

任意画面サイズに対する映像・関連情報提示バランスの最適化

Optimizing Presentation Balance of Video and its Related Information to Any Screen Size

粕谷 英司†
Eiji Kasutani

山田 昭雄†
Akio Yamada

1. はじめに

放送のデジタル化やネットワークのブロードバンド化に伴い、映像を視聴しながら、関連する情報をインターネットやデータ放送等を通じて参照する視聴者が増えてきている。筆者らは、映像とその関連情報とあわせて同一画面に提示する際に、端末の画面サイズおよび利用者の嗜好にあわせて変換・提示する方法について検討してきた[1]。しかしながら、この方法は番組全体を通じて映像・関連情報の配置バランスを決定するため、映像区間に応じて表示画面を十分に活用できないという問題点があった。本稿では、利用者に提示される映像および対応する関連情報の価値を評価するための尺度を定義し、映像区間毎にその価値を最大化するように映像および関連情報を変換・提示する方法を提案する。

2. 映像・関連情報の同一画面への提示

利用者がある映像区間を視聴中に端末から関連情報を要求した際に、映像区間と、対応する関連情報とが同一画面に提示される。図1に処理手順を示す。本方式では、表示画面、映像解像度および関連情報の構造を参照して、映像区間および関連情報の変換・配置方法を決定する。これらの変換・配置方法は、利用者に提示される映像および関連情報の価値が表示画面サイズの範囲内で最大になるように決定する。ここで価値とは、原映像および関連情報が、変換・配置後にどれだけ忠実に再現されるかを示す指標とする(コンテンツ価値と呼ぶ)。決定された条件に基づき映像および関連情報を変換・配置する。

映像と関連情報の配置は、表示画面を $p:1-p$ ($0 \leq p \leq 1$) に分割し、一方の領域を映像の表示に、他方を関連情報の表示に割りあてることにより行う。配置方法および変換方法の最適化問題は、コンテンツ価値を最大にする p を求める問題に帰着することができる。各 p におけるコンテンツ価値は、表示される映像の価値と、表示される関連情報の価値を変数とする関数として記述できる。

3. 評価関数の定義

利用者に提示される映像・関連情報の価値を評価するためには、それぞれの価値を定量的に表現する必要がある。変換・提示された映像の価値 $g(p)$ と、関連情報の価値 $h(p)$ を同一次元で評価することでそれぞれの価値を算出する。合計値をコンテンツ全体の価値 $V(p)$ とする。

$$V(p) = g(p) + h(p) \quad (1)$$

なお、表示画面は縦、横のどちらにも分割できるため、それぞれの分割に対するコンテンツ価値を算出し、より高

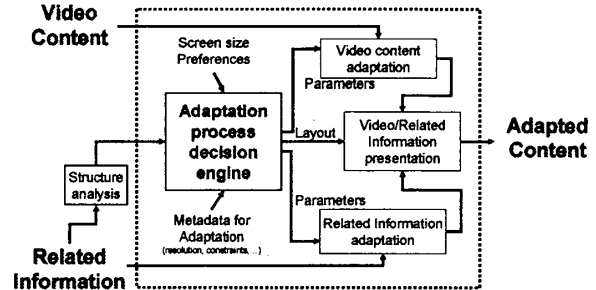


図1. 映像・関連情報の同一画面への提示処理

表1. 各モダリティの価値評価軸および評価項目

評価軸	映像		テキスト
	空間	時間	
信号レベル (物理量)	画素数	フレーム数	文字数
知覚レベル (見やすさ)	画質, 視認性等	再生のスム ーズさ, etc	必要ページ数, 表示時間等
意味レベル (重要度)	各領域の 重要度	各映像区間 の重要度	各表示対象項目 の重要度

い価値を持つ分割方法を採用する。

価値算出のための評価軸を信号レベル、知覚レベル、意味レベルの3つの次元に分類する。表1に、価値評価軸と評価項目を示す。映像・関連情報の同時視聴という目的を考慮して、限定条件を設けて評価関数を設定する。具体的には、映像は表示領域内に収まること、関連情報は表示領域内に収まらない場合には次ページに提示されること、映像の画質変化がないこと、また関連情報としてはテキストのみを常に同じ大きさで表示するものとする。以降、解像度が横 x 画素、縦 y 画素の表示領域における映像・関連情報の価値をそれぞれ $g(x,y)$, $h(x,y)$ とする。

