

2つの学会発表録画を同時視聴するためのシステム

太田 佳 敬[†] 宮 下 芳 明^{†,††}

本稿では、学会の登壇発表を録画した動画を、2つ同時に視聴できるシステムを提案する。発表で重要な部分が同時刻に再生されないように、重要でない部分を高速再生し、重要な箇所の再生位置をずらす。重要箇所の検出は学会中に行われたチャットログを用いた。これにより2つの動画を同時に、かつ注目されている部分は見逃さずに視聴することができる。

Simultaneous Viewing System for Two Video Recordings of Academic Presentations

YOSHIAKI OTA[†] and HOMEI MIYASHITA^{†,††}

In this paper, we propose a system that enables the user to watch two recorded videos of academic presentations simultaneously. The system avoids playing important parts at the same time by controlling playback speed; it shifts the important part by accelerating the insignificant part. The system detects the important parts from the chat logs by the audience. In this manner, the user can understand what are presented on two videos at the same time without missing interesting parts.

1. はじめに

1.1 背景

学問に携わる人数は年々増加している。このまま人数が増加していくと、1つの学問領域に多くの人が参加する時代が来る。その結果、研究結果である論文の数、発表の場である学会の数や、学会に参加する延べ人数も増加する。しかし研究者の時間は有限であるため、全ての学会に参加し、全ての発表を見ることは不可能である。

最近では発表のオンライン中継を行う学会や、さらにその録画をアーカイブして公開する学会が増えた¹⁾²⁾。これによりその時その場に居合わせなくても、発表の内容を後から参照する事が可能になっている。今後この流れはより加速し、多くのオンライン中継や、発表を録画した動画がアーカイブされるようになる。しかし、動画を鑑賞するためには再生時間と同じだけの時間が必要であり、見ようとする動画の数が増えると、鑑賞する時間も増加していく。そのため効率よく

発表を見ることは、今後より大きな問題になる。

1.2 目的

後述するように、動画を効率的に見るための研究は多く存在する³⁾⁴⁾⁵⁾⁶⁾。これらの大半は、動画の重要でない箇所をフィルタリングし、重要な箇所からなる要約動画により、短時間で効率的に動画を見る手法である。この手法は動画一本あたりの視聴時間を大幅に減らせるが、話の流れが要約されたことによって、動画のコンテキストが解りにくくなるという問題がある。学会発表のような前提条件や過程の理解が必要な場合には、重要な部分だけが含まれる要約された動画を見ても、意味が分からなかったり、誤解する可能性がある。栗原は、扱わなければならない情報の総量が増えたときに、このようなフィルタリングにより適正量に削減する手法だけではユーザが理解する情報の量は増えず、ユーザが情報を理解する速度を上げる必要もあると述べている⁷⁾。またそのための手法として、映画鑑賞時に字幕の無い所を高速再生し、音声があるために字幕が表示される所をそれが読める速度で再生することで、内容を把握しつつ通常よりも短時間で映画を鑑賞できるシステムを提案している。この手法は、間がたくさんあるドラマや映画などの場合には有用である。しかし、学会発表のようにほぼ常時説明が行われているような場合には、省略する間の部分が少なく高速再生ができないために、短時間に要約して視聴する

[†] 明治大学大学院 理工学研究科新領域創造専攻 デジタルコンテンツ系

Program in Digital Contents Studies, Program in Frontier Science and Innovation, Graduate School of Science and Technology, Meiji University

^{††} 独立行政法人 科学技術振興機構, CREST
JST, CREST

ことは難しいと思われる。これらのことから、学会発表の内容を把握しつつ、動画を効率的に見るための手法は確立されていないと言える。

ところで、WISS (インタラクティブシステムとソフトウェアに関するワークショップ)¹⁾ は、国内で先駆的にオンラインとの連動を試みている学会の1つである。そのため、昔から学会発表中にチャットを並行して行っており、多くのチャットシステムが研究されている⁸⁾⁹⁾¹⁰⁾¹¹⁾ 当初はオンラインチャットによって実世界の議論が不活性化する懸念もあったが、むしろ大幅な議論の活性化に貢献し、現在では WISS のスタイルとして定着している¹²⁾ つまり WISS では学会発表を聞きつつ他の作業を行うことが推奨され、実際に効果を上げているのである。

