

メタデータを用いた映像コンテンツ検索候補提示手法

紺家 裕子^{†‡} 中茂 睦裕[†] 草地 良規[†]

[†]NTT コミュニケーションズ 〒108-8118 東京都港区芝浦 3-4-1

[‡]お茶の水女子大学 〒112-8610 東京都文京区大塚 2-1-1

E-mail: [†]{y.konya, m.nakashige, yoshinori.kusachi}@ntt.com, [‡]{yuko.konya}@is.ocha.ac.jp

あらまし 映像サービスの多様化により映像コンテンツの数は膨大になっており、視聴者が目的のコンテンツを容易に見つけることは困難となってきている。特に、映像受信端末にはコンテンツ検索機能があるが、限られたリモコン操作では満足な検索結果を得られない場合が多々ある。本稿では、検索文字入力時に候補を提示することでコンテンツ検索の操作性を改善する一方法について提案する。具体的には、検索文字の一部入力に対し、逐次絞込み候補を提示する。その際、提示の際にはインターネット上での露出度を定義/算出し、露出度の高いコンテンツ候補を上位に提示する。最後に、提案手法を実装し、映像コンテンツ検索を支援する実験を行った結果について報告する。

キーワード 映像サービス, メタデータ, 映像検索, 文字入力

Search term suggestion tool for video search using metadata

Yuko KONYA^{†‡} Mutsuhiro NAKASHIGE[†] and Yoshinori Kusachi[†]

[†]Innovative IP Architecture Center, NTT Communications 3-4-1 Shibaura, Minato-ku, Tokyo, 108-8118 Japan

[‡]Ochanomizu University 2-1-1 Otsuka, Bunkyo-ku, Tokyo, 112-8610 Japan

E-mail: {y.konya, m.nakashige, yoshinori.kusachi}@ntt.com, {yuko.konya}@is.ocha.ac.jp

Abstract: There is a number of video content because of recent video market growth. It is hard for audiences to reach a content that an audience wants to watch or he/she is interested in. However, many video services have search window in their services, he/she feels irritating because most of the audience uses a legacy remote control as an input device. In this paper, we propose a new method for improving the operability of searching content. Our method suggests content's list limited by input text order of the number of their appearance in the Internet. Experimental results show our method is prior to previous works.

Keyword: TV/Video Service, metadata, video search, text-input

1. まえがき

1.1. 映像配信サービスの現状

映像サービスの多様化により映像コンテンツの数は膨大になっている。地上波、衛星波によるTV放送やインターネットを通じて配信される放送では100Chを越える放送もあり、視聴したいときに視聴を開始できるオンデマンド型コンテンツでは、VODやWEBのストリーミングビデオなどのサービスがあり、数万タイトルを売り文句にしているサービスもある。これらの映像サービスは、有料で提供されるもの、無料で提供されるものがあるが、どちらも多くのユーザーに継続して視聴してもらうことでサービスが成立している。

これらの映像サービスの多くにはメタデータと呼ばれる情報が付与されている。映像に付与されているメタデータとは、映像のタイトル、出演者名、内容、放送時間等の情報のことである。例えば、放送型のサービスで提供されている電子番組表(EPG: Electric

Program Guide)では放送時間、放送チャンネルの一覧を表示し、番組メタデータと呼ばれる番組タイトル、概要、出演者名などの情報が記載されている。VODなどオンデマンド型のサービスでは、コンテンツメタデータと呼ばれるタイトルや出演者名、内容に加えて映像の属するジャンルやキーワード、放送時間ではなく配信可能期間等の情報が付与されているものが多い。オンデマンド型のコンテンツリストのことをECG (Electric Content Guide)という。放送型サービスにおける、番組メタデータ、VODサービスにおけるコンテンツメタデータをあわせて映像メタデータと呼ぶ。映像メタデータは多くのコンテンツの中から、視聴者が目的のコンテンツを発見するために用いられる。

