

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

Development and Evaluation of a System for Watching Videos at Very High Speed

栗原 一貴*

Summary. 本論文では動画を高速に鑑賞する技術について検討する。映画 DVD など一般的な字幕付きの動画を対象として、字幕のない箇所は高速再生し、字幕のある箇所については字幕を読むことが可能のように再生することで、内容を把握しつつ全体として通常よりも短時間による鑑賞を可能にする。さらに高速鑑賞時の負荷軽減のための字幕表示インターフェースとしてセンタリング、フェーディングを実装する。また再生速度、総鑑賞時間の指定により動画を出力でき、モバイル機器などの一般的な動画プレイヤーで再生可能なフォーマットに変換可能な汎用性の高いエンコーダを実装し、評価実験により有効性を示した。

1 はじめに

インターネットの普及により、我々個人の扱う可能性のある情報の量は飛躍的に増加し続けている。例えば資料 [1] によると動画投稿サイト Youtube に 1 分間に投稿される動画数は 600 にのぼり、それは 25 時間分にも及ぶという。

また、我々の扱う可能性のある情報メディアの多様化も著しい。テキスト文書情報に加えて、Podcast などによる音声情報、及び動画共有・中継サービスによる動画情報など多岐に渡ってきており、それぞれの情報量が増加の一途を辿っている。

扱わなければならない情報の総量が増えたとき、一般的にその解決策として二つの戦略が検討可能である。一つは、情報を適切にフィルタリングして適正量に削減する戦略である。もう一つのアプローチは、ユーザの情報受容速度を向上する戦略である。前者については、情報要約、検索、推薦などの分野で研究が今なお盛んであり、その革新的な成果は広く我々の社会に既に浸透している。しかし今後どれだけそれらが高機能化しても最終的にはある意味のある単位にまとまった情報がユーザに提供され、受容されるという構図は変わらないだろう。従って結局はユーザ個人の情報受容速度がボトルネックとなるため、後者の戦略も必要となる。

本論文では後者の、ユーザの情報受容速度向上を支援するアプローチをとる。テキスト情報については「速読」と呼ばれる技術が従来から提唱されてきており [2]、そのような啓蒙書は多数存在する。一方で音声および動画についても、市販のビデオレコーダー、iPod 等の音楽プレイヤーや PC 上の一般のメディアプレイヤーで搭載されているように、再生速度を任意にコントロールすることにより高速な情報受容が可能になる。しかし情報として理解可能な再生速度には上限があり、これらの機器においても通常は 2 倍程度の高速再生が設定上の限度である。

本研究は、動画の高速な情報受容の研究の第一歩

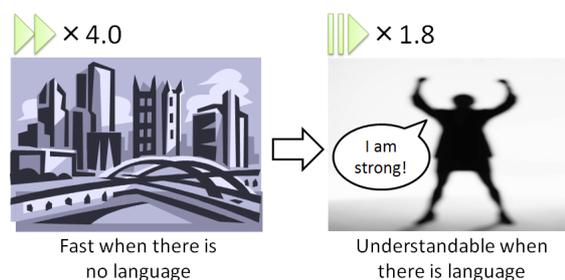


図 1. システム概要

として、映画 DVD など一般的な字幕付き動画を対象とし、画像としての動画の受容速度と音声言語および文字言語の受容速度の違いを考慮に入れた高速な動画鑑賞手法を検討する。具体的には字幕のない箇所は高速再生し、字幕のある箇所については字幕を読むことが可能のように再生することで、単純な高速再生と比べて内容を把握しつつ全体として通常よりも短時間による鑑賞を可能にする (図 1)。

本論文では、まず関連研究との関係性を論じる。その後、動画の高速鑑賞方法の検討を行い、プロトタイプの実装について詳細に記述する。さらに、評価実験について報告する。

2 関連研究

本研究は動画メディアの高速鑑賞を検討するものである。メディアの高速再生に関しては、一般的なメディアプレイヤーで既に再生速度の調整は広く実装されている。Foulke らは、音高を変えない音声の高速再生が理解の上で有効であることを示した [3]。Vemuri らは音声情報の高速再生時に、その音声の音声認識結果のテキスト情報を提示することによるユーザの情報処理速度の向上の試みを検討した [4]。また青木ら [5] は聴覚のみを用いた音楽検索インターフェースとして高速再生を導入している。我々は音声だけではなく映像も同時に鑑賞するという、マル

