

インタラクティブ・パノラマ映像配信システム における配信方式の検討

井上雅之[†] 木全英明[†] 深澤勝彦[†] 松浦宣彦[†]

本報告ではハイビジョンを超える超高精細パノラマ映像を一般家庭に普及している回線およびハイビジョンテレビで視聴するためのパノラマ映像配信方式について検討する。従来、超高精細パノラマ映像をタイル分割し、配信する方式が提案されているが、各タイルを一律なビットレートで符号化すると映像品質に空間的なムラができ主観品質が低くなるという課題があった。そこで、本稿では制限帯域以下で高品質パノラマ映像を配信可能とするデータフォーマットと配信方式について提案する。提案システムでは、データフォーマットに MVC を利用し、各タイルを複数のビットレートで符号化、多重化する。提案する配信方式は、目標配信レート以下で主観映像品質を最大化することができる。主観評価実験の結果、従来方式より提案方式のほうがより高い主観品質を達成できた。

A Study on Delivery Method for Interactive Panoramic Video Delivery System

Masayuki Inoue[†], Hideaki Kimata[†],
Katsuhiko Fukazawa[†], and Norihiko Matsuura[†]

In this paper, we study delivery methods for the interactive panoramic video delivery service that enables users to watch high-resolution panoramic images beyond HD quality using available network resources and the TV sets in users' homes. The typical existing system divides the high-resolution panoramic video into tiles and the sender transmits a set of tiles, the partial panoramic video. Coding each tile at a uniform bitrate yields poor video quality because each tile has different visual quality. This paper proposes a new data format and delivery method to achieve high-quality partial panoramic video transmission even over restricted bandwidth networks. The proposed data format is based on the MVC standard. Each tile is encoded at multiple bitrates and multiplexed synchronously. Proposed delivery method maximizes the partial panoramic video quality even in restricted bandwidth networks. An experiment shows that the proposed delivery method can achieve higher subjective video quality than the conventional approach.

1. はじめに

近年、スーパーハイビジョンシステム¹⁾や 4K デジタルシネマ²⁾などハイビジョンを遥かに越える超高精細映像システムの研究開発が盛んである。視聴の形態として、高精細なディスプレイでそのまま表示する形態と一部を切り出して表示する形態がある。我々は後者のアプローチについて研究開発を進めており、現在普及しているハイビジョンテレビと光回線（たとえばフレッツ光³⁾）を利用して、超高精細パノラマ映像を視聴可能とするためのシステム開発を行っている⁴⁾。このシステムは、サッカースタジアムやコンサートホールの S 席の臨場感を家庭に届けることを目的とし、①一般家庭にあるハイビジョンテレビで視聴、②S 席からの超高精細映像の一部を視聴、③ユーザごとに異なる映像をインタラクティブに視聴可能とすることを特徴とする。その他、テレビ会議⁵⁾や遠隔教育⁶⁾などの様々なシステムが提案されている。これらのシステムでは、超高精細映像は映像サイズが大きいため、現在普及している回線では十分な映像品質が得られないという課題がある。

従来、目標帯域で伝送するために、MB（マクロブロック）ごとの QP 値変更などが通常の映像符号化では行われている。これは、ユーザ全員が同じ映像を視聴する、従来の TV 放送などでは適切な方法であった。しかしながら、本システムでは、各ユーザが超高精細パノラマ映像の異なる部分を視聴することを特徴とするため、視聴領域に対して適応的な配信制御が必要となる。

一般的な配信制御方法として、超高精細映像を小さなタイル映像に分割し、ユーザが注視している部分を含むタイル映像のみを伝送する方法がある。文献 7) ではパノラマ映像を 6 つの垂直スライスに分割し、各スライスを MPEG2 で符号化し、ユーザが要求した注視領域を含むスライス映像のみを伝送する。文献 8) では、複数のクライアントがアクセス可能なパノラマ・ストリーミング・システムを提案している。このシステムではサーバは全パノラマ映像を符号化せず、クライアントからの要求に応じて限られた領域のみ符号化・伝送する。しかしながら、これらのシステムではユーザの注視領域の符号化後の品質や利用可能な回線帯域などを考慮せずに、それぞれのタイルが符号化・伝送されているため、超高精細パノラマ映像を伝送するには不十分であった。ゆえに、我々の目的は制限帯域のもとで部分的なパノラマ映像を高品質で伝送することにある。

