一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report
LOIS2019-67(2020-03)

メディア授業に適した授業動画配信方式の検討

原田　明成† 　張　　徳鵬†† 　檜垣　泰彦†††

† 千葉大学工学部　〒263–8522 千葉市稲毛区弥生町 1–33
†† 千葉大学大学院融合理工学府　〒263–8522 千葉市稲毛区弥生町 1–33
††† 千葉大学アカデミック・リンク・センター　〒263–8522 千葉市稲毛区弥生町 1–33
E-mail: ††higaki.yasuhiko@faculty.chiba-u.jp

あらまし　メディア授業で用いる教材用の動画を配信するのに適した動画配信方式とそのパラメータの検討を行った．説明用スライドに話者の動画を添えた映像を想定し，スマートフォン等のモバイル環境で利用できること，スライドの内容を読めること，必要最小限のビットレートとすることを要件として検討した．その結果，マルチビットレートの HLS 配信方式とすることとし，通信状況の良い状況，悪い状況，その間の 3 つのビットレートで構成することとした．SSIM による動画の画質の定量評価，電車移動中のビットレートの変化の確認を行った．ビットレートの上限を指定することで，通信制限のかかった状況でも動画を停止させることなく再生できることを確認した．

キーワード　HLS，動画配信，H.264，SSIM，画質評価

A study of a video distribution system suitable for media classes

Akinari HARADA†, Depeng ZHANG††, and Yasuhiko HIGAKI†††

† Faculty of Engineering, Chiba University　1–33 Yayoi-cho, Inage-ku, Chiba-shi, 263–8522 Japan
†† Graduate School of Science and Engineering, Chiba University　1–33 Yayoi-cho, Inage-ku, Chiba-shi, 263–8522 Japan
††† Academic Link Center, Chiba University　1–33 Yayoi-cho, Inage-ku, Chiba-shi, 263–8522 Japan
E-mail: ††higaki.yasuhiko@faculty.chiba-u.jp

Abstract　We studied a video distribution method and its parameters that are suitable for distributing videos of teaching materials used in media courses. Assuming that the video is attached to the slide together with the lecturer's video for explanation, it is required to be able to be used in a mobile environment such as a smart phone, to be able to read the contents of the slide, and to have the minimum required bit rate. As a result, we decided to use a multi-bit rate HLS distribution system, which consists of three bit rates between good and bad communication conditions. The quantitative evaluation of the video image quality and the change of the bit rate when the train is running are confirmed by SSIM. By specifying the upper limit of the bit rate, you can confirm that the video does not stop playing even when communication restrictions are imposed.

Key words　HLS, Video distribution, H.264, SSIM, Image quality evaluation

1. はじめに

千葉大学では動画配信を用いたメディア授業の本格的導入を進めている．このような教材用動画の多くは，スポーツやドキュメンタリーの動画と比べると，動きが少ない傾向がある．そこで，授業で用いる教材用動画として，図 1 に示すような説明用スライドに話者の動画を添えた映像を想定し，スマートフォン等のモバイル環境で利用できること，スライドの内容を読めること，必要最小限のビットレートとすることを要件として検討する．

2. 動画配信方式

現在利用されている他の動画配信方式としてはプログレッシブダウンロード，RTMP や MPEG-DASH 等がある．ストリーミング配信に RTMP のように専用の RTMP サーバーを作る必要がなくストリーミングサーバとして普通の HTTP サーバーで実現ができる等構成がシンプルなところから HLS (HTTP Live Streaming) 方式 [1] を採用することとした．これはセグメント化された H.264 MPEG-2 TS 動画（.ts）およびインデックスファイル（.m3u8）を使って，ライブ配信やオンデマンド動画配

This article is a technical report without peer review, and its polished and/or extended version may be published elsewhere.
Copyright ©2020 by IEICE

