

放送コンテンツ間の関連性に基づいた検索システム “ローミングナビ”TM

"Roaming Navi"TM Content-Centric User Interface

山内 康晋 鈴木 優 安次富 大介

■ YAMAUCHI Yasunobu

■ SUZUKI Masaru

■ AJITOMI Daisuke

インターネットやデジタルテレビ (TV) 放送などで膨大な数のコンテンツが提供されるようになり、従来の番組表や録画リストなどの一覧表示型では、目的のコンテンツの発見に時間が掛かるという問題が出てきている。

東芝は、コンテンツ間の関連度を基にしたコンテンツの探索や発見が可能なユーザーインターフェースを開発し、この概念に基づいたコンテンツ検索システム ローミングナビを実現した。コンテンツ間の関連性の根拠や強さに応じて、表示する方向や距離を変化させることで、より直感的な探索が可能となり探索効率を大幅に改善できる。また、ユーザー評価に基づいた関連度算出手法の開発により、妥当性の高い関連コンテンツの提示ができるようになった。

Users are increasingly demanding a means of performing effective searches for media contents of interest to them from among the enormous volume of broadcast and Internet contents available, instead of the conventional list-based type of graphical user interface.

In response to this need, Toshiba has developed Roaming Navi, a relevance-driven content search system utilizing a content-centric user interface based on relationships among media contents to access or find new media contents of interest to the user. Relevant contents are displayed on the screen according to the relevance information, with the distance and direction of the relevant content from the center content indicating the strength and attribute of the relevance, respectively. We have also developed a feasible relevance calculation scheme based on user evaluation tests, and confirmed its effectiveness.

1 まえがき

デジタル機器の処理性能の向上とブロードバンド (高速大容量通信) 接続の普及による、放送コンテンツ, IPTV (Internet Protocol Television) やVoD (Video on Demand) などの映像配信サービスの増加を背景に、一般の家庭でも膨大な数のコンテンツを扱えるようになってきた。このような状況のなか、番組表やリスト表示といった従来のユーザーインターフェースでは、目的のコンテンツにたどり着けない、あるいは膨大なページめくりが必要であるといった課題が見えてきている。

一方、日常生活の中でTV離れが進んでいると言われており、その理由として“ほかにすることが増えた”、“見たい・おもしろい番組がない”、などが挙げられている⁽¹⁾。近年の生活スタイルの変化により、TV視聴の優先順位が低下していることもあるが、膨大な映像の中に見たい番組があるにもかかわらず、ないと思いきや、膨大なコンテンツの中から、ユーザーの興味のある番組や映像コンテンツを限られた時間の中で効率よく選別して提示する技術が求められている。

2 コンテンツ指向ユーザーインターフェース

東芝は、視聴中、あるいはユーザーの好みのコンテンツを手

がかりに関連するコンテンツを提示することで、興味のあるコンテンツの探索や発見を促すナビゲーションインターフェースを開発した⁽²⁾。関連度の算出、関連度の高さに基づいた関連コンテンツの選別と提示、そしてユーザーによるコンテンツの選択という一連の処理を特徴としていることから、“コンテンツ指向ユーザーインターフェース”と呼んでいる。このユーザーインターフェースの概念に基づき、TV番組間の関連性を基に膨大なコンテンツの中からユーザーの興味に沿ったコンテンツを効率よく探索できるコンテンツ検索システム“ローミングナビ”を開発した。

ローミングナビでは、図1に示すように、ユーザーが目しているTV番組コンテンツ (以下、注目コンテンツと呼ぶ) が画面中央に配置され、これと関連するコンテンツ (以下、関連コンテンツと呼ぶ) 群が関連性の強さに応じて同心だ円状に配置される。ユーザーは現在視聴している番組や番組表画面で選択可能な番組、更には録画リスト画面で選択可能な録画番組などを注目コンテンツとし、ローミングナビに移行することができる。注目コンテンツに近い位置ほど関連性が強く、遠くなるほど関連性が弱い。画面は分類軸により上 (タイトル)、下 (キーワード)、左 (ジャンル)、右 (人物) の四つのエリアに分かれており、関連コンテンツは関連性の根拠によって分類されて配置される。

