性が向上し、テレビ番組の概要を早見したり、好きなシーンを効率よく探すことができるようになる。また、関連情報との統合によって、付加価値が向上し、キーワード検索の手間をかけることなく関連ウェブページを簡単に閲覧できるようになると考えられる。本稿では、番組録画後のオフライン利用、および、放送を受信しながらのオンライン利用の両方に対応できるように改良した、次世代蓄積型テレビを想定したプロトタイプシステム WA-TV (Webifying and Augmenting TV-content system) を実装、評価することで、提案方式の有効性を確認した。

A method is proposed for enhancing the browsability and the understandability of TV programs that converts TV programs into Web content and integrates the results with complementary information retrieval using the Internet. Converting the programs into Web pages enables the programs to be skimmed over to get an overview and for particular scenes to be easily explored. Integrating complementary information enables the programs to be viewed efficiently with value-added content. Implementing and testing a prototype system for next-generation storage TV, “WA-TV” (Webifying and Augmenting TV-content system), which has been improved for usage of both offline and online browsing, validated the approach taken by the proposed method.

1. はじめに

近年、HDレコーダの性能向上に伴い、一般の利用者が録画できるテレビ番組のデータ量も飛躍的に増加している。現在、600GBの容量をもつHDレコーダが登場し、画質モードによっては1070時間以上の録画が可能である。しかし、録画した映像が増大しても、人間が一日に視聴できる時間には物理的に限界がある。このため、膨大な映像データから興味のある部分を効率よく探したり、内容の概要を手軽に理解したり、限られた時間でハイライトだけを見るといった、情報を効率よく取得する仕組みが重要となる。

また、テレビ番組は、従来からプロの番組制作者によって作成された映像コンテンツであり、一般にリアリティや質が極めて高い。しかし、番組制作のためには公共性や視聴率を確保する必要があるため、一定時間内の番組で提供される情報の詳細さや多様さについては限界があり、利用者にとって必ずしも十分とはいえないことがある。番組内で紹介された関連情報を必要に応じて詳しくあるいは幅広くいろいろな視点で見たいという要求は、人々の生活スタイルの多様化に伴い、今後ますます増加していくと考えられる。このため、番組中に利用者が興味をもった情報や関連する情報を効率よく見ることのできる仕組みが重要となる。

従来、膨大な映像データから興味のある部分を効率よく探したり、内容の概要を手軽に理解するといったさまざまな視聴方法についての研究は活発に行われてきた。しかし、ウェブコンテンツをはじめとする関連情報と関連付けることで、アクセス可能な情報のバラエティを補強するというアプローチはほとんど行われていない。

そこで本稿では、テレビ番組をウェブコンテンツにメディア変換し、その結果を、インターネット上での補完情報検索の結果と統合することで、番組の一覧性と理解性を向上させる方式を提案する。TV番組からウェブコンテンツへのメディア変換によって、テレビ番組の一覧性が向上し、テレビ番組の概要を早見したり、好きなシーンを効率よく探すことができるようになる。また、関連情報との統合によって、付加価値が向上し、キーワード検索の手間をかけることなく関連ウェブページを簡単に閲覧できるようになると考えられる。本稿では、番組録画後のオフライン利用、および、放送を受信しながらのオンライン利用の両方に対応できるように改良した、次世代蓄積型テレビを想定したプロトタイプシステムWA-TVを実装、評価することで、提案方式の有効性を示す。

本稿の構成は以下の通りである。2章では、WA-TVの概要と処理手順について説明する。3章では、実験と考察を示し、4章では関連研究を整理する。最後に5章でまとめを述べる。