チモーダルな高速鑑賞に関する諸問題を検討する。特に高速再生時の複数モダリティ間の理解可能速度の違いの検討が重要である。

別のアプローチでの高速化の事例として、並列化を扱ったものが挙げられる。音声・音楽情報の検索・選択の際、三次元的な音像提示を行ない一度に複数の音源を扱えるような手法はこれまで提案されてきているが、Vazquez-Alvarez らはその認知的負荷とインタフェース設計指針について議論している [6]。また、動画の鑑賞の効率化のために同時に複数の動画を再生したいというニーズがあり、そのようなソフトウェアが公開されている¹。

我々は動画鑑賞の並列化は一般のユーザへの負担が大きいと考える。単一のストリームの動画メディアの高速鑑賞のみを取り扱う。一方で我々の高速動画鑑賞方式用いた上で複数の動画メディアを並列的に提示することも技術的には容易であるため、将来的には並列化のアプローチとの統合も可能である。

映像の鑑賞を扱った研究として、清山らはテレビなどのリアルタイム放送の鑑賞時に高齢者等にも理解が容易になるよう、音声再生速度を通常より下げて再生する機能を開発している [7]。Cheng らは映像の変化率に合わせて再生速度を自動調整するアプローチを提案しているが、消音状態でコンテンツを扱っており、音声モダリティについての検討がなされていない [8]。青木ら [9] はニコニコ動画における閲覧者コメントの盛り上がりをもとに動画メディアの重要度を推定し、これをもとに要約を行なう手法を提案している。

3 高速鑑賞方式

本章では動画の高速な鑑賞を可能にする方式を総合的に検討する。

3.1 高速鑑賞方式の分類と検討

映画 DVD コンテンツを念頭においた場合、その構成要素として、「主映像」、「言語情報 1:字幕映像」、「言語情報 2:字幕映像に対応する音声」、「字幕映像に対応しない音情報」、の 4 種に分解可能である。主映像とは、視覚情報として提示されるもののうち、字幕以外の「地の映像」のことである。字幕映像とは、主映像に重畳して表示される字幕の映像のことである。これは言語情報のひとつである。字幕映像に対応する音声とは、すべての音情報のうち、字幕に対応付けられている「声」の情報である。これも言語情報のひとつである。それ以外の、BGM や効果音等は字幕映像に対応しない音情報となる。

映像の高速鑑賞方法として、主映像と言語情報の同期（言語情報の提示開始と終了のタイミングが主映像と一致していること）を保存するかどうか、および言語情報を選別し完全に削除することを許すかどうかによりいくつかのアプローチが考えられる。

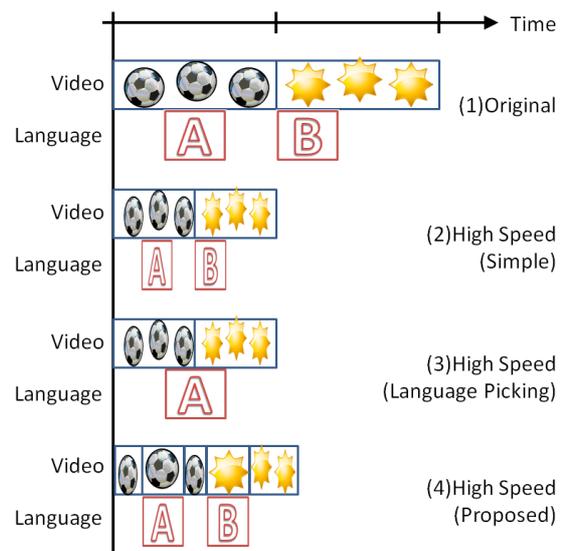


図 2. 高速鑑賞方式の模式図

図 2 は、考えられるアプローチを時系列の模式図で可視化したものである。(1) は通常の再生を表している。(2) の単純高速再生は単純に再生速度を増加させる鑑賞方式である。この際、音声の周波数は変化しないように調整される場合もある。主映像と言語情報は常に同期している。近年のビデオレコーダーには 1.3 倍もしくは 2 倍程度の単純高速再生が実装されている。さらなる高速化を本方式で図る場合、音声聞き取れなかったり対応する字幕が読み取りきれなくなる可能性がある。