ニコニコ生放送¹³⁾ を使い、中継を行った WISS やニコニコ学会¹⁴⁾ では、視聴者がニコニコ生放送の機能によりリアルタイムにコメントを残していた。さらに WISS や Augmented Human 2011²⁾ はコメント機能の無い Ustream¹⁵⁾ でも配信を行ったが、それを見ていたと思われる人々が Twitter のハッシュタグを使ってコメントをリアルタイムに書き込んでいた。

このように、発表を聞きながらオンラインメディアに発言ができる余裕があるなら、学会発表を録画した動画を見ているときにも、もう1つ動画を見ることができるのではないだろうか。これができれば、内容を把握しつつ全体としておよそ半分の時間で動画を見ることができる。

だが、動画を2つ同時に再生する場合、重要な箇所が同時に再生されてしまいどちらも見られずに見逃す可能性が生じる。これに対して、重要でない箇所の再生速度を調整することで、重要な箇所を再生するタイミングを調節することが可能である。それにより、2つの動画を同時に再生しても、重要な箇所が同時に再生されることを回避できる。

前述のように、WISS やオンライン中継をした学会では、発表と並行してリアルタイムにコメントが書かれている。本稿第二著者は、ニコニコ動画のコメントをメタ情報として扱い、それを元に映像から重要な箇所を抽出して要約を行った¹⁶⁾ このような手法を学会で書き込まれたコメントに適用すれば、学会発表の重要な部分を抽出できる。

以上をふまえ本稿では、2つの発表動画を1つの画面に表示させ、かつ重要な箇所が同時に再生されないように、発表時のオンラインメディアのログを利用して調整するシステムを提案する。

2. 関連研究

2.1 動画処理に関する研究

2.1.1 映像特性を元にした動画要約

動画の特徴点などを利用し、動画圧縮を行う手法はかつてから様々な手法が試みられてきた。森山らは、カッティングや登場人物の台詞、BGM、効果音などのトラック構造からドラマ映像の心理的な表現を検出し、それを元に映像を圧縮した³⁾ 永橋らは、ニュース映像の字幕や映像的特性からメインピックの部分を検出し、それを元に映像を要約した⁴⁾ 出口らは、映画の文法と呼ばれる編集上強調された区間と従属する区間を抽出することで、強調された区間だけでなくそれに至る経緯も含んだ映像要約を可能にした⁵⁾

2.1.2 メタ情報を元にした動画要約

動画に付随するメタ情報を利用し、動画の要約を行う手法はいくつか存在する。伊藤らは、動画と用意されたメタ情報を元に、ユーザが指定する任意の要約率で音声テキストを要約することで、それに対応する映像も要約する手法を提案した⁶⁾ 本稿第二著者は、ニコニコ動画のコメントを元に音楽のサビ検出や、映像から重要な箇所を抽出し、映像を要約する手法を提案した¹⁶⁾ 栗原は、字幕がある部分を文字が読める速度で、無い部分を高速再生することで、通常的高速再生と比べて内容を把握しつつ短時間で閲覧を可能にした⁷⁾ SmartPlayer¹⁷⁾ では、映像の変化率によって再生速度を自動調節し、短時間での動画閲覧を可能にした。

