モバイルネットワーク上におけるメディアメタデータの付与とその活用

古田恒志 † 河口信夫 ‡* 稲垣康善 †

† 名古屋大学大学院工学研究科

: 名古屋大学情報連携基盤センター

* 名古屋大学統合音響情報研究拠点

furuta@inagaki.nuie.nagoya-u.ac.jp

Tagging and Application of Media Meta-data in Mobile Network

Hisashi Furuta[†] Nobuo Kawaguchi^{‡*} Yasuyoshi Inagaki[†]

†Graduate School of Engineering, Nagoya University

[‡]Information Technology Center, Nagoya University *Center for Integrated Acoustic Information Research(CIAIR), Nagoya University

1 はじめに

近年,PDA や携帯電話などに代表される小型情報端末が急速に普及し,誰もが手軽に持ち歩く環境が整いつつある.近い将来,無線通信網の発展によって,情報端末が場所や状況を選ぶことなく他端末と通信し連携動作を行うようになると予想される.また一方で,DVD などのデジタル動画像メディアの普及や,ブロードバンド・常時接続環境の整備,HDD の大容量・低価格化などにより,マルチメディアコンテンツに対する需要が高まってきている.このように,マルチメディアコンテンツの利用が活発になるに伴い,コンテンツ内のある特定のワンシーンを対象としたユーザの要求が生じることがある.これは例えば,以下のようなものである.

- サッカー観戦に行って,素晴らしいファインプレイを見たので,そこだけをもう一度見たい。
- DVD を見ていたら、友人が熱を上げている俳優がノンクレジットで出演しているのを見つけた。そのシーンを友人に見せてやりたい。

しかし、ユーザのこうした要求を満たす仕組みは今まで存在し なかったため、ユーザがこれらのことを行うのは困難であったり、 可能であったとしても手軽に行えるものではなかった. 本研究では, ユーザが望む任意のシーンを記録し,その記録情報の配布・交換・ 共有・集積によって,上述のような要求を満たし,従来にない柔軟 なコンテンツ利用を行うための手法を提案する. 任意のシーンを記 録するために,コンテンツに対するメタデータを定義する.実際に コンテンツに対しメタデータを付与する方法として,携帯端末によ るマーキング方式を採用する.ユーザが携帯する端末は,コンテン ツを提供しているシステムとネットワーク接続し,利用中のコンテ ンツの情報を取得する.ユーザは,コンテンツを利用しながら,任 意のシーンに差し掛かった時に手元にある端末のボタンを押すだけ で、煩雑な操作をすることなく手軽にメタデータを付与することが できる.さらに,今後ユビキタス環境の普及が進み,どのような場 所でもコンテンツ提供機器と携帯端末の連携が可能になれば,自宅 で DVD を鑑賞しているといった状況のみならず,映画館,競技場, 街角やテーマパークなど, あらゆる状況下でマーキングが実行可能 となる.生成したメタデータは,マーキングしたシーンの再生や, そのシーンの関連情報のダウンロードなどに利用できる.

なお本稿では「シーン」という用語を,コンテンツ中の一部分を指すものとして用いることとする.シーンの長さについては特に明確な指標を設けることはしないが,慣習的にワンシーンと呼ばれる程度の長さを想定している.

本稿の概要を示す.2章で,マルチメディアコンテンツ中のあるシーンを特定するメディアメタデータを提案し,3章では,既存のマルチメディアコンテンツに対しメディアメタデータを効率的に付与するための手法を提案する.4章では,得られたメディアメタ

データの活用例を挙げ、5 章でそのプロトタイプシステムの実装について述べる。6 章では本稿をまとめ、今後の課題を記す。

2 メディアメタデータ

「あのサッカーの試合のあのプレイが見たい」「2時間ある映画 の中の,あのワンシーンだけ見られればいい」など,ユーザの要求 しているコンテンツが、マルチメディアコンテンツ全体ではなく、 コンテンツ中の特定シーンを指している場合がある.このような場 合,コンテンツそのものをユーザに提供するのはコストが高く,ま たシーンを探す作業も煩わしい.これを解決するためには,コンテ ンツ中のワンシーンを参照・記録する仕組みが必要となる、本研究 では、コンテンツ中のワンシーンとなる動画像や音声などをその ままのデータとして保存するのではなく、メタデータによってシー ンを記述する.メタデータとは,一般には"データに関して記述し たデータ"である.文書・映像・音楽などのデータに対してタグ付 けされ,元データの意味や構造,内容を記述するデータであり,元 データの検索・ブラウズ・ナビゲーション・編集作業などに利用さ れる . ISO/IEC JTC1 SC29/ WG11 によってメタデータの表記 方法に関する国際標準規格が策定されており [3], 近年注目を集め ている.オンラインオブジェクト検索のためのメタデータ表現方法 [4] や, コンテンツに対するメタデータを登録した XML DB の利 用によるシーン検索 [5] などの研究がされている.