2. WA-TVの概要と処理手順

WA-TV (Webifying and Augmenting TV-content system) は、テレビ番組をウェブコンテンツにメディア変換し、インターネット上での補完情報検索の結果と統合する新しい番組コンテンツの視聴方式を提供する。図1, 2に、WA-TVの概念を示す。利用者は、通常のテレビ視聴ができるだけでなく、ウェブコンテンツにメディア変換された番組内容一覧を示す画面へ自由に行き来することができる(図1)。また、変換されたウェブコンテンツは番組内容の一覧を示すだけでなく、番組内容に関連したウェブページへのリンクを統合することで番組内容の情報を補強している(図2)。

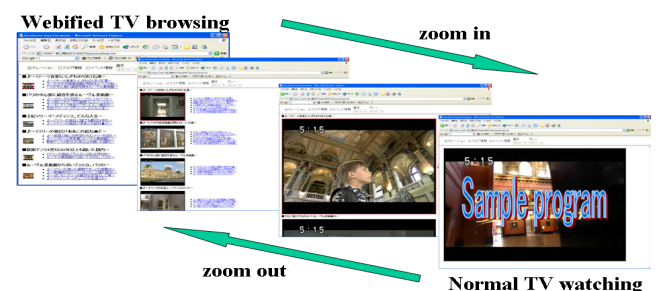


図1 WA-TVの概念(1)
Fig.1 Concept of WA-TV (1)

[†]正会員 情報通信研究機構 miya.qiang@nict.go.jp

[‡]正会員 京都大学 ktanaka@i.kyoto-u.ac.jp

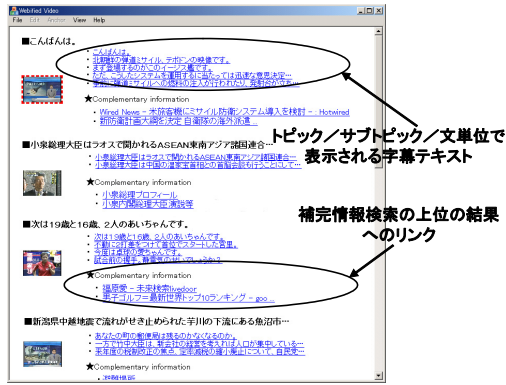


図2 WA-TV の概念(2)
Fig.2 Concept of WA-TV (2)

WA-TV では以下のように処理を行う。

まず、テレビ番組を録画し、クローズドキャプション(以後、字幕データと呼ぶ)を用いた階層的なトピック分割を行う。

番組録画後にオフラインで処理をする場合においては、トレーニングデータを特に必要とせず、任意のドメインで利用できるという利点のある、単語分布に基づくトピック分割手法[1]を用いることにした。このトピック分割を字幕データ s に適用させることでトピック単位に分割された結果 S を得る。分割された各トピックの字幕データを再び s として同様のトピック分割を行うことで、サブトピック単位に分割された結果 S を得る。映像データは、字幕データの分割結果から得られる時刻に基づいて分割される。

以上より、番組映像は、トピック単位、サブトピック単位、文単位の3つの異なる詳細度データに分割される。これを図3のように階層的に構造化する。

トピック 1		トピック 2		トピック i	
サブトピック 1	サブトピック 2	サブトピック j	...
文 1	文 2	文 k	...

図3 異なる詳細度で分割された字幕および映像の構造化

Fig.3 Structuring of closed captions and video segmented into different levels of details

番組録画後ではなく、放送時に受信しながらオンラインで処理する場合においては、字幕データは、番組全体のテキストとしてではなく、連続的に受信される断片的なテキストとして処理される必要がある。ここでは、受信した字幕データに対する、文の識別を行わず、一回分の受信データをブロックと呼び、分割処理の最小単位とする。

ここでは、同じストーリーや話題について述べている字幕データの類似度と語の共起関係は高いと考え、類似度と共起関係の値およびその変動幅に注目して、セグメントの切れ目を決める手法をとることとした。本手法は、受信された字幕データの中でキーワードペア間の共起関係を計算し、共起関係が強ければ、それらの字幕データは一つ的话题を述べてい