2.1.3 動画閲覧の支援

動画要約以外にも様々な手法で動画を処理し、効率的に閲覧するためのシステムはいくつか提案されている。谷らは、ニコニコ動画のコメント数の動的な変化を利用してキーフレームの抽出を行い、動画の要約を目指した¹⁸⁾ 田中らは、動画からシーンの自動抽出を行い、それをブラウジング可能にすることで、目的の動画により早くたどり着けるようにした¹⁹⁾ 當麻らは、映像からカットデータを抽出し、カットされた部分の意味を推定し、複数のカットから映像のシナリオ構造を再構成することで、映像の自動的構造化を行った²⁰⁾ 磯貝らは、動画から要約動画を作る際に、コメント数の多いシーンから必要なシーンのみを抽出するための手法を提案した²¹⁾ 栖閑らは、バッファに溜めた2つのTV動画を短縮再生し、それらを短時間で切り替えることで、リアルタイムに2つのTV映像を見るシステムを提案した²²⁾ また、動画を2つ同時に見るためのソフトウェアは既にいくつか存在する²³⁾²⁴⁾²⁵⁾ し

かし、これらのソフトウェアは動画自体の速度を変えたり、複数の速度によって盛り上がっている箇所の再生時間を調節する機能は無い。

2.2 チャットに関する研究

Lock-on-Chat⁸⁾では、参加者間で画像を共有し、それら画像の特定部分を会話に結びつけることで、複数の話題に分散した議論を促進している。Lock-on-Chat IKKI⁹⁾では、匿名で発言をし一定以上の賛同が得られた場合にのみ、賛同者全員の名前を連判状として表示することで匿名の発言を促している。KairosChat¹⁰⁾では、発言内容に応じて異なる流速の発言を可能にし、逸脱発言のしやすさと自然な発言履歴の精練化の両立を目指した。On-Air Forum¹¹⁾では、参加者の負荷を抑えつつ、コンテンツとコミュニケーションを同時並行的に把握できることを目指した。藤本らは、プレゼンシステムにマンガ表現を取り入れるとともに参加者にチャットを提供し、かつその盛り上がり具合や人気の発言をプレゼン上に割り込ませるシステムを提案した。²⁶⁾

3. 事前調査

今回システムを作成するにあたり、まずは登壇発表と並行してチャットを行っている WISS のログを分析した。分析に使用したログは WISS2010 のものである。

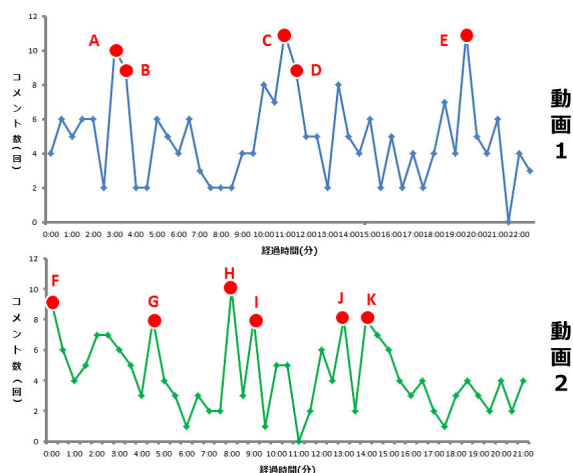


図 1 発表動画におけるコメント数

図 1 は、WISS2010 で発表されたある 2 つの発表についてチャットログを 30 秒ごとに区切り、その区間のコメント数をグラフにしたものである。赤い丸をつけた部分は、発言数の上位 3 位までの地点である。どちらのグラフも、コメント数がとても多くなる部分と、そうでない部分が分かれている。このチャットは発表の内容に連動しているものがほとんどである。そのた

め、発表には視聴者が強く興味を示す部分と、そうでない部分があることが分かる。言い換えれば、発表にはメリハリがあり、興味がある部分ではチャットが盛り上がっていると言える。

次に、どのような時に発表が盛り上がっているかを調べた。そのために、図 1 で赤い丸をつけた、コメント数が多い 11 カ所について、実際の発表ではどのような部分なのかをまとめた。視聴者がコメントを書き込むまでの時間には多少のばらつきがあるため、その区間の前一分まで遡った。

表 1 コメントが多いところの内容
Table 1 Contents of the excessive comments

箇所	動画 1	箇所	動画 2
A	デモ開始	F	発表開始
B	デモ開始	G	目的の発表
C	機能説明	H	機能説明
D	機能説明	I	機能説明
E	質疑応答	J	デモ開始
		K	発表終了