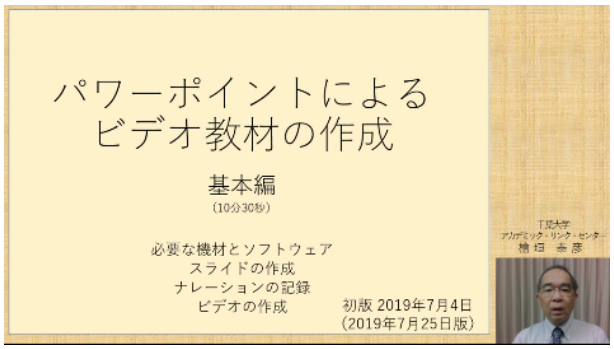

図 1　教材用動画の例

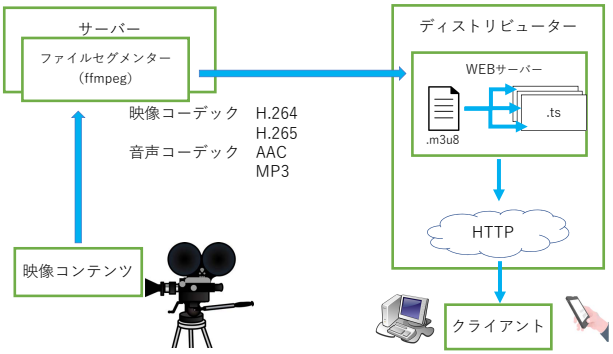

図 2　HLS の構成

信を可能にするインターネット動画配信技術である．AES-128 暗号化によるコンテンツの保護機能や，画質ごとにビットレートを変えた動画を準備してクライアント側回線の通信速度に合わせて自動的に映像品質の切り替えを行うアダプティブビットレートの機能を有する．しかし，HLS をネイティブで再生できる環境が限られていることや，アダプティブビットレートを実現するためにはマルチビットレートのコンテンツを用意する必要がありストリーム配信サーバーのストレージを圧迫する等の問題がある．

本研究では，教材用動画を対象とし，ffmpeg を用いて映像コーデック H.264，音声コーデック AAC でセグメントファイルにトランスコードする際の実用的なパラメータを検討し，比較検証する．トランスコード前後の動画を SSIM による定量評価と実際に視聴した際の主観評価で比較し，教材用動画に適した HLS 動画配信パラメータを検討する．

3.　HLS (HTTP Live Streaming)

HLS(HTTP Live Streaming) は，Apple Inc. によって開発され，2009 年に発表された HTTP ベースのアダプティブビットレートストリーミング通信プロトコルである．プロトコルのサポートは，メディアプレイヤー，WEB ブラウザー，モバイルデバイス，ストリーミングメディアサーバーで広く普及しており，毎年行われる映像産業の調査では，最も人気のあるストリーミング形式であることが確認されている [2]．HLS は，RTP などの UDP ベースのプロトコルとは異なり，標準の HTTP トランザクションに基づいており，標準の HTTP トラフィックを通過させるファイアウォールやプロキシサーバーを通過できる．これによりメディアコンテンツを従来の HTTP サーバーから提供し，インターネットのように一般的に利用されている HTTP ベースのコンテンツ配信ネットワークで配信することができる．コンテンツの保護は AES-128 暗号化と HTTPS を使用した安全な復号化鍵配布により行う．

HLS はサーバーコンポーネント，ディストリビューションコンポーネントおよびクライントソフトウェアの 3 つで構成されている．サービスアーキテクチャの構成を図 2 に示す．HLS の動画配信の流れとしては，まず映像コンテンツを作成する．ファイルセグメンターでその映像コンテンツの H.264 ビデオ及び AAC オーディオのメディアセグメント（.ts）とインデックス