本研究が提案するメディアメタデータは、コンテンツ中のワンシーンに関する情報をまとめてこれをひとつのメタデータとしたものである。シーンをメディアメタデータとして扱うことは、シーンのデータをそのまま保存することに比べて少量のデータで済み、またメタデータを付与することも容易な仕組みで実現できる。メディアメタデータが保持する情報をまとめたものを表1に示す。これらの情報のうち、必須となるものはメディアIDとシーン特定子の二つである。

2.1 メディア ID

メディア ID は、マルチメディアコンテンツを特定するための情報である.これは、対象としているコンテンツによって、二種類のタイプが考えられる.第一に、対象としているコンテンツが、スタジアムで観戦しているスポーツの試合や、コンサート会場での演奏会、またテレビやラジオの番組などの場合である.このような場合

表 1: メディアメタデータが保持する情報

情報		役割
メディアID **		マルチメディアコンテンツを特定
シーン	特定子 ※	コンテンツ中 のシーンを特定
	クリエイタID	メディアメタデータ作成者
	属性	シーンのタイプ
	コメント	注釈

では,基本的に同一コンテンツの提供は一度しか行われない.そのため,そのソース(例えば会場や,受信しているチャンネルなど)と,コンテンツを利用している時間によって,コンテンツを特定することが可能である.

もう一方は , 対象としているコンテンツが DVD や音楽 CD な ど、時間・場所を問わずユーザによって自由に幾度と利用されるメ ディアの場合である.このような場合,前述したようなソースでは コンテンツを特定することは不可能であるから, コンテンツを識別 するために ID を設定する必要がある.これらのコンテンツのデー タ自体は,同一のコンテンツであれば同一のデータを持つため,こ れによって ID を決定することが可能である. 例えば, Gracenote が提供する CDDB[1] 音楽認識サービスでは , CD の TOC^{-1} デー タにより CD を認識しており, コンピュータに音楽 CD を入れる と, Gracenote のサービスを受けたソフトウェアプレイヤアプリ ケーションは認識情報をデータベースに問い合わせて CD の情報 を得ている.ただし,このようにコンテンツデータ自体から ID を 決定する方式の場合, ID がユニークとなる保証はないため, 極め て稀にではあるが ID が衝突する問題が残る.また, ID からコン テンツを参照するためには , 既存コンテンツの ID を蓄積したデー タベースを用意しておく必要がある.

2.2 シーン特定子

シーン特定子は、メディア ID によってマルチメディアコンテンツが特定された上で、求められるシーンがコンテンツ中のどのシーンなのかを特定する情報である。何を以ってシーン特定子とするかは、コンテンツの形態によって異なる。例えば、コンテンツが動画像や音声であれば、タイムコードに基づき開始時からの経過時間を指定すればよいし、競技場で観戦するサッカーの試合であれば、時間情報に加えて座席番号も指定するなどが考えられる。

2.2.1 シーン特定の形式

シーンを特定する際,その形式はさらに二種類に分類される.すなわち,単一のシーン特定子を用いる場合と,複数のシーン特定子を用いる場合である.前者の形式は,コンテンツ中のある特定のポイントを指定したい場合に利用される.後者の形式は,シーン特定子を二つ用いることによりシーンを特定する.すなわち,シーンの開始(イン)とシーンの終了(アウト)の二つのポイントを指定し,「ここからここまで」というワンシーンを決定する.

2.3 その他の情報

メディア ID とシーン特定子によってマルチメディアコンテンツのシーンを特定することが可能であるが,実際にアプリケーションによってメディアメタデータが利用されることを考えた場合,これに加えていくつかの付加情報があることが望ましい.それらの一部を以下に挙げる.