表 1 が、図 1 中の赤い丸と実際の動画内容との対応表である。動画 2 の発表の開始と終了直後を除くと、どれもがデモや提案システムの説明など、発表で最も注目される箇所である。このことから、コメント数が増えている場所を見ることが、発表そのものが盛り上がる部分を知ることができる。

4. 提案システム

前章で述べた、コメント数から発表動画の盛り上がり状況を抽出できるということを活用し、それを元に盛り上がっている部分をできる限り重ならないように速度を調節し、2 つの動画を同時に見るシステムを提案する。提案するシステムは、

- 動画の盛り上がり度を計算
- 盛り上がり度を利用した動画合成

の 2 段階の処理から成り立つ。以下に各機能の説明を述べる。

4.1 動画の盛り上がり度計算

本システムは、チャットデータが含まれているログファイルを読み込み、一定時間単位に分解したのちに盛り上がり度を計算する。この盛り上がり度の計算を行う時間の長さのことをユニットと呼ぶ。さらに、パラメータとしてユニットに区切る間隔、盛り上がり度を判別するための閾値、動画再生速度の上限 (最高速度) を指定できる。(図 2)

システムは、ユニット単位でログファイル中のデータを取り出し、盛り上がり度を計算する。今回はログ

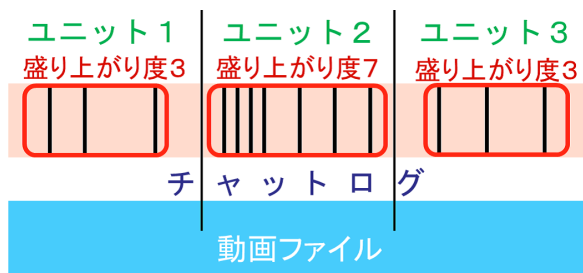


図2 ユニットと盛り上がり度

ファイルとして事前調査で使った WISS のチャットログを使い、ユニットのコメント数を盛り上がり度として定義した。

4.2 盛り上がり度を利用した速度調整

本システムでは動画合成手法として、2つの動画を左右に並べる形で合成する。合成にあたっては前述の機能が求めた盛り上がり度を利用し、二つの動画が左右に並んで再生され、かつ可能な限り盛り上がっている部分が被らないように動画を調節する。動画の合成時に、盛り上がり度による調整を行わないと、図3の調整前のように盛り上がっている部分が重複する。そこで、矢印の区間のようにあまり盛り上がっていない部分の再生速度を速くし、その後ろの部分にずらすことで盛り上がり度の衝突を回避する。

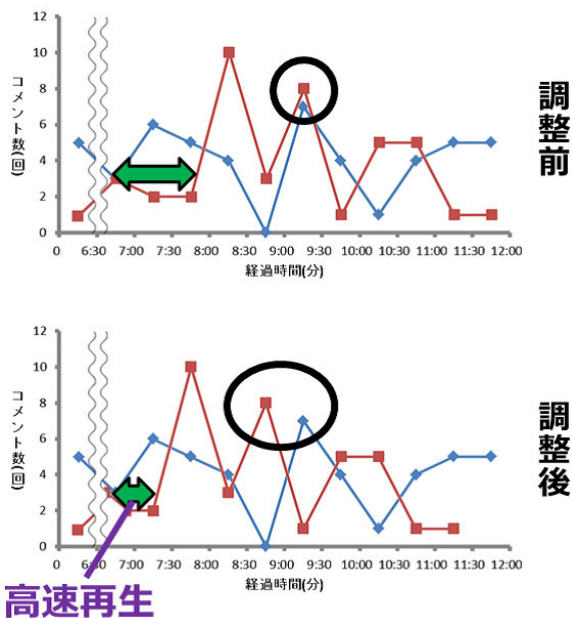
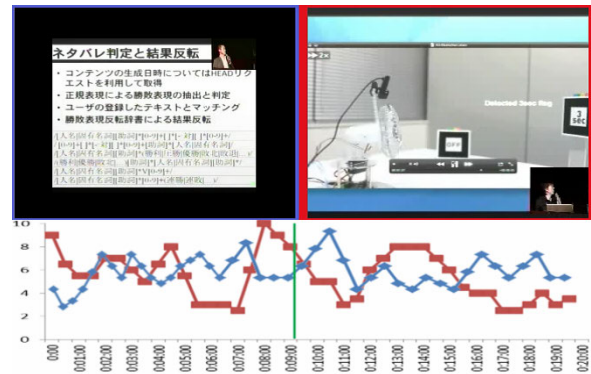


