

複数動画同時視聴による視聴時間の短縮手法 ～複眼～

宇佐美寛^{†1} 濱川礼^{†2}

本論文は、数多くの動画からより短時間で、ユーザが見たいと思う動画を見つけ出す支援を行う手法について提案する。近年、YouTube やニコニコ動画のような動画共有サイトの需要が高まり、多くの動画コンテンツが蓄積されている。これら蓄積された動画コンテンツの中にはユーザの興味・関心が沸くコンテンツが多く埋もれている。そのため、すべての動画を視聴することができれば数多くのすばらしいコンテンツに出会うことが可能である。しかしながら、すべての動画を見ることは不可能である。そこで、本研究では、複数の動画を同時に視聴することで視聴時間の短縮を図る手法と、「複数同時視聴」に「ながら見」（ながら作業）を組み合わせることで視聴時間を短縮する手法について提案する。また、『複眼』を実際に開発し、評価実験を行うことで本手法の有効性を示した。

Reduction Method of Video Viewing Time by Multiple Simultaneous Viewers ～ COMPOUND EYE ~

HIROSHI USAMI^{†1} REI HAMAKAWA^{†2}

In this paper, we propose the technique that we support to find out many videos to think to want to watch many videos in shorter time a user. In recent years, the demand for video-sharing sites like YouTube and NIKONIKODOUGA, a lot of video content is stored. Considered in these stored video content to the content that the user's interests boil and many are buried. Therefore, it must meet a number of great content can be used to watch all videos are possible. However, it all videos see is impossible. In this paper, we propose a method to shorten the viewing time by viewing multiple videos at the same time as a method to shorten the viewing time by "NAGARAMI" to "multiple watching" a combination of (work with). Have shown the effectiveness of the proposed method by actually developing a "compound eye", and also conducted experiments.

1. はじめに

本論文では、複数動画を同時視聴することで多数の動画からより短時間で、ユーザが見たいと思う動画を見つけ出す支援を行う手法について提案する。近年、YouTube やニコニコ動画などの動画共有サイトの需要が高まるにつれて、多くの動画コンテンツが蓄積されている。YouTube では 1 分間に 60 時間分の動画がアップロードされており[1]、ニコニコ動画では総計 830 万件以上アップロードされている[2]。これらの蓄積された動画コンテンツの中にはユーザの興味・関心が沸くコンテンツなのに会う機会がなかった動画コンテンツが多く埋もれている。そこで、動画共有サイトでは「キーワード検索」「カテゴリ検索」「動画の推薦」といった機能や「ユーザ指定のタグ」「評価値」等ユーザ参加型の検索方法を提供することで、よりユーザが興味・関心が沸く動画に出会う機会を増やしている。また近年、ソーシャルネットワークを多くの人が利用しているが、動画がリンクとテキストのみの投稿よりも 12 倍共有されるとの報告がある[3]。これは、ユーザがソーシャルネットワーク上の「知り合い」から多くの動画を薦められる可能性が高いことを示している。

しかし、一人のユーザは多くの投稿・共有された動画が本当に自身に興味・関心を持てるかどうか分からない。そのため、「自分自身で見て」興味・関心を持てる動画であると判断する必要があるが、提供される動画が多いため全てを見る時間を確保するのは従来の動画共有サイトが提供している諸機能だけでは不十分である。そのためユーザ自身の見たい動画をいかに効率よく見つけ出すかは重要な課題となる。

本論文では、多量の動画をより短時間で効率的に視聴できるようにするため、個々の動画の再生倍率を変化させることで、効率的な視聴を可能とし、加えて、それらの動画を複数同時に表示し視聴することで、より効率的な視聴を可能とする手法を提案する。

2. 関連研究

動画の効率の良い視聴方法には大きく分けて二つのアプローチがある。視聴時間を短縮する手法と、複数の動画を同時に視聴する手法である。視聴時間を短縮する手法として、ダイジェスト生成と再生倍率の変更の二種類の手法。

ダイジェスト生成では重要度の低いシーンをカットし、ユーザが見たい場面のみで動画を構成する研究[4][5][6][7][8]などがなされており、視聴時間を大幅に減らすことが可能であるが、シーンカットを行い視聴したい可能性の高い箇所の抽出をするため、話の流れがわからなく

^{†1} 中京大学大学院 情報科学研究科
Graduate School of Information Science, Chukyo University.

^{†2} 中京大学 情報理工学部
School of Information Science and Technology, Chukyo University.

なる問題がある。また、多くの研究は特定のジャンルに特化しているため、ジャンルに応じた最適な選別手法も必要となる。

再生倍率の変更を行う場合であれば、視聴可能な倍率が限られるため、大幅な短縮はできなくなるが、適切に変更した場合内容の把握を行うことができる。この倍率変更を行う研究として、古くは[9][10]、最近では栗原の研究[11]があげられる。この研究は字幕情報を持つ動画において「速読」の技術を適応し、適切に再生倍率を変更することで視聴時間の短縮を実現している。しかし、近年増えてきている多くの動画には必ずしも字幕はついていないため、この効果的な視聴時間の短縮が利用できる範囲が限られてしまう。そのため、今回は処理速度が速く、様々な種類の動画に利用できる動画中の変化量を用いた短縮手法[10]を利用する。

複数の動画を同時に視聴する手法として、例えば水野らの研究[12]があげられる。この研究は、大画面に多数の動画を同時に表示し、ユーザが指定した動画が三次元的に強調表示されるようになっている。複数表示は表示分だけユーザの視聴時間を削減することが可能だが、数が大きくなりすぎた場合、すべてを見る事ができず内容が把握できなくなったり、数十の動画を同時に見る事が可能なユーザが居たとしても、一つ一つの動画の表示サイズが小さくなるため、巨大な画面が必要になるなどの問題点がある。

そこで、本研究では、動画の変化度を用いて再生倍率を変化させることで再生時間を短縮し、さらに複数同時に表示することでより視聴時間を短縮する手法について提案する。また、画面への表示する方法として、単純に複数個並べた場合の「複数動画同時視聴」と複数動画同時視聴を利用し興味・関心を持った動画を1.0倍速で「視聴しながら」他に興味・関心が沸く動画がないか探索する「ながら見」の二種類を提案する。

3. 利用イメージ

本論文の提案する手法の利用場面として想定しているのは、

- YouTubeなどの動画共有サイトで興味・関心が沸く動画の探索の場合(ある程度絞り込んだ場合も含む)
- ソーシャルネットワークのタイムライン上に表示された動画をすべて確認したいと考える場合
- 自分でカメラなどを用いて録画した整理されていない動画から目的の動画を探す場合

など多くの興味・関心が沸くかどうか分からない動画を視聴する場合である。図1の様にユーザは動画共有サイトから興味のある単語・ジャンルから検索を行い、その検索結果を「複数動画同時視聴」を利用し視聴する。動画が終了した際、終了させた際には検索結果の残りの動画が次々と再生されていく。その中から興味・関心の持てる動画を「な

がら見」を利用して、中心に表示する。「ながら見」状態で興味を持っている動画に意識を向けていなくて良い内容の時に、周りに表示されている動画から興味・関心が沸く動画がないかを探し、楽しむ。そして、興味・関心を持った動画の視聴が終わったり、興味・関心が無くなった場合は、「複数動画同時視聴」で再度動画を探していく。ソーシャルネットワークの場合は、タイムライン上に流れてきた動画が「複数動画同時視聴」の画面に表示されていくため、動画共有サイトと同様に利用し興味・関心の沸く動画を楽しく視聴する。録画の場合は、「複数動画同時視聴」の画面にフォルダ内の動画が次々と表示されていくため、動画共有サイトと同様に利用し、思い出の動画を眺めたり、タイトルなどからは判断しづらい動画の探索を行う。

この様に「複数動画同時視聴」と「ながら見」を交互に利用することで、よりよい視聴環境になると考える。

また、本手法では動画の再生倍率を変更するが、その変化度がユーザ毎に異なる可能性が高いため、ユーザ自身に重みをたった一度設定してもらう事で、その差異を吸収している。

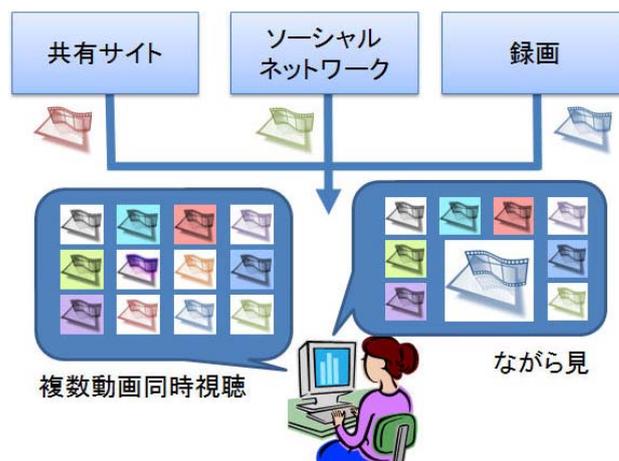


図1 複数動画同時視聴・ながら見の利用イメージ図

4. 再生倍率の変更手法

本手法では各フレームにおいて、変化量を求め、その変化度合いに応じて、そのフレームにおける再生倍率を決定し、時間の短縮を行う。変化量は[10]に示されるように、一部を除き、変化度合いと人が認識出来る再生倍率が反比例することが分かる。そのため、変化量が多い場合には低い倍率を設定し、変化量が少ない場合は高い倍率を設定する。

図2は本手法の流れを示している。まず、動画中の変化量を算出する。算出された変化量に対して平滑化処理を行った後、変化量に応じた倍率の決定を行う。その後、複数動画同時視聴の場合は各動画は各々の変化量に応じた再生倍率に変更しながら再生される。ながら見の場合は視聴したい動画(メイン動画)を1.0倍で再生し、それ以外の動画は再生倍率の変更を行った後、各々の変化量に応じた倍

率に変更しながら再生される。

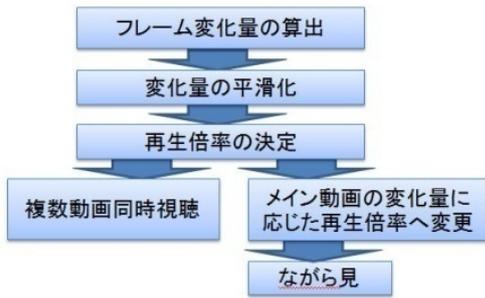


図 2 手法の流れ図

また、動画の音声においては、ユーザが興味を示した（マウスカーソルをあわせた動画）場合に1つ、ながら見の場合はメイン動画が常に再生されるため、興味を示した物と併せて最大2つの動画において音声をを再生し、他の動画では再生しない様にした。

4.1 フレーム変化量の算出

変化量の算出は[10]に基づいて以下の方法で算出する。動画はフレーム単位で処理を行っていく。一つのフレーム画像を分割 ($A \times B$) し、分割を行った一つ一つに対しピクセル単位で求めた色ヒストグラムを合計する。その後、分割したフレーム画像に対し分割 χ^2 検定を適用し、その値を変化量とする。これを全フレームに対して行い、動画全体の変化量を求める。

4.2 変化量の平滑化

動画中の映像は場面の切り替え時を除き、連続している。しかし、算出した変化量にはノイズが混じり、変化量の遷移がなだらかな変化とはならない場合がある。そのため、以下の式を用いて変化量の平滑化を行う。前後 T 秒間のフレーム（最大フレームを $MAXF$ ）の平滑化を行うとし、平滑化前の変化量を OV_i 、平滑化後の変化量を V_i ($i=1, \dots, TF$)。 x は現在のフレーム番号を指している。

$$V_x = \frac{\sum_{\max(0, i-T)}^{\min(MAXF, i+T)} (OV_i)}{\min(T, i) + \min(T, MAXF - i) + 1}$$

この式を利用して平滑化した例を図 3 に示す。この図を見ると変化量の差がかなり大きい箇所が多い事が分かる。人はフレームで画面を見ているわけではないため T 秒間で平滑化することで、人が見て感じる変化に違和感なく近づけることが可能になる。

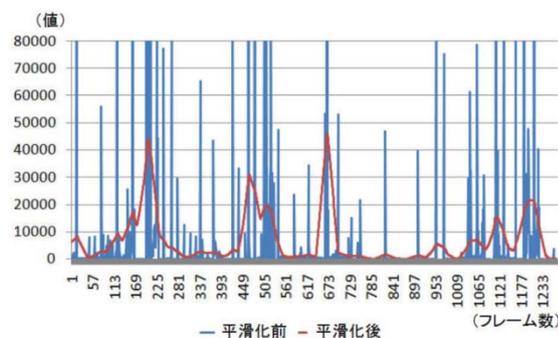


図 3 変化量の平滑化例

4.3 再生倍率の決定

平滑化した変化量に応じた再生倍率の決定を行うため、平滑化した変化量から、最大変化量 $MAXV$ と最小変化量 $MINV$ を抽出する。この変化量に応じて再生倍率を決めるが、全ての変化量に対して割り当てる事は現実的ではない。本手法では N 段階に分割して割り当てを行う。まず $MAXV$ と $MINV$ の差を $N-1$ 段階に分割し値 (NV) を求める。

$$NV = \frac{(MAXV - MINV)}{N - 1}$$

また、再生倍率も同様にシステム上の最大値を $MAXR$ 、最小値を 1.0 、 $N-1$ 分割した値を NR とする。

$$NR = \frac{(MAXR - 1.0)}{N - 1}$$

その後、変化量に応じた各フレームの再生倍率 (OR_x) を割り当て、再生倍率を決定する ($0 \leq i \leq N - 1$)。

$$(NV \times i) \leq OV_x < (NV \times (i + 1)) \quad \text{となる} i \text{ に応じて}$$

$$OR_x = iNR$$

4.4 メイン動画の変化量に応じたサブ動画の倍率変更

本論文で提案する「ながら見」とは、3章の様にユーザ自身が興味・関心を持った動画を視聴しながら他に興味・関心を持つ動画を探ることが前提のため、興味・関心を持った動画（メイン動画）を中央に配置し、他の8つの動画（サブ動画）をその上・左右に表示する。このイメージは図 4 となり、メイン動画の変化量（ただし再生倍率は 1.0 倍に固定）に応じてサブ動画の倍率を個々に変化させる。これは、メイン動画をユーザが視聴する場合、変化度合いが高い場合にメイン動画に注視してしまうため、周りのサブ動画が面白いかどうか判断することは非常に困難になるためである。また、メイン動画をより大きく表示することで、より内容を把握しやすくし、楽しめる状態にする。

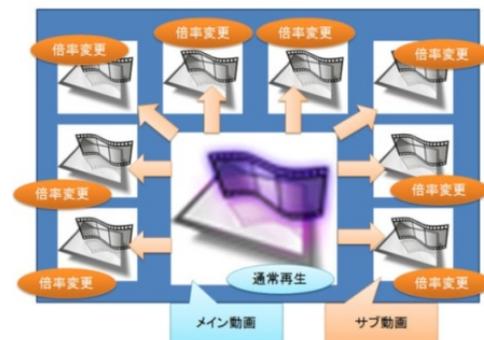


図 4 ながら見時の変化

本手法では、3章の様な利用を想定しており、必ずしもすべての動画が同時に開始されないことを仮定する必要がある。この例を図 5 に示す。



図 5 メイン動画とサブ動画 2 画面の例

そのため、それぞれの動画の現在視聴している時刻を時間的処理開始位置として設定している。メイン動画の再生倍率を SVR 、サブ動画の再生倍率を $OCVR$ 、サブ動画の変化後の再生倍率を CVR 、最大再生倍率を $MAXR$ 、メイン動画の再生位置を x 、サブ動画それぞれの現在の再生位置を $y_i (0 \leq i < 8)$ 、とする。また、この場合において、 CVR は 0 にならないものとして、サブ動画の再生倍率を以下の式を用いて変更する。

$$CVR_{y_i} = \left(1.0 - \left(\frac{SVR_x - 1.0}{MAXR} - 1.0 \right) \right) \times OCVR_{y_i}$$

この式を適応した例を図 6、メイン動画とサブ動画の激しさにに対する対応表を表 1 に示す。

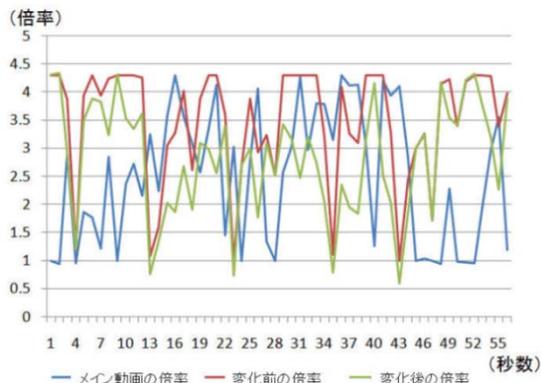


図 6 SVR と CVR の比較グラフ

表 1 影響の表

メイン サブ	変化が激しい	変化が激しくない
変化が激しい	影響が少ない	影響が少ない
変化が激しくない	影響が多い	影響が少ない

表 1 が示すように、メイン動画が激しく変化し、周りのサブ動画を見る余裕がないとき、再生倍率の変化を抑える。その影響例を図 7 に示す。

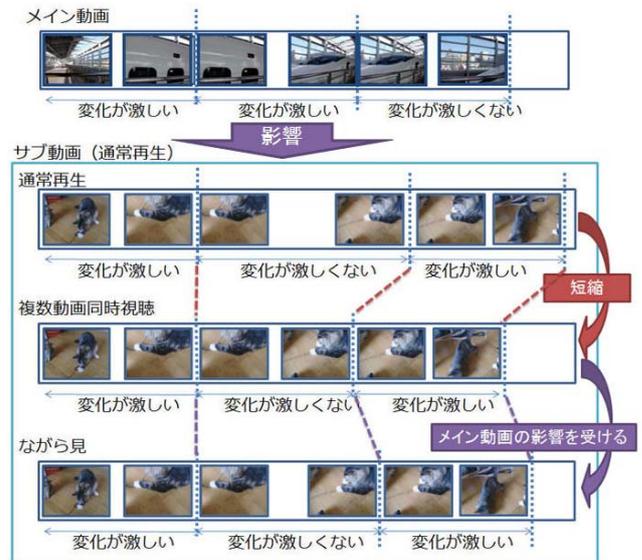


図 7 メイン動画のサブ動画への影響例

5. 実装

5.1 システム構成

今回、本手法の実用性評価のため、『複眼』システムの開発を行った。『複眼』は、ユーザに対し一つの GUI を提供し、「ながら見」(図 8) と「複数動画同時視聴」(図 9) 二つの表示方法を提供し相互にマウス操作で遷移可能である。



図 8 ながら見の画面



図 9 複数動画同時視聴の画面

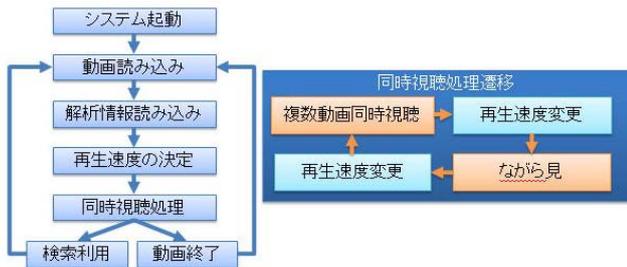


図 10 システム遷移図

『複眼』のシステムと UI の遷移を図 10 に示す。

UI 機能として「一時停止」「シークバー」「動画終了」「お気に入り・後で見るリストへ登録」「リストに登録されている動画のみを再生」「タイトル検索機能」がある。動画を 4 章の手法を用いて再生倍率を変更し表示している。選択している動画は通常の再生倍率で再生する。マウスカーソルで選択した動画は 1.0 倍速で再生され、音声が出る。

5.2 各種パラメータ

『複眼』では、変化量抽出の際に画像サイズにより処理時間が大きく異なるため、入力した画像サイズをディスプレイ上で表示されるサイズと同一である 480×360 に変更して処理を行っている。今回解析に利用した計算機の性能は Core i7 820QM で、解析時間は平均して動画の再生時間の約 7 分の 1 であった。4 章で定義した変数の値は以下のように設定した。A=4, B=4, T=1, N=11

また、今回 MAXV と MINV を決定する際、多くの動画で異常に高い値や低い値が見られたため、ソートを行い、最大値 (MAXV) を 9 割の位置の数値、最小値 (MINV) を 1 割の位置の数値に設定している。また、評価時間が長引くことを控えるためデータの解析(動画の変化量の解析のみ)は事前に行い、データベースへ保存している。この時、保存されるデータはフレーム毎の変化量である。また、動画の再生時間が 1 秒進むごとに処理を行うかどうか判定し、再生倍率の変更を行っている。この時、人により最適な再生倍率が変化すると考えられるため、ユーザが自由に設定し、再生倍率に変化を与える倍率である重みを用いている。

6. 評価

以下 6 種類で比較評価を行った。サムネイルとリンクで構成された Web ページ(検索機能あり)を利用した場合 A, 1 画面で再生倍率が固定(評価者指定)の場合 B と変動した場合 C, 複数画面で再生倍率が通常(1.0 倍)の場合 D, 固定の場合 E, 変動の場合 F. 本論文の手法は F に該当し、変動とは変化量に応じて再生倍率が変化する手法を指す。また、F に関しては「複数動画同時視聴」と「ながら見」を好みで利用してもらった。評価は A~F それぞれ 15 分利用してもらい、学生 8 名に結果を記入してもらった。2811 件の未知の動画の中から A は一覧表示、それ以外はランダムで表示される動画を視聴していく。また、一度視聴した動画は再度テストの中で視聴することはない(A を除く)。視聴中に続きをみたいと思わない動画に関しては途中で視聴を止め、次の動画を見ていただくようにした。図 11 は利用した動画の再生時間とその件数を示している。ジャンルは「コメディ」「ブログ」「スポーツ」など多岐にわたる。

0~1分	1~10分	10~30分	30分~1時間	1時間以上
286件	2102件	357件	27件	39件

図 11 再生時間

各種類における動画の視聴数は図 12 のようになった。

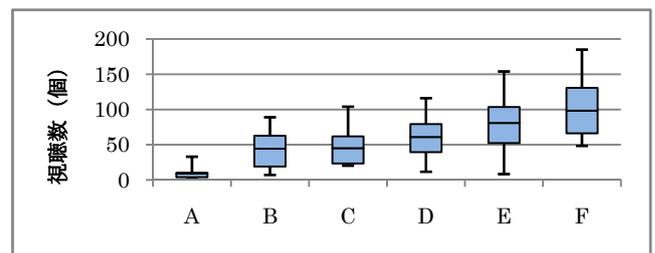


図 12 動画視聴数の評価

視聴数に大きな開きが見られた。これは、一つの動画に対してじっくりと視聴するし判断するか、短時間の視聴で判断するかの違いがあったためと考えられる。しかし、評価者それぞれに見た場合、視聴数の比はほぼ同様であった。ながら見の利用時間が長い場合は視聴数が減る傾向が見られた。

各種類において、「内容が把握できたと感じたか」を 4 段階で回答してもらった。結果は図 13 の様になった。

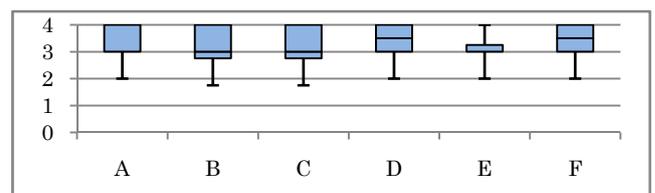


図 13 内容把握の評価

ほとんどの種類において、一部の評価者を除き、内容が把握

握できる事が分かった。しかし、「最適な重みがよく分からない」との意見があるため重みをよりユーザに合う値へ簡易に設定できる必要がある。

お気に入りに登録した動画数と後で見る予定の動画数の平均は図 14 のようになった。

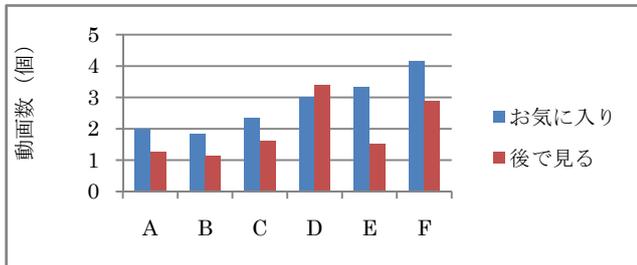


図 14 お気に入りに・後で見る数の評価

今回、評価者が好むかどうか分からない動画ばかりを集めたにもかかわらず、15分間で一定数登録された。無数の動画から検索などを利用すればより興味がわく動画が増えると考えられるので、登録数は増えると考えられる。

6種類において総合評価を行い、最も良い動画を6点、最も悪い動画を1点とした。平均点は図 15 の様になった。

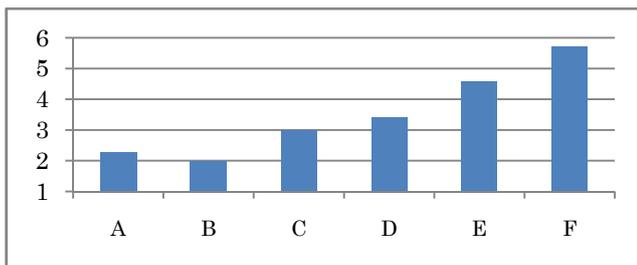


図 15 総合評価

本手法を用いた物が最も良い結果となった。また、便利さに関してを今後も使いたい(4点)から二度と使わない(1点)で回答してもらったところ、平均が3.25点、UIに関して利用しやすい(4点)から利用しづらい(1点)で回答してもらったところ、平均点が2.75点となった。「動画がちょうど良いぐらいに切り替わるのが良いと感じた」「1つの動画を見ながら、次の動画を探せるのはうれしい」などのコメントが寄せられ、通常倍速の複数動画同時視聴より『複眼』がよりユーザの支援を行える事が分かる。しかし、「同時表示数を変えられると良い」「動画をマウスで好きな位置に配置したかった」などの意見があるため、ユーザに最適なUIの設計が必要である。

7. まとめ

本研究では、動画中の色ヒストグラムの変化量を算出し、その変化量から算出した倍率を元に、動画の再生倍率を変更するタイミング(時間)を決定し、再生倍率を変化させることで、効率的な視聴を可能とし、それに複数動画同時視聴とながら見を組み合わせることで多量の動画をより短

時間で効率的に視聴でき、ユーザが興味・関心を持てる動画の探索の支援を行う事ができた。しかし、動画毎の再生倍率の決定に関してよりよくする必要性や、重みをより簡易に・適切に最適な倍率を設定できる仕組みを作ることの重要性、ユーザ毎の最適なUIの必要性が示された。

8. 今後の課題・展望

本研究の複数動画同時視聴による視聴時間の短縮手法は、様々な動画において、実際より短時間で効率よく視聴することができ、ユーザが興味・関心を持てる動画の探索の支援を行う事ができた。しかし、同時に表示する動画の数と対応する最適な再生倍率の設定、及びユーザ毎のより簡易な・適切な倍率を設定できる仕組みが必要であることに加え、ユーザにより見やすい表示方法が異なるため、柔軟なUIを設計する必要がある。これらの課題を解決すると、ユーザが現在の動画検索の代わりとして利用することでより多くの動画を楽しむことが可能となる。

参考文献

- 1) YouTube 統計情報, 2012/10/19.
http://www.youtube.com/t/press_statistics
- 2) ニコニコ動画, 2012/10/19.
http://www.nicovideo.jp/video_top/
- 3) M ブース, メディア計測サービスのシンプリーメジャー社が実施した調査結果
<http://www.blogherald.com/2012/08/21/photos-and-videos-are-best-storytelling-devices-for-social-media-infographic/>
- 4) Brett Partridge Dan R. Olsen and Stephen Lynn. Time warp sports for internet television. Transactions on Computer-Human Interaction (TOCHI), 17(4), 2010.
- 5) Ling-Yu Duan, Jinqiao Wang, Yantao Zheng, Jesse S. Jin, Hanqing Lu, and Changsheng Xu. Segmentation, categorization, and identification of commercial clips from tv streams using multimodal analysis. In Proceedings of the 14th annual ACM international conference on Multimedia, pages 201-210, 2006.
- 6) Martha Larson Stevan Rudinac and Alan Hanjalic. Visual concept-based selection of query expansions for spoken content retrieval. In Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval, 2010.
- 7) Alexander Haubold, Promiti Dutta, and John R. Kender. Evaluation of video browser features and user interaction with vast mm. In Proceeding of the 16th ACM international conference on Multimedia, 2008.
- 8) Anthony Tang, Sebastian Boring, "#EpicPlay: Crowd-sourcing Sports Video Highlights." CHI'12, 2012.
- 9) Rei Hamakawa and Jun Rekimoto, "Object composition and playback models for handling multimedia data", Multimedia Systems, ACM, 1994.
- 10) 伊藤秀和, 濱川礼, "限られた視聴時間内における動画の効果的な時間短縮手法", 電子情報通信学会技術研究報告. HIP, ヒューマン情報処理, 2008.
- 11) 栗原 一貴, "動画の極限的な高速鑑賞のためのシステムの開発と評価", WISS2011.
- 12) 水野慎士, 平野良介, 堤幸彦, "個と群衆が共存するインタラクティブシステム", 情報処理学会 インタラクシオン 2012, 2012.