

インタラクティブ・パノラマ映像配信システム におけるマルチ解像度配信方式の一検討

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。筆者ら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部分配信方式について提案してきた。しかしながら、この部分配信方式では、クライアントが配信されていない部分映像を新規に要求してから表示するまでに映像が途切れ、黒表示になってしまう問題があった。本稿では、オリジナルのパノラマ映像を複数の解像度に解像度変換し、高解像度映像を部分配信し、低解像度映像を全配信することで、上記黒表示を回避するマルチ解像度配信方式について提案し、主観評価実験の結果を示す。

A Study on Multi-Resolution 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. We have proposed a new data format and partial delivery method to achieve high-quality partial panoramic video transmission even over restricted bandwidth networks. However, this partial delivery method yields visual interruption because the desired view is displayed some time after the user's request is submitted. Our solution, proposed in this paper, transforms the original panoramic video into multi-resolution panoramic video and transmits the high-resolution partial video and a complete low-resolution video in order to avoid this visual disruption. We show the experiments conducted on the multi-resolution delivery method.

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ユーザー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。過去に、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が、対象とする映像はHD以下であり、我々が目標とするHDを遥かに超える超高精細映像を伝送することは困難である。

一般的な配信方法として、超高精細映像を小さなタイル映像に分割し、すべてのタイル映像を伝送する方法と、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) は前者の方法であり、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、すべてのスライス映像を伝送する。そして、ユーザの注視領域を含むスライス映像のみを復号化・表示する。文献 8) は後者の方法であり、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、これらの方法では、タイル映像の特徴によらず、一様なビットレートで符号化されるため、動物体を含むタイル映像と背景などの静止領域のみのタイル映像では符号化後の映像品質が大きく異なるという問題があった。この問題に対処するため、我々は各タイル映像を複数のビットレートで符号化・多重化することで作成するデータフォーマットとユーザの注視領域の主観映像品質を最大化する部分配信方式の提案を行ってきた⁹⁾。しかしながら、この部分配信方式では、クライアントが配信されていない部分映像を新規に要求してから表示するまでに映像が途切れ、黒表示になってしまう問題があった。

本報告では、オリジナルのパノラマ映像を複数の解像度に解像度変換し、高解像度映像を部分配信し、低解像度映像を全配信することで、上記黒表示を回避するマルチ解像度配信方式について提案し、主観評価実験の結果を示す。

[†] 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. 提案システム概要

2.1 システム構成

図1に提案システム構成を示す. 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, オリジナル映像を複数の解像度に変換した後に, タイル分割する. ここでは最低の解像度の映像をナビ映像とし, それ以外の解像度の映像をメイン映像とする. その後, メイン映像における各解像度の各タイル映像を複数のビットレートで符号化したメイン・ストリームと, ナビ映像を単一のビットレートで符号化したナビ・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, 符号化レートと PSNR の対応テーブル (以下, 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 Coding)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てからサーバへ格納される. 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.

図2は, クライアントの GUI を示しており, ナビ映像, 部分パノラマ映像, コントロールダイアログから構成される. クライアントでは, パノラマ映像ストリームを受信する前にレート品質対応テーブルを受信し, この対応テーブルを利用し ROI を含むタイル領域の映像品質が最大となるようにタイル映像を要求する. ナビ映像上でのユーザの ROI 指定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 ROI を含むタイル映像と符号化レートにより関連づけされた 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.

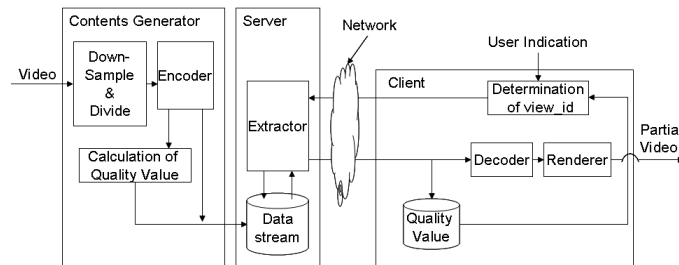


図1 提案システム構成図

Figure 1 Proposed partial video streaming system architecture.

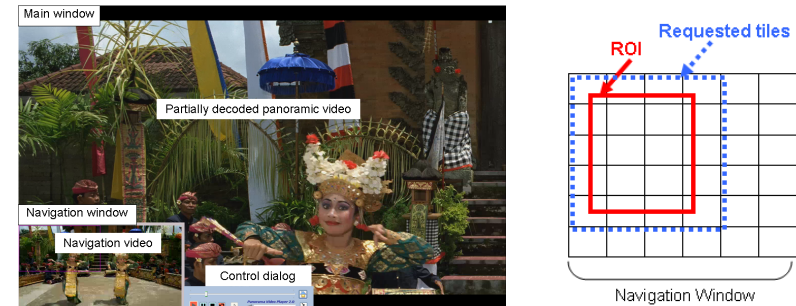


図2 クライアントの GUI

Figure 2 GUI image of the client.