図3 盛り上がり度を利用した速度調整

また、単に動画を並べて再生する場合、ユーザはどちらの動画を見るべきか判断できない。そこで、図4のように、その時盛り上がっている方の動画を表示・音声共に大きくし、盛り上がっていない方の動画を小

さく表示し、音声は流さない。さらに画面下にどちらが盛り上がっているかを提示することで、ユーザが注視すべき動画を指示する。



5. 運用実験と課題

WISS2011のWISS Challengeにて、WISS2011のチャットログと発表の録画を使い、期間中に上映をした。その結果、重要な部分の抽出と提示には成功した。しかし、いくつかの問題点も判明した。以下にその中から抽出した現状の課題とその改善案を述べる。

5.1 動画選択インターフェース

現在のシステムは2つの動画をあらかじめエンコードし、それを閲覧するものである。しかし複数の動画がある場合、見たい動画は人によって違うため、ユーザに合わせてエンコードする2つの動画を選ばなければならない。そのため、ユーザが任意の2つの動画を選び、その場でエンコードして見られるようなインターフェースが必要である。

5.2 コメント数の計算

谷も述べているが、チャットなどに発言が書き込まれた時間は、常にユーザが書き込もうと思ったシーンよりずれている¹⁸⁾。その差を補正する必要があるが、打つのが早かったり、実況をしているためにすぐ話題を書き込む人と、打つのが遅かったり、熟考したために前の話題を後から書き込む人がいる。そのため、一律に補正をすると不具合がでる可能性がある。またニコニコ動画のようなサイトでは、該当の箇所に戻って書き込むといった事もあるため、その場合はずれを補正するわけにはいかない。

現状のシステムでは、コメント数をある程度の長さのユニットとして見ることで多少のずれを吸収している。しかし、この手法ではユニットの切れ目の所で同じような問題が出てしまう。これに対する解決案とし

では図5のように、後続のユニットから、現在のユニットに対するコメントと考えられる部分のコメントも、計算対象とする手法が考えられる。

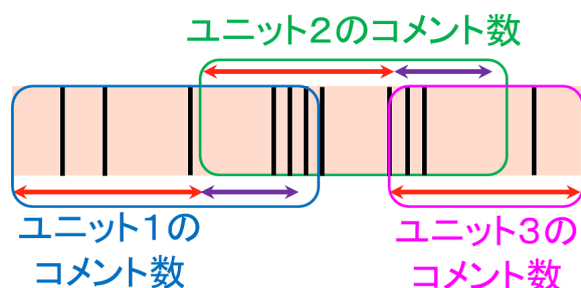


図5 コメント数の計算

これにより、盛り上がっている箇所が2つに分断されて検出されなくなるという事が防げる。この方式ではコメント数が全体的に大きくなってしまいが、本システムではコメント数が相対的に多い少ないで判断しているため影響はないと思われる。実際に、図1で使ったチャットログに対してこの手法で補正したものが図6の動画1である。全体的にグラフは増加しているが、盛り上がっている部分が大きく増加しているのに対して、盛り上がっていない部分はほとんど増加していないことが解る。