(3) は [7] などで実装されている機能で応用可能な高速鑑賞方式である。主映像の高速再生中、言語情報の一部もしくは全部を切りだして、速度を変換し提示する。主映像の速度は固定であるため、言語情報を理解可能な速度にするために主映像との同期を諦めている。情報を選別し完全に削除することを許容せず、言語情報のすべてを理解可能な速度にして主映像に重ねると、両者の間に深刻なズレが生じる可能性がある。

我々が本論文で提案する (4) 「変速再生」は、主映像と言語情報の同期を保存し、かつすべての言語情報を保存するアプローチを取る。すなわち、主映像のみの箇所については理解が可能な限界の速度で再生し、言語情報の提示箇所ではその言語情報が理解可能な限界の速度で再生する、というアプローチである。映像の再生速度を構成要素別に制御することで、単純高速再生を超える速度での高速鑑賞を可能とする。また情報を選別し完全に削除することがなく、主映像と言語情報は同期がとれている。

本論文では字幕映像、およびそれに対応する音声という 2 種の利用可能な言語情報のうち、学習を前提とすればより高速な理解が可能となる文字情報モダリティである字幕映像に特に注目する。字幕映像は一般に流通している映画 DVD などで広く利用可能であり、DVD の規格からその表示タイミング情

¹ PluralMediaPlayer, MultiWindowMediaPlayer, Wmp-Sxunl など

報および発話された内容のテキスト情報が容易に抽出できる利点がある。

3.2 変速再生の定量的検討

本節では変速再生方式における再生速度や鑑賞に必要な時間などを定式化し検討する。対象の映像の自然速度における鑑賞時間を $L[\text{sec}]$ とする。主映像のみの箇所は S_m 倍速、字幕の付与された箇所は S_s 倍速で再生されることとする。映像中の主映像のみ箇所と字幕映像箇所の時間比を $r : 1 - r$ とすると、対象映像の鑑賞に必要な総時間 $L_Q[\text{sec}]$ は、

$$L_Q = \frac{rL}{S_m} + \frac{(1-r)L}{S_s} \quad (1)$$

と表せる。比 L_Q/L を圧縮率と定義すると、全体としてどの程度時間圧縮が実現したかがわかる。

上式でわかるように、本再生方式は S_m と S_s がユーザ依存パラメータだが、 L_Q への寄与が独立している点の特徴である。適切なインタフェースを備えることで、 S_m と S_s それぞれについてユーザが高速化に向けた学習を行うことが実現できると考えられる。

式 1 における主映像のみ・字幕映像比率 r は対象映像に依存するパラメータである。実際の映画から r の分布を見積もった。対象は日本語字幕のある映画 DVD で、以下の 15 本である。アバター、クライマーズハイ、デスノート 1、カイジ、南極料理人、ペイフォワード、流星の絆 1、三文オペラ、SAW、SP 野望篇、シンレッドライン、月の輝く夜に、パンデジポイント、夕凧の街桜の国、ドラえもん・のび太の結婚前夜である。新旧さまざまなジャンルのものを選んだ。

その結果、 r の平均値は 0.578 であり、標準偏差は 0.118 であった。ここから、およそ 5~7 割程度は主映像のみの箇所であることがわかる。すなわち、調査対象の映画に限定すれば S_s よりも S_m の方が圧縮率への寄与率が平均的には大きいことがわかる。

4 インタフェース

4.1 鑑賞方法指定インタフェース

ここでは変速再生を前提とした動画の鑑賞方法の指定方法について検討する。提案システムでは (1) S_m と S_s の直接指定、(2) 総鑑賞時間と S_s による指定、の 2 種類のインタフェースを開発した。