1.2. コンテンツ検索の現状

目的のコンテンツ発見方法として、映像メタデータ内のキーワード検索、ジャンル等の階層検索、レコメンド、EPG/ECGの一覧表示、試視聴・ザッピング(放

送コンテンツにおける試視聴)がある。映像視聴を開始するシーンとしては2パターンあり、一つはユーザが視聴したいコンテンツを決めている場合(目的作品有り)、もう一つは映像を視聴したいという欲求のみで希望のコンテンツが決まっていない場合(目的作品無し)である。表1に示すように、ジャンル等の階層検索やレコメンド、EPG/ECGによる一覧表示、ザッピングによるコンテンツの発見は後者の場合に利用されることが多い。また、サービスの制限などで前者でも階層検索等を用いることもあるが、階層検索では絞り込みに限界があるため、コンテンツの量が多い場合は目的に達するまで多くの時間かかってしまう。よって、大量のコンテンツから目的作品を検索する場合はキーワード検索を用いることが適している。

表1 大量コンテンツからの検索目的別の検索方法とその評価。

検索方法	目的作品有り	目的作品無し
キーワード	○	×
ジャンル	△(絞り込みに課題有り)	○
レコメンド	×	○
EPG/ECG 一覧	△(時間等の情報が必要)	○
試視聴/ザッピング	×	○

○:可能, △:条件付き可能, ×:不可能, 時間がかかる

しかし、目的のコンテンツがわかっており、キーワードによる検索インターフェースが備えられているにもかかわらず、目的のコンテンツにたどりつけなかったり、またはたどりつくまでに時間がかかったりしてしまうことが発生している。MMD 研究所の実施した「2010年テレビ商戦、及びTV VODサービスに関する実態調査」[1]によると、TV上のVODサービスがわかりにくい理由として、「見たいコンテンツにたどりつくまでに時間がかかる:74.2%」、「検索が複雑:37.9%」という調査結果が出ている。(図1)これは、映像サービスの提供者、視聴者ともに不幸な状況である。

検索しても目的のコンテンツにたどりつけない場合、目的を達成できないことから、検索することに疲れてしまい興味を失ってしまうという状況もある。この課題を解決するためには、時間をかけず、目的のコンテンツに到達できる検索方式を提供する必要がある。

TV VODサービスが使いづらい理由

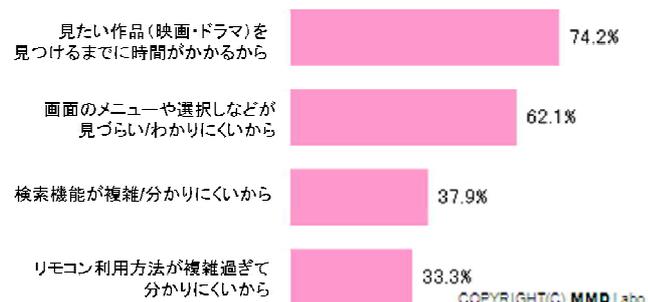


図1 調査結果 TV VOD サービスが使いづらい理由

また、図1の調査結果にもあるように、特に、TVやセットトップボックス(STB)といった映像受信端末では、入力操作がテンキーのリモコンであり、キーボードとマウスが付属したパソコン等とは操作性が異なるため、映像発見の行動に少なからず影響を与えている。本稿では、映像受信端末におけるキーワード検索の支援方法について提案する。

1.3. 映像受信端末における映像検索の課題

映像受信端末において目的のコンテンツを発見する際、キーワードで検索する場合またはそれ以外の方法から見つける場合がある。大量のデータから決まったコンテンツを見つける場合、キーワード検索が有効であるが、キーワード検索で発見できない状況としては、①検索してもヒットしない、②検索でヒットしていても絞り込めず、探しだすのに時間がかかる(途中であきらめる)、③操作が面倒なので検索を利用しない、の3パターンが存在している。

上記のような状況における要因として、表2の状況が想定される。

表2 コンテンツ発見時の状況と想定される要因

No.	状況	想定される要因
①	検索してもヒットしない。	検索語句, 検索ロジック, 検索対象となる映像メタデータに問題がある。 →メタデータの課題
②	検索でヒットしているが、絞り込めず見つからない。	検索語句, 検索ロジック, 結果の表示順に問題がある。 →並び順の課題
③	操作が面倒なので検索を利用しない。	入力デバイス, 入力方法に問題がある。 →操作性の課題