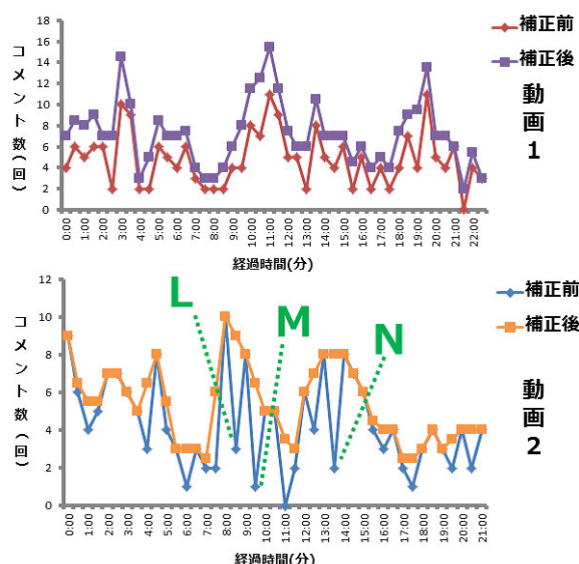


図6 補正後のコメント数

また、図6のL、J、Kのように、一瞬だけコメント数が少なくなる場合も盛り上がっていないと見なしている。そのため、動画エンコード時に短い時間で何度も速度が変わってしまい、ユーザにとって視聴しづらいものになってしまう可能性がある。そこで、一瞬だ

け落ち込んでいるようなコメントに対しては、その前後を繋ぐように補完を行う手法も必要である。図6の動画2の補正後は、その場所のコメント数を元のデータ、もしくは元のデータの前後のコメント数の平均のどちらか大きい方を採用したグラフである。これにより、一瞬だけコメントが少なくなる場合が改善され、かつ前後が共に盛り上がっていない部分への影響は殆どないことが解る。今後は、この2つの補正をかけた上でエンコードを行っていく予定である。

5.3 コメント数以外の計算手法

現状では、あるユニットの盛り上がり度計算には、コメント数しか反映させていない。例えば WISS で利用されている On-Air Forum ではコメントを書かずに送信ボタンを押すと、その数に応じて画面全体を赤くすることができる Excite 機能が備わっている¹¹⁾。そのため、Excite 機能により画面が赤くなった箇所も、盛り上がっている箇所として定義することができる。

また、長い文字列が連続されて投稿されているような箇所は、コメント数が少なくても重要な会話をしている可能性が高い。そのため、その会話を引き起こしたような箇所は盛り上がり始めた箇所として考える必要もある。

6. ま と め

本稿では、発表時のオンラインメディアのログを利用し、2つの発表動画を1つの画面に表示させ、かつ重要な箇所が同時に再生されないように調整するシステムを提案した。発表時のチャットログにはメリハリがあり、発言数が多く盛り上がっている部分を重要箇所と考え、盛り上がっていない部分を高速再生することで、重要な箇所の再生位置を調節している。今後は評価実験に加え、より利用者が重要な動画を見やすい表示手法を調べていく。

参 考 文 献

- 1) Workshop on Interactive Systems and Software: URL: <http://www.wiss.org/> (2011年12月10日に参照)。
- 2) Augmented Human International Conference: URL: <http://www.augmented-human.com/> (2011年12月10日に参照)。
- 3) 森山剛, 坂内正夫: ドラマ映像の心理的内容に基づいた要約映像の生成, 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理, Vol.84, No.6, pp.1122-1131 (2001)。
- 4) 永橋功丞, 宮岡伸一郎: 映像中の情報を用いたニュース映像要約手法の研究 (動画像・メディア理解, 学生セッション, 人工知能と認知科学), 全国大