最も単純なのは (1) の S_m と S_s について、通常速度を 1.0 とした時の倍率として直接的に指定するものである。これは各ユーザが自分に最適な速度をそれまでの使用経験から知っている場合に有効である。

(2) は「どうしてもこの映画を 30 分で見たい」のように動画の鑑賞に必要な時間をもととの動画の総時間 L とは無関係にユーザが指定するインタフェースである。任意の時間を分単位で入力すると、式 1 から字幕情報速度をなるべく変えないよう主映像速度を変化させる。本インタフェースは日常生活におけるユーザと情報との新たな接点を模索した

CastOven[10] に適用することもでき、システムの機能拡張が可能である。

4.2 字幕表示インタフェース

我々は更に字幕付き動画の高速鑑賞時の負荷軽減のため、以下の 2 つのインタフェースを開発した。

センタリング: 高速で字幕付き動画を鑑賞する際に問題となると考えられるのは、字幕映像と主映像との間の視線移動である。通常速度の鑑賞であれば字幕映像の表示時間は十分あるため問題とならないが、高速鑑賞時は頻繁な視線移動により疲労が蓄積するか、もしくは字幕のみに注目してしまい主映像を快適に鑑賞できなくなる可能性がある。

本研究では栗原ら [11] の提唱した「聴衆の視線移動の局所化」の概念に基づき、冗長な視線移動を最小化する。通常、主映像は画面中央部を中心としてコンテンツが構成される。一方で字幕はそのような主映像の邪魔にならないよう、画面周辺部、主に画面下部に表示される。センタリングは、字幕を画面中央部に表示することで主映像と字幕の中心を一致させる手法である。鑑賞者が画面中央部を常に注視するだけで鑑賞が可能になる効果があると予想される。一方で本手法は主映像コンテンツを結果的に「邪魔する」状態になるため、鑑賞の快適さに与える影響について評価する必要がある。

フェーディング: フェーディングは、ある字幕の表示が終わった際、次の字幕が表示されるまでの区間、直前に表示していた字幕を半透明化し継続表示する手法である (図 3)。これにより各字幕の表示時間が延長されるため、高速鑑賞が快適になると考えられる。また半透明化することで既にその言語情報の提示が本来終了していることが可視化されるため、主映像との同期が主観的には保たれると考えられる。延長される表示時間は S_m と S_s との比の組み合わせにより異なるが、両者が一致している場合、3.2 節で挙げた 15 映画における総数 10913 の字幕の平均で 2.62 倍の延長を実現できた。

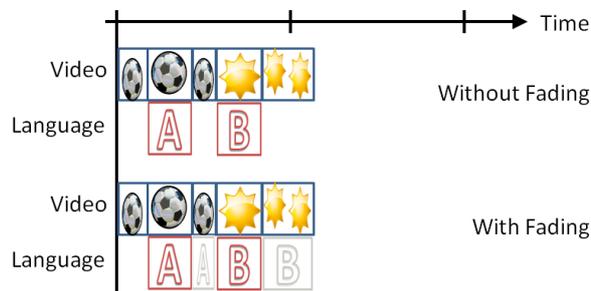


図 3. フェーディング

5 システムの実装

我々は変速再生を実現するプロトタイプとしてビデオエンコーダを実装した。当初はリアルタイムに再生速度を変更する動画プレーヤーとしての実装を試みたが、2 倍を超える任意の S_m , S_s をリアルタ

イム制御しながら映像と音声を安定して出力することが難しかったため断念した。一方で実装した動画エンコーダを用いれば任意の S_m , S_s の組み合わせの動画ファイルを瞬時に出力できるため、あらゆる再生倍率の動画ファイルを事前準備し切り替えることでリアルタイム倍率制御と同等のことが通常の動画プレーヤーで実現可能である。

実装した動画エンコーダを使用する前処理として、字幕付き動画を用意する必要がある。DVD をソースとして用いる場合、DVD Fab²などの GUI による DVD リッパーのオプション設定により簡便に行なうことが可能である。CUI では VSRip³を用いて行なうことができる。前処理の結果、動画ファイルと一般的なフォーマットの字幕ファイル (.idx および .sub) が生成される。