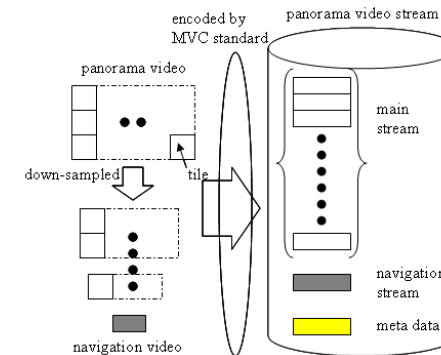


図3 パノラマ映像のためのデータ・フォーマット

Figure 3 Data format for panoramic video.

2.2 データフォーマット

図3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, オリジナルのパノラマ映像とオリジナルを複数の解像度に変換した複数解像度のパノラマ映像から生成され, ナビゲーション・ストリームは解像度変換後の最低の解像度のナビ映像からのみ生成される. ここで, ナビ映像はユーザの ROI 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.

メイン・ストリームでは、複数解像度のパノラマ映像は、解像度ごとに $M \times N$ にタイル分割され、図 4(a)に示すように各タイル映像に `tile_id` が付与される。きれいにタイル分割できない複数解像度のパノラマ映像の両端部分は黒でパディングされる。メイン・ストリームには、この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 4(b))。

メタデータには、図 4(a)の `tile_id` マップ、図 5 に示す `tile_id` と `view_id` の対応テーブル、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。

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、黒パディング部分は PSNR の算出には含まない。また、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。

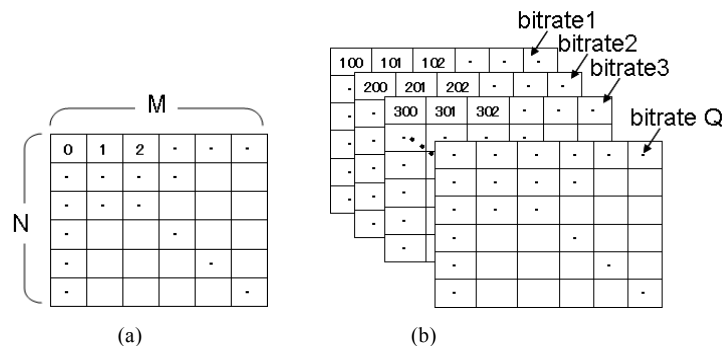


図 4 (a) パノラマ映像の `tile_id` マップ
 (b) 各ビットレートの `view_id` マップ

Figure 4 (a) `Tile_id` map of panoramic video.
 (b) `View_id` map at each bitrate.

view_id					
tile_id	bitrate1	bitrate2	bitrate3	...	bitrateQ
0	100	200	300
1	101	201	301
2	102	202	302
...					...
...

(a)

PSNR					
tile_id	bitrate1	bitrate2	bitrate3	...	bitrateQ
0	PSNR100	PSNR200	PSNR300
1	PSNR101	PSNR201	PSNR301
2	PSNR102	PSNR202	PSNR302
...					...
...

(b)

図 5 (a) タイル・ビュー対応テーブル
 (b) レート品質対応テーブル

Figure 5 (a) `Tile-view` mapping table.
 (b) `Rate-quality` mapping table.

3. 配信方式の提案

3.1 クライアント処理

ここではクライアント処理について説明する。クライアントは、映像ストリームを受信する前に、まず図 4(a)の `tile_id` マップ、図 5 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 ROI 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できる。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。この要求と同時に、ROI 移動時の黒表示を回避するため、低解像度の全体パノラマ映像を最低のビットレートでサーバへ要求する。

3.2 `View_id` の決定方法

図 6 に `View_id` の決定フローを示す。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、ROI 全体の主観映像品質を上げる。これは、最も映像品質の悪いタイルは ROI 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からである。

提案する配信方式では、まず、低解像度のパノラマ映像を構成するすべてのタイル映像について最低のビットレートを割り当てる。次に、ROI に含まれるタイルを特定する。続いて、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する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を実現する。

ここでは、提案方式 1 を部分的な高解像度パノラマ映像をマルチビットレートで配信する方式、提案方式 2 を部分的な高解像度パノラマ映像をマルチビットレートにより配信すると同時に、低解像度の全体パノラマ映像も配信する方式とする。つまり、前者の view_id の決定フローは、図 6 から「低解像度に最低のビットレートを割り当て」の処理を除いたものとなる。後者の view_id の決定フローは図 6 となる。

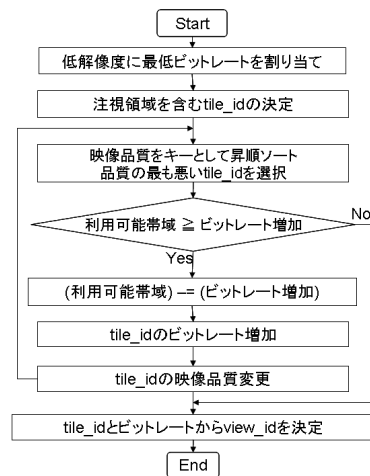


図 6 view_id の決定フロー

Figure 6 Determination flow of view_id.