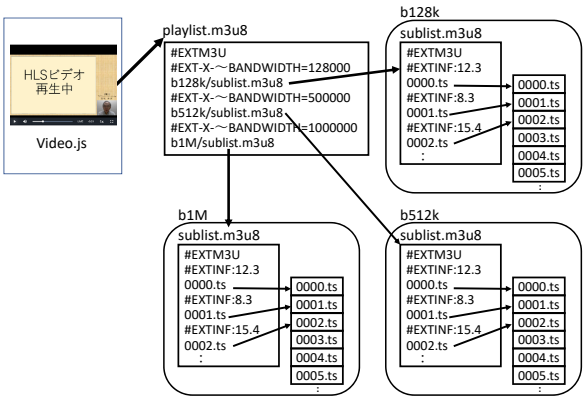

図 3　マルチビットレート HLS の構成

ファイル（.m3u8）を作成する．図 3 に HLS を用いた動画再生時のインデックスファイルとメディアセグメントの関係を示す．クライアントはマスタープレイリスト（図 3 の playlist.m3u8）を読み取り，メディアプレイヤーの帯域幅に適したメディアプレイリスト（図 3 の sublist.m3u8）のセグメントファイルを順番に要求し，分割されたファイルをギャップなしで再生する．また，複数ビットレートのメディアセグメントを用意することでモバイル環境のような通信回線の状況が悪くても異なるビットレートの動画に切り替えることで途切れることなく再生するアダプティブビットレートストリーミングを可能にする．本研究ではネイティブでサポートされていないブラウザーでも HLS 動画配信できるように，video.js(注1) を使用し多くの環境に対応できるようにした．

4.　FFmpeg

本研究では動画のトランスコードとセグメント化を FFmpeg(注2) で行った．FFmpeg はビデオ，オーディオおよびその他のマルチメディアファイルとストリームを処理するための膨大なライブラリとプログラムをひとまとめにしたソフトウェアで構成される無料のオープンソースプロジェクトである．その中核となるのはビデオおよびオーディオファイルのコマンドラインベースの処理用に設計された ffmpeg プログラムであり，フォーマットトランスコーディング，トリミングや連結などの

(注1)：https://videojs.com/
(注2)：https://www.ffmpeg.org/

表 1　ffmpeg のパラメータ

ピクセル数	映像 [kbps]	音声 [kbps]
1024 × 576	1000	128
854 × 480	256	64
640 × 360	96	32

基本的な編集，拡大縮小などのビデオスケーリング，ビデオポストプロダクションエフェクト（補正や修正などの動画全体の編集）などの多彩なオプションを使用可能なため，幅広く利用されている．

5.　画質評価手法

SSIM（Structural SIMilarity）[3] は Wang らによって提案された指標で，画像構造の類似度が人間の画質劣化の知覚に寄与するとの仮定に基づき定義されている．画像構造の類似度には，ピアソンの相関係数を用いており，これにトランスコード前後の動画の輝度・コントラストで評価値を構成している．SSIM は 0–1 の範囲の指標で，高ければ高いほど精度が良いことを示している．JIIMA ガイドライン [4] によると，SSIM が 0.98 以上あると主観評価としては元の画像と圧縮画像の区別がつかない，0.90–0.98 の間では拡大すれば劣化がわかるレベル，0.90 以下では明らかに劣化がわかるとなっている．

本研究では，トランコード後の動画と元の動画の画質を目視による主観評価と ffmpeg の ssim フィルタを用いて SSIM 値を算出し定量的に評価した．

6.　各種動画配信パラメータの設定

HLS 動画配信で考慮するべきパラメータとしては，ffmpeg でトランスコードする際の映像と音声のビットレート，映像のピクセル数，マスタープレイリストの BANDWIDTH 等が挙げられる．インターネット動画配信を行う場合，動画のデータ量を極力小さくする努力が必要である．必要以上のピクセル数やビットレートにすることで，膨大な動画データを視聴者が受信しなければならなくなることや，動画をサーバーに保存するためのストレージの確保，トランスコードする際の負荷や処理時間が増えてしまう．また，映像授業は PC だけでなくスマートフォンやタブレットなどのモバイル環境での視聴も想定されるので，アダプティブビットレートストリーミングが実現できる教材用動画に適したピクセル数とマルチビットレートの設定を検討した．設定したパラメータをまとめたものを表 1 に示す．

