

リッチメディア解析と配信技術

Rich media analysis and delivery technology

要 旨

【キーワード】

リッチメディア、動画要約、動画マーケティング、適応的動画配信、MPEG-DASH、サムネイル画像、高画質ズーム

リッチメディアの利用が普及し、単に動画を視聴するというだけの機能提供では、価値が十分ではなくなっている。我々は、新たな価値提供を目指し、多くの動画がある中で特定の動画を注目させるためのサムネイル画像を自動で選ぶ「自動選定技術」と、動画視聴に伴うデータトラフィックの増加とその影響で、円滑に動画視聴できなくなる問題に対応する「適応的動画配信技術」を提案する。この技術は、従来人手により行っていたサムネイル画像の選定を自動化し、個々人の嗜好に合わせて動画内容への注意を喚起することを可能にする。また、通信インフラ利用のコスト増とネットワーク逼迫による動画再生停止などの体感品質の劣化を同時に抑制し、いつでもどこでも快適に動画視聴を行いたいという要望に応える。

Abstract

【Keywords】

rich media, video summarization, video marketing, adaptive video streaming, MPEG-DASH, thumbnail images, high-quality zooming.

As the use of rich media gains in popularity, simply making video content available for viewing online is no longer sufficient in maintaining a competitive edge in the market. To further enhance the effectiveness of video content, Fuji Xerox proposes two supplementary features. The first is an automated thumbnail selection technology that extracts relevant thumbnail images from videos to draw users' attention to specific video content. This technology automates the thumbnail selection process by extracting images that are determined to be appealing to users based on their individual preferences, thus eliminating the need to manually select images. The second feature is an adaptive video streaming technology that alleviates the negative effects of network congestion on the user viewing experience by suppressing increases in data traffic caused by video playback. This feature also reduces communication costs while maintaining continuous video playback, making it possible to provide a stressless viewing experience to users, irrespective of their location or time of viewing.

執筆者

内橋真吾 (Shingo Uchihashi)
藤居 徹 (Akira Fujii)
ムラリ・スレシュ (Suresh Murali)

研究技術開発本部 インキュベーションセンター
(Incubation Center, Research and Technology Group)

1. はじめに

動画に代表されるリッチメディアは、個人向けエンターテイメント用途にとどまらず、企業内の社員教育や技術伝承、お客様に対する広告宣伝や商品の詳細説明、動画マーケティングなどさまざまな場面で利用が広がっている。背景として、インターネットの高速化や、スマートフォンなどのモバイル機器の普及により、いつでもどこでも動画を視聴できる環境が整ってきたことが大きい⁹⁾。リッチメディアの利用拡大に伴い、企業が動画共有サイトに登録する動画数も近年大きく増加している。

増え続ける動画に対し、見たい箇所を短時間で探し出せるようにする研究が盛んに行われてきた⁴⁾。こうした研究事例の多くは、利用者が知りたい情報についてキーワードを入力すると、関連する動画をデータベースより検索して提示することを想定している。動画とキーワードを自動で対応づけるために、動画中のナレーションの音声認識、映像中の文字領域の検出、映像中の人物の顔認識やオブジェクト、状況の自動認識などさまざまな技術が提案されてきた¹⁾。

その一方で、動画を利用するために必要な技術は検索にかぎらないことがわかってきた。たとえば、広告宣伝のための動画は、検索されるのを待つのではなく、興味を喚起する魅力的なサムネイル画像を人手で選び提示し、視聴を誘導している。コンテンツ自体も、ターゲットに合わせて人手で作成されている。このような人手の作業はコストがかかり、対応できる数も限られる。モバイル機器で動画を視聴する場面では、不安定なモバイル通信回線のため動画再生が停止したり、待ち時間が発生したりして視聴体験の品質が確保できていない。また、膨大なデータ量に起因する通信コストの負担も無視できない。

本稿では、従来の動画検索のほか、新たに必要性が見えてきたリッチメディアの解析と配信技術について紹介する。

2. リッチメディア解析と配信

2.1 富士ゼロックスのこれまでの取り組み

富士ゼロックスでは、リッチメディアの持つ伝達力や説得力を活用して人と人とのコミュニケーションを支援する研究技術開発に取り組んできた。利用者からの要望の全体像を図1に示す。社員教育と情報共有を主な用途とした動画管理システムMediaDEPOを開発した¹⁵⁾。ユーザーが動画の内容を簡単に把握し、見たいところを早く見つける支援をする技術を利用したユーザーインターフェイスを提案した^{3),5),14)}。一部はMediaDEPOの機能として組み込まれている。また、講義動画を解析しスライド中の文字を検索できるようにする技術¹⁰⁾や、少ないサンプル画像から効率よく共通項を見つけ出して学習し自動でラベルづけする技術¹⁷⁾を提案し、テキスト入力による動画や静止画の検索を可能としてきた。