2.3.1 クリエイタ ID

マルチメディアコンテンツ中のシーンを特定することを行ったユーザ,すなわちメディアメタデータの作成者を示す情報である.シーンを特定したユーザを参照したり,シーンを作成者によって分類するなどの際に利用される.この情報によって示されるユーザは,マルチメディアコンテンツを利用しているエンドユーザの場合もあれば,コンテンツ提供側のユーザの場合もあり得る.

また,年齢や性別といった各クリエイタの個人情報が入力されたデータベースを用意し,クリエイタ ${
m ID}$ からこれらに参照できるようにすれば,メディアメタデータを活用するアプリケーションによっては有益である.

2.3.2 シーン属性

マルチメディアコンテンツに付与されたメディアメタデータが 複数あり、それらを分類して利用したいという場合がある.シーン の属性情報を付加しておくとこのようなケースに有効である.以下 に一例を挙げる.映画館で映画を見るユーザに,印象に残ったシー ンにメディアメタデータを付与する操作をさせ,これを繰り返して

```
<?XML version="1 0" encoding="Shift JIS"?>
<!DOCTYPE MediaMetadata [
<!ELEMENT Contents|D (#PCDATA)>
        <!ELEMENT Source (Location|D | Channel | ...)>
   <!ELEMENT Scene (Specifier, Creator?, Attribute?, Comment?)>
      <!ELEMENT Specifier ((SceneTime, StadiumInfo?), ...)>
<!ELEMENT SceneTime (Pointer | Clip)>
            <!ELEMENT Pointer (#PCDATA)>
            <!ELEMENT Clip (InTime, OutTime)>
        <!ELEMENT StadiumInfo (SeatNo...)</pre>
      <!ELEMENT Creator (Creator | D. Name?.</pre>
      <!ELEMENT Attribute (UserEmotion | Genre | ...)>
         <!ELEMENT UserEmotion EMPTY>
         <!ATTL|ST UserEmotion value (Exciting | Sad | ...) #REQU|RED>
        <!ELEMENT Genre EMPTY>
      <!ELEMENT Comment (#PCDATA)>
  <!ELEMENT Editor (Editor | D. Name?. ...>
```

図 1: メディアメタデータの DTD (一部)

```
《Media》をはははは

《Media》をMedia ID》

《Contents ID》 0011223344 〈Contents ID》

《Media ID》 〈Media》

《Scene》

《Specifier》

《SceneTime》 〈Pointer》 1h13m27s 〈Pointer》 〈SceneTime》

〈Specifier》

《Creator》 〈Creator ID》 56789 〈Creator ID》 〈Creator》

《Attribute》 〈UserEmotion value="Exciting"/》 〈Attribute》

《Comment》 このシーンは必見! 〈Comment》

《Scene》

《Editor》 とEditor ID》 56789 〈Editor ID》 〈/Editor》

〈Media Metadata》
```

図 2: メタデータ例

充分な数のメディアメタデータを収集する.この場合「印象に残った」とひとつに括られているシーン群にも,感動したシーンであったり,笑ったシーンであったり、驚いたシーンであったりなど,実際には様々な意味のシーンが混在する.これらのシーンに,メディアメタデータを付与したユーザの感情を属性として保持させておくと,感動したシーンのみをまとめて表示させたいといった要求にも容易に対応することができる.

2.3.3 コメント

メディアメタデータによってシーンが特定されていても,利用するユーザにとってシーンの内容が把握しにくければ利便性が低い. コメントによって利用効率を向上させることができる.

XML におけるメディアメタデータの DTD の一部を図 1 に , そのインスタンス例を図 2 にそれぞれ示す .

3 メディアメタデータの付与手法

前章で提案したメディアメタデータを実際に生成する手法について述べる。メディアメタデータの要素の中で必須のものはメディアIDとシーン特定子であるが、シーンを特定する毎にこれらをユーザが入力するのは非常にコストがかかり現実的でない。メディアメタデータを生成する手法には、ユーザからの入力が容易であり、簡素なインターフェースが求められる。これは、コンテンツを利用する立場である一般エンドユーザがメディアメタデータを手軽に活用できるようにするという観点からも重要である。

3.1 マーキングによるメタデータ付与

本研究では,マルチメディアコンテンツに対しメディアメタデータを付与する手法として,マーキングによるメタデータ付与を提案する.この手法は,コンテンツが動画や音声などタイムベースのメディアである場合に効果的である.