ユーザーは、リモコンやマウスを用いて、配置された関連コ



図1. ローミングナビにおけるコンテンツの配置 — 注目コンテンツとの関連性の強さを中央からの距離に、関連性の根拠を上下、左右の配置方向に対応づけてコンテンツを配置することで、効率よくコンテンツを探索できる。

Layout of contents in Roaming Navi display

コンテンツ間をカーソル移動できる(図2)。このカーソル移動動作により、フォーカスが画面範囲を超えた場合、現在フォーカスが当たっている関連コンテンツが存在する分類軸方向に画面がスクロールし、特定軸内での関連コンテンツの配置画面に遷移する。

人物軸方向へカーソルを進めた場合の例を図3に示す。人物軸に配置された関連コンテンツが、メイン画面と同様に関連性の強さに応じて配置されている。更にカーソルを進めると、この画面内で注目コンテンツの表示を残したまま、関連コンテンツ群の提示部分がスクロールされ、より下位の関連コンテンツにアクセスできるようになる。

関連コンテンツにフォーカスが当たると、注目コンテンツとの関連情報(関連性の根拠)がポップアップ表示される。関連情報として、番組タイトルの類似具合、注目コンテンツと関連コンテンツで共に出演している人物や共に扱っている話題を表すキーワード、共通しているジャンルなどが表示される。また、選択した関連コンテンツを注目コンテンツに変更して再検索できるため、ユーザーは、この動作を繰り返しながら、自分の興



図2. ローミングナビのメイン画面 — ユーザーは、リモコンやマウスを用いて配置された関連コンテンツ間をカーソル移動できる。フォーカスが当てられた関連コンテンツは拡大表示される。

Main screen of Roaming Navi



図3. ローミングナビにおける関連情報の提示 — 人物軸方向にカーソルを進めた例である。フォーカスの当たっているコンテンツは、注目コンテンツと共通に登場する人物をポップアップで表示することで、関連性の根拠を確認できる。

Display of relevance information in Roaming Navi

味のあるTV番組を発見できる。

このように、ローミングナビでは、従来のキーワード入力型のTV番組検索システムと異なりキーワードを入力する必要がない。TV番組コンテンツを起点として、上下、左右の方向感覚で、それぞれにマッピングされたタイトル、人物、キーワード、ジャンルといった関連性を基に効率よく探索できる。