コンテンツ作成支援として、講義やプレゼンテーションを簡単に記録し配信用コンテンツを作成するツール¹⁹⁾や、専門の動画作成者でないユーザーによる撮影編集を支援するツール^{11),16)}を提案した。さらに、会議室や建物内の活動を撮影し動画コンテンツとしてあとから参照できるシステム^{12),18)}を構築した。動画コンテンツを紙媒体から参照できる技術も提案した¹³⁾。

2.2 動画視聴者の体験向上を支援する技術

動画の利用を包括的に支援するソリューションを提供するために、これまでの取り組みに加えて、動画視聴者の体験の質を高める技術を提案する。具体的には、視聴者の嗜好に応じて代表画像を選定する技術と、モバイル機器における動画視聴の快適性と通信データ抑制を両立する技術に新たに取り組んだ。

個人の嗜好に合わせて動画内容を提示する技術により、従来は人手により行っていたサムネイル画像の選定を自動化し、個々人の嗜好に合わせて動画内容への注意を喚起する。これに

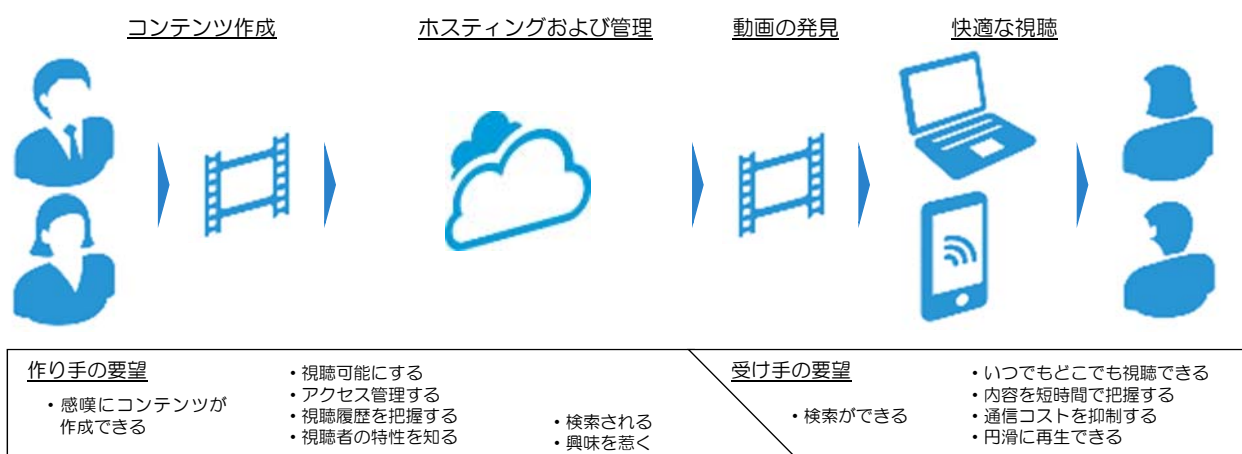


図1 リッチメディア解析および配信技術のスコープ

より、視聴者は興味を惹く動画を発見しやすくなる。

また、適応的動画配信技術により、増大するデータトラフィックに伴い発生する通信インフラ利用費の増加とネットワーク逼迫による動画再生停止などの体感品質の劣化を同時に抑制し、いつでもどこでも快適に動画視聴を行いたいという要望に応える。

3. 個人の嗜好に合わせた動画内容提示技術

3.1 背景

動画マーケティングが企業における動画利用目的の主要な一つとなっている。潜在的なお客様に対して商品やサービスに関する情報をプッシュ型で提供することで、購買行動を促している。動画は、他のメディアによる情報提供と比べて、訴求力が高く購買行動につながる確率が高いことが示されており、いかに視聴してもらえるかが鍵となっている。

提供する商品やサービスについて知らない潜在的なお客様がそれらを検索することは考えにくく、動画提供側から情報を発信する必要がある。そのため、興味を喚起する魅力的なサムネイル画像を選定し提示する手法がよく採用されている。しかし、サムネイル画像の選定は人手で行っているためコストがかかり、個々の潜在的なお客様が持つ嗜好に合わせることは困難である。個人ごとの嗜好に合わせて動画を選定し、さらにサムネイル画像の選定まで自動で行う技術が必要となる。