前処理で得られたファイルを入力として変速再生動画を出力するのが実装したエンコーダであり、C#を用いて実装した。これは GUI からの S_m および S_s の直接指定、もしくは総鑑賞時間の指定に基づき、動画ファイルを適切に伸縮・結合した AviSynth スクリプト⁴として生成するものである。映像および音声の伸縮には AviSynth の組み込み関数である TimeStretch 関数、AssumeFPS 関数、ChangeFPS 関数を用いて、音高が変化しないようにした。字幕情報は Layer 関数を用いて画像として重畳した。AviSynth スクリプトはテキストファイルでありながら、直接 Windows Media Player で動画として再生できる。さらに ffmpeg⁵を用いて MP4 や WMV などの一般的な動画フォーマットに変換し、スマートフォンなどの広範なデバイスで再生可能とする機能を内蔵した。

6 評価実験

本節では提案システムを用いた評価実験について報告する。実験には視聴覚が健全な 10 名 (30 代男性 2 名, 20 代男性 7 名, 20 代女性 1 名) が実験協力者として参加した。

6.1 方法

全ての実験協力者が (1) 音声言語受容速度測定、(2) 字幕映像受容速度測定、(3) 主映像受容速度測定、(4) 総合的映像鑑賞、の 4 つのタスクを遂行した。タスクの実行順は (4) が最後に固定であるが、(1),(2),(3) については順序効果の相殺のため実験協力者ごとにランダム化した。環境としてノートパソコン Sony VAIO VPCF1 (スクリーンサイズは縦 20.0cm, 横 36.4cm) を机の上に置き、椅子に実験協力者が座り通常の PC 操作時と同様の距離を液晶画面からとった。音声についてはヘッドフォンを用い、音量を各実験協力者が適切に設定した。

それぞれのタスクにおいて液晶画面に Windows Media Player をフルスクリーン表示した。そこで様々に速度等を変更した上でエンコード (映像:mp4 VGA 8bit AVC/h264 23.98fps 1200kb/sec, 音声: AAC 48kHz 2.0ch 128kb/sec) したアニメ映画「ドラえもん・のび太の結婚前夜」の冒頭 7 分間分 (通常速度再生時) を再生した。

(1) では字幕のない映像について主映像と音声を鑑賞させ、言語として音声内容の理解が可能な上限の速度を求めた。速度の変化範囲は 1.0 倍から 10.0 倍まで 0.5 倍刻みで、音高の変化のない単純高速再生である。

(2) では字幕のついた映像について字幕のみに注目させ、言語として字幕内容の理解が可能な上限の速度を求めた。ここでも速度の変化範囲は 1.0 倍から 10.0 倍まで 0.5 倍刻み、音高の変化のない単純高速再生である。

(3) では音声をミュートし字幕のない映像を鑑賞させ、主映像が映像として理解可能であり流れを追うことができる上限の速度を求めた。この場合発話される言語の内容は判別できないため、発話内容の理解については行わなくてもよいように教示を行なった。ここでも速度の変化範囲は 1.0 倍から 10.0 倍まで 0.5 倍刻みの単純高速再生である。

(4) では (2) および (3) で求めた主映像と字幕の理解可能速度の上限を組み合わせた字幕付き映像を鑑賞させて、鑑賞の快適さを主観評価させた。たとえば (2) で 2.5 倍が上限と答え、(3) で 6 倍が上限と答えた場合は、 S_m が 6 倍で S_s が 2.5 倍であるような音高の変化のない変速再生映像を鑑賞する。この際、字幕の表示位置について、オリジナルの字幕と同様に画面の下部に表示する O 条件、センタリングを行なう C 条件、センタリングかつフェーディングを行なう CF 条件の 3 条件を試行した。主観評価は 1 をとても不快、3 を普通、5 をとても快適とする 1 から 5 までの実数値の申告により行なった。また本タスクでは一部 (2 名) の実験協力者について、画面上での視線の注視点座標を非装着型視線計測器 Tobii X60 で計測して 60Hz で記録した。

