番組実況チャットに基づく視聴者視点を利用した放送番組のビュー生成

宮森　恒[†]　　中村 聡史[†]　　田中 克己[†‡]

†情報通信研究機構けいはんな情報通信融合研究センター　〒619-0289 京都府相楽郡精華町光台 3-5
‡京都大学 大学院情報学研究科 社会情報学専攻 〒565-0456 大阪府吹田市河田 4-5-6
E-mail: †miya@nict.go.jp, †gon@nict.go.jp, ‡ktanaka@i.kyoto-u.ac.jp

あらまし　本稿では，視聴者によって書き込まれる番組実況チャットに基づく視聴者視点を利用した放送番組のビュー生成手法を提案する．番組内容と同期して書き込まれる番組実況チャットのデータを統計処理・認識処理することで，番組の盛り上がり場面や，各シーンに対する特定の視聴者のリアクション等を効果的に抽出することができると考えられる．本稿では，これら抽出結果によって，視聴者の盛り上がりや自分と類似した嗜好をもつ他人が興味を示す部分など，さまざまなビューを生成でき，これを利用することで放送番組をこれまでにないさまざまな観点から多角的に視聴可能であることを示す.

キーワード　番組実況チャット，通信放送コンテンツ融合，メタデータ生成，視聴者，視点，ダイジェスト

View generation of TV content using TV viewer's viewpoint based on live chat on the web

Hisashi MIYAMORI[†]　Satoshi NAKAMURA[†]　Katsumi TANAKA[†‡]

† National Institute of Information and Communications Technology (NICT)

3-5 Hikari-dai Seika-cho Souraku-gun, Kyoto, 619-0289 Japan

‡ Kyoto University　Yoshida Honmachi, Sakyo, Kyoto, 606-8501 Japan
E-mail: †miya@nict.go.jp, †gon@nict.go.jp, ‡ktanaka@i.kyoto-u.ac.jp

Abstract　This paper proposes a view generation method of TV program based on viewer’s viewpoint obtained from live chat on the web where messages are posted by the TV viewers. Important scenes in the program, responses by a particular viewer, etc. can be extracted effectively by computing statistically and/or recognizing data in a live chat which are obtained in sync with the broadcast content. This paper shows that by using the computed results, the various views can be generated indicating the momentum of reactions by viewers, the scenes which are interested by a particular viewer who has similar preferences with the user, etc., realizing a new way of viewing the original content from various perspectives.

Keyword　Live chat, fusion of broadcast and web content, metadata generation, viewer, viewpoint, digest

1. はじめに

近年，ハードディスク（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は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．また，複数チャンネルの同時連続録画が行える HD レコーダも登場しており，6 チャンネル分の番組を 1 週間連続して録画することも可能である．さらに，放送の多チャンネル化により，数十チャンネルの放送サービスが各地域で普及しつつある．HD の記憶容量は飛躍的な速度で増加しているため，近い将来，ペタバイトの容量をもつ HD レコーダが登場し，地域に関係なく数百チャンネルの番組を月単位・年単位ですべて録画・蓄積できるようになるだろう．これにより，テレビ番組はもはや生で視聴するだけでなく，普段は録画・蓄積しておき必要に応じて後で視聴するといったスタイルが一般的になるだろう．

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．また，テレビ番組は実時間で見るように作られているため，２倍速や３倍速などで見るのは困難である．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多様な視聴方法を提供

する仕組みが重要となる．

そのような仕組みを実現するための重要な基礎技術として,映像のインデキシングが挙げられる.従来，動画像中の色やテクスチャ，カメラ操作，人物の顔などの特徴や，字幕テキスト，音の種別や大きさなどをマルチモーダルに利用した種々のインデキシング手法が提案されてきた．しかし，これら従来手法が利用してきたデータは，番組の配信側（放送局）から提供される情報のみであったため，インデキシングされた情報は，基本的に，番組作成者・配信側の意図を反映しているだけであった．よって，従来手法においては，シーン探索や内容の早見，ダイジェスト視聴などの内容に，自分以外の他の視聴者の視点や反応といった要素を取り入れることはできなかった．

一方，近年番組実況チャットに対する注目が高まっている．番組実況チャットは，テレビ番組の視聴者がインターネット上のチャットコミュニティに集まり，テレビ番組の内容について感じたことなどやテレビ番組の内容自体のことをリアルタイムに書き込み，会話を楽しむシステムである．利用者は，テレビ番組と並行してチャットを楽しむことにより，視聴者との擬似的な感情の共有による一体感を味わうことができる．