3.2 課題および技術

近年、SNSへの書き込みなど利用して個人の嗜好を推定できるようになってきている²⁰⁾。こうして得られる個人の嗜好を用いることで、より効果的に情報提供を行うことができると考えられる。たとえば、観光旅行を計画している人は、観光名所、食事、宿、アクティビティなど個人によって興味の対象が異なる。観光名所の中でも、景色を重視するか、歴史的建造物を重視するかはさまざまである。観光地の動画を推奨する場合は、

こうした多様な嗜好を考慮し、個人に合わせることが望ましい。複数の概念を含む個人の嗜好情報を、そのまま既存の動画検索エンジンへ入力しても良い結果は得られない。通常検索エンジンに対し、概念の異なる複数のキーワードを入力しても、キーワードのいずれかに対応した結果しか返さない。この場合は、検索結果の上位が特定のキーワードにしか対応していなかったりするなど期待と異なる挙動を示す。

この問題に対応するために、上位概念と下位概念の二層により個人嗜好を表現する手法を導入した。個人の嗜好を、上位概念への関心度 W_i ($i=1, \dots, M$) と、 i 番めの上位概念に属する下位概念への関心度 w_j ($j=1, \dots, N$) として表現する。個人の嗜好と動画内容の合致度を判定する場合に、上位概念と下位概念双方を考慮することで、特定の事象に偏らない選定が行えることが期待できる。この表現を用いた動画およびサムネイル画像選定の手順を図2を用いて説明する。各関心度の求め方については、前述の嗜好推定やアンケートからの直接入力などが考えられるが、本稿では扱わない。

- ① 動画をセグメントに分割する。データベースに登録された各動画に対し、等間隔でフレーム画像を取り出し画像特徴量を算出する。この特徴量に基づきフレーム画像をクラスタリングし、連続したフレーム画像が異なるクラスタに属する箇所を切れ目として動画をセグメントに分ける。
- ② セグメント内容を解析する。各セグメントの中間点に最も近いフレーム画像を代表画像として抽出する。この代表画像に対し、前出の下位概念に対応する画像識別器を適用し、 N 次元のセグメントベクトル $(s_{i1}^k, s_{i2}^k, \dots, s_{iN}^k)$ を作成する。ここで s_{ii}^k は、 k 番めの動画における i 番めのセグメントに対し i 番めの下位概念が含まれる度合いを示す。
- ③ 1つの動画に含まれるすべてのセグメントの情報を用いて $s_i^k = \max_l s_{il}^k$ により統合した N 次元の動画ベクトル $(s_1^k, s_2^k, \dots, s_N^k)$ を作成する。
- ④ 嗜好情報と合致する動画をデータベースから選定する。式(1)を用いて k 番めの動画に対する合致度スコア v_k を算出

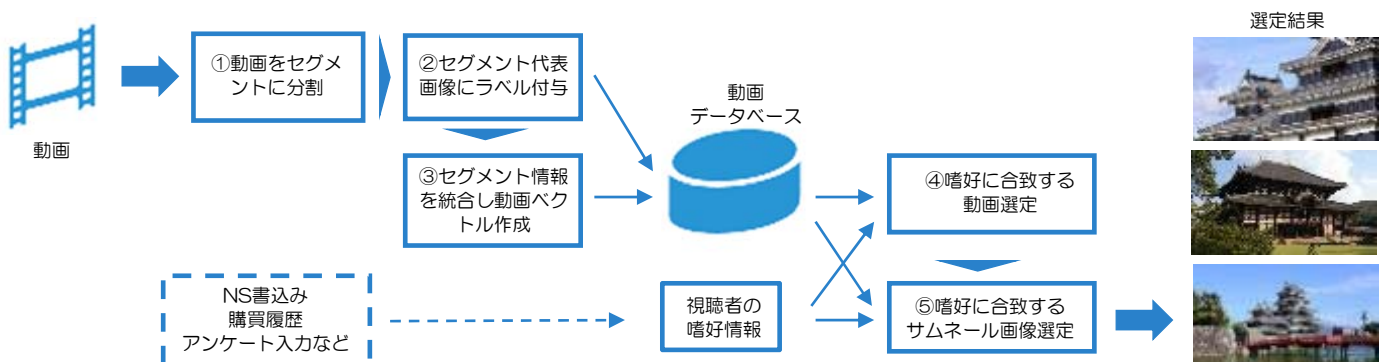


図2 個人の嗜好に合わせた動画およびサムネイル画像の選定

る。ここで第一項は下位概念の合致度を表し、第二項は上位概念の合致度を表す。それぞれが合致度スコアに寄与する割合を $\alpha:1-\alpha$ としている。

$$v_k = \alpha \sum_j \max_i w_{ij} s_j^k + (1-\alpha) \sum_i W_i \max_j w_{ij} s_j^k \quad (1)$$