3.1 信号レベルの評価

物理量を尺度に評価する場合、変換後の映像の画素数、および関連情報の文字数が評価対象とする。映像については、アスペクト比を変えないことなく、映像表示領域にできるだけ広く表示できるように変換する。映像の拡大縮小率を b とおくと、 b は画素数の割合を反映するため、 b を変数とする関数で表現できる。 b を評価値として利用する場合、映像価値 $g(x,y)$ は以下のように定義できる。

$$g(x,y) = b, \quad b = \min(b_1, b_2), \quad b_1 = x/x_{orig}, \quad b_2 = y/y_{orig} \\ \text{if } (b > b_{max}) \quad b = b_{max}; \text{ if } (b < b_{min}) \quad b = 0 \quad (2)$$

ただし、 b : 映像の拡大縮小率、

(x_{orig}, y_{orig}) : 原画像の高さ、幅、

b_{max}, b_{min} : 拡大縮小率の最大値、最小値。

関連情報価値 $h(x,y)$ は、変換後の関連情報に含まれる文字数 m' を、変換前の文字数 m で正規化することで求める。

$$h(x,y) = m'/m \quad (3)$$

3.2 知覚レベルの評価

知覚を尺度に評価する場合、映像の見やすさ、関連情報の見やすさを評価対象とする。関連情報については、表示に必要なページ数が少ないほど情報全体が見やすくなるため、ページ数に反比例して高い評価を割り当てる。ページ数を評価値とした場合、以下のように定義できる。

$$h(x, y) = h(l) = \begin{cases} 1 & (l < L) \\ (l_{\max} - (l - L)) / l_{\max} & (L < l < l_{\max}) \\ 0 & (l > l_{\max}) \end{cases} \quad (4)$$

l : 選択された項目全てを表示するのに必要な行数

L : 画面内に表示できる行数

l_{\max} : 最大行数 (最大許容ページ数 $P_{\max} \times L$)

映像については、符号化条件やノイズによる画質の変化を考慮する。ただし本稿では画質は一定、映像は常に画面に収まるように変換されるため、 $g(x, y)$ は次式を満たす。

$$g(x, y) = 1 \quad (5)$$

3.3 意味レベルの評価

情報の重要度を尺度に評価する場合、表示できる映像または映像領域の価値、関連情報として表示される各項目の情報価値を評価対象とする。まず映像については、評価の対象となる映像区間に登場する人物や物体それぞれに重要度が割り当てられているものとする。登場物などの注目領域(Region-of-Interest (ROI))が n 個存在し、各 ROI の重要度を z_k ($1 \leq k \leq n$)、各 ROI の表示の有無を $c_1, c_2, \dots, c_k, \dots, c_n$ ($1 \leq k \leq n$)、 k 番目の項目が映像全体を通して表示される場合には $c_k = 1$ 、全く表示されない場合には $c_k = 0$ とする。定められた表示領域の範囲内にできるだけ重要度の合計が高くなるように ROI の組合せを選択する。

$$g(x, y) = \max \left\{ \sum_{k=1}^n \{c_k \cdot z_k\} / \sum_{k=1}^n z_k \right\} \quad (6)$$

一方、関連情報については、関連情報として表示できる項目が m 個あり、各項目の情報価値を u_k ($0 \leq u_k \leq 1$) とする。また、表示の有無を $a_1, a_2, \dots, a_k, \dots, a_m$ ($1 \leq k \leq m$) とする。 k 番目の項目を表示する場合には $a_k = 1$ 、しない場合には $a_k = 0$ とする。 $h(x, y)$ は以下の式で表現できる。

$$h(x, y) = \max \left\{ \sum_{k=1}^m \{a_k \cdot u_k\} / \sum_{k=1}^m u_k \right\} \quad (7)$$