4. 評価実験

今回の評価実験では、配信方式と ROI 移動の観点から評価した。具体的には、3 種類の配信方式を用いて、ROI パンニング移動時の映像品質について主観評価を実施した。

4.1 被験者

被験者は、男性 10 名、女性 10 名の合計 20 名。年齢は 20 代から 50 代を対象とし評価実験を実施した。

4.2 実験環境

表 1 に示すように、映像配信のためのネットワークには LAN、テスト映像の評価には 37 インチ LCD を用いた。視距離は 3H (138cm) とした。評価室の照明は一般的な白色蛍光灯を用いた。

4.3 実験方法

表 1 に示すように、目標配信レートを 15Mbps に設定して 3 種類の配信方式について比較実験を行った。テスト映像は 2 種類、注視領域サイズは HD であり、テスト映像注視時はスクリプトにより ROI を自動的にパンニング移動させた。

テスト映像については、図 7 に示す 'Dance' と 'Soccer' を用いた。Dance では、2 人の踊り子がカメラの正面で踊るシーンであり、Soccer では、サッカーの試合を俯瞰できる位置から撮影しているシーンである。つまり、被写体のサイズが大きく異なるテスト映像を用意した。表 2 にテスト映像作成時の実験パラメータ設定内容を示す。Dance のオリジナル (高解像度) のパノラマ映像サイズは 3840×1920 画素であり、低解像度のサイズを 1920×960 画素とした場合、タイルサイズ 640×320 画素で分割すると、高解像度のパノラマ映像は横 6 タイル×縦 6 タイル、低解像度は横 3 タイル×縦 3 タイルで表現できる。高解像度のパノラマ映像を表 1 の目標配信レート 15[Mbps]以下とするため、従来方式によるタイルのビットレートは、416[kbps]とした。提案方式 1,2 によるタイルのビットレートは、500 [kbps],1 [Mbps],1.5 [Mbps]とした。Soccer のオリジナル (高解像度) のパノラマ映像サイズは 6400×1280 画素であり、低解像度のサイズを 3200×640 画素とした場合、タイルサイズ 640×320 画素で分割すると、高解像度のパノラマ映像は横 10 タイル×縦 4 タイル、低解像度は横 5 タイル×縦 2 タイルで表現できる。高解像度のパノラマ映像を表 1 の目標配信レート 15[Mbps]以下とするため、従来方式によるタイルのビットレートは、375[kbps]とした。提案方式 1,2 によるタイルのビットレートは、Dance と同様に 500 [kbps],1 [Mbps],1.5 [Mbps]とした。

配信方式については、高解像度の全体パノラマ映像 (メイン・ストリーム) を 15Mbps 以下となるように単一ビットレートで全配信する従来方式、高解像度のパノラマ映像

を部分配信する提案方式 1, 高解像度の部分映像と低解像度の全体映像を配信する提案方式 2 を比較に用いた.

各被験者は, テスト映像を 10 秒間注視した後に, 表 3 の質問紙を用いて主観評価を行った. なお, 注視時の ROI サイズは HD 固定, ROI 位置は右へ 1920 画素移動するパンニングとした. 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 に近い評価とした.



(a) Dance

(b) Soccer

図 7 テスト映像のユーザ注視領域
 Figure 7 User's view of test sequences.

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]
比較方式	従来方式 (単一解像度配信, 単一ビットレート方式) 提案方式 1 (単一解像度配信, マルチビットレート方式) 提案方式 2 (複数解像度配信, マルチビットレート方式)
テスト映像	Dance, Soccer
ROI 移動	パンニング
ディスプレイ	シャープ AQUOS37 インチ LCD(LC-37GX5)
ROI サイズ	HD(1920x1080[pixels])
視距離	3H (138cm)

表 2 (a)Dance の実験パラメータ設定
 (b)Soccer の実験パラメータ設定

Table 2 (a) Experimental parameter sets for Dance.
 (b) Experimental parameter sets for Soccer.