まず、①検索してもヒットしない場合、検索語句が検索対象の映像メタデータに含まれていないというこ

とがある。そもそも検索対象のサービス内に存在しないということも考えられるが、多くは、検索者が正式名称を覚えていない場合である。映像コンテンツの名称は奇抜なものや冠言葉がついたものなど正式タイトルが長い場合もあり、検索者は略称やタイトルの一部分のみを覚えているからである。人名等についても愛称や別名で検索するとヒットしなくなってしまう。同様に、検索ロジックについても前方一致で検索している場合も先頭からの正式名称を知らないと検索にヒットしないという状況が発生する。これは「メタデータの課題」といえる。

また、メタデータの課題の一つとして、映像メタデータに読みがなが振られていない点がある。映像受信端末にはかな漢字変換機能を保持していないもの、保持していてもユーザが正式な漢字表記を覚えていないため、仮名で検索してもヒットしないという事態が発生する。海外の作品でも同様の問題がある。例えば、「ミッションインポッシブル2」という作品があるが、映像メタデータ上のデータとしては「M:I-2」と表記されているとする。読みがながふられていない場合は、英文字入力をしなければ検索にヒットしない。「Mission impossible 2」と検索した場合もヒットしないこととなる。

次に、②検索でヒットしているが絞り込めず見つからない場合、指定語句が短すぎて絞り込めない、検索対象の幅が広すぎる、50音順、管理コード順など、検索者のニーズと無関係な表示のため複数ページを遷移しなければならない等の要因がある。これは「並び順の課題」といえる。



図 2 テンキーのリモコン (左) とソフトウェアキーボード (右)

3 点目に③操作が面倒なので検索を利用しないという場合であるが、図 1 の市場調査にもあるように、リモコン操作が影響している。入力デバイスとなるリモコンはテンキー (数字ボタン)、4 色カラーボタン、上下左右 (十字) キー、決定キーを備えたリモコン (図 2

左) で赤外線通信を用い受信機とやりとりするものが主流である。文字入力に関しては、50 音の各行を 1 つのボタンに割り当てるトグル方式 [2] や画面にソフトウェアキーボードを表示し、十字キーで選択する方式 (図 2 右) などがある。これは、「操作性の課題」といえる

上記の状況をまとめると、映像受信端末におけるキーワード検索の課題は、①メタデータの課題、②表示並び順の課題、③操作性の課題の 3 つに分類される。

2. 従来技術

上記の課題において、それぞれの従来技術を以下に示す。

2.1. メタデータの課題

まず、メタデータの課題について、略称やよみがな付与に関する対応策として、良く知られているコンテンツについて、略称や別名と正式名称とを紐付けるデータベースを準備する方法がとられている。手動でデータベースへ登録しており、コンテンツの量が膨大なので全てのコンテンツに対応させるとなると、登録コストが大きくなってしまふ。例えば、1 時間番組のみ 24 時間放送、100ch だったと仮定すると、2400 コンテンツ分のメタデータのくくりつけや補完を手動で行うこととなる。実際には人気のコンテンツやメジャーなものにのみ適用されている場合が多い。

また、かな漢字変換については、ふりがなを振ることと逆の発想として日本語入力 Front End Processor (FEP) や Input Method Editor (IME) といったかな漢字変換機能が付属している端末が発売されている。漢字の候補がたくさん存在した場合にそれらを選択するという、新たな操作が必要となる。また、これらに履歴情報を用いる仕組みなども存在するが、家族で利用する TV において履歴情報が本当に有効かは定かではない。

2.2. 操作性の課題

操作性の課題に関しては付属しているリモコンとは別のデバイス (図 3 に示すようなキーボード付きのリモコン、スマートフォン等、他媒体のリモコン化) の利用や音声認識の導入などの研究がされている。[3] しかし、子供や高齢者など、キーボードやタッチパネル型の入力デバイスになれていない世代もいる。また、別途それらの装置を準備しなければならないというコスト面の問題も存在する。

また、WEB サイトでは、入力語から想定されるキーワードや検索履歴などを用いて候補語を提示するサジェスト機能 [4][5] であるが、一般の映像受信端末には導入されていない。



図 3 キーボード付きリモコン

2.3. 並べ方の課題

並べ方の課題については、WEB サイト上の技術として、検索履歴などにより上位に表示するものや[5]、映像サービスでは、視聴履歴などを用い上位に表示する研究がある。[6] しかし、この場合、検索履歴や視聴履歴がない場合に対応できないというコールドスタート問題や、新規作品に対応できないという課題がある。