本報告では、超高精細パノラマ映像配信の課題について考察し、この課題を解決するための提案システムおよび配信方式について述べ、主観評価実験結果を示す。

[†] 日本電信電話株式会社 NTTサイバースペース研究所

[†] NTT Cyber Space Laboratories, NTT Corporation

2. 超高精細パノラマ映像配信の課題

制限帯域以下で超高精細パノラマ映像をユーザの注視領域に従い配信するための方法としてタイル分割する配信方法が一般的である。しかし、この配信方法において、タイル映像を一様のビットレートで符号化するとその PSNR は大きく異なる値を持つことが予備実験より確認できた。

たとえば、パノラマ映像を 36 個のタイルに分割し、タイル映像に 0 から 35 の tile_id を付与した様子を図 1 に示す。図 2 は、横軸にフレーム番号、縦軸に PSNR を取り、タイル映像を同一のレートで符号化、代表的なタイル映像 {0,7,14,21,28,35} の復号画像の PSNR の変化をグラフ化したものである。タイル映像 35 は動きのある物体がない部分であり、PSNR は約 35dB 付近を推移しているが、タイル映像 14,21 は動きの激しい人物が存在する部分であり、PSNR が最悪な場合は 25dB 以下となり、動きのある部分とない部分で 10dB 以上の差があることがわかる。以上、従来のようにパノラマ映像をタイル分割し一様なレートで符号化することは映像品質に大きなムラを発生させ、主観的な映像品質劣化の原因となる。ゆえに、この映像品質のムラを軽減することで、主観的な映像品質を向上させることができると考えることができる。



図 1 パノラマ映像のタイル分割例
Figure 1 Example of tile division.

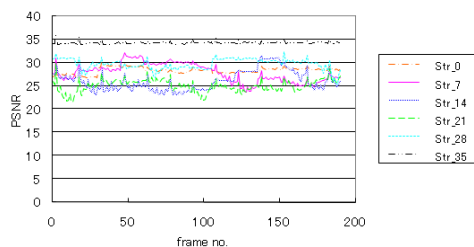


図 2 代表的なタイルの PSNR の変化
Figure 2 Changes in PSNR at the representative tiles.

3. 提案システム概要

前章の課題を解決するため、制限帯域以下で映像品質を最大化する仕組みを導入する。本提案システムでは、各タイル映像を複数のビットレートで符号化し、符号化レートとローカル復号画像の映像品質の対応テーブル（以下、レート品質対応テーブル）をパノラマ映像配信前にクライアントへ送信し、注視領域に対して適応的な配信制御をする。以下、システム構成とデータフォーマットについて述べる。

3.1 システム構成

図 3 は提案システム構成を示している。このシステムは、コンテンツ・ジェネレータ、サーバ、クライアントから構成される。コンテンツ・ジェネレータでは、パノラマ映像をタイル分割およびダウンサンプリングし、タイル映像とナビ映像を作成する。その後、各タイル映像を複数のビットレートで符号化したストリームと、ナビ映像を符号化したストリームを作成する。なお、この符号化処理と並行して、ローカルデコード画像から各ビットレートの PSNR を計算することで、レート品質対応テーブルを作成する。最後に、H.264/AVC の拡張規格である MVC(Multiview Video Codec)によって各ストリームとレート品質対応テーブルがデータストリームとして多重化される。サーバでは、クライアントからの要求に従いデータストリームからタイル映像ストリームを取り出す。クライアントでは、図 4 に示すように、各ストリームを受信し、部分的なパノラマ映像とナビ映像が表示される。ナビ映像上でのユーザの注視領域指定の変化によりパノラマ映像の見ている部分も変化する。ここで、複数のビットレートにより MVC 符号化されたタイル映像には view_id が付与され、クライアントは、ユーザの指定する注視領域を含むタイル映像に関連づけられた view_id を用いて、サーバへ部分的なパノラマ映像の要求を行う。

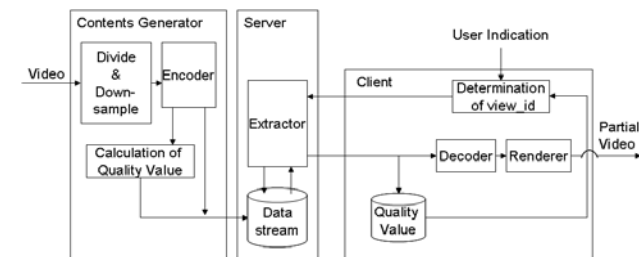


図 3 提案システム構成図

Figure 3 Proposed partial video streaming system architecture.