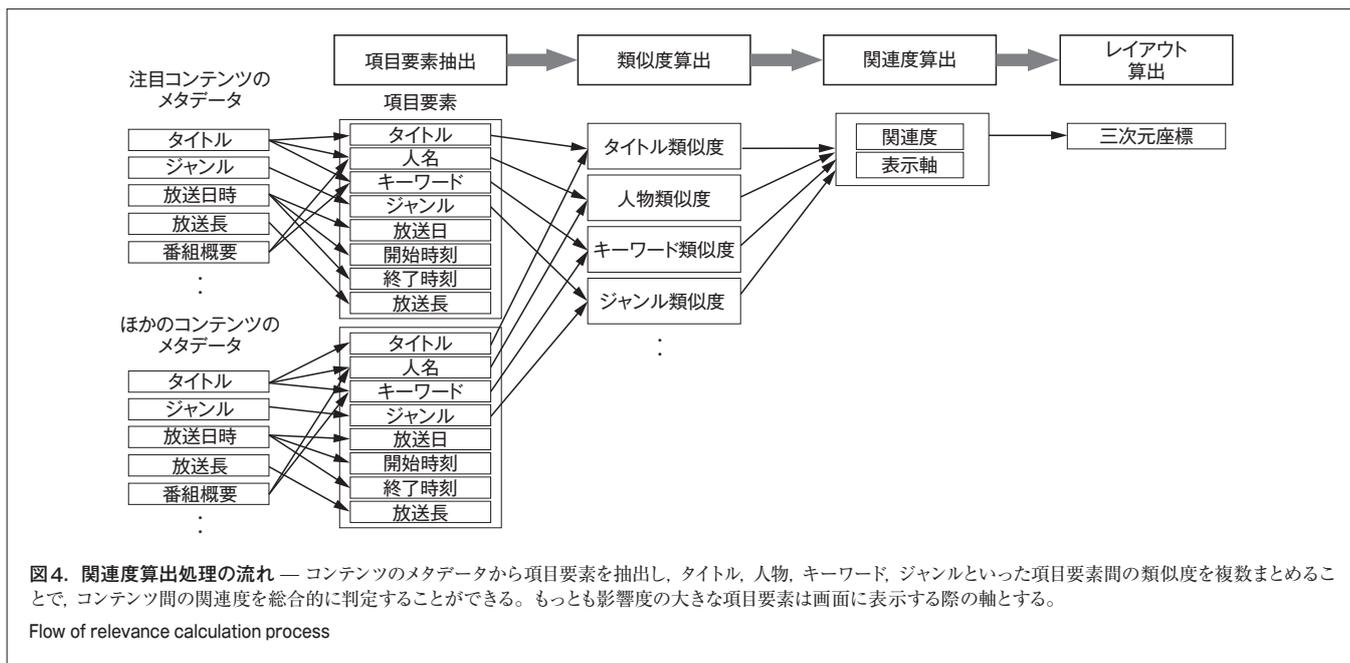
更に、起点となるTV番組コンテンツを切り替えることで関連コンテンツの再検索が行われるため、興味のあるコンテンツをたどることを繰り返しながら様々なコンテンツにアクセスできる。目的のコンテンツを探すだけでなく、今まで自分が見たことのない番組、思いがけないキーワードでつながっていた番組などを探すことが可能となり、新たな番組の発見も期待できる。

3 ローミングナビにおける処理の流れ

ローミングナビにおいて、ある注目コンテンツの周囲に関連コンテンツが配置されるまでの一連の処理を図4に示す。この処理は項目要素抽出と、類似度、関連度、及びレイアウトの算出の4段階から成る。関連度算出の基になるのは放送波などから取得できるEPG(電子番組表)などのメタデータ(番組の付加情報)である。メタデータには各コンテンツのタイトル、ジャンル、放送日時、放送長(放送時間)、及び自然文で記述された番組概要が含まれる。

項目要素抽出処理では、メタデータから図4に示すような要素を抽出する。更に、形態素と語句意味の解析技術⁽³⁾により、タイトルと番組概要から、番組の内容を表すキーワードと、出演者などの人名の情報を抽出する。このときキーワードと人名には語句意味の解析技術により、地名、組織名、スポーツ用語、歴史上の人名、キャラクター名、一般の人名、著名人名など約100種類の意味カテゴリが付与される。

類似度算出処理では、注目コンテンツとそのほかのコンテンツ間の項目要素ごとの類似度を求める。例えば、注目コンテン



つとあるコンテンツの項目要素に共通する人物が3人であれば、これらのコンテンツ間の人物類似度は3.0となる。このとき、共通する人物の意味カテゴリによって類似度の値を補正する。例えば、著名人名が共通する場合は一般の人名が共通する場合よりも人物類似度の値は大きくなる。キーワード、タイトル、ジャンルの類似度なども同様に求める。

関連度算出処理では、複数の項目要素の類似度を総合して注目コンテンツとそのほかのコンテンツ間の関連度を算出する。これは項目要素間の客観的な類似性を、ユーザーが感じるコンテンツ間の感性的な関連性に変換するために必要な処理である。関連度算出処理については4章で詳細に述べる。