る可能性が高いと考える。この処理でサブトピック単位の分割結果が得られる。そして、この分割結果を適宜結合することでトピック単位の分割結果を求める。結合は、各セグメントの話題構造の結合に基づいて行なう。話題構造は主題語と内容語の集合のペアから構成され、ここでは、主題語は、対象テキスト中、出現頻度が高く、かつ、その他のキーワードとの共起関係の強いキーワードとし、内容語は、対象テキスト中、主題語との共起関係の強いキーワードとした。話題構造の結合は、主題語・内容語を表す頂点の集合と、キーワード間の主題語・内容語関係を表すエッジの集合を用いて表現される、一つの連結成分からなる有向グラフ(話題グラフと呼ぶ)の和によって実現される。話題グラフの和は、その和集合が一つの連結成分からなる有向グラフである場合はその和集合を結果とし、それ以外の場合は空集合で定義される。

以上により、番組映像は、トピック単位、サブトピック単位、ブロック単位の3つの異なる詳細度データに分割される。

次に、分割された各セグメントの内容を補完する関連情報を補完情報検索によって取得する。補完情報検索は、上述の話題構造に基づき構造化質問を生成することで、従来の類似検索とは異なった、より詳細なあるいは別の観点からのウェブページを検索できる手法である[2]。

最後に、分割された字幕データ、番組映像、および、対応する関連情報は、ストーリーボード形式に表示されるようなウェブコンテンツとして出力される(図2)。このウェブコンテンツの表示は、ズーム操作によって、分割映像の大きさが連続的に変化するように、また、異なる詳細度データへの遷移が適宜行われるように制御される。これにより、通常のテレビ視聴画面とストーリーボード閲覧画面をシームレスに移動することができ、番組中の興味のあるシーンを効率よく探索することができる(図1)。また、関連情報へのリンクが統合されているため、状況に応じてより詳しい/幅広い情報へ効率よくアクセスすることができる。

3. 実験および考察

3.1 オフライン分割手法

ニュース、ドキュメンタリー、ドラマ、スポーツのテレビ番組を対象として、字幕テキストを用いたオフライン階層トピック分割の精度について調べた。

表1 オフライン分割結果

Table 1 Result of off-line segmentation

	適合率	再現率	F 値
1次分割	0.63 (32/51)	0.43 (32/74)	0.51
2次分割	0.55 (97/177)	0.37 (97/262)	0.44

その結果、ニュースについては1次分割の適合率・再現率がそれぞれ約0.6、約0.4、2時分割の適合率・再現率がそれぞれ約0.5、約0.4であった(表1)。ただし、これは分割位置の一致を厳密に計算した場合の結果であり、主観的にはトピック境界でほぼ正確に分割されていることを確認した。正確としない例としては、1トピックが20文以上の長い冒頭ニュースの場合、1次分割で2つ以上のトピックに分割されることがあることが挙げられる。しかし、長いニュースは2つ以上に分割されると内容を把握しやすくなるため、実効的にはあまり問題にはならないことを確認した。しかし、トピック境界があいまいになりがちで、ドキュメンタリー、バラエティ、ドラマ、スポーツについては、単語分布だけに頼ら

ない改良が必要であることがわかった。文脈の転換点等を示す単語・音響・動画等から境界を推定する分割方法を考える必要がある。

3.2 オンライン分割手法

まず、28ヶ月間のNHKニュース7の字幕データを利用して、共起度辞書を作成した。この共起度辞書を利用して、字幕データのセグメンテーションに必要な語の共起度を調べた。なお、共起度辞書に登録されてない語のペアの共起度は0とし、ストップワードを省くため、940ワード（そのうち593個は英単語）のストップワードリストを作成した。

正解の判断モデルは、Informedia[3]で利用されているモデルを修正したものを利用する。すなわち、システムが判定したセグメントの境(identified boundary)と、人間が判定したセグメントの境(reference boundary)との距離(ブロックの数)が1以内であれば、システムが正しく判定したとする。