図4 クライアントの GUI

Figure 4 GUI image of the client.

3.2 データフォーマット

図5に示すように、提案するデータフォーマットは2種類の映像ストリームとメタデータからなる。

メイン・ストリームは、全体のパノラマ映像から生成され、ナビゲーション・ストリームは全体のパノラマ映像を縮小した低解像度パノラマ映像から生成される。ここで、低解像度パノラマ映像（ナビ映像）はユーザの注視領域の指定に用いられる。これら2種類の映像ストリームは、MVCにより複数の同期したストリームとして、符号化、多重化される。

例えば、パノラマ映像の分割数が36個の場合、メイン・ストリームには図6(a)に示すように36個のtile_idを付与したタイル映像を複数のビットレートで符号化した映像ストリームが含まれる。なお、同じタイル映像であってもビットレートごとに異なるview_idを付与する(図6(b)(c)(d))。

メタデータには、図6(a)のtile_idマップ、図7に示すtile_idとview_idのマッピングテーブル、符号化時に作成されるレート品質対応テーブルが含まれる。なお、図7におけるL,M,Hは低レート、中レート、高レートを順に示している。

本報告では、レート品質対応テーブルにおける品質にはPSNRを用いている。PSNRは原画像と符号化時に発生するローカルデコード画像との誤差より算出される。なお、このレート品質対応テーブルを含むメタデータはパノラマ映像配信前に送信される。

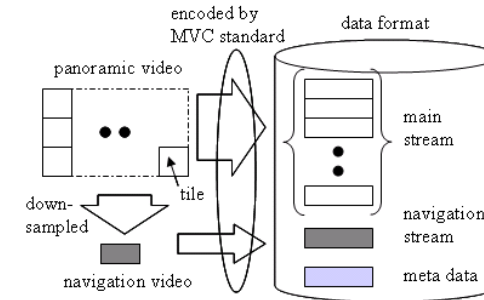


図5 パノラマ映像のためのデータ・フォーマット

Figure 5 Data format for panoramic video.

0	1	2	3	4	5
6	7	8	9	10	11
12	13	14	15	16	17
18	19	20	21	22	23
24	25	26	27	28	29
30	31	32	33	34	35

(a)

100	101	102	103	104	105
106	107	108	109	110	111
112	113	114	115	116	117
118	119	120	121	122	123
124	125	126	127	128	129
130	131	132	133	134	135

(b)

200	201	202	203	204	205
206	207	208	209	210	211
212	213	214	215	216	217
218	219	220	221	222	223
224	225	226	227	228	229
230	231	232	233	234	235

(c)

300	301	302	303	304	305
306	307	308	309	310	311
312	313	314	315	316	317
318	319	320	321	322	323
324	325	326	327	328	329
330	331	332	333	334	335

(d)

図6 (a)パノラマ映像のtile_idマップ
(b)低ビットレート時のview_idマップ
(c)中ビットレート時のview_idマップ
(d)高ビットレート時のview_idマップ

Figure 6 (a) Tile_id map of panoramic video.

(b) View_id map at the lowest bitrate.

(c) View_id map at the middle bitrate.

(d) View_id map at the highest bitrate.

tile_id	view_id		
	L	M	H
0	100	200	300
1	101	201	301
2	102	202	302
3	103	203	303
4	104	204	304
5	105	205	305
6	106	206	306
7	107	207	307
8	108	208	308
9	109	209	309
10	110	210	310
11	111	211	311
12	112	212	312
13	113	213	313
14	114	214	314
...			
35	135	235	335

図7 タイル・ビュー対応テーブル

Figure 7 Tile-view-mapping table.

4. 配信方式の提案

ここではクライアント処理の概要と、view_id の適応的な選択による主観映像品質を最大化する配信方式について述べる。

4.1 クライアント処理

図8に提案方式による復号化とレンダリング処理の基本部分について示す。ここでは、ユーザの注視領域の指定に応じてどのように復号とレンダリングされているかについて述べる。メイン・ストリームから得られる映像は図4に示されるメイン・ウィンドウに表示され、ナビゲーション・ストリームから得られる映像は図4に示されるナビゲーション・ウィンドウに表示される。図9にクライアントの全体フローを示す。映像ストリームを受信する前に、まず図6(a)の tile_id マップ、図7のタイル・ビュー対応テーブルやレート品質対応テーブルを含むメタデータを要求・受信する。次に、ナビゲーション・ストリームを要求・受信し、復号・レンダリングする。ユーザはナビゲーション映像上に表示された注視領域を拡大・縮小、左右移動させることで、希望の映像を見ることが出来る。ユーザが見ている映像の品質を最大化するように、

view_id を適応的に選択して、映像ストリームをサーバへ要求する。その後、この view_id に基づきメイン・ストリームをクライアントが受信する。コントローラは tile_id マップ、タイル・ビュー対応テーブルより、デコードされるべきタイルおよびレンダリング位置を決定し、デコード、レンダリングを行う。