すべての動画について合致度スコア v_k を計算し、大きい順に並べ替えて上位の動画を選定する。

- ⑤ サムネイル画像を選定する。前ステップで得られた各動画中のセグメントについて、前ステップと同様に式(2)を用いて嗜好情報との合致度スコア v_i を計算する。合致度スコア v_i が最大となるセグメントを特定したあとに、最も関連の強い上位概念に対する嗜好 W_i を半減し、合致度スコアを再計算する手続きを必要なサムネイル画像の数だけ繰り返す。嗜好値の半減と合致度スコア計算を繰り返すことで、同一概念の画像が上位を占めることが回避できる。

$$v_i = \alpha \sum_j \max_i w_{ij} s_j^k + (1-\alpha) \sum_i W_i \max_j w_{ij} s_j^k \quad (2)$$

3.3 検証方法と結果

前節に示した手順を実際の動画に適用し、個人の嗜好に基づく動画選定およびサムネイル画像選定を行うシステムを構築した。対象としてインターネット上に公開されている日本の観光振興に関する動画約200本を用いた。観光に関する紹介を行うことを前提とし、観光目的に関する上位概念となる単語15個と、下位概念となる単語を100個をあらかじめ定義した。表1に一部を示す。下位概念として定義した単語それぞれに対して、インターネット上の検索エンジンに入力して画像を100枚取得し、得られた画像を教師データとして機械学習を用いて画像識別器を作成した。画像識別器は、入力画像と学習した単語が合致すると合致度に応じた正の値を返す。教師データとは別に単語ごとに10枚用意したテスト画像を用いて識別器の性能を評価したところ、正当率は8割弱となった。

動作検証のため、嗜好情報を直接入力するインターフェイスを用意し、入力した嗜好情報の細かな違いに応じて期待どおりの動画選定とサムネイル画像の選定が行われることを確認した。また、同じ入力を用いて、一般的なベクトル内積計算による結果と提案手法の差異を検証した。嗜好情報に複数の概念が

表1 観光目的に関する単語群（抜粋）

上位概念	下位概念
グルメ	カフェ、スイーツ、寿司、そばうどん、お酒
アクティビティ	スキー、スキューバダイビング、スポーツ観戦
歴史	遺跡、寺、神社、城
自然	花見、紅葉、雪景色、山、鍾乳洞
温泉	露天風呂、温泉



図3 「歴史」と「グルメ」の組み合わせに対する画像の選定例

含まれている場合、提案手法のほうが良好な結果が得られることを確認した。例として、「歴史」と「グルメ」を入力した結果を図3に示す。図3で上の二段は提案手法による選定結果であり、「歴史」と「グルメ」の2つの概念が含まれている。下の二段は一般的な内積計算による選定結果で、「グルメ」に偏った画像選定がされている。

3.4 議論

提案手法を実環境で利用するには、用途に応じた上位概念と下位概念の設定と、対応する画像識別器の作成を都度行わなければならない。しかし、適用するアプリケーション領域ごとに共通化ができることが想定されるため、利用が増えるほど初期作業の負荷は減少すると考えられる。また、ImageNet²⁾など既存の画像と関連づけたオントロジー研究が進んでおり、階層構成を持つ一般的な概念に対応する画像識別器の構築に必要なデータが入手できるようになってきている。こうした研究の成果を利用することによって、適用分野に依存する特殊な概念へのみ対応すればよくなり、初期作業負荷の軽減が期待できる。

3.5 今後の展望

先に述べた評価結果は定性的な性能確認にとどまっている。今後、ユーザー評価実験により定量的に性能を確認することを予定している。さらに、実際のマーケティングデータを用いて、動画視聴への誘導にどの程度効果があるかも検証しなければならない。

今回は視覚情報にのみ基づき動画の選定を行った。実際には音声など視覚情報以外の情報も用いて個人の需要にあった推薦を行うことも求められる。たとえば教育用途の動画など、視覚的には必ずしも特徴的ではない動画に対しても、視聴へ誘導できる魅力的なサムネイル画像を自動で作成することは、新たな挑戦となる。

4. 適応的動画配信技術

4.1 モバイル動画視聴の課題と適応的配信

動画利用の普及とともにデータトラフィックの増加がモバイル通信ネットワークを逼迫している。これによりインフラ増強が必要となりその費用を補うためモバイル基本利用料金の増加、あるいは事実上従量制となっている通信費の増大が生じている。また、視聴者への直接的な影響として体感品質の劣化も招いており、いまだモバイル環境で自由に動画を視聴できる環境が十分に整っているとは言えない状況である。