4 つのタスクの終了後、実験協力者に自由回答で感想を得た。

6.2 結果

タスク (1),(2),(3) から、音声情報、字幕情報、主映像の速度変化に対する実験協力者の鑑賞可能人数比率がプロットできる (図 4)。図中の実線は、2 パラメータのロジスティック関数 $y = 1/(1 + \exp(a(x - b)))$ を最小二乗法で当てはめたものである。なお、主映像速度については、今回の最大速度である 10 倍でも 10 名中 7 名が限界と感じていない。10 倍以上の映像は試験しなかったため、真の限界は不明なままである。

タスク (4) から、限界的な速度の変速再生における映像鑑賞時の快適さが得られる。図 5 は 3 種類の字幕表示条件の快適さの平均と ± 標準偏差を示したものである。O, C, CF 条件の平均値はそれぞれ

² <http://www.dvdfab.jp/>

³ <http://sourceforge.net/projects/guliverkli/files/VSRip/>

⁴ http://avisynth.org/mediawiki/Main_Page

⁵ <http://ffmpeg.org/>

1.95, 3.43, 3.75であった。一元配置分散分析を行った結果、3種の条件の間には有意な差が見られた ($F(2, 27) = 19.46, p < .001$)。Bonferroni法による事後検定を行ったところ、O条件とC条件の間、およびO条件とCF条件の間にそれぞれ有意な差が得られた (いずれも $p < .001$)。

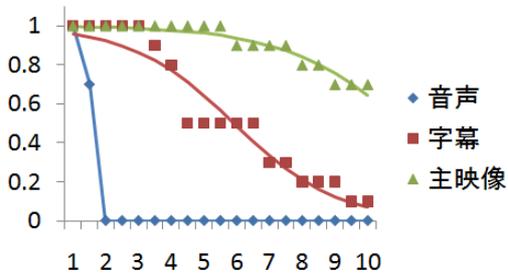


図 4. 鑑賞の速度倍率 (横軸) と実験協力者の鑑賞可能人数比率 (縦軸) の関係

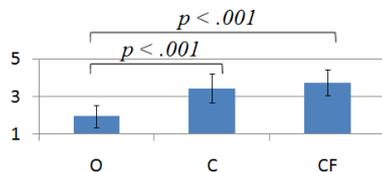


図 5. 3つの字幕表示条件における快適さの平均

6.3 考察

まず図 4 について、音声情報、字幕情報、主映像の順に鑑賞可能性の速度耐性が高くなることが確認された。音声情報鑑賞可能速度は特に個人差が小さく、次いで字幕情報、主映像の順で個人差が大きくなっていくことが確認された。パラメータ b (鑑賞可能人数比率が 0.5 になる速度) を鑑賞可能速度の代表値として捉えると、それぞれ音声情報で 1.55 倍、字幕情報で 5.91 倍、主映像で 11.1 倍であり、音声言語の耳からの理解を諦めることによって式 1 により変速再生方式は鑑賞時間の 85.5% の削減を実現できる可能性があるという著しい結果を得た。

しかし、今回得られた結果は実験に使用した映像素材に依るところが大きい。まず、使用した映像素材は短時間のものであるため、極限的な高速鑑賞時に疲労が蓄積されるほどの分量ではなかった⁶。また全体的にシーンの切り替え速度やカメラ視野の移動速度が小さく、言語も比較的明瞭であり高速にならないような配慮がされている可能性のあるアニメ映像である。したがって今回の実験で得られたのは提案システムを用いて実現可能な動画の高速鑑賞の上限に近い結果が得られるケースであったと考えられる。より長時間の鑑賞時、およびアクション映画やサスペンス映画など、より激しい映像変化や速い会話のやりとりが多いコンテンツでは、今回と異なる結果が得られると考えられる。我々のインフォーマルな調査ではそのような会話の速いコンテンツであっても音声言語情報よりも字幕言語情報の方が高