注目コンテンツとその他すべてのコンテンツ間の関連度を求めた後、レイアウト算出処理によって注目コンテンツの周囲に関連コンテンツを配置する。中央からの方位方向については関連度算出処理で関連度の値の向上にもっとも寄与した項目要素で判定し、タイトルであれば上、人物であれば右、キーワードであれば下、ジャンルであれば左方向に配置する。また、中央からの距離については算出された関連度の値が大きい関連コンテンツほど中央近くに配置されるようレイアウトを決定する。関連度の値をそのまま中央からの距離に割り当てることもできるが、リモコンなどの十字キーでの操作を可能とするため、関連度の値が大きいものから順に中央から埋めていく方式とした(図2)。

4 外部被験者実験に基づく関連度算出式の精度向上

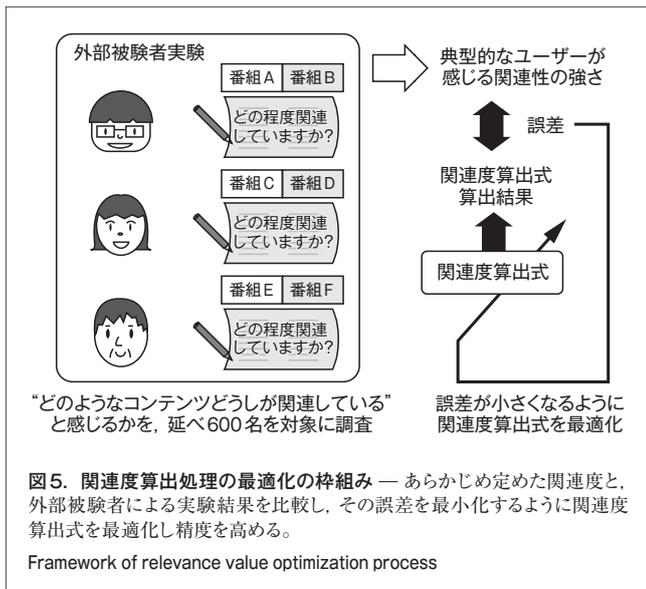
ユーザーの感じるTV番組間の関連性は主観性の高いものであり、関連度算出式の導出には、評価尺度の明確化が重要

である。ここでは、外部被験者実験で得られた、ユーザーが主観的に評価した関連性の値と、関連度算出式により算出された関連度の一致度合い、つまりユーザーごとの回答誤差を考慮して補正した順位相関係数がある一定の値以上であるユーザーを、納得しているとみなす。全ユーザーのうち、納得しているユーザー数の割合をユーザーカバー率と定義し、(1)式に表す。

$$\text{ユーザーカバー率} = \frac{\text{納得しているユーザー数}}{\text{全ユーザー数}} \times 100 (\%) \quad (1)$$

一般に主観評価に基づく統計的な解析手法としてはコンジョイント分析が用いられるが、多種多様なTV番組データを対象とした場合、高精度化には大量の被験者データが必要であった。そこで、項目応答理論 (IRT : Item Response Theory) を用いた新たなアンケート調査手法を適用することで、ひとり当たりの設問数を抑えながら、効率よくコンテンツ間の関連性の強さを表す評価データを取得した。IRTは、TOEFL (Test of English as a Foreign Language) などの能力テストで用いられる手法であり、被験者間の個人差などに起因するばらつきをなくすことで異なる被験者による回答を、同一尺度上で比較評価することを可能にするものである。

今回は被験者ひとり当たりの回答負担は増加させることなく、評価設問数を効率よく増加させることが目的である。被験者600人を20グループに分け、各グループに20問ずつコンテンツ間の関連性の比較評価を割り当てて実施することで、比較評価設問を合計で400個取得できた。また、関連度算出式についても、コンジョイント分析で得られる線形式から(2)式に