実験では、二日間の字幕データ(NHKニュース7)を用いた(821ブロック)。人間によって判定されたセグメントは73個であった。実験結果を表2に示す。これより、Informediaのインクリメンタル・クラスタリング手法の最善のF値(0.367)[3]より、提案手法がよい結果を出していることがわかる。ただし、表1と比較すると、オンライン分割よりもオフライン分割の精度の方がよいことがわかる。

表2 オンライン分割結果

Table 2 Result of on-line segmentation

	適合率	再現率	F 値
ストーリー分割	0.330	0.426	0.372
トピック分割	0.387	0.360	0.373

3.3 補完情報検索の妥当性

補完情報検索で得られたウェブコンテンツの妥当性に関して、トピック分割がほぼ正確に行えたニュースに関して調べた。関連する補完情報を2名の評価者が選択して判断することで調べた結果、詳細な内容の補完の適合率は約0.6、幅広い内容の補完の適合率は約0.7程度であった(表3)。補完情報検索の結果を複数、利用者に提示するようにすれば、補完情報としてある一定の精度で適切な内容を提供できると考えられる。今後より多くの被験者による評価実験を行う必要があると考えられる。

表3 補完検索結果の適合率

Table 3 Precision of complementary retrieval results

	適合率
詳細な内容の補完	0.63 (32/51)
幅広い内容の補完	0.71 (36/51)

3.4 ズーミングインタフェースの評価

テレビ番組を早送り・巻戻しする従来のインタフェースと比較する形で、ズーミングインタフェースの使用感に関する評価実験を行った。年齢と性別で分類した計126名を対象とし、実験データは、30分のニュース番組5本を用いた。

被験者の方には、番組内容の概要や詳細をたずねる質問にズーミングインタフェースおよび従来インタフェースを用いて回答して頂き、画面表示を見たときの内容把握のしやすさ、内容理解の深さ、内容把握までにかかる時間、操作のしやすさといった項目に対するアンケートに5段階評価で記入して頂いた。比較に用いた操作画面を表4に示す。

表4 オフライン分割結果

Table 4 Result of off-line segmentation

操作画面	構成
操作画面1	階層データ(図4で示した3階層)、ズームあり、自動トピック分割
操作画面2	階層データ(図4で示した3階層)、ズームあり、人手によるトピック分割
操作画面3	階層データ(図4で示した3階層)、ズームなし
操作画面4	階層データ(図4で示した文単位のみ)、ズームあり
操作画面5	従来の早送り・巻戻しインタフェース

図4の各操作画面の使用感の結果(各プロットに95%信頼区間を合わせて示す)をみると、階層データありに相当する操作画面1~3の結果はほぼ同等の結果となった。自動および人手によるトピック分割やズームの有無による違いは、操作画面の内容把握のしやすさ、内容理解の深さ、内容把握までにかかる時間、操作のしやすさに対して実質的に大きな影響を与えないことが示されていると考えられる。また、操作画面1~3に比べ、階層なし(操作画面4)や従来インタフェース(操作画面5)の結果はいずれも低くなっており、階層データを構成する有効性が示されている。

一方、図5の今後使っていきたい操作画面のアンケート結果を見ると、ズームありの操作画面1 or 2の方がズームなしの操作画面3より、より多くの利用者に受け入れられる可能性がある点が示されている。要因の一つとして、ズームありの方が階層の切り替わりをスムーズに表現できるため、切り替わりによる文脈変化を容易に維持でき、より多くの人にわかりやすいインタフェースを提供できるためと考えられる。

また、操作画面1 or 2と操作画面5の結果を比較すると、ほぼ同程度の数の被験者が従来インタフェースを今後も使いたいと感じていることがわかる。これは、使用感としてはある程度優れていると感じるズーミングインタフェースに興味を示す一方で、これまで慣れ親しんでいるインタフェースがいかに根強く利用者に浸透しているかを示す結果となっている。今後、新しいインタフェースをデザインする際には、従来インタフェースをある程度配慮した操作環境を考える必要があることを示している。インタフェースを設計する上での一つの興味深いヒントになると考えられる。