ネットワークが輻輳している状態における動画視聴は、配信サーバーからクライアントへ再生に十分なデータが供給されず、そのために再生の開始が遅れたり、途中で止まったり（ストール）、画像乱れが生じたりする。これらは視聴者の体感品質を大きく劣化させる。加えて、動画表示領域の解像度が大きすぎる場合、ダウンロード量が増加し、通信量の増大を招く。これらの原因は、ネットワーク状態や視聴デバイスに対し動画の仕様が適していないケースが多い。これらの課題を解決する手法として、変化するネットワーク状態やデバイスの情報から、都度適切な動画サイズを選択する「適応的動画配信 (Adaptive Video Streaming)」が提案されている。

適応的動画配信は、Apple社²¹⁾やマイクロソフト社²²⁾等が提案する方式と、オープンソースとして開発が進められているMPEG-DASH (Dynamic Adaptive Streaming over HTTP)^{7),8)}等が挙げられる。これらは対応するOS、ウェブブラウザ等に違いがあり、互換性を有しないが、次に示す共通した特徴と機能を有している。プロトコルにはHTTPストリーミングを採用し、従来のプッシュ型専用ストリーミングサーバーを不要とした。動画ファイルは、同一コンテンツで、段階的に異なるビットレートを複数用意し、さらにそれぞれ数秒単位の同一間隔のセグメントに分割する。これらの分割ファイルの特性やURL等を記述したマニフェストファイルも合わせてhttpサーバーに設置し、クライアントは、そのマニフェストファイルを取得してビットレートの選択を行い、サーバーよりセグメントデータを順にダウンロードし、再生する (図4)。

この複数提案されている適応的動画配信方式の中で近年活発に標準化が進んでいるのがMPEG-DASHである。MPEG-DASH (単にDASHとも称される) は、2012年4月に国際標準化機構ISO/IEC23009で標準化され、MPEG-DASH再生および制御が可能となるHTML5のMSE (Media Source Extensions) が主要ブラウザに搭載された。さらに、オープンソースであることから、独自の機能やアルゴリズムを加えることが可能であり、加えて、知的財産権の共同ライセンス化の動き²³⁾から商用利用の障壁が下がり、体感品質の良好な動画視聴がさまざまなデバイス、OS等、幅広い環境のもとで実現できるものと期待されている。

我々は、このMPEG-DASHに着目し、モバイル機器での動画視聴を意識した動画配信および視聴の検討を行った。

4.2 MPEG-DASHによるモバイル向け動画配信

動画視聴においてユーザー体感品質を劣化させる代表的な例は、再生開始が遅い、画質が悪い、再生中に止まる等が挙げられる。これらの原因は、配信サーバーとクライアント間の帯域が不十分なためであることが多い。このため、MPEG-DASHでは、ネットワーク状態、とりわけ帯域を検知しつつ、それに合わせてビットレートを変更する手法がよく使われている。すなわち、帯域が狭いケースには低ビットレートを、十分広いケースには、高ビットレートを選択する。これにより、視聴中の一時停止や開始の遅延の問題に一定の効果を得られる。しかし、単にこれをモバイルネットワークに適用すると、下記の問題に直面する。

1. モバイルネットワークの帯域の変化は大きく、追従できないケースがある
2. 十分な帯域が確保できるとき、通信量が不必要に増大する
3. 低ビットレートでは、画像が劣化するシーンが発生する
4. 早送り時に再生が停止する

これら4つの問題を解決するため、我々はそれぞれに次の施策を講じた。

1. ネットワーク帯域は短周期で大きく変動するため、帯域を検知した時点でその値はすでに変化してしまっていることがしばしばある。そのため、動画のクライアント側にバッファを設け、帯域変動をある程度吸収させた。また随時バッファ量をチェックし、量が減少したときには、動画停止を予測し低いビットレートの動画に切り替えた。
2. 視聴しているウィンドウサイズを検知し、それに適した解像度と対応するビットレートを選択させた。これにより、帯域が十分に確保できているケースにおいても過大な解像度を選択せず、画質を保持しつつ低いビットレートを選択可能となる。