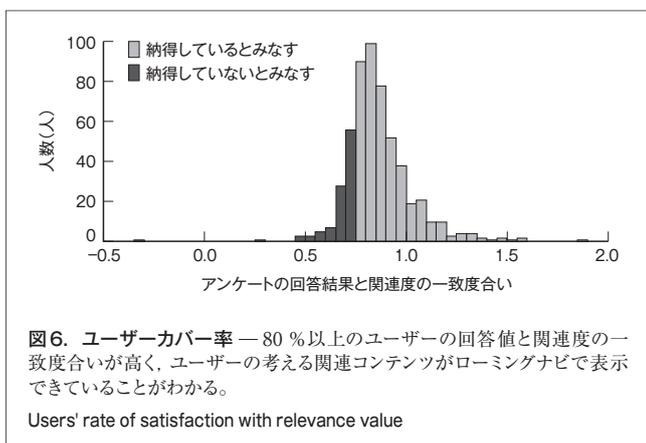


示すような非線形で属性（項目要素）間の相互作用を含む式に拡張を行った。これらの被験者実験の回答結果400問をユーザーの感じるコンテンツ間の関連度として用い、ユーザーカバー率をもっとも高くなるように関連度算出式の最適化を行う（図5）。

$$R = a_1 S_1^{\beta_1} + a_2 S_2^{\beta_2} + \dots + a_M S_M^{\beta_M} + a_{1,2} S_1^{\beta_1} S_2^{\beta_2} + a_{1,3} S_1^{\beta_1} S_3^{\beta_3} + \dots + a_{M-1,M} S_{M-1}^{\beta_{M-1}} S_M^{\beta_M} \quad (2)$$

R：関連度算出値
a：項目要素ごとの影響を制御するための重み係数
S：項目要素ごとの類似度
 β ：項目要素ごとの類似度の影響を制御するための重み係数

最適解探索には実験計画法、遺伝的アルゴリズム、重回帰分析を組み合わせた手法を用いることで、計算量を約1/1,000以下に抑えることができ、従来法で数週間必要であった最適



化が1日以内でできた。以上により最適化した関連度算出式を、前述の400問とは別の24個の評価設問でユーザーカバー率を評価した結果80%以上になり、実用的な精度であることが確認できた（図6）。

5 あとがき

インターネットやデジタルTV放送で提供される膨大なコンテンツの中から、放送種別をユーザーに意識させることなく、コンテンツ間の関連性に基づいた関連コンテンツの提示と選択を繰り返すコンテンツ指向ユーザーインタフェースの概念を述べた。その概念に基づいたコンテンツ検索システムであるローミングナビを開発し、関連性を基にユーザーが関心のあるコンテンツを選択しながら、探索や発見を楽しむという新しい視聴スタイルを提供することができるようになった。

今後は、ユーザーの状況や好みをより反映した関連度の算出及び、対象となるコンテンツの拡充を図っていく。

文献

- 朝日大学マーケティング研究所. “テレビ番組表に関するマーケティングデータ(1)”. <<http://repose.jp/2165/8/66.html>>, (参照2009-12-17).
- 三原功雄, ほか. “テレビ番組間の関連性に基づくコンテンツ検索システム「ローミングナビ」の開発”. 情報処理学会ヒューマンコンピュータインタラクション研究会報告. 2010-HCI-137, 8, 2010, p.1-8.
- 山崎智弘, ほか. 話題抽出エージェントを用いた番組検索システムの実装. コンピュータソフトウェア. 25, 4, 2008, p.41-51.



山内 康晋 YAMAUCHI Yasunobu

研究開発センター ヒューマンセントリックラボラトリー主任研究員。コンピュータグラフィックス及びヒューマンインタフェースに関する研究・開発に従事。情報処理学会、ACM会員。Human Centric Lab.



鈴木 優 SUZUKI Masaru

研究開発センター 知識メディアラボラトリー主任研究員。情報検索技術の研究・開発に従事。情報処理学会、人工知能学会会員。Knowledge Media Lab.



安次富 大介 AJITOMI Daisuke

研究開発センター ネットワークシステムラボラトリー。IPTVを中心としたデジタル家電ネットワークの研究・開発に従事。Network System Lab.