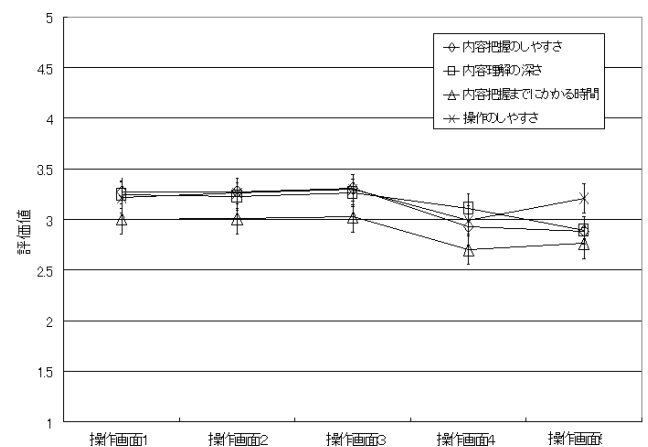


図4 各操作画面の使用感の比較

Fig.4 Comparison of usability of each user interface

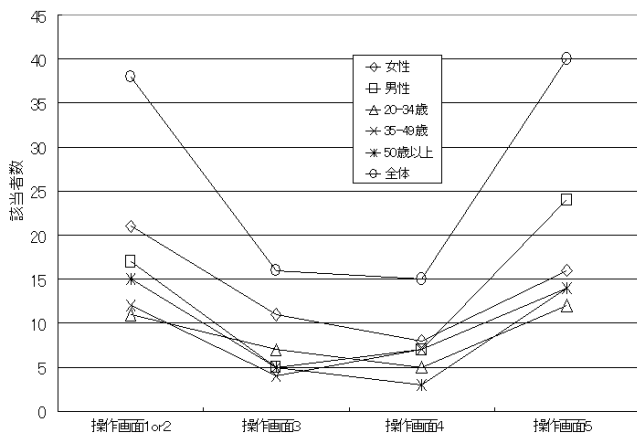


図5 今後使いたい操作画面

Fig.5 User interface one wants to use in the future

3.5 考察

WA-TV の重要な点は、テレビ番組をウェブコンテンツにメディア変換することにより、これまで受動的に視聴するスタイルが一般的だった番組映像を、より能動的に「閲覧」することができるようになった、という点である。本稿では、番組内の関連箇所へジャンプするハイパーリンクと、より詳しいあるいは幅広い関連情報へジャンプするハイパーリンクをウェブ上に統合することで能動的閲覧を実現した。

さらに、ウェブコンテンツに変換されていることを利用すると、複数番組の情報統合も可能になると考えられる。例えば、複数の番組を同一の日付でまとめたグループに対する一覧を表示したり、同一番組・同一トピックを時系列的にまとめたグループに対する一覧を表示すること等が考えられる。これらの表示は、その階層レベルによって、動画像と音響データで再生されたり、テキストデータと音響データのみで再生されたりしてもよい。さまざまな提示方法が提案法の枠組みの中で実現できると考えられる。

番組の分割については、本稿では、字幕データを用いたオフライン分割、および、オンライン分割の2通りの手法を実装した。実験の結果からもわかるように、一般に、オフライン分割の方がオンライン分割に比べて精度の高い結果が得られる。オフライン分割は、番組録画が終了した後に番組視聴をする場合に適用され、見たいシーンを探するためのトピック一覧の正確性を重視した見方を提供する際に有効と考えられる。一方、オンライン分割は、番組放映中に視聴する場合に適用され、より高い頻度で多くの補完ウェブページを提示する見方を提供する際に有効であると考えられる。