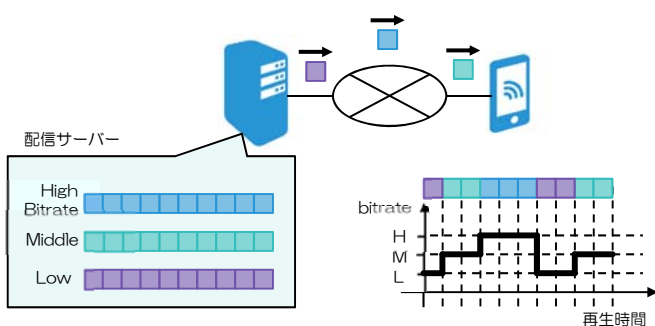


図4 適応的動画配信



図5 MPEG-DASHのビットレート選択に適用したパラメーター

3. 元の動画からビットレートの異なる複数の動画へ変換する工程で、短周期な変化や、色情報を多く含むシーンでは、フレーム間情報差異が大きくなり、符号化効率が低下する。そのためビットレートを固定すると、特に低ビットレートで画質が劣化し、体感品質の低下を招く。これを回避するため、低ビットレート用に変換した動画の画質が体感品質に与える影響を評価し、劣化が大きい箇所は帯域が許容数する範囲でより高いビットレート用の動画を用いて画質を確保する手法を導入した。変換前の元動画ファイルとビットレート変換後の動画ファイルから同一時刻のフレームを周期的に抽出し、静止画の構造的比較を行う。元動画フレームからの構造的な差異に応じた画質指数を用いて各時刻の体感品質の低下を予測し、あらかじめ設定した閾値を下回らないようにする。

4. 再生速度（PBR：Play Back Rate）によってもビットレートを変化させた。早送り（PBRが1以上）では、単位時間あたりのダウンロード量は増加し、またクライアント側の復号処理も高速になるため、再生停止を招きやすい。一般的に、早送り時には、視聴者目的のシーンを探したり、内容の概略を把握したりすることが目的であることが多く、画質は重要視されない。そのため、早送り時には前記画質指数を利用して最低限必要な画質を確保しつつビットレートを減少させた。一方、スローモーション時（PBRが1.0未満）のケースでは、視聴者が細部まで詳細に視聴することを求めており、またサーバーからの単位時間あたりの通信量も減少するため再生が停止する確率も下がることから、ビットレートを増加させた。

以上をまとめると、我々は図5に示す5つのパラメーターを取得し、その情報から体感品質を確保しつつ、通信量の抑制できるビットレートを選択する手法を考案した。

4.3 高画質ズーム機能

視聴支援機能として高画質ズーム機能を付加した。ズームは、ユーザーが視聴中に任意のタイミングで高解像度静止画を取得できるもので、サーバー側が高解像度動画を有していることを利用している。ユーザーが視聴中に画面上をダブルクリック

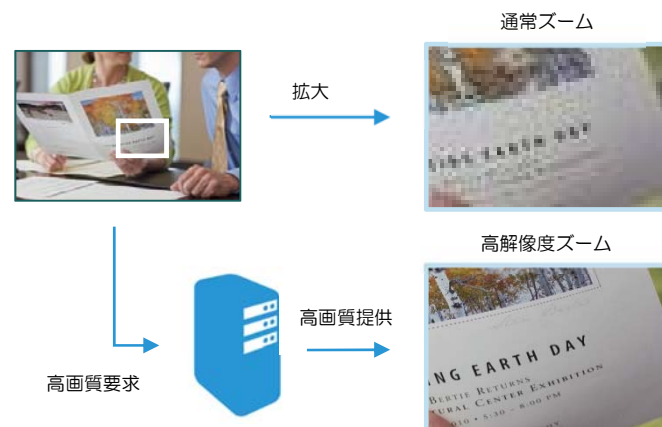


図6 高画質ズーム

すると、該当フレームをサーバーに要求し、サーバーは、保有する最高画質の動画から、該当フレームを抽出しクライアントに送信し表示する（図6）。一般的には、動画の拡大は表示ビットレートを拡大したものであり、鮮明な画質が得られないことがあるが、本手法を用いることにより、高解像度が得られ、取得できなかった新たな情報を得ることができる。これは、さまざまなアプリケーション、たとえば、セキュリティーやオンラインラーニング等に適用可能と考えられる。