そこで，本稿では，番組実況チャットのデータを利用することで，放送番組に対する視聴者の視点や反応を反映したさまざまな見方を可能とするビュー生成手法を提案する．番組実況チャットのデータを統計処理・認識処理することで，番組の盛り上がり場面や，各シーンに対する特定の視聴者のリアクション等を効果的に抽出することができると考えられる．これにより，視聴者が盛り上がっている場面や自分と類似した嗜好をもつ他の視聴者が興味を示す場面などをランキング表示したり時間順に表示する，様々なビューを生成することができる．提案手法により，視聴者の生の視点や反応を反映させた形での番組視聴が可能となり，自分以外の大勢の視聴者とともに視聴しているという一体感や感情の擬似的共有感が生まれることにもなると期待される．

本稿の構成は以下の通りである．2 章では番組実況チャットについて述べる．3 章では番組実況チャットを利用した多用なビュー生成の処理概要について説明し，4 章では考察を行う．5 章では関連研究について述べ，最後に，6 章でまとめを述べる．

2. 番組実況チャットとは

2002 年に開催されたサッカーのワールドカップでは，日本の躍進も手伝い日本中が沸きかえった．試合の入場チケットはすぐに売り切れてしまい，試合の開催されるスタジアムから溢れた人のために，実際に試合が行われていないスタジアムなどでパブリックビューイング（スポーツの試合を大型ビジョンで無料放送するというもので，みんなでスポーツ観戦を楽しむということが目的)が行われたのは記憶に新しい(図１)．多くの日本代表のサポーターがスタジアムだけでなく，バーやビアガーデンなど様々な場所で開かれたパブリックビューイングにて日本代表に声援を送り，日本代表の活躍を楽しんでいた.2003 年にプロ野球阪神タイガースが優勝したときにも，関西の多くの居酒屋やビアガーデンにテレビが設置され，阪神タイガースのファンが集まり，阪神タイガースの快進撃を楽しんでいた．こうした，一人ではなくみんなで試合を観戦するというスタイルは，欧米などでは一般的である．イタリアではサッカーの試合のたびに多くの人がスポーツバーに集まって熱狂し，アメリカでは大リーグの試合があるたびにスポーツバーに集まって試合観戦を楽しんでいる．

図１．パブリックビューイングの例

この，多くの人が同じ空間に集まりテレビ番組を楽しむというスタイルの利点は様々にあると思われるが，その最たるものは感情を擬似的に共有したいということであると考えられる．ともに喜び，ともに悲しみ，ともに笑い，ともに泣き，ともに怒る．こうした視聴者の感情の擬似的共有による一体感が楽しいために，そうしたスタイルの視聴が行われるのであろうと考えられる．また，テレビ番組の内容について同意を求めたいとか，他人の反応を見たいというのも一つの要素であろう．

テレビ番組ではこうした点を考慮し，色々な趣向を凝らしている．例えば，バラエティ番組などで笑い声を挿入したり，テレホンショッピングにおいて客の驚きの声を挿入したりしているのは，テレビ番組の視聴者に対する感情の一提案であろう．しかし，テレビ放送で提供されるものは画一的なものであり，視聴者全員が同様に感じるとは限らない．例えば，A チームと B チームによるサッカーの試合を観戦する場合，視聴者は「A チームの応援」「B チームの応援」「中立」の 3 つの立場に大別される．この立場の違いは大きく，同じ試合を見ていても A チームが勝っているとき，B チームが勝っているときでそれぞれ，視点が異なり，感情の起伏も違ったものになると考えられる．また，

同じ A チームを応援している場合でも，好きな選手の違いや，チームに対する思い入れの違いなどにより，視点の違いは生じるであろう．

一方，近年番組実況チャットに対する注目が高まっている．番組実況チャットというのは，テレビ番組を観ている視聴者がインターネット上のチャットコミュニティに集まり，テレビ番組の内容について感じたことなどやテレビ番組の内容自体のことをリアルタイムに書き込み，会話を楽しむシステムである．テレビ番組と並行してチャットを楽しむことにより，擬似的な感情の共有および一体感を味わうことができる．これはまさに，従来の物理的な空間に人々が集まってテレビ放送を楽しんでいたものを，仮想的な空間に応用したものであるといえる．

巨大掲示板サイトである 2 ちゃんねるでは，この番組実況チャットを大規模サポートしている．2 ちゃんねるでは，NHK 専用，日本テレビ専用，TBS 専用，フジテレビ専用，テレビ朝日専用，テレビ東京専用，ラジオ専用など各カテゴリ（ここでは放送局）に応じた番組実況チャット用の掲示板が設置されており，その中に番組自体について語り合うスレッド（ある話題に関して投稿されたメッセージの集まりで，書き込まれた時間順に並べられている）が多数存在する．同一の番組であっても，視聴者ごとに興味や嗜好が異なるため，2004 年の日本シリーズにおいては，阪神タイガースファン専用スレッド，西武ライオンズファン専用スレッド，福岡ダイエーホークスファン専用スレッド，野球初心者専用スレッドなど様々な視点に基づくチャットコミュニティが作られ，棲み分けられていた．これは，現実において阪神タイガースファン専用のパブリックビューイング，西武ライオンズファン専用のパブリックビューイングという形で棲み分けられていたのと酷似している．テレビ番組放送中には番組実況チャットから人が絶えることがないほど人気が高い．