 $^{^1}$ table of contents: CD の時間情報が書き込まれている領域 .

3.1.1 概要

マーキングによるメタデータ付与は,ユーザがマルチメディアコンテンツを利用しながら,任意のシーンに差し掛かった時に手元の機器のボタンを押すこと(マーキング)により,そのシーンを示すメディアメタデータを付与する手法である.本手法を行うユーザは,細かい設定や入力を考えることなく,ボタンを押すという手軽かつ簡素な動作のみによってメディアメタデータを付与することができる.これは丁度,ウェブブラウザで気に入ったウェブページをブックマークして保存しておく感覚に似ている.

本手法は,ユーザにコンテンツを提供する機器と,ユーザがマーキング動作を行う機器とがアドホックネットワークで接続し,互いに通信することによって行われる.ひとつのコンテンツ提供機器に対し,複数のマーキング機器が接続される場合もあり得る.観客を収容した映画館での利用などがこれにあたる.

3.1.2 コンテンツ提供機器

マーキング手法を行うためには,動画プレイヤなどといったコンテンツを提供する機器にも,特有な機能を備えさせる必要がある.それは,現在利用中であるマルチメディアコンテンツのメディア ID を認識し,マーキング機器に通知する機能である.

3.1.3 マーキング機器

本研究では,マーキング機器として携帯電話や PDA などの携帯端末を使用することを考えている.これは,携帯端末の持つ以下のような特徴による.

- 小型かつ軽量であり,手軽に利用できる.
- 一般に普及している.また今後普及が拡大すると予想でき, 導入にかかるコストが小さい.
- 時間や場所を選ばずに使用できる.利用場所を自宅など特定の場所に限定することなく,映画館やテーマパークなど,ユーザの行く先々でのマーキングを可能とする.
- 個人専用の端末として用いられることが多い.これはクリエイタ ID を付与するのに有益である.

3.1.4 ローカルマーキングとリモートマーキング

実際のマーキング動作の流れを述べる.マーキング機器にはその所有者のクリエイタ ${
m ID}$ が予め与えられているものとする.以下の手法をローカルマーキングと名付ける.

- 1. マーキング機器がコンテンツ提供機器に接続要求.
- 2. コンテンツ提供端末は内部リストにマーキング機器を追加し, 状態やメディア ID などをマーキング機器に返す.以降,コ ンテンツ提供端末の状態が変化した場合は,状態変化をリストの全端末に通知.
- コンテンツの提供が開始されていれば、マーキング機器内部のタイムカウンタをコンテンツのタイムコードと同期させ、マーキング開始。
- 4. ユーザが任意の個所でマーキングを行う.マーキング機器は メディア ID, タイムコード,クリエイタ ID に基づきメディ アメタデータを生成し記録.
- 5. コンテンツ提供終了とともにマーキング終了.

ただし,この形式では,付与されたメディアメタデータがマーキング機器にしか残らず,コンテンツ提供側はメタデータを即時に取得できない.また,コンテンツ提供機器の状態変化に伴い大量の通信が行われる.これを考慮した形式として,以下に示すリモートマーキングを提案する.

- ユーザがマーキングを行う.マーキング機器は,コンテンツ 提供機器にクリエイタ ID やシーン属性を通知し,メディア メタデータ付与を要求.
- 2. コンテンツの提供が開始されていれば,コンテンツ提供機器 はメディアメタデータを生成し記録.また,メディア ID やタイムコードをマーキング機器に送信.

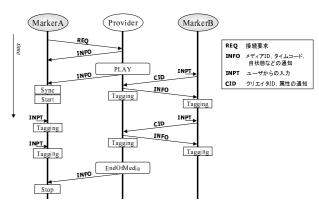


図 3: マーキング手法のシークエンス図

3. マーキング機器がデータを受信してメディアメタデータを生成し記録.

この形式では、メディアメタデータがコンテンツ提供側にも残る.またマーキング端末にタイムカウンタ機能を必要としない.一方で、マーキングを行う毎に通信がされるため、マーキングの頻度が大きいとコストが増す.またインとアウトによりシーンを特定する場合、ある端末からイン入力があると、同端末からアウト入力があるまでイン入力の情報を保持しなければならない.

各形式のタイムシーケンスを図 3 に示す. マーキング機器 A はローカル, B はリモートの形式でマーキングを行っている.