速に鑑賞できるという関係は保存されているため、提案手法は有効であると予想している。また [8] と組み合わせ、映像の時間当たりの変化率を定量的に求め主映像速度に反映させることにより、映像変化の激しさを補正した見やすい速度での鑑賞も実現可能であると考えられる。これは今後の課題である。

次に図 5 に示した 3 種類の字幕表示条件での快適さについて考察する。C および CF 条件において、鑑賞の快適さは 3 (普通) を上回っており、映像の変速再生とセンタリングの組み合わせが高速鑑賞方式として有効であることが示唆された。O 条件よりも他の 2 条件が有意に快適であるという結果が得られた点について、実験協力者のコメントから、限界に近い速度での動画鑑賞においては O 条件では字幕・主映像間の視線の移動が不快さにつながっている可能性が示唆された。一方字幕が画面中央に表示される C および CF 条件では、基本的に視線は画面中央付近に向けられるため、そのような不快さは軽減されていると考えられる。その確認のため一人の実験協力者について、タスク (4) における O 条件と CF 条件での画面上の視線の注視点座標分布を、 y 座標 (鉛直方向) のみに限定した注視時間の相対頻度のヒストグラムとしてプロットした (図 6)。確かに CF 条件では注視点は画面中央 ($y = 600$) 付近に集中しており、O 条件では字幕表示位置である画面下部 ($y = 350$) と画面中央 ($y = 600$) 付近に注視点が分散していた。これは CF 条件により「聴衆の視線の局所化 [11]」が達成され有効に機能していることを示唆するものである。

また、C 条件と CF 条件で有意な差が得られなかった点については、CF 条件がフェーディングにより 1 字幕あたりの表示時間の増加に寄与するものの、矢継ぎ早な会話のやり取りのある箇所では字幕と字幕の間の沈黙区間が短く、フェーディング時間がほぼゼロに近くなってしまいう問題、会話のない長い区間に直前の字幕がずっとフェーディング表示されてしまい主映像にそぐわなくなる問題などが実験協力者のコメントから示唆された。前者については原理的に対処不可能であるが、後者についてはフェーディング表示時間の上限を設定し字幕を消去することで対応が可能であると考えられる。

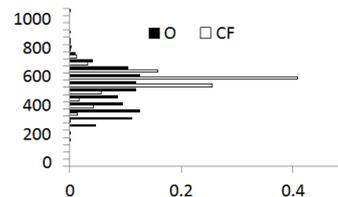


図 6. O 条件と CF 条件における画面上の注視点分布ヒストグラム (横軸は注視時間の相対頻度、縦軸は y 座標 [pixel])

7 まとめ

本論文では動画の高速な鑑賞方法として主映像と言語情報の再生速度を独立して制御する変速再生方

⁶ ただし本映像作品を最後まで視聴したとしても $S_m = 11.1$, $S_s = 5.91$ のときわずか 3 分 44 秒である。

式を提案し、字幕付き DVD を対象としたビデオエンコーダとして実装した。評価実験により字幕のセンタリングが高速鑑賞に有効であることが示され、また鑑賞時間を平均 85.5%削減できる可能性のある事例が示された。今後はよりフォーマルなユーザスタディを行い、読書速度と字幕読み速度の相関、映像のジャンルに依存するユーザの高速鑑賞の限界速度の調査等を行いたい。また、音響情報処理や音声認識技術により、字幕のない任意の動画の高速鑑賞を可能にすることも検討したい。