教材用動画を想定し Microsoft office PowerPoint のスライド 20 枚を 1 分ごとに切り替え，スライドショーの記録機能で録画した動画 (.mp4 1440 × 1080 20 分 映像ビットレート 2.8Mbps　音声ビットレート 198kbps 映像フレームレート 30fps，音声サンプリングレート 44.1kHz) を映像の変化が少なく動画の容量をそれほど必要としない，話者の声が聞こえる程度の音質があれば十分な点などを考慮して表 1 のようなパラメータに設定した．各ピクセル数でそれぞれ映像ビットレート 1M，256k，96kbps と音声ビットレート 128k，64k，32kbps でトランスコードしたものを比較した．

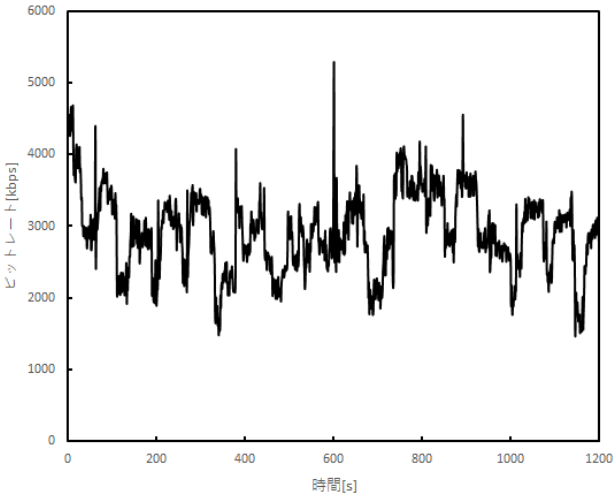

図 4　トランスコード前（入力）動画のビットレート

ピクセル数に関しては，ある程度の画面サイズを想定した HD（1280 × 720），WSVGA (Wide-SVGA 1024 × 576) と YouTube 等動画配信サイトの中〜低画質の動画に多い FWVGA (Full wide VGA 854 × 480)，HVGAW (Half VGA Wide 640 × 360) に設定し，ビットレートに関しては，YouTube での HD 画質のビットレートが 2Mbps 程度なのでデータ容量の節約も兼ねて 1 Mbps，音質に関しては 128kbps を最高にして，幅をもたせて複数のビットレートを設定した．映像ビットレート 96kbps，音声ビットレート 32kbps については，スマートフォンの通信制限時の回線速度が 128kbps であること，固定電話や PHS のビットレートが 32kbps であることからこの数値にした．

7.　結果と考察

7.1　動画全体のビットレート

ピクセル数は一定とし，ビットレートを変えてトランスコードしセグメント化せずに出力した動画ファイルのフレーム単位のビットレートを調べた．ffmpeg の showinfo オプションを使用し，フレームごとの情報を出力した．出力した情報をまとめたトランスコード前後の動画のビットレートを図 4–6 に示す．グラフの横軸は動画の時間，縦軸はフレームごとのビットレートをフレームレートで合計したビットレートである．

図 4–6 より元の動画のビットレートのピークが 600 秒あたりなのに対して，トランスコード後の動画のビットレートは 750 秒あたりである．また，ffmpeg の-b:v オプションによるビットレートの指定は ABR(平均ビットレート) であることが図 5，6 のように各ビットレートでビットレートが高いところと低いところがあることで確認できる．平均すると指定したビットレートに近い数値になり，ffmpeg の CRF オプションよりもトランスコード後のファイル容量を予測しやすい．

7.2　SSIM による画質評価

図 7 は横軸が動画の時間，縦軸が各ビットレートのフレームごとの SSIM 値である．

図 7 から，トランスコード後の動画は 800 秒あたりで品質が

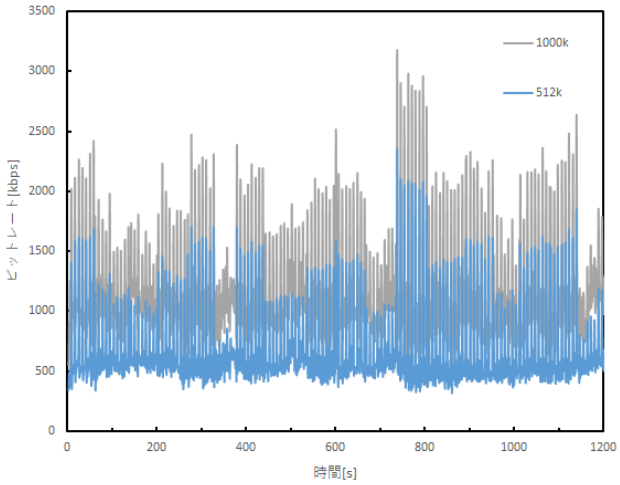

図 5　トランスコード後の動画のビットレート (1M, 512kbps)