番組実況チャットでは，設定による違いはあるが，書き込みの内容には時間と，書き込み者を一意に識別する ID が付加されていることが多い．この識別 ID を利用することで，どの書き込みは誰が書いたのかということを区別することができる．2 ちゃんねるの番組実況チャットでは，一日経過するごとに書き込み者の ID が変化するが，本研究で想定する環境では，すべての番組実況チャットについて，同一書き込み者の ID は一意であるものとする．

3. 番組実況チャットを利用した多用なビュー生成

著者らは，番組実況チャットから得られる情報を利用し，テレビ番組中のどのような場面でどのような感情の変化が生じたのかということを書き込み内容から取得することで，視聴者の独自の視点に基づくテレビ番組コンテンツのインデキシング手法を提案している [13]．本節では，このインデックス情報を利用することで，例えば，A チームを応援している視聴者の視点，中立の立場で視聴している視聴者の視点，ある選手の熱狂的なファンの視点など，様々な視点に立って録画しておいたテレビ番組を視聴することができるようになることを示す．

まず，番組実況チャットを利用した番組コンテンツのインデキシング手法について説明する．

番組実況チャットは，設定による違いはあるが，基本的に，書き込み時刻，書き込み者の ID，書き込み内容から構成されている（図２）．

80 名前：名無しさん 04/10/19 08:48:22 ID:R9C+YqFS
ものすごいブリッツの中からTD
ｷﾀ━━━━━━(ﾟ∀ﾟ)━━━━━━!!

81 名前：名無しさん 04/10/19 08:48:28 ID:+FNfUHkk
50点以上入りそうな勢い

82 名前：名無しさん 04/10/19 08:48:30 ID:BZZg6/Cd
1stだけで･･･

83 名前：名無しさん 04/10/19 08:48:48 ID:7iYg967m
マニングはこういうとき凄くムキになる。
しかしただでさえ背が高いのにジャンプ投げとは反則だー

84 名前：名無しさん 04/10/19 08:49:15 ID:3bLzDEmh
初心者には空中戦がわかりやすくてよいかもな。

85 名前：名無しさん 04/10/19 08:49:23 ID:R9C+YqFS
ブリッツ多かったけどバック何人残ってたんだろ？

図２．番組実況チャットの書き込み内容の例

例えば，このデータをもとに，ある単位時間ごとに書き込みエントリ数を計測し，その時間変化をグラフに表すと図３のようになる．この図より，番組実況チャットに参加している視聴者のうちどのくらいの人数が番組内容に対して反応を示したかという度合い（視聴者の反響）を各時刻について把握することができると考えられる．

また，番組実況チャットへの書き込みでは，アスキーアートを用いた特徴的な表現がしばしば用いられる．例えば，スポーツ番組で盛り上がった場面などでは，以下のようなアスキーアートがよく利用される．

1. ｷﾀ━━━━(ﾟ∀ﾟ)━━━━!!!!
2. ｷﾀ━━━━━∩(ﾟ∀ﾟ)∩━━━━━!!!!
3. TD━━━∩(ﾟ∀ﾟ∩)━━━ !!
4. ｷﾀ━━━━ヽ(ﾟДﾟ)ﾉ━━━━!!!!
5. ｱｯﾋｬｯﾋｬ!ヽ(ﾟ∀ﾟ)ﾉｱｯﾋｬｯﾋｬ!

同様に，落胆した場面では，以下のようなアスキーアートがよく用いられる．

1. (ノ∀｀) アチャー
2. ヽ(｀Д´)ﾉｳｳｧｧｧﾝ
3. (´・ω・｀)ｼｮﾎﾞｰﾝ
4. ｳｧｰ('A`)
5. 工ｴｴｴｴｴ(;´Д｀)ｴｴｴｴｴ工

このような特徴的な表現の出現回数をパターンマッチ

ングにより検出することで，書き込みを行った視聴者の盛り上がり（喜び・興味）や落胆（悲しみ）といった感情の度合いを抽出することができると考えられる．

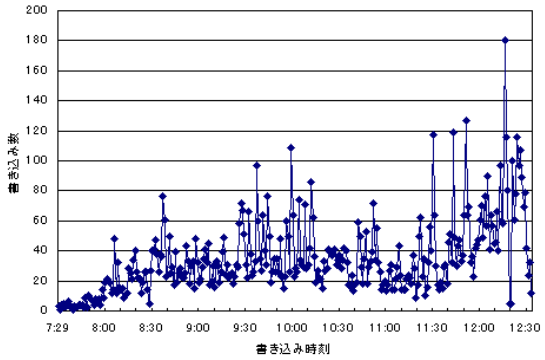

図３．番組実況チャットへの書き込み数の時間変化

以上のような考えに基づき，ここでは図４のようなインデキシング処理を行う[13].