4.4 社内動画配信トライアル

前述の適応的動画配信の効果を検証するため、社内動画配信システムを構築した。通信トラフィックをより現実近づけるため、社内複数拠点、部門から選ばれた合計100人程度を対象とし、期間を設定してトライアルを実施した。コンテンツとして、長さや場面が変わる頻度の多様性を考慮し、社内の技術紹介や、国内外の部門紹介、経営層からのメッセージ等を掲載した。動画変換パラメーターは全コンテンツに対して固定とし、ビットレートは、200kbpsから2000kbpsまで5段階を設定した。ユーザーが動画をアップロードすると、サーバー側ではビットレート変換、セグメント分割、マニフェストファイルの作成、変換された動画の画質指数取得、サムネールの作成を自動で行う。また、ユーザー側インターフェイスは、提案した適応的動画配信に対応する専用のプレーヤーを作成した（図7）。作成したMPEG-DASHプレーヤーにログ機能を加え、視聴時の再生開始所要時間、ストール頻度、データ通信量を保存した。加えて、アンケートから体感品質等を取得した。

動画再生が開始するまでの所要時間が体感品質に影響を与える閾値は2秒とされている⁶⁾。従来の配信方式では、通信が逼迫する時間帯にそれを超過するケースがしばしば発生していた。本トライアルでは、動画の再生開始は97%で2秒以下（図8）であったことがログ解析からわかった。また、ダウンロード量は従来の配信方式で配信している同一コンテンツのビデオ



図7 作成した動画プレーヤー

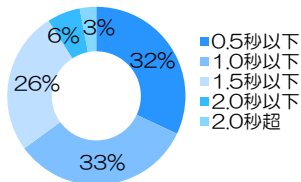


図8 再生開始所要時間

オと比較しておおよそ60%減少していた。このとき画質による体感品質の低さを訴えた例はなかった。選択されたビットレートの画質指数を解析したところ、体感品質が低いとされるビットレートを選択された例はなかった。これらの結果より、体感品質の確保と通信量の抑制が両立できていることを確認した。

5. おわりに

拡大するリッチメディアを利用する場面を想定し、個人ごとの嗜好に合わせたサムネイル画像の自動選定技術と、増大するデータトラフィックに対応する適応的動画配信技術とを提案した。個人の嗜好に合わせることで、視聴者は興味を惹く動画を発見しやすくなるのが期待できる。体感品質を考慮して通信データを随時選択することにより、通信インフラを利用するコストを抑制しながら、いつでもどこでも快適に動画を視聴できるようにする。

本稿で提案した技術を、これまでの取り組みと合わせることで、リッチメディアの持つ伝達力や説得力を活用した人と人とのコミュニケーション支援がより包括的に行えると考える。今後はさらにお客様視点で、品質の向上並びに機能の拡張を図り、お客様の業務改善に貢献することができるリッチメディアの解析と配信について研究技術開発を進めていく。