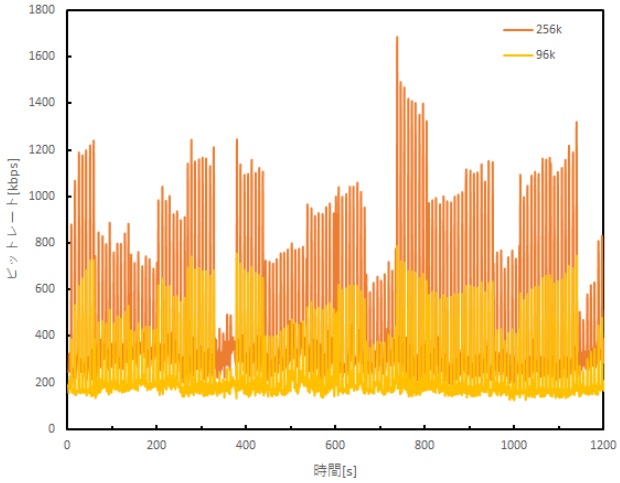

図 6　トランスコード後の動画のビットレート（256k, 96kbps）

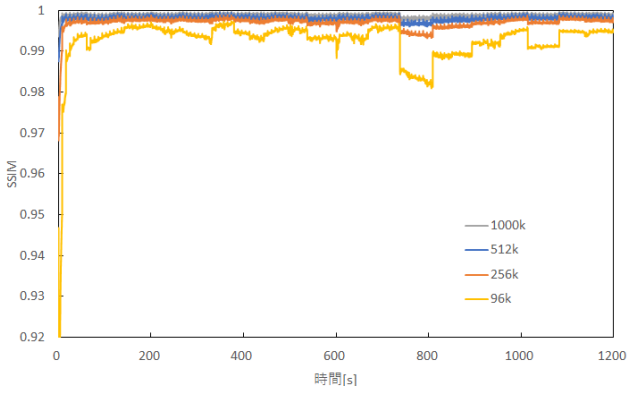

図 7　各ビットレートの SSIM 出力結果

劣化していることが確認できる．この箇所は動画内でグラフや表を表示しており情報量が多く 96kbps のように低ビットレートでは元の動画を再現できずノイズ等が発生しているからである．右下の話者を映している映像の品質はビットレートの違いにより目に見えて劣化を確認できるが，画像全体で見ると拡大しなければ気にならないくらいほとんど劣化がなく，グラフ，表や式など教材用動画として必要な情報を得ることが可能であった．