ただし、いずれの形式においても、シーン属性を記録するためには何らかの工夫が必要となる。各属性毎にそれぞれボタンを用意するなどの改良は考えられるが、一方でユーザビリティは低下する。また、コメントを付加するにはこの手法は適さない。これらのデータ入力には、マーキング手法とは別のエディタのようなシステムが求められる。

4 メディアメタデータの活用

本章では,ユーザからのマーキングによって得られたメディア メタデータを活用したコンテンツ利用方法を述べる.

4.1 単一のメディアメタデータでの利用

メディアメタデータに記録されたメディア ID をもとに、インターネットなどを経由してサーバからマルチメディアコンテンツをダウンロードし、シーン特定子により該当シーンを表示する、といった機能を動画プレイヤに備えることにより、「このシーンが見たい」というユーザの要求を手軽な操作で実現可能にする・ユーザは、マーキング動作を行うことで気に入ったシーンをダイレクトに記録できる・そのシーンの収められているコンテンツメディアを探し出して用意する必要はなく、そのシーンを探すためにサーチ再生をする必要もない。

本手法を用いれば、映画館で映画を見ながら、ユーザは手元の携帯端末で気に入ったシーンを自由にマーキングすることができる、ユーザによっては、映画が終わった後、気に入ったシーンをもう一度見たり、あるいはよく理解できなかったシーンを見直したりしたいということもある、このようなエーズに対し、マーキングしたワンシーンを比較的低料金で提供するサービスがあれば便利である、これは映画館に留まった考えではない、サッカーの競技場において、試合会場に設置された複数のカメラから試合を録画しておき、それをデータ配信するシステムを用意することにより、上述の例と同様にしてゴールシーンやファインプレイ、その他各ユーザがそれぞれ気に入ったシーンをユーザに提供することができる。

加えて,携帯端末に記録されたメディアメタデータは,通信によって他の端末に配布することが可能である.これによって,出会ったユーザが互いの携帯端末に記録されている気に入ったシーンを見せ合ったり,趣味の合う仲間の間でマーキングしたシーンを共有するといった利用法もできる.

4.2 複数メディアメタデータの集積的利用

メディアメタデータの付与を行うユーザが多数いた場合を考え る.様々なマルチメディアコンテンツに対して,多数のユーザが マーキングを繰り返し行うに従い,メディアメタデータが蓄積され る.これにより,前節で提案したメディアメタデータの単独利用と は異なり、複数メディアメタデータを集積物として捉え、そこから 得られる統計情報を利用したシステムが構築できる.映画館で上 映される映画に対して、観客から「楽しかったシーン」でのマーキ ング動作を募り,充分な量のメディアメタデータを収集する.これ をタイムチャート上に整理すれば, 各シーン毎のメタデータ付与数 が獲得できる.これによってシーンに重要度を設定し,重要度の高 いシーンを優先的にピックアップしたダイジェスト映像を作成する ことができる. さらに, メディアメタデータに記録されたクリエイ タ ID から年齢・性別といった情報を参照できれば , メディアメタ データをユーザ情報によってカテゴリに分類し, それぞれで集計す ることにより、カテゴリ毎のマーキングの傾向を明らかにすること が可能となる.その結果,そのコンテンツを高く評価する傾向のあ るユーザカテゴリを選出してマーケティングに利用することができ る.また重要度の高いシーンも各カテゴリ毎で異なったものになる と予想されるため,ダイジェスト映像がユーザから要求されるその 都度,最も興味を引くと思われるダイジェストを作成できる.

他にも、ユーザが毎回のマーキングを同一の携帯端末で行い、付与されたメディアメタデータの履歴が残されるものとすると、各々のユーザ毎の趣味・嗜好がマーキングの傾向として履歴に現れる、複数ユーザ間で互いの履歴を比較することにより、似た趣味を持つユーザを探したり、出会ったユーザ同士に共通するシーンを探すことができ、コミュニケーション支援が行える、また、他ユーザの評価に基づいてコンテンツを推薦する協調フィルタリング技術についての研究がなされており[6][7][8]、メディアメタデータをこれらに応用して新しい推薦システムを構築することも考えられる。