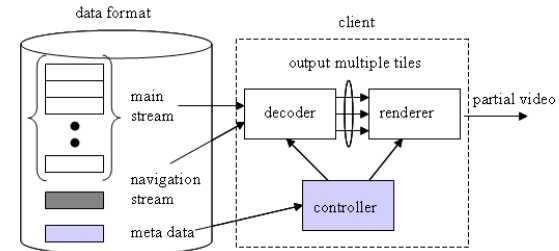


図8 クライアントのデータ処理コンポーネント

Figure 8 Client's components for data handling

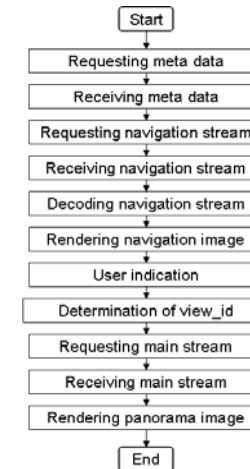


図9 クライアントの全体フロー

Figure 9 Client's general flow

4.2 View_id の決定方法

各タイルのビットレートをタイル・ビュー対応テーブルに従い View_id によりサーバへ要求することで、配信する際のレート制御がされる。基本的な配信制御の方針は、メタデータに含まれるレート品質対応テーブルから、最低のレートで最も SNR の低いタイルから順にビットレートを段階的に上げることで、注視領域全体の主観映像品質を上げるとのことである。なぜなら、最も映像品質の悪いタイルは注視領域で示される部分的なパノラマ映像の主観品質に大きく影響するからである。

提案する配信方式では、まず、注視領域に含まれるタイルを特定する。次に、特定されたタイル映像を対象として、レート品質対応テーブルより、最低のビットレートで符号化の際にできるローカルデコード画像の PSNR をキーとして昇順ソートを行い、最低の PSNR を示すタイルを特定する。その後、該当タイルを品質改善するために利用できる帯域があれば、該当タイルの符号化ビットレートを1段階上げる。これにより、該当タイルの PSNR が変化するため、再度 PSNR をキーとして昇順ソートを実施し、最低の PSNR を示すタイルを特定する。利用できる帯域がなくなるまで、上記処理を反復することで、目標ビットレート以下で主観映像品質を最大化する配信制御方式を実現する。

5. 評価実験内容

今回の評価実験では、図 10 に示すテスト映像'Dance'のユーザ注視領域に該当する部分を 30 秒間見てもらった後に、表 3 に示される質問紙を用いて、主観評価実験を実施した。ここでは 5 段階評価を行い、5 に近づくほど主観品質評価が高くなる。つまり、Q1 については「きれいな」ほど、Q2 については「鮮明な」ほど、Q3 については「ブロック歪が気にならない」ほど 5 に近い評価結果となる。なお、今回の評価実験では、図 10 に示すように注視領域サイズ、注視領域の位置は固定とし評価を実施した。

表 1 に実験条件を示す。ネットワークは LAN を用い、目標配信レートを 15Mbps に設定して配信を行った。この実験では、目標配信レート以下で、すべてのタイルを単一のビットレートで符号化する従来方式と、注視領域を含むすべてのタイルの符号化レートを適用的に選択する提案方式との比較評価実験を実施した。

表 2 に実験パラメータ設定内容を示す。パノラマ映像サイズ 3840×1920 画素をタイルサイズ 640×320 画素で分割すると、パノラマ映像は横 6 タイル×縦 6 タイル=36 個のタイルで表現できる。実際の提案システムでは注視領域は自由に拡大・縮小可能なため、注視領域が 3840×1920 画素と最大となる場合も想定し、表 1 の目標配信レート以下となるようにするためには、従来方式では各タイルのビットレートを 400kbps

で要求し、提案方式では、各タイルのビットレートを適的に 400kbps, 1 Mbps, 1.5 Mbps で要求するため、表 2 のパラメータ設定とした。



図 10 テスト映像'Dance'のユーザ注視領域
 Figure 10 User's view of test sequence 'Dance.'