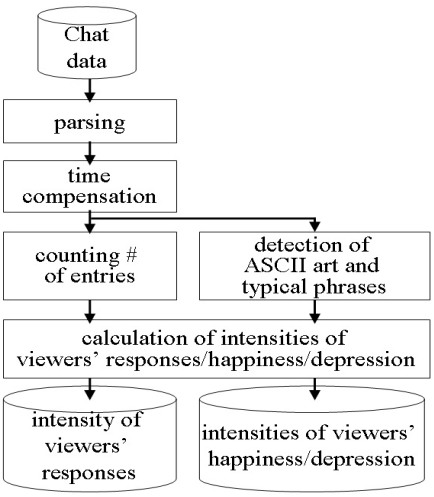

図４．インデキシング手順の概要

まず，記録された実況チャットのデータをパースして，書き込み時刻，書き込み者の ID，書き込み内容を得る．

なお，書き込み時刻と，書き込み内容に対応する番組映像シーンの時刻にずれが生じる場合がしばしばある．これは，映像中のあるイベントが起きた後に，視聴者が書き込みを行うためであり，ここでは，書き込み内容のテキスト長に応じた補正秒数の算出を行い，書き込み時刻の補正をする．

次に，ある単位時間ごとの書き込みエントリ数を計測することにより，反響の大きさ度数を計算する．また，アスキーアートによる特徴的な表現の出現回数を単位時間ごとにパターンマッチングで検出することで，視聴者の盛り上がり度数および落胆度数を抽出する．実際には，書き込み内容に出現する「ナイス」，「すげー」，「凄い」，「すっげぇ」，「キター」といった盛り上がりを表す単語や，「あーあ」，「アホ」，「ださださ」，「だせえ」，「なんだかなぁ」といった落胆を表す単語を用いた学習などと並行して検出器を構成することにより，それぞれの感情をより効果的に抽出することができる．

以上のような処理の結果，以下のような数値の組からなるインデックスが生成される．

1. 視聴者（書き込み者）全体について
   (ア) 計測時刻，計測単位時間
   (イ) 書き込みエントリ数
   (ウ) 盛り上がり・落胆の感情度数
2. 視聴者（書き込み者）個別について
   (ア) 書き込み者ＩＤ
   (イ) 計測時刻，計測単位時間
   (ウ) 書き込みエントリ数
   (エ) 盛り上がり・落胆の感情度数

以上のインデックスおよび実際の書き込みデータを利用することで，録画番組を様々な視点や見方で視聴可能なビュー生成が可能となる．

ビューの表示インタフェースとしては，種々の方式が考えられるが，ここでは，ウェブ化ビデオ[12]で採用されているストーリーボード形式のインタフェースを用いて説明する．ウェブ化ビデオ[12]は，次世代蓄積型テレビを想定したプロトタイプであり，テレビ番組がハイパーリンクを用いたウェブコンテンツに変換されているため，様々なビュー設定による表示インタフェースとして適していると考えられるためである．

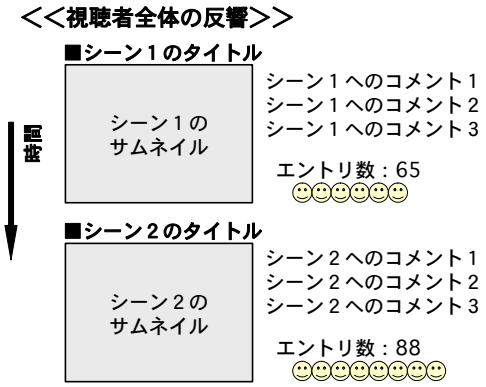

図５．反響の大きさの時間順によるビュー

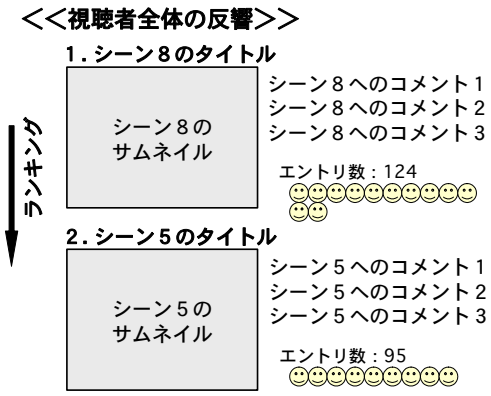

図６．反響の大きさのランキングによるビュー

まず，さまざまなシーンに対する時間順およびランキング順によるビューが可能になる．ここでシーンとは，単位時間毎や内容に沿ったトピック毎など適当な分割基準や手段を介して番組映像を分割して得られる部分映像という意味であり，以下の図５～図９では，単位時間＝１分としたときの各部分映像をシーンとしている．

