

多画面映像システムにおける多視点映像符号化の一検討

木全英明 志水信哉 上倉一人 八島由幸

日本電信電話株式会社 NTT サイバースペース研究所 〒239-0847 横須賀市光の丘 1-1

E-mail: {kimata.hideaki, shimizu.shinya, kamikura.kazuto, yashima.yoshiyuki}@lab.ntt.co.jp

あらまし 臨場感の高い映像体験を実現する手法として複数の画面を利用した映像表現が目ざされている。例えば複数の画面を利用したテレビ会議システムが実用化されている。画面数が多くなると、処理する画像情報も多くなり、映像信号を効率よく圧縮符号化することが望まれる。また、高い圧縮効率だけでなく、ROI といった機能を備えることも望まれる。本稿では、現在国際標準化が進められている多視点映像符号化 MVC を拡張した通信システムと映像符号化の提案を行う。

キーワード 多画面, 多視点, 映像符号化, MVC, ROI

A Study on Multi-view Video Coding for Multiple Display System

Hideaki KIMATA Shinya SHIMIZU Kazuto KAMIKURA and Yoshiyuki YASHIMA

NTT Cyber Space Laboratories, NTT Corporation 1-1 Hikari-no-oka, Yokosuka, 239-0847 Japan

E-mail: {kimata.hideaki, shimizu.shinya, kamikura.kazuto, yashima.yoshiyuki}@lab.ntt.co.jp

Abstract The multiple display system is becoming an expected solution for providing realistic visual experiences. For instance, a video conference system which uses multiple displays is produced in the market. High efficient video coding method is expected for such a multiple display system because much more image information should be transmitted. In addition to the compression efficiency, a particular functionality such as ROI (Region Of Interest) is also expected. We propose a visual communication system and video coding method which extends the upcoming International Standard video coding MVC (Multi-view Video Coding) for multiple display systems.

Keyword multiple display, multi-view video, video coding, MVC, ROI

1. まえがき

私達の生活の中で複数の映像ディスプレイが使われるようになってきた。家にはテレビがあり、常時持つ携帯電話にも大型のディスプレイが載るようになり、また携帯型のゲーム機でも映像を楽しむことができるようになってきた。更には、複数のディスプレイが連動する多画面映像システムが用いられるようになってきている。例えば、PC は 2 画面表示が出来るようになり、ゲーム機においても 2 画面を使って楽しむ機器が世の中に出てきている。複数のディスプレイを結合して大画面化や広角化の効果を得られる他、ゲーム機ではタッチパネル型ユーザインタフェースとしても利用されるなど、多画面映像ディスプレイは臨場感を向上する効果があると言える。今後、映像コンテンツもこのような多画面ディスプレイを利用して楽しむようになることが期待される。

多画面ディスプレイを使った映像通信や映像配信も提案されている。映像通信については 3 画面のディスプレイを使うテレビ会議システムが提案されている [1]。また映像配信においても複数の画面を結合して一つの画面として利用するシステムが利用されている [2]。このようなシステムでは画面を増やすことで見ら

れる映像シーンの範囲を広げる効果があり、これにより臨場感の高い映像を提示する。更に多画面映像システムにおいて、受信者がインタラクティブ操作をシステムに対して行うことでより臨場感を高めることができる。例えば追加された画面の方に特定の領域の拡大画像を表示しても良い。実際にスポーツ中継では一部の拡大画像が要望されたとの報告がなされている。本研究では、このような受信側がインタラクティブに特定の領域を切り出して表示する映像システムの実現を目指す。

このような画面の一部のみを表示する際に、画面全体を復号するのではなく所望の一部の領域の画像情報のみを復号する機能を ROI (Region Of Interest) と呼ぶ。ROI 機能を持つ映像符号化については、一つの映像が複数の分割映像で構成される仕組みが例えば [2] で提案されている (図 1)。映像符号化には JPEG2000 が提案されており、JPEG2000 のタイル機能により ROI を実現している。また大画面を広角にした全方位映像においても同様に ROI 機能が必要とされる。ROI 機能を備えた全方位映像の符号化にも JPEG2000 が提案されている [3]。また空間スケーラブル符号化と組み合わせた方法が提案されている (図 2) [4][5]。図 1 のように小領