3.4 評価関数の統合

3.1-3.3.で示した物理量・見やすさ・重要度による評価式の積を求めることにより、3つの要素を統合した評価関数を生成する。映像については、与えられた映像表示領域に対して表示する ROI とその拡大縮小率の最適組合せを求める問題として記述できる。ROI が定められていない場合には、フレーム全体を1つの ROI とする。式(2)(5)(6)を統合することにより、 $g(x, y)$ は以下のとおり表現できる。

$$g(x, y) = \max \left\{ \sum_{k=1}^m b_k \cdot \{c_k \cdot z_k\} / \sum_{k=1}^m z_k \right\} \quad (8)$$

関連情報については、提示される項目全体の持つ重要度と、見やすさとのトレードオフを考慮した関数になる。式(4)は関連情報に含まれる文字数から算出できるため、式(3)を実質的に含むため、式(4)(7)から $h(x, y)$ を算出する。

$$h(x, y) = \max \left\{ h(l) \cdot \left(\sum_{k=1}^m \{a_k \cdot u_k\} / \sum_{k=1}^m u_k \right) \right\}, h(l) = \begin{cases} 1 \\ (l_{\max} - (l - L)) / l_{\max} \\ 0 \end{cases} \quad (9)$$

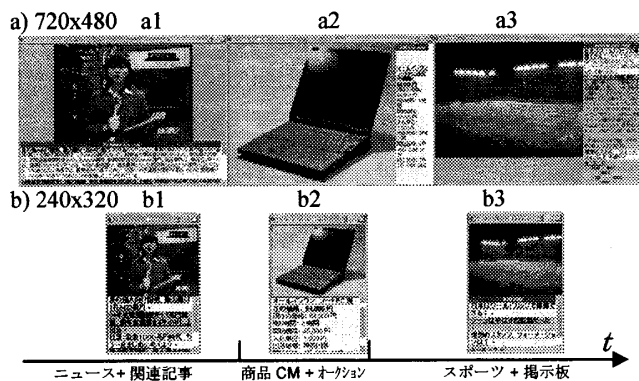


図2. 複数の画面サイズへの映像・関連情報提示例

4. 検証

任意の画面サイズで映像を見ながら関連情報を参照できるテストベッドを、式(1)(8)(9)を評価関数として実装した。複数種類の場面を含む映像と、各映像場面に対応する関連情報を用いて実験を行った。映像の解像度は 640x480 画素である。関連情報としては、各種 Web ページからタイトルと詳細テキストの組合せを抽出して用いた。重要度は全項目同一にし、 P_{\max} は 16 と設定している。

図2に、各映像区間と関連情報の、複数の画像サイズへの表示例を示す。ニュース映像、スポーツ映像については、対応する関連記事見出し一覧、および掲示板の投稿タイトル一覧のみが b)240x320 の画面サイズで表示されている。これに対して a)720x480 ではニュース映像の見出しのみならず詳細記事まで表示され、スポーツ映像に対しても掲示板の内容全体が表示されている。一方、商品 CM 区間に対しては、a)b)ともにタイトル・詳細の双方が表示されている。関連記事や掲示板と比べて、提示する関連情報が少ないためである。映像表示領域については、関連記事や掲示板情報など提示すべき関連情報が多い場合には関連情報の表示を妨げない程度に確保されている。また、オークション情報のように情報量が少ない場合には広く確保されることで映像区間毎にコンテンツの視認性が高められている。従来方式[1]では図2における a1,b1 の配置バランスで常に映像・関連情報が提示されるのに対して、提案方式では、利用者端末の持つ画面サイズが映像区間に応じて有効に活用できていることがわかる。

5. まとめ

利用者に提示される映像および対応する関連情報の価値を評価するための尺度を定義し、映像区間毎にその価値を最大化するように映像および関連情報を変換・提示する方法を提案した。本方式の特徴は、各映像区間に対応する関連情報の内容および表示画面に応じて映像・関連提示バランスを決定する点、アダプテーションの最適化が、利用者に提示されるコンテンツ価値を最大化することにより行われる点である。主観評価実験および利用場面に応じた評価尺度の設定が今後の課題である。

参考文献

- [1] E.Kasutani, T.Ebrahimi: "Preference-based Multimedia Content Adaptation for Screen Size Independent Access", Special session on UMA in WIAMIS 2005, April 2005.