- 会講演論文集, Vol.70, No.2, pp.357-358 (2008).
- 5) 出口嘉紀, 吉高淳夫: 映画の文法に基づく要約映像の生成 (セッション 2: 映像・コンテンツの生成), 情報処理学会研究報告. データベース・システム研究会報告, Vol.2004, No.3, pp.33-40 (2004).
 - 6) 伊藤一成, 酒井康旭, 斎藤博昭: メタデータ解析と自然言語処理を併用した動画要約 (セッション 2: 映像・コンテンツの生成), 情報処理学会研究報告. データベース・システム研究会報告, Vol.2004, No.3, pp.41-48 (2004).
 - 7) 栗原一貴: 動画の極限的な高速視聴のためのビデオエンコーダ, エンタテインメントコンピューティング 2011 (2011).
 - 8) 西田健志, 五十嵐健夫: Lock-on-Chat: 複数の話題に分散した会話を促進するチャットシステム, 第 13 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2005), pp.117-120 (2005).
 - 9) 西田健志, 五十嵐健夫: 「あと一步の勇氣」を引き出すコミュニケーションインタフェース, 第 48 回プログラミングシンポジウム報告集, pp.153-160 (2007).
 - 10) 小倉加奈代, 松本遥子, 山内賢幸, 西本一志: Kairos Chat: 主観的時間の概念を導入したチャットシステム, インタラクティブ 2010 論文集, pp.259-266 (2010).
 - 11) 西田健志, 栗原一貴, 後藤真孝: On-Air Forum: リアルタイムコンテンツ視聴中のコミュニケーション支援システム, 第 17 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2009), pp.95-100 (2009).
 - 12) 綾塚祐二, 河口信夫: 参加者が作る会議支援システム: WISS Challenge, Vol.23, No.4, pp.76-81 (2006).
 - 13) ニコニコ生放送: URL: <http://live.nicovideo.jp/> (2011 年 12 月 10 日に参照).
 - 14) ニコニコ学会 : URL: <http://niconicogakkai.jp> (2011 年 12 月 10 日に参照).
 - 15) USTREAM: URL: <http://www.ustream.tv/> (2011 年 12 月 10 日に参照).
 - 16) 青木秀憲, 宮下芳明: ニコニコ動画における映像要約とサビ検出の試み (セッション 2), 情報処理学会研究報告. HCI, ヒューマンコンピュータインタラクション研究会報告, Vol.2008, No.50, pp.37-42 (2008).
 - 17) Cheng, K.-Y., Luo, S.-J., Chen, B.-Y. and Chu, H.-H.: SmartPlayer: User-Centric Video Fast-Forwarding, *ACM CHI 2009 Conference Proceedings* (2009).
 - 18) 谷直紀, 山崎俊彦, 相澤清晴: コメント数の動的な変化を利用した CGM 動画要約 (D-12. パターン認識・メディア理解 A(パターンメディアの認識・理解・生成), 一般セッション), 電子情報通信学会総合大会講演論文集, Vol.2009, No.2, p.114 (2009).
 - 19) 田中聡, 脇本浩司, 神田準史郎: シーン検出による動画情報の自動要約・閲覧技術の開発, 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, Vol.99, No.181, pp.53-58 (1999).
 - 20) 當麻徹, 宮森恒, 富永英義: デジタル動画画像の自動的構造化及び要約作成手法に関する研究, 電子情報通信学会技術研究報告. IE, 画像工学, Vol.95, No.582, pp.9-14 (1996).
 - 21) 磯貝佳輝, 齊藤義仰, 村山優子: 視聴者コメントを用いた動画検索支援のためのダイジェスト動画作成アルゴリズムの検討, 全国大会講演論文集, Vol.2011, No.1, pp.347-349 (2011).
 - 22) 栖閑邦明, 高田 格, 杉山阿葵, 岡田謙一: 短縮再生を利用した二重 TV 視聴支援手法 (セッション 7: テレビ), 情報処理学会研究報告. データベース・システム研究会報告, Vol.2008, No.7, pp.163-170 (2008).
 - 23) PluralMediaPlayer : URL: <http://www2.ocn.ne.jp/xtaczx/home.html> (2011 年 12 月 10 日に参照).
 - 24) MultiWindowMediaPlayer : URL: <http://www1.odn.ne.jp/fripper/> (2011 年 12 月 10 日に参照).
 - 25) WmpSxunl : URL: <http://wmpsxunl.s317.xrea.com/> (2011 年 12 月 10 日に参照).
 - 26) 藤本雄太, 宮下芳明: プレゼンとプレゼンの場をマンガ表現するインタラクティブシステム, 技術報告 (2010).