本稿では、テレビ番組とその字幕データに基づき、異詳細度のコンテンツ間をズーム遷移可能なウェブコンテンツを動的に生成した。映像の適当な分割手段と内容に関するメタデータ抽出/利用が可能であれば、テレビ番組以外の映像にも提案手法の対象を広げていくことができると考えられる。今後、適宜実験を行い、さらなる有用性の向上を目指したい。

4. まとめと今後の課題

本稿では、テレビ番組をウェブコンテンツにメディア変換し、インターネット上での補完情報検索の結果と統合する新

しい放送コンテンツの視聴方式を提案した。また、次世代蓄積型テレビを想定したプロトタイプシステム WA-TV を実装した。これにより、TV 番組からウェブコンテンツへのメディア変換によって、テレビ番組の一覧性が向上し、テレビ番組の概要を早見したり、好きなシーンを効率よく探すことができるようになることを確認した。また、関連情報との統合によって、付加価値が向上し、キーワード検索の手間をかけることなく関連ウェブページを簡単に閲覧できるようになることを示した。

今後は、より多くの被験者による利用実験を行うことにより、一覧性と理解性の向上がもたらす新しい番組視聴方式の有効性、実用性等をさらに踏み込んだ形で実証することができると期待される。

[文献]

- [1] Utiyama, M., Isahara, H.: A Statistical Model for Domain-Independent Text Segmentation. ACL/EACL-2001, pp. 491-498.
- [2] 馬 強, 田中 克己: 話題構造に基づく放送と Web コンテンツの統合のための検索機構. 情処論文誌: データベース Vol.45 No.SIG 10 (TOD23), pp.18-36, 2004.
- [3] Huauptmann A., Chang J.C., Hu N.N., and Wang Z.R.: Text Segmentation in the Informedia Project, <http://www-2.cs.cmu.edu/~hnn/project/ML-project/ml-report.htm>.
- [4] Christel, M.G., Huang, C.: Enhanced access to digital video through visually rich interfaces. ICME, MD-L5.1, 2003.
- [5] Sumiya, K., Munisamy, M., Tanaka, K.: TV2Web: generating and browsing web with multiple LOD from video streams and their metadata, ICKS2004, pp.158-167, 2004.
- [6] Uchihashi, S., Foote, J., Girgensohn, A., Boreczky, J.: Video Manga: generating semantically meaningful video summaries. Proc. ACM Multimedia 99, 1999.
- [7] 橋本隆子, 白田由香利, 真野博子, 飯沢篤志: TV 受信端末におけるダイジェスト視聴システム, 情処論文誌, Vol.41, No.SIG3(TOD6), pp.71-84, 2000.

宮森 恒 Hisashi MIYAMORI

独立行政法人情報通信研究機構(NICT)メディアインタラクショングループ主任研究員。1997 年, 早稲田大学大学院後期博士課程了。1996~1997 年, 同大理工学部助手。工学博士。主に、マルチメディアコンテンツ処理に関する研究に従事。ACM, 電子情報通信学会, 情報処理学会, 映像情報メディア学会, 日本データベース学会各会員。

馬 強 Qiang MA

独立行政法人情報通信研究機構(NICT)メディアインタラクショングループ専攻研究員。京都大学大学院情報学研究科博士後期課程修了。博士(情報学)。主に、情報検索, マルチメディア情報システム, 情報統合の研究に従事。IEEE, ACM 等各会員。

田中 克己 Katsumi TANAKA

京都大学大学院情報学研究科社会情報学専攻教授。独立行政法人情報通信研究機構メディアインタラクショングループリーダ。1976 年京都大学大学院前期博士課程修了, 工学博士。主にデータベース, マルチメディアコンテンツ処理の研究に従事。IEEE Computer Society, ACM, 人工知能学会, 日本ソフトウェア科学会, 情報処理学会, 日本データベース学会各会員。