表 2　動画全体の SSIM

映像ビットレート [kbps]	SSIM
96	0.9927
256	0.9971
512	0.9981
1000	0.9989

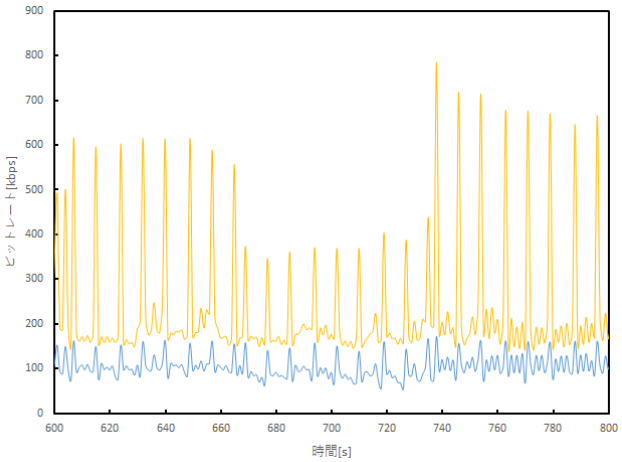

図 8　上限を 96kbps に制限した場合（青）としない場合（橙）のビットレート

表 2 に映像ビットレート別の動画全体としての SSIM を示す．また，表 2 より動画全体の SSIM 値は各ビットレートで 0.99 以上であり，教材用動画のように動きの変化が少ない映像では 96kbps のように小さい映像ビットレートで圧縮しても良いことがわかる．しかし，一部 SSIM 値が小さくなる箇所があることからわかるように，グラフや写真，スライドのアニメーション機能などは減らして動画全体の動きと情報量を少なくして動画を作成したほうが良いと考えられる．

オリジナルの動画の容量は 426MB で，トランスコード後の動画の容量はビットレート 1Mbps = 164MB, 512kbps=93.9MB, 256kbps=56.7MB，96kbps=19.6MB である．アダプティブビットレートストリーミングを実現させるためには，マルチビットレートのストリームを用意する必要があり，中程度の画質に用いるビットレートは動画ファイルの容量やクライアント側の帯域幅等を考慮して本研究では 256kbps を選択した．

7.3　ビットレート上限指定

パラメータの設定において，スマートフォンの通信制限時の回線速度が 128kbps であること，固定電話や PHS のビットレートが 32kbps であることからこの数値にしたが，映像ビットレートが平均ビットレート（ABR）であることからストリームの中で 128kbps を超えるところがある．そのため通信制限時にスムーズに再生できず動画が中断してしまう．ffmpeg でトランスコードする際に-maxrate を指定することでこの問題を解決できる．図 8 にビットレート上限を指定した場合としない場合のビットレートのグラフを示す．-maxrate オプションと-bufsize オプションを指定することにより上限の制限が機能

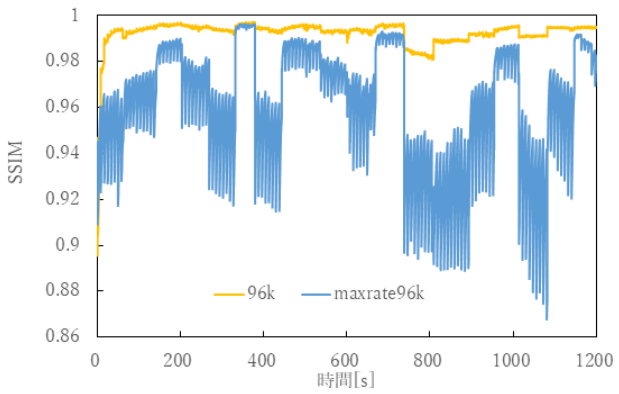

図 9　上限を 96kbps に制限した場合の

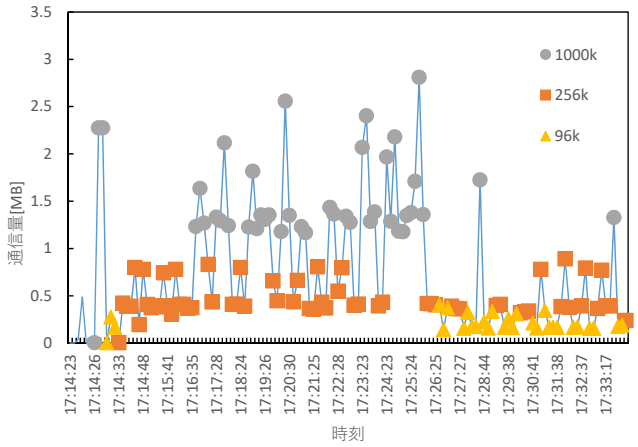

図 10　電車乗車時における通信状況が比較的悪い回線の通信状況

する．これにより回線速度が 128kbps に制限されてもほとんど途切れることなくストリーミング可能となった．最大ビットレートが制限された場合の SSIM 出力結果を図 9 に示す．グラフより SSIM の値が 0.9 以下のところが確認できる．この部分で他の 96kbps と比べるまでもなく画質が劣化していることがわかる．動画全体の SSIM は 0.9617 と低い．