Fig. 1: Large display by multiple display systems

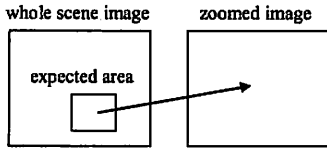


Fig. 2: Spatial scalable image system

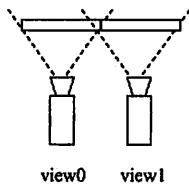


Fig. 3: Segmented images for large image

域に分割して符号化すれば一部のみを復号することが可能となるが、分割することで画面全体の符号化効率が低下する問題がある。また空間スケーラブル符号化と組み合わせる方法では、必ず全画面に相当するベースレイヤを復号する必要があるため、所望の領域を得るために必要な符号量が多い問題がある。

本稿では、多画面システムにおいて図1に示すような分割映像を利用したROIを実現するために、多視点映像符号化 MVC (Multi-view Video Coding)を拡張したシステム構成と符号化方式の提案を行う。提案方式では分割映像を一つの視点映像(view)とみなして符号化する(図3)。図1における1画面を1つのviewとみなして符号化することとなり、図3に示すようにちょうどview間でオーバーラップする領域がなくなった形が多視点映像に相当する。

2. 想定する多画面映像通信システム

本研究で想定する多画面映像通信システムを説明する。送信側からはライブ映像を配信し、受信側では通常の映像(受信者が領域を指定しない映像)と“所望”の領域の映像を共に受信して、多画面で視聴するようなシステムを考える。ここで“所望”の領域は映像送信側で複数の候補を用意しておき、受信側が候補を選択するものとする。図4に想定するシステムの概要を示す。撮影された映像信号は送信側で符号化され、サーバで配信される。受信側は2つの画面を備えており、それぞれ映像符号化データを受信して再生する。こ

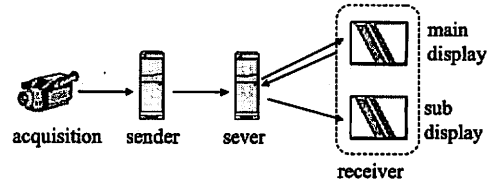


Fig. 4: Supposed multiple display transmission system

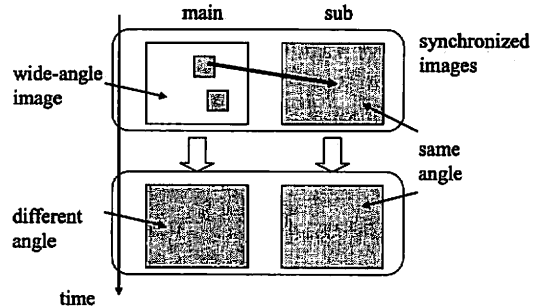


Fig. 5: Usage of multiple display images

で再生される映像は画面間で同期が取れている。

この2画面のうち一つはメイン画面であり通常の映像を表示し、もう一つはサブ画面であり“所望”の映像を表示する。すなわちサブ画面に表示するのは視聴者が選択した領域の拡大映像である。このように2画面に表示する映像は同じコンテンツの一部とする。メイン画面とサブ画面では、図5に示すように、あるシーンの広角映像と拡大映像の組み合わせを表示したり、あるいは異なる視点からの映像を表示する。多画面映像システムでは、このようにサブ画面に“所望”の映像を表示することを特徴とする。

またサーバから受信側へ送信される映像符号化データは1ストリームで構成されており、このような2画面の映像を表示するためにスケーラブルな構成になっている。

3. 提案映像符号化方式

3.1. 多視点映像符号化方式 MVC の概要

本稿で提案する符号化方式は多視点映像符号化方式 MVC を応用する。MVC は H.264/AVC の拡張規格として、現在 MPEG(ISO/IEC)と ITU-T が共同で標準化を進めている。ここで MVC の概要を説明する。MVC では複数のカメラ映像を同期して並列に復号できるような仕組みが採用されている。符号化効率を改善する符号化方式も提案されているが、詳細は本稿では割愛する。ここでは同期して並列に復号するために採用されている技術を説明する。

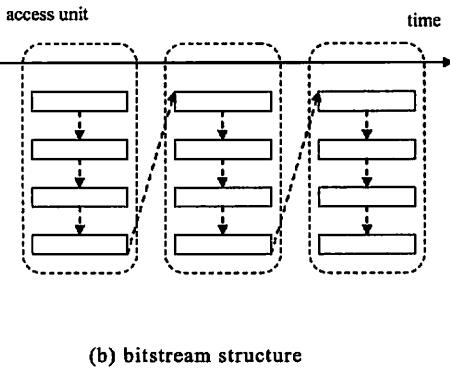