例えば，各単位時間当たりの書き込みエントリ数から視聴者全体の反響をシーンの出現順に従って表示すると図５のようになる．シーン毎に，サムネイルと書き込みコメント,書き込みエントリ数の数値,および,エントリ数の大きさを表した顔アイコンが表示され，それらが縦方向に時間順に配置されている．このビューは，シーンの時間経過に沿った流れを把握する際に役立つ．

一方，書き込みエントリ数でソートすることにより，視聴者全体の反響が大きかった順に各シーンを並び替えて表示することができる（図６）．このビューは，一般的に反響の大きかったシーンをいくつかかいつまんで視聴する際に役立つ．特に，このランキングによるビューは，番組実況チャットから視聴者の反応を得ることではじめて可能になる表示であり，これまでにはない新しい表示法の一つであると考えられる．

また，視聴者全体あるいは特定の視聴者に基づくビューが可能である．例えば，各シーンに対する任意の視聴者のコメントを表示すると図７のようになる（すべてのコメントを表示することはあまり現実的でないので，通常はいくつかのコメントを適宜選択することにより表示する）．このビューは，視聴者全体の反響を把握する際に有効となる．

一方，指定された書き込みＩＤに対応する視聴者のコメントのみを表示すると図８のようになる．このビューは，個別の視聴者の反響に注目して視聴する際に特に有効である．

これら全体／個別の視聴者に基づくビューでは，それぞれいずれも時間順あるいはランキング順による表示が可能であり，利用者は，状況に応じてさまざまな見方を選択することが可能となる．

さらに，視聴者の反響や盛り上がり・落胆の度合いに基づくビューが可能である．例えば，書き込みエントリ数から視聴者全体の反響を表示すると図５あるいは図６のようになる．このビューは，視聴者の一般的な反響（喜びや悲しみの感情などを含めてある内容によって書き込みを行った結果の反応）を総合的に把握する際に役立つ．

一方，盛り上がりの感情度数から，視聴者Ａの盛り上がったシーンだけを時間順に表示すると，例えば，図９のようになる．ランキング表示にすれば，盛り上がり度数の大きい順に並ぶので，シーン２はシーン１よりも上部に表示されることになる．このビューは，特定視聴者の盛り上がりや喜びの感情が表れた部分だけを把握する際に役立つ．もし落胆の感情度数から表示を行えば，視聴者Ａが落胆したシーンだけを表示することができる．

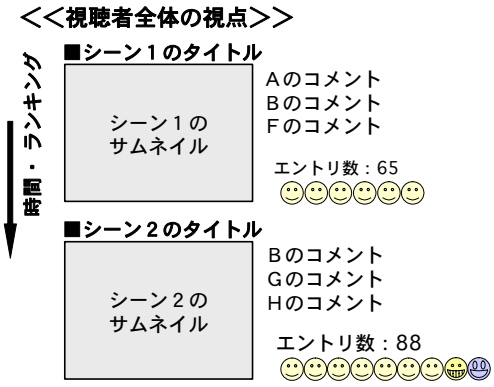

図７．視聴者全体の視点を表すビュー

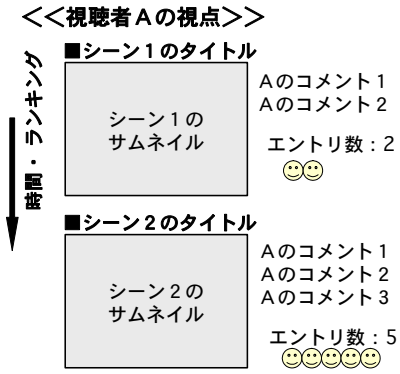

図８．個別の視聴者の視点を表すビュー

また，図９では，盛り上がり度数の大きさに応じてサムネイルやテキストのサイズ，アイコンを大きく表示している．また，盛り上がりの感情を表わすためにアイコンの種類を変えて表示している．このようにして，利用者にとってよりインパクトのある多様な表示を提供することが可能になる．