(a) Dance		
パノラマ映像サイズ	オリジナル(高解像度)	3840x1920 [pixels]
	低解像度	1920x960 [pixels]
コンテンツ長		10 [sec]
タイルサイズ		640x320 [pixels]
タイル分割数 (高解像度, 低解像度)		6x6, 3x3 [tiles]
従来方式によるタイルのビットレート		416 [kbps]
提案方式によるタイルのビットレート		500 [kbps],1 [Mbps],1.5 [Mbps]

(b) Soccer		
パノラマ映像サイズ	オリジナル(高解像度)	6400x1280 [pixels]
	低解像度	3200x640 [pixels]
コンテンツ長		10 [sec]
タイルサイズ		640x320 [pixels]
タイル分割数 (高解像度, 低解像度)		10x4, 5x2 [tiles]
従来方式によるタイルのビットレート		375 [kbps]
提案方式によるタイルのビットレート		500 [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?

5. 評価実験結果と考察

被験者 20 名を対象とした, 2 種類のテスト映像, 3 つの配信方式による比較評価実験の結果, 各質問に対して 20x2x3=120 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.

テスト映像 Dance に関して, 表 4 より, Q1,Q2,Q3 において提案方式 1 の評価結果が最もよく, 次に提案方式 2, 続いて従来方式の順であった. 提案方式 2 では ROI のパンニング時に低解像度の映像が一部表示されることや, 高解像度映像と同時に低解像度映像が配信されているため, 高解像度のビットレートが不足し, 提案方式 1 より映像品質が劣化するためと思われる. 表 2 より, 低解像度の映像を全て配信するためには $500[\text{kbps}] \times (3 \times 3[\text{tiles}]) = 4.5[\text{Mbps}]$ のビットレートが必要となり, 高解像度の映像の配信レートは $15[\text{Mbps}] - 4.5[\text{Mbps}] = 10.5[\text{Mbps}]$ となることが分かる.

テスト映像 Soccer に関しては、Dance とは異なる傾向を示している。これは、Dance の被写体サイズが大きく、Soccer が小さいことが評価に影響していると思われる。

Q1 に関して、提案方式 1 の評価結果が最もよく、次に提案方式 2 と従来方式が同程度の評価結果を得た。1920x1080 画素固定の ROI 移動がある場合、ROI の位置によりサーバへの要求タイル数が増える。ベストケースの場合、要求タイル数は 12 タイル、ワーストケースの場合は 20 タイルとなる。表 2 のパラメタ設定では、低解像度の映像を全て配信するためには $500[\text{kbps}] \times (5 \times 2[\text{tiles}]) = 5.0[\text{Mbps}]$ のビットレートが必要となり、高解像度の映像の配信レートは $15[\text{Mbps}] - 5[\text{Mbps}] = 10[\text{Mbps}]$ となることが分かる。よって、ワーストケースの場合は高解像度のタイル映像の要求時のビットレートは 500kbps となり、ビットレートの点でも従来方式の 375kbps と同程度となってしまふ。ゆえに、従来方式と提案方式 2 が同程度の評価になったと思われる。

Q2 に関して、提案方式 1 の評価結果が最もよく、次に従来方式、提案方式 2 の順であった。Q3 に関しては、提案方式 1 と従来方式が同程度によく、次に提案方式 2 の順であった。提案方式 2 のみ ROI 移動中に、被写体が小さい、低解像度の映像が表示されるため、鮮明度が失われブロック歪が目立つ結果になったと思われる。

表 4 評価結果の平均値
 Table 4 Mean rating results.

配信方式	Dance	Soccer
Q1		
提案方式 2 (低解像度配信あり)	2.10	2.80
提案方式 1 (低解像度配信なし)	3.20	3.20
従来方式	1.65	2.75
Q2		
提案方式 2 (低解像度配信あり)	2.60	2.50
提案方式 1 (低解像度配信なし)	3.30	3.05
従来方式	2.25	2.80
Q3		
提案方式 2 (低解像度配信あり)	1.55	2.65
提案方式 1 (低解像度配信なし)	3.00	3.55
従来方式	1.30	3.35

6. おわりに

本報告では、オリジナルのパノラマ映像を複数の解像度に解像度変換し、高解像度映像を部分配信し、低解像度映像を全配信することで、黒表示を回避するマルチ解像度配信方式について提案した。主観評価実験の結果、被写体が大きい場合、提案方式は従来方式より主観映像品質に関して高い評価を得ることができた。しかし、低解像度配信をする場合は低解像度配信にビットレート割当がされてしまうと十分な映像品質が得られなくなることがわかった。

今後は、高解像度映像と低解像度映像とのビットレート比率、最適な低解像度映像のサイズや MVC 符号化時のビットレート段階について検討する予定である。また、提案システムでは、クライアント側で ROI が移動することを特徴としており、ROI 移動中の映像品質は ROI 静止時の映像品質と比較して劣化する。ROI を移動できるメリットと ROI 移動中の映像品質劣化はトレードオフの関係にあり、従来とは異なる観点の評価手法が今後必要となる。

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に謹んで感謝の意を表す。

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.
- 9) 井上雅之, その他: 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, 情処研報 Vol.2010-AVM-69, No.9 (2010)