参考文献

- M. G. Christel: Automated Metadata in Multimedia Information Systems: Creation, Refinement, Use in Surrogates, and Evaluation, Morgan and Claypool Publishers, (2009).
- J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei: "ImageNet: A Large-Scale Hierarchical Image Database", Proc CVPR, pp.248-255, (2009).
- A. Girgensohn, J. Boreczky, L. Wilcox: "Keyframe-Based User Interfaces for Digital Video", 富士ゼロックステクニカルレポート No.14, pp.71-79, (2002).
http://www.fujixerox.co.jp/company/technical/tr/2002/02_09.html (参照日: 2016.12.20)
- W. Hu, N. Xie, L. Li, X. Zeng, and S. Maybank: "A Survey on Visual Content-based Video Indexing and Retrieval", IEEE Trans on Systems, Man, and Cybernetics, Part C, Vol.41, No.6, pp.797-819, (2011).
- S. Uchihashi, J. Foote, A. Girgensohn, J. Boreczky; "Video Manga: Generating Semantically Meaningful Video Summaries", Proc. of ACM Multimedia, pp.383-392, (1999).
- S. S. Krishnan and K. Sitaraman; " Video stream quality impacts viewer behavior: inferring causality using quasi-experimental designs", Proc. of Internet Measurement Conference, pp.211-224, (2012).
- 平林光浩, "次世代動画配信技術「MPEG-DASH」技術概要と標準化・関連技術動向", 映像情報メディア学会誌, Vol.67, No.2, pp.109-115, (2013).
- 藤沢 寛: "MPEG-DASHとハイブリッドキャスト", NHK技研 R & D, No.156, pp.14-21, (2016).
- "スマートフォン・エコノミー"〜スマートフォン等の普及がもたらすICT産業構造・利用者行動の変化, 平成24年版 情報通信白書, 第2節, 総務省, (2012).
<http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h24/pdf/n2020000.pdf> (参照日: 2016.11.15)
- J. Adcock, M. Cooper, L. Denoue, H. Pirsiavash, L. A. Rowe: "Talk Miner: A Lecture Video Search Engine", 富士ゼロックステクニカルレポート No.21, pp.118-128, (2012).
http://www.fujixerox.co.jp/company/technical/tr/2012/t_06.html (参照日: 2016.12.20)
- S. Carter, M. Cooper, J. Adcock: "ShowHow: Supporting Expository Video Capture and Access," 富士ゼロックステクニカルレポート No.23, pp.34-42, (2014).
http://www.fujixerox.co.jp/company/technical/tr/2014/s_04.html (参照日: 2016.12.20)
- A. Girgensohn, D. Kimber, J. Vaughan, T. Yang, F. Shipman, T. Turner, E. Rieffel, L. Wilcox, F. Chen, T. Dunnigan: "DOTS: Support for Effective Video Surveillance", 富士ゼロックステクニカルレポート No.17, pp.83-100, (2007).
http://www.fujixerox.co.jp/company/technical/tr/2007/t_03.html (参照日: 2016.12.20)
- Q. Liu, C. Liao, L. Wilcox, A. Dunnigan, B. Liew: "Embedded Media Markers: Marks on Paper that Signify Associated Media", 富士ゼロックステクニカルレポート No.20, pp.96-106, (2011).
http://www.fujixerox.co.jp/company/technical/tr/2011/t_04.html (参照日: 2016.12.20)
- J. Pickens, J. Adcock, M. Cooper, M. Back, P. Qvarfordt, G. Golovchinsky, A. Girgensohn: "Interactive Multimedia Search: Systems for Exploration and Collaboration", 富士ゼロックステクニカルレポート No.18, pp.112-130, (2008).
http://www.fujixerox.co.jp/company/technical/tr/2008/t_03.html (参照日: 2016.12.20)
- 市村 哲, 菅野英介, 小川正和, 鈴木理敏: "MediaDEPO: Webにおけるビデオ映像の一覧性を高めるビデオアクセス技術", 富士ゼロックステクニカルレポート No.13, pp.4-12, (2000).
- 内橋真吾: "Improving Camera Control for Capturing Meeting Activities using a Floor Plan", 富士ゼロックステクニカルレポート No.14, pp.46-55, (2002).
http://www.fujixerox.co.jp/company/technical/tr/2002/02_07.html (参照日: 2016.12.20)
- 加藤典司, 福井基文, 坪下幸寛, 尾崎良太: "確率的生成モデルを用いた自動画像アノテーション", 富士ゼロックステクニカルレポート No.22, pp.1-9, (2013).
http://www.fujixerox.co.jp/company/technical/tr/2013/t_03.html (参照日: 2016.12.20)
- 新宮 淳, 栗原陽子, 竹内孝行, Jim Vaughan, Don Kimber: "リアルタイムコミュニケーション状況抽出&解析技術", 富士ゼロックステクニカルレポート No.23, pp.64-71, (2014).
http://www.fujixerox.co.jp/company/technical/tr/2014/s_07.html (参照日: 2016.12.20)
- 田丸恵理子, 埴 雅典, 森澤正之, 安藤英俊, 日永龍彦, 平野敦資, 永峯猛志, 篠崎謙吾: "音声同期スクリーンキャプチャシステムを用いた学生の主体性を引き出す反転授業の試み", 富士ゼロックステクニカルレポート No.23, pp.43-52, (2014).
http://www.fujixerox.co.jp/company/technical/tr/2014/s_05.html (参照日: 2016.12.20)
- 三浦康秀, 服部圭悟, 大熊智子, 増市 博: "Distant Supervisionによる感性

トピックの抽出”，富士ゼロックステクニカルレポート No.23, pp.72-80, (2014).

http://www.fujixerox.co.jp/company/technical/tr/2014/s_08.html (参照日: 2016.12.20)

- 21) Apple Inc. “HTTP Live Streaming”
<https://developer.apple.com/streaming/> (参照日: 2016.11.15)
- 22) Microsoft Corporation. Smooth streaming
<https://technet.microsoft.com/ja-jp/library/dd775200.aspx> (参照日: 2016.11.15)
- 23) MPEG LA Announces Call for Patents to Organize Joint License for MPEG-DASH
<http://www.mpegla.com/Lists/MPEG%20LA%20News%20List/Attachments/96/n-15-07-27.pdf> (参照日: 2016.11.15)

筆者紹介

内橋真吾

研究技術開発本部 インキュベーションセンターに所属
専門分野：マルチメディア、動画要約、画像処理

藤居 徹

研究技術開発本部 インキュベーションセンターに所属
専門分野：マルチメディア配信解析技術

ムラリ・スレシュ

研究技術開発本部 インキュベーションセンターに所属
専門分野：マルチメディア配信解析技術