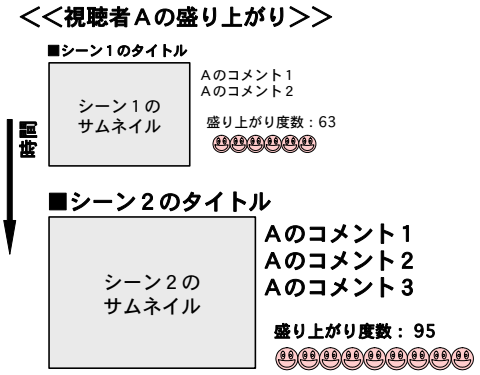

図９．視聴者Ａの盛り上がりを表すビュー

以上で示した全体／個別視聴者ごとのビュー，あるいは，反響／盛り上がり／落胆の度合いに基づくビューについても，番組実況チャットから視聴者の反応を得られることではじめて可能になる表示であり，これまでにはない新しい表示法の一つになると考えられる．

さらに，これらのビューを応用するとさまざまなイ

ンタラクションや提示を利用者に提供できると考えられる．例えば，番組実況チャットのインデックスや書き込みデータと並行して，スポーツの試合などにおけるスコア情報を利用することにより，特定のチームが点を入れた際にだけ盛り上がりの感情度数が高い視聴者や，チームに関係なく点が入った際に盛り上がりの感情度数が高い視聴者を抽出することができる．これにより，例えば，Ａチームを応援している視聴者の視点，中立の立場で視聴している視聴者の視点を表すビューを提供することができると考えられる．

また，特定の選手名が含まれており，その書き込み内容の感情度数が高いものを抽出することができる．この情報を利用することで，ある選手について強い思い入れを持つ人を推定し，熱狂的なファンの視点に基づくビューを提供することができると考えられる．

迷惑な書き込みをする人がいる場合は，書き込み者ＩＤやそのような迷惑な書き込み内容の特徴を学習した識別器を用いることで，迷惑な特定の人や書き込み内容を検出し適宜表示内容から除くことも可能になる．

図１０．視聴者全体の反響を表す時間順ビュー

図１１．視聴者全体の反響を表すランキングビュー

一方，提案手法を利用することで，自分と同じ（あるいは異なる）興味（価値観）をもつ他の視聴者を発見することができると考えられる．例えば，自分の ID をもとにしたビューを生成し，その場面で同様の反応をしている視聴者や真反対の反応をしている視聴者の ID を記憶しておく．同じ興味をもつ視聴者をグループ管理しておくことで，他の番組を視聴する際に同じ興味をもつ視聴者によるビューが手軽に生成できる．この視聴形態は未視聴の番組をダイジェスト視聴するときに特に有効となるであろう．

図１０～図１３に提案方式によるビューを実装した例を一部示す．図中，見出しはトピック毎に分割して得られた字幕データであり，これに対応する部分映像がサムネイルとして表示されている．サムネイル右のリンクも字幕データを表示したものである．その下部に，視聴者の反響や感情度数が表示され，感情を表わす顔アイコンとともに実況チャットの書き込み内容が表示されている．

図１２．視聴者Ａの視点を表示した時間順ビュー

図１３．視聴者Ａの盛上りを表示した時間順ビュー

4. 考察

ここでは，提案手法の利点や応用例，今後の展開などについて考察する．

番組実況チャットを利用したビュー生成による最大の利点は，ビューを生成する際の条件に，視聴者の反応といった要素を新たに含めることができるようになったという点である．図１４は，提案法により従来の番組視聴の形態には存在していなかった視聴者の反応という新たな要素が導入されたことを示している．

これにより，

- 視聴者による反応を利用したランキングによる提示
- 視聴者の反響の大きさや盛り上がり・落胆の度合いを共有した視聴
- 自分と同じ（あるいは異なる）興味を示す他の視聴者の価値観の発見・共有

といったさまざまな付加価値をもつ視聴が実現できるようになったと考えられる．

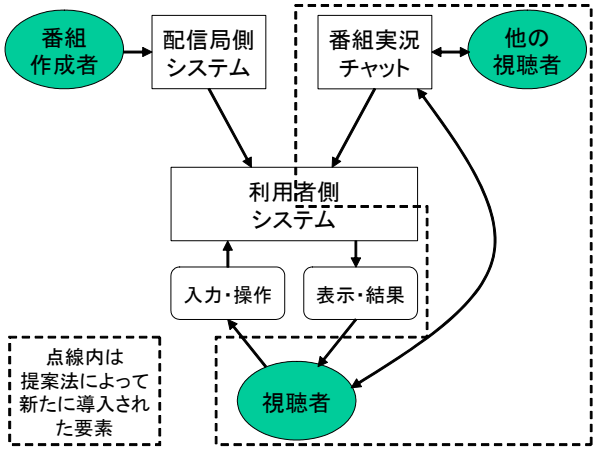

図１４．提案方式の概要

さらに，視聴者の反応を利用したランキングは，１つの番組だけでなく，他の番組やある一定期間に放送された番組をまとめた複数の番組に対して行うことができる．これにより，例えば，この一週間での視聴者の反響や盛り上がりの感情度数が大きかったシーンのランキングといった提示に応用することができる．