5 実装

以上の手法に基づき,そのプロトタイプシステムを Java 言語 (JDK1.3, JMF2.1) を用いて実装した.コンテンツを提供する機器である動画プレイヤ,リモートマーキングを行う機器であるマーカーをソフトウェア的に実装し,機器間通信に cogma[2] システムを利用した.コンテンツとして,MPEG2 形式の動画ファイルを用意し,メディア ID は,動画ファイルのファイル名とファイルサイズから決定するものとした.プレイヤとマーカーはそれぞれ通信機能を備えており,プレイヤはコンテンツのデータベースからコンテンツを入手し再生する機能を持つ.プレイヤとマーカーは互いに連携し,マーキングによってメディアメタデータ付与を行う.メディアメタデータによって記録されたシーンはプレイヤで再生できる.ひとつのプレイヤに対し複数マーカーが接続することも可能であり,またマーカーがあるプレイヤと接続して付与したメディアメタデータを別のプレイヤで利用することも可能である.

また , ブラウザからの要求に応じてマーカーを制御するサーバを実装し , 第三世代携帯電話 FOMA を用いてマーキングを行うシステムを構築した . マーキングによって記録されたシーンは , FOMA の i モーション 2 コンテンツとして閲覧することができる . この外観を図 5 に示す . またプロトタイプシステムの構成図を図 4 に示す .

6 おわりに

本研究では,多種多様なマルチメディアコンテンツ中のワンシーンに対するユーザの要求に応じるため,メディアメタデータによってシーンを特定し,マーキング動作によってこれを付与する手法を提案した.この手法では,手軽な動作によって任意のシーンを容易に記録することが可能である.

今後の課題として,メディアメタデータを活用するシステムの 充実と実装が挙げられる.本研究で提案されるマーキング手法は,

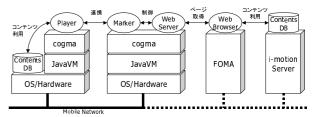


図 4: プロトタイプシステム構成図



図 5: 携帯電話によるマーキングの外観

メディアメタデータを付与するためのインターフェースに過ぎず, メディアメタデータを如何に活用していくかが重要となる.本研究 で実装されたプロトタイプシステムでは,提案したメディアメタ データ活用法の一部しか実装しておらず,まだ充分とは言えない. 特に,メディアメタデータを集積的に利用したアプリケーションの 実装により,従来にない新しいサービスを提供することが重要であ ると考えている.加えて,実装したシステムを実用して複数ユーザ によるメディアメタデータ付与を行い,提案した手法の有用性を確 認することも必要である.

また,ユーザの意図を忠実に反映したメディアメタデータ付与手法の実現も考えている。本手法では,ユーザが記録したいシーンを正確にマーキングできるとは限らない。マーキングした瞬間と記録したかったシーンがずれている可能性がある。これを解決するためには,マーキングしたメディアタイムを調整する,コンテンツを予めいくつかのチャプターに区切る,などの方法が考えられる。これら手法の問題解決能力やユーザビリティについて検討したい。

他にも、コンテンツ提供機器がマーキング機器を発見する仕組みを本稿では述べていない.さらに、マーキング機器が発見できたとしても、その機器のユーザがコンテンツ利用者なのかどうかを判別しなければならない.これらを考慮した手法も考えていきたい.参考文献

- Gracenote, Gracenote CDDB Music Recognition Service, http://www.cddb.com/
- [2] こぐまプロジェクト, Cooperative Gadgets for Mobile Appliances, http://www.cogma.org/
- [3] José M. Martínez, MPEG-7 Overview, http://mpeg.telecomitalialab.com/standards/mpeg-7/mpeg-7.htm
- [4] DCMI, Dublin Core Metadata Element Set, Version 1.1: Reference Description, DCMI Recommendation(1999), http://dublincore.org/documents/dces/
- [5] 東芝, MPEG-7 を利用した XML マルチメディア・データベース, http://www.toshiba.co.jp/mmlab/tech/d03.htm
- [6] P.Resnick; N.Iacovou; M.Suchak; P.Bergstrom; and J.Riedl, "GroupLens: Open Architecture for Collaborative Filtering of Netnews", Conference on Computer Supported Cooperative Work, pp. 175-186 (1994)
- [7] GroupLens Research Group, http://movielens.umn.edu/
- [8] Jacobi; Jennifer A.; Benson; Eric A., "System and methods for collaborative recommendations", United States Patent 6, 064, 980(2000)

 $^{^2}$ 受信速度最大 384kbit のパケット通信により映像や音などのコンテンツを提供するサービス .