次章では、有効な候補提示と提示語句の並べかたの改善で、入力デバイスの不足を補い、操作性を改善し、コンテンツ発見の支援をする方法を提案する。

3. 提案方式

並べ方の課題解決手法として以下 3 点をコンセプトとした検索方法を提案する。

- ・変換予測機能を付与し、入力の手間を省く。
- ・変換候補表示、検索結果表示に人気度などの尺度（露出度）を用い、ユーザが、今現在調べたいと思う項目を上位に表示する。
- ・人気度等の尺度は当該サービスや類似サービス、またインターネット上の情報などを利用して算出する。

図 4 に提案方式による検索時の画面遷移例を示す。

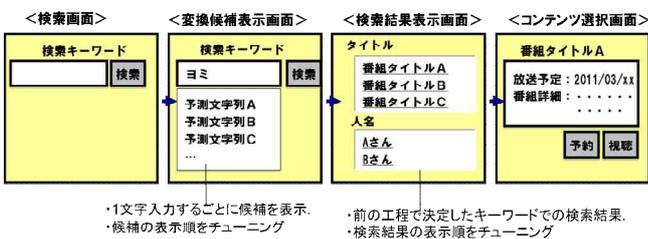


図 4 検索実施時の画面遷移例

変換候補、検索結果表示の順位付けをする尺度を露出度として複数定義しそれらに加重をかけて算出した値（加重露出度）を表示の順位付けに利用する。

3.1. 露出度と加重露出度

本提案方式においてキーポイントとなる露出度の算出方法について以下に述べる。

結果表示の順番を決定するために利用する露出度であるが、視聴者の興味は時間や外部状況などに左右されるため、基準となるデータが1つであるとは限らない。そのため、時間、内部状況、外部状況を反映できる情報を基とした複数の露出度を算出する。仮に n 個の情報源を用いるとして、一つの候補語句に対して露出度 1-露出度 n の N 個の露出度が存在する。露出度の算出に用いる情報として利用するのは、当該サービス内の映像メタデータ、類似他サービスにおける映像メタデータ、インターネット上の評判情報、検索履歴情報などを用いる。一番簡易に算出する場合は当該サービスのみを利用し、時期的な情報や話題性などのある情報を考慮したい場合は、インターネットの情報などを用いると良い。

これらの複数の露出度を混合して利用する場合に利用する値が加重露出度である。加重露出度は一つの語句に対して 1 つであるが、加重露出度を算出するルールは複数設ける事が可能であり、時間帯、検索者の属性などで使い分ける事ができる。また、露出度単体で利用する場合は、他の露出度の加重割合を「0」としてものと同じである。

4. 評価

提案システムを使用し入力効率について評価を実施した。インターネットをクロールし、約 4 万件の映像コンテンツのメタデータを拾集し、メタデータ DB に格納し、そのメタデータ DB の内容を当該サービスとして定義した。

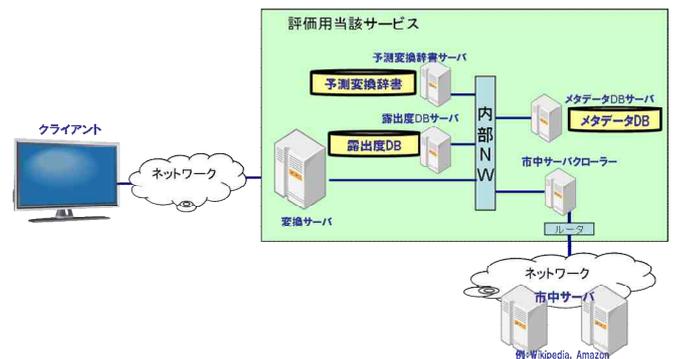


図 5 評価用システム構成

クライアントは既存の映像受信端末、リモコン入力、ソフトウェアキーボードによる構成とした。(図 5)

その当該サービスから著者が見たいと思うタイトルを 35 件抽出し、タイトル検索を実施し、従来方式と提案方式による操作数をカウントし、コンテンツ発見までの時間が短縮されることを操作数で評価する。ここで、従来手法とは露出度を加味しない、データベースからの検索結果順で表示されるものである。