なお，提案手法では，リアルタイム視聴および録画後の視聴のいずれの視聴形態にも適用できるものである．テレビの実況中継を見ながら実況チャットを生で見る使い方や,録画テレビ映像と録画実況チャットを，後から様々なビューを設定して楽しむという使い方が考えられる．また，テレビ（HD レコーダ）だけでなく，PC や携帯電話といった様々な機器を用いた利用形態が考えられる．

番組実況チャットの書き込みが行われた時刻と，その内容に対応する番組映像シーンの時刻にずれが生じる場合がしばしばある．これは，映像中のあるイベントが起きた後に，視聴者が書き込みを行わざるを得ないためであり，ここでは，書き込み内容のテキスト長に応じた補正秒数の算出を行い，内容の同期をとるように配慮している．

さて，本稿では，番組実況チャットの基本的なデータである書き込み時刻，書き込み者の ID，書き込み内容に基づいたインデキシングおよびビュー生成を行ってきた．これに加えて，番組実況チャットに参加している視聴者プロファイルが利用できると，さらに多様なビュー生成が可能になると予想される．例えば，男性・女性による視点，年齢による世代別の視点，関東や関西といった地域別の視点，日本，韓国，中国，米英仏独といった国別の視点などを反映させたビューなどが実現できると考えられる．こうした利用スタイルが受け入れられると，いわゆる 2ch などの掲示板だけでなく色々な番組実況チャットコミュニティができていくのではないかと期待される．

本稿では，番組実況チャットのみを利用したが，性別，年齢，地域，国籍などを記述した視聴者プロファイルも含めた形で視聴者情報を積極的に流通・管理・利用可能な仕組みができることになると，より多様な番組視聴が可能になると考えられる．なお，この件については，プライバシー情報の保護など多くの問題がある．

5. 関連研究

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したり，限られた時間でハイライトだけを見るといった多様な視聴方法についての研究は活発に行われてきた．

例えば，映像コーパスから興味のあるショットを探索するために視覚的要素を利用したインタフェースが提案されている[1]. 映像セグメントを表す複数のキーフレームから構成されるストーリーボードと視覚的特徴によるフィルタリング機能を組み合わせることで効率的な探索が可能だとしている．

また，番組概要の把握やシーン探索を効率行うための表示インタフェースとして TV2Web[2]や Video Manga[3]が提案されている．TV2Web では，映像セグメントのサムネイルと字幕テキストから構成されるストーリーボードを異なる詳細度レベル(セグメント，シーン，ショット）で構造化し，これら構造データをズームメタファによってスムーズにつなぎながら（切り替えながら）表示することで効率的な概要把握やシーン探索が可能だとしている．Video Manga では，クラスタリングされた映像セグメントが長くて稀なものの重要度を高く計算し，その値に応じて表示するキーフレームの大小を制御する．異なるサイズのキーフレームをマンガのような表示形態に配置しサマリーとして表示することで，より重要な場面を視覚的に把握しやすいインタフェースを実現している．

限られた時間でハイライトを視聴するために，テレビ番組のためのダイジェスト視聴システムが提案されている[4]. 番組映像中に配信側で用意された汎用的な番組索引と検索用ルールが含まれることを前提とし，これらを利用して利用者の嗜好を考慮したハイライトをダイジェスト視聴することができる．

これらの手法は，いずれもキーフレームや重要な映像セグメントの時間順によるビューを基本としていた．一方，提案手法は，このような時間順によるビューだけでなく，他の視聴者による盛り上がりや反応を反映した番組シーンのランキングビューといった提示を可能としている．このような見せ方は従来手法には存在

しておらず，提案手法によって初めて可能になったものであると考えられる．

また，シーン探索や概要把握，ハイライト視聴の仕組みを実現するための基礎技術である映像インデキシング技術の研究も活発に行われてきた．

例えば，色を用いた例としては，色ヒストグラムや特徴的な色の集合，隣接する特徴的な色対などを利用した手法が提案されている[5]．これにより，問合せで指定された画像中の物体が，動画中に存在するかどうかの判定などが可能になる．また，動きを用いた例としては，動きベクトルを用いた手法が提案されている[6]．ブロックマッチング法などで得られた動きベクトルを，ズームやパンなどを特徴づけるパラメータでモデル化し分類することにより，各カメラ操作の有無や強さの判定が可能になる．