7.4　マルチビットレートの効果

設定したビットレートで HLS 動画配信を行った場合における，モバイル環境でのアダプティブビットレートストリーミングの効果の検証を行った．電車内で教材用動画を視聴することを想定して，実際に電車に乗った状態で動画を視聴し，サーバ側の apache のログから，どの時刻にどのビットレートのセグメントが使われたかを調べた．図 10，11 に時刻と，どのビットレートのセグメントが使われたかをグラフで示す． 図 10 は通信状況が比較的悪い回線の場合，図 11 は比較的通信状況の良い回線の場合である．何れも 17:15 頃に千葉大学正門前の京成線みどり台駅を出発し，京成津田沼駅で乗り換え，17:30 頃に八千代台に到着している．通信状況がよくない図 10 では頻繁に 256 k や 96k に落ちていることがわかる．それに対して，総じて通信状況の良い図 11 では最速の 1000k が使われている．通信環境が良くない場合は，マルチビットレートの構成をとることが有効であることが確認できた．

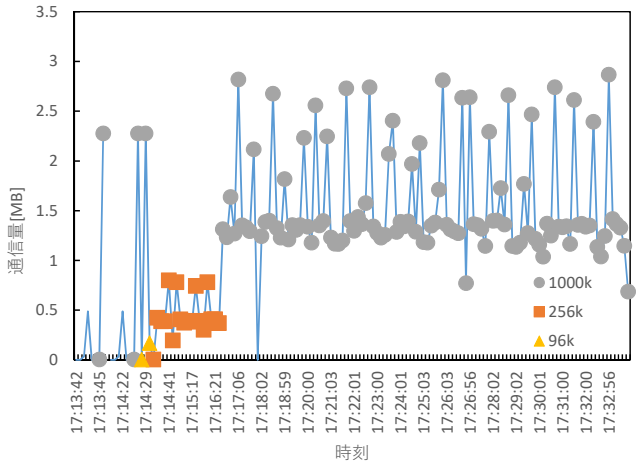

図 11　電車乗車時における比較的通信状況が良い回線の通信状況

8.　ま　と　め

本研究では，教材用の動画を配信する仕組みとして HLS を採用することとし，トランスコード用のツールとして ffmpeg を用いて映像コーデック H.264，音声コーデック AAC でセグメントファイルにトランスコードする際の実用的なパラメータを検討し比較検証した．トランスコード前後の動画を SSIM による定量評価と実際に視聴した際の主観評価で比較し，教材用動画に適した HLS 動画配信パラメータを検討した．

文字が書かれたスライドを想定した教材用動画を 1M，256k，96kbps でトランスコードしても，目視による主観評価，SSIM による定量評価において画質の劣化はわずかであることを確認した．

通信制限を対象に映像ビットレート 96kbps，音声ビットレート 32kbps を平均ビットレート（ABR）で指定すると，通信速度を遅くした場合スムーズに再生できなかったが，最大ビットレートも指定することによって制限下でも途切れることなく再生できることを確認した．

マルチビットレートの HLS を用いればモバイル環境のように通信回線が不安定な状況でも途切れることなく再生が可能であり，低ビットレートのコンテンツを用意することで通信容量の節約にもなる．

文　　献

[1] Draft-pantos-http-live-streaming-07 HTTPLive Streaming, https://tools.ietf.org/html/draft-pantos-http-live-streaming-07 （2020 年 2 月 1 日　閲覧）

[2] 2019 Video Developer Report, https://cdn2.hubspot.net/hubfs/3411032/Bitmovin\%20Magazine/Video\%20Developer\%20Report\%202019/bitmovin-video-developer-report-2019.pdf （2020 年 2 月 1 日　閲覧）

[3] Image Quality Assessment - Center for Neural Science - NYU, https://www.cns.nyu.edu/pub/eero/wang03-reprint.pdf （2020 年 2 月 1 日　閲覧）

[4] 符号化画質-電子通信学会知識ベース http://www.ieice-hbkb.org/files/02/02gun_05hen_09.pdf （2020 年 2 月 1 日　閲覧）