入力する文字列は一般にそのタイトルの略称となっているものを使用した。タイトル先頭からのものもあれば、中間を利用したものもある。タイトル入力には従来手法においても候補提示機能を利用し露出度による並べ替えのみに差異が出るようにした。露出度としては当該サービスのデータとインターネット上のWikipediaのデータ、AmazonDVDでの検索結果を用い、加重割合は当該サービス外の露出度を利用の意味を検証するため、1:1:1と1:0:0の2パターンで測定した。

4.1. 評価結果

評価の結果を表3、表4に示す。表3には、検索対象のコンテンツにたどりつくまでにリモコンのボタンを押下した回数をカウントし、35件の平均値、最大値、最小値、中央値を取ったものを示す。

表4には、検索対象コンテンツが予測候補5件に表示されるまでに入力した文字数を示す。これはソフトウェアキーボードによる実験のため、操作数が入力タイトルの文字の並び順に影響していることを除外して評価するためである。

表3 従来手法と提案手法による、検索操作数平均値、最大・最小・中央値（単位：回）

	提案手法		従来手法
	1:1:1	1:0:0	
平均値	14.8	17	22.4
最大値	27	42	51
最小値	2	5	8
中央値	15	17	21

表4 従来手法と提案手法による、入力文字数の平均値、最大・最小・中央値（単位：文字）

	提案手法		従来手法
	1:1:1	1:0:0	
平均値	2.2	2.4	3.1
最大値	4	5	8
最小値	1	1	2
中央値	2	2	3

4.2. 考察

上記の結果より、ソフトウェアキーボードでの操作数、候補表示が現れるまでの入力文字数共に、提案手法のほうが少なくなっている結果が得られた。また、わずかながらではあるが、外部の情報を加味した加重割合（当該サービス：Wikipedia：Amazon DVD）＝（1：1：1）のほうが、加味しなかったものと比較して操作数、入力文字数共に減少していることがわかった。また、操作数の減少により操作性の課題も若干解消される。

5. まとめ/今後の展開

本稿では、大量の映像コンテンツから目的の映像コンテンツの発見を支援するため、キーワード検索時の候補表示およびその表示順序に着目し、コンテンツ発見に時間がかかっている課題を解決する方法を提案した。提案方法では、露出度という内部・外部の状況を反映する値を複数用い、それらを混合した加重露出度を用いることで検索候補文字の提示順を変更した。表示順を変更することで、ユーザが目的のコンテンツ発見までの操作数が少なくなり、検索にかかる時間を短縮することを提案した。

また、提案方法の評価のため、評価システムを用いて検索時の操作数をカウントし、加重露出度を用いた候補の提示順により、検索対象コンテンツが上位に表示され、操作数が減少すること確認でき、リモコンとソフトウェアキーボードを用いた入力においても、入力時の操作数が少なくなり、必ず目的のコンテンツにヒットすることが確認できた。

今後、他社サービス、インターネット上の情報等、さまざまな露出度を組み込み評価するとともに、ユーザ嗜好に関する値を用いチューニングすることでより視聴者のニーズに合わせた方式と出来るように検討を進める。また、視聴者の満足度についてインタビュー等を実施したい。

文 献

- [1] MMD 研究所, 「2010 年テレビ商戦、及び TV VOD サービスに関する実態調査」, 2010, http://mmd.up-date.ne.jp/news/detail.php?news_id=717
- [2] Detweiler, Mark C.; Schumacher, Robert M.; Gattuso, Nicholas L., 'Alphabetic Input on a Telephone Keypad', Human Factors and Ergonomics Society 34th Annual Meeting Proceedings, pp. 212-216(5), 1990.
- [3] Anna L. Cox, Paul A. Cairns, Alison Walton and Sasha Lee, 'Tlk or txt? Using voice input for SMS composition', Personal and Ubiquitous Computing, Volume 12, Number 8 / November, 2008.
- [4] 市村由美, 齋藤 佳美, 木村和広, 平川 秀樹, '入力予測機能を組み込んだ仮名漢字変換システム' 信学論(D-2), Vol. J85-D-2 pp.1853-1863 Dec.2002.
- [5] Google Search, <http://www.google.com>
- [6] 井川一樹, 福原知宏, 藤井秀樹, 武田英明, "テレビ番組の視聴履歴と電子番組表を用いた番組推薦システムの構築と評価", 24th Annual conference Japanese Society for Artificial Intelligence, Japan, June, 2010.