さらに，マルチモーダルなインデキシングの例としては，シーン分割，カメラ操作検出，顔検出，映像キャプション検出等を利用して，シーンの重要度を判定・選択する手法[7]や，テキスト解析で得られたキーセンテンスと画像解析で得られたキー画像を，動的プログラミングを使って関連付ける手法[8]が提案されている．マルチモーダルな情報に加えて，特定分野のドメイン知識を利用することにより，画像中の人物動作等の，より詳細な内容に関するインデキシングが可能であることも示されている[9][10][11]．

しかし，これら従来手法が利用してきたデータは，番組の配信側（放送局）から提供される情報のみであったため，インデキシングされた情報は，基本的に，番組作成者・配信側の意図に基づく画一的なものであった．よって，従来手法においては，シーン探索や内容の早見，ダイジェスト視聴などの内容に，自分以外の他の視聴者の視点や反応といった要素を取り入れることはできなかった．提案手法では，番組実況チャットを利用することで，放送番組に対する視聴者の視点や反応を反映したインデキシングが可能となり，これを利用することでさまざまな見方を可能とするビュー生成が可能になったと考えられる．

6. まとめ

本稿では，番組実況チャットに基づく視聴者視点を利用した放送番組のビュー生成手法を提案した．番組実況チャットのデータを統計処理・認識処理することで，番組の盛り上がり場面や，各シーンに対する特定の視聴者のリアクション等を効果的に抽出することができる．本稿では，これら抽出結果によって，視聴者の反応を利用した時間順／ランキングによるビュー，全体／個別の視聴者の視点を反映したビュー，反響の大きさ・盛り上がり・落胆の感情度数を反映したビューといったさまざまなビューを生成できることを示した．これらはいずれも，視聴者による反応を利用することで初めて可能になったものであり，従来手法にはない新しい付加価値をもたらすメリットがある．さらに，提案手法は，他の視聴者の返す反応を共有した視聴や自分と同じ（あるいは異なる）興味を示す他の視聴者の価値観の発見・共有といった新しい番組視聴を提供するものであると考えることができる．

今後は，書き込み内容の統計処理やパターンマッチング，番組映像との同期補正，インタフェースの改善など今後プロトタイプの改良を行っていく予定である．また，性別，年齢，地域，国籍などを記述した視聴者プロファイルを積極的に流通・管理・利用可能なサイトの構築などを行うことによって，より多様な番組視聴が可能になることを実証できると期待される．

文　献

[1] Christel, M.G., Huang, C.: Enhanced access to digital video through visually rich interfaces. ICME, MD-L5.1, 2003.

[2] 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.

[3] Uchihashi, S., Foote, J., Girgensohn, A., Boreczky, J.: Video Manga: generating semantically meaningful video summaries. Proc. ACM Multimedia 99, 1999.

[4] Hashimoto, T., Shirota, Y., Mano, H., Iizawa, A.: Prototype of Digest Viewing System for Television. IPSJ, Vol.41, No.SIG3(TOD6), pp.71-84, 2000.

[5] Nagasaka, A., Tanaka, Y.: Automatic video indexing and full-video search for object appearances. IPSJ, Vol.33, No.4, pp.543-550, 1992.

[6] Akutsu, A., Tonomura, Y., Hashimoto, H., Ohba, Y.: Video indexing using motion vectors. In SPIE Proc. VCIP '92, pp.522-530, 1992.

[7] Smith, M., Kanade, T.: Video Skimming and Characterization through the Combination of Image and Language Understanding Techniques. CVPR, 1997.

[8] Nakamura, Y., Kanade, T.: Semantic analysis for video contents extraction - spotting by association in news video. ACM Multimedia, pp.393-401, 1997.

[9] Intille, S.S., Bobick, A.F. Closed-world tracking. Proceedings of the Fifth International Conference on Computer Vision, pp. 672-678, 1995.

[10] Saur, D.D., Tan, Y-P., Kulkarni, S.R., Ramadge, P.J.: Automated analysis and annotation of basketball video. Storage and Retrieval for Image and Video Databases V, SPIE-3022, pp.167-187, 1997.

[11] Miyamori, H.: Automatic annotation of tennis action for content-based retrieval by integrated audio and visual information. CIVR2003, LNCS2728, Springer-Verlag, pp.331-341, 2003.

[12] Miyamori, H., Tanaka, K.: Webified video: media conversion from TV program to Web content and their viewing method integrated with related information, DEWS2005 2C-i6, 2005.

[13] Miyamori, H., Nakamura S., Tanaka, K.: Automatic indexing of broadcast content using its live chat on the web. IEICE NLC2004-123, PRMU2004-205, pp.43-48, 2005.