表 1 実験条件

Table 1 Experimental conditions.

ネットワーク	LAN(100Mbps)
目標配信レート	15 [Mbps]
テスト映像の長さ	30 [sec]
比較方式	従来方式 (単一ビットレート方式), 提案方式
ディスプレイ	24 インチ LCD
注視領域サイズ	HD(1920x1080[pixels])
視距離	3H (87cm)

表 2 実験パラメータ設定

Table 2 Experimental parameter sets.

パノラマ映像サイズ	3840x1920 [pixels]
タイルサイズ	640x320 [pixels]
タイル分割数	6x6 [tiles]
従来方式によるタイルのビットレート	400 [kbps]
提案方式によるタイルのビットレート	400 [kbps], 1 [Mbps], 1.5 [Mbps]

表 3 テスト映像についての質問紙

Table 3 Questionnaire about the test image.

	質問
Q1	映像はきれいですか？
Q2	映像は鮮明ですか？
Q3	映像の中のブロック歪は気になりましたか？

6. 評価実験結果と考察

被験者 10 名（男性 5 名，女性 5 名）を対象とした，2 つの配信方式の比較評価実験の結果，各質問に対して 20 個の評価結果が得られた．表 4 に各質問に対して得られた評価結果の平均値を示す．t 検定の結果，Q1 について $t(18)=3.97$ ($p<0.05$)，Q3 について $t(18)=5.982$ ($p<0.05$)であった．Q1 および Q3 について，配信方式による平均値の差は有意であった．つまり，従来方式と比較して，提案方式によるパノラマ映像の方が「きれい」で「ブロック歪が気にならない」ということが言える．特に，踊り子が激しく動く顔部分の品質改善が顕著であり，提案方式の有効性が確認できた．

表 4 評価結果の平均値

Table 4 Mean rating results.

配信方式	Q1	Q2	Q3
提案方式	3.3	3.3	3.7
従来方式	1.6	2.5	1.1

7. おわりに

本報告では，ハイビジョンを遥かに超える超高精細パノラマ映像を一般家庭に普及している回線およびハイビジョンテレビで視聴するためのパノラマ映像配信方式について提案した．従来，超高精細パノラマ映像をタイル分割し，配信する方式が提案されているが，各タイルを一様なビットレートで符号化すると映像品質に空間的なムラができ主観品質が低くなるという課題があった．そこで，本稿ではデータフォーマットに MVC を利用し，各タイルを複数のビットレートで符号化，多重化し，目標配信レート以下で主観映像品質を最大化することができるシステムの提案を行った．主観評価実験の結果，従来方式より提案方式のほうがより高い主観品質を達成できた．特に，踊り子が激しく動く顔部分の品質改善が顕著であり，提案方式の有効性が確認できた．

今後は，注視領域が拡大・縮小，左右移動する場合についての検討を進める予定である．また，スポーツ映像など他の映像についても適用範囲を広げていく予定である．

謝辞 本評価実験に協力いただいた NTT サイバースペース研究所の田中康暁氏と山口好江氏に謹んで感謝の意を表する．

参考文献

- 1) http://av.watch.impress.co.jp/docs/news/20090519_169530.html
- 2) <http://allkyoto.picky.or.jp/pdf/fujii.pdf>
- 3) http://flets.com/campaign/campaign_opt.html?banner_ID=200001
- 4) 木全英明，石橋聡：インタラクティブ・パノラマ映像システム，信学誌 Vol.93, No.5, pp. 387-391 (2010)
- 5) J. Foote and D. Kimber, “Flycam: Practical panoramic video and automatic camera control,” in IEEE International Conference on Multimedia and Expo. ICME 2000, NY, vol.3, July 2000, pp.1419-1422.
- 6) J. Foote and D. Kimber, “Enhancing Distance learning with Panoramic Video,” in Proc. of the 34th Hawaii International Conference on System Sciences, 2001, pp.1-7.
- 7) K. Ng, S. Chan, and H. Shum, “Data Compression and Transmission Aspects of Panoramic Videos,” IEEE Transactions On Circuits and Systems for Video Technology, vol.15, no.1, pp.82-95, Jan. 2005.
- 8) B. Y. Kim, K. H. Jang, S. K. Jung, “Adaptive Strip Compression for Panoramic Videos,” Proc. of the Computer of the Computer Graphics International, Jun. 2004.