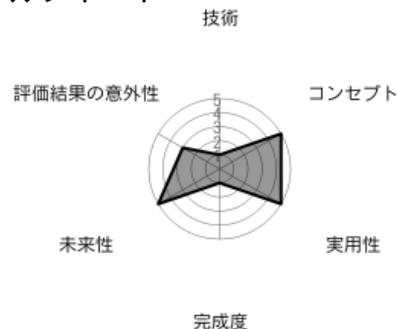
謝辞

本研究は科研費 (23700155) の助成を受けた。

参考文献

- [1] NFOGRAPHIC: What Happens Online Every 60 Seconds. <http://www.scribball.com/2011/06/infographic-what-happens-online-every-60-seconds/>.
- [2] 速読を始めよう。 <http://www.ponp.jp/info/speed.html>.
- [3] Foulke, W., and Sticht, T.G. "Review of research on the intelligibility and comprehension of accelerated speech," *Psychological Bulletin*, 72, pp.50-62, 1969.
- [4] Vemuri et al. "Improving speech playback using time-compression and speech recognition," In *Proceedings of CHI'04*, pp.295-302, 2004.
- [5] 青木秀憲, 宮下芳明. 視覚を用いない状況下での高速楽曲探索インタフェースの設計と検証, *情報処理学会論文誌*, Vol.51, No.2, pp.356-364, 2010.
- [6] Vazquez Alvarez Yolanda, and Brewster A. Stephen. "Designing spatial audio interfaces to support multiple audio streams," In *Proceedings of MobileHCI '10*, pp.253-256, 2010.
- [7] 清山 他. 高品質リアルタイム話速変換システムの開発, *電子情報通信学会誌*, Vol.J84-D-II, No.6, pp.918-926, 2001.
- [8] Cheng et al. "SmartPlayer: User-Centric Video Fast-Forwarding," In *Proceedings of CHI'09*, pp. 789-798, 2009.
- [9] 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, *情報処理学会研究報告 2008-HCI-128/2008-MUS-75*, Vol.2008, No.50, pp.37-42, 2008.
- [10] Watanabe et al., "CastOven: a microwave oven with just-in-time video clips," In *Proceedings of Ubicomp '10*, pp.385-386, 2010.
- [11] Kurihara et al., "Toward localizing audiences' gaze using a multi-touch electronic whiteboard with sPieMenu," In *Proceedings of IUI'11*, pp.379-382, 2011.

アピールチャート



未来ビジョン

本研究の目指す未来は、ネット上に爆発的に増加し続ける各種情報に対し「情報受容の納得できる形での効率化」を実現することである。現代における情報検索技術、推薦技術、自動要約技術の恩恵は筆舌に尽くしがたい。一方でこれらの技術が提供するものは、基本的には機械的もしくは集合知的なアルゴリズムに基づき情報を選別し、選ばれなかったものを非表示化した結果である。すなわち、「捨てる」判断が他者任せになっている。これらの技術なしではもはや我々は生きていけないが、せめてこれらを経て最終的に自分の目の前に与えられた情報の受容は納得した形で行いたいと願うのは筆者だけではないはずだ。提案した高速鑑賞方式は結局は映像の間引きであり、情報を捨てることには変わらないが、鑑賞者の主観的な「目を通した」という納得に対する影響が小さいように思われる。自分はこの情報を受容した、と自信を持って公言できるためには、本研究のように「とにかく全ての情報を一度脳内に通す」というプロセス、もしくはそう思わせてくれるプロセスが必要なのではないだろうか。なお筆者のこのようなビジョンは、毎

年主催している ACM SIGCHI 勉強会 (<http://qwik.jp/userstudy/88.html>) にも根付いており、毎年多数の賛同者を得て開催されている。

未来ビジョン謝辞 私は本研究を進めるにあたり、果たしてこの技術が映画を守るのか、壊すのかをずっと自問自答し続けた。提案手法はコンテンツ過多の時代において確実により多くの良い映画に触れる機会をもたらすものであり、良い映画を探し得たら場合によっては改めて通常速度で鑑賞する選択肢もある。一方で本技術により本来の良さが歪曲して解釈されてしまう映画も必ず出てくるだろう。そのような中、「ドラえもん・のび太の結婚前夜」という短編アニメを主映像 5.5 倍、字幕部 1.7 倍で鑑賞した。技術的な問題点を洗い出している最中であり、また昔見たことのあるコンテンツであったために内容には全く興味を持たず始めた鑑賞であったが、気がつく私は物語に引きこまれており、時にドキドキし、ハラハラし、ラストシーンではいつのまにか涙を流していた。私は確信した。極限的な高速鑑賞でも物語の感動は保存しうる。この場を借りてドラえもん和故藤子不二雄 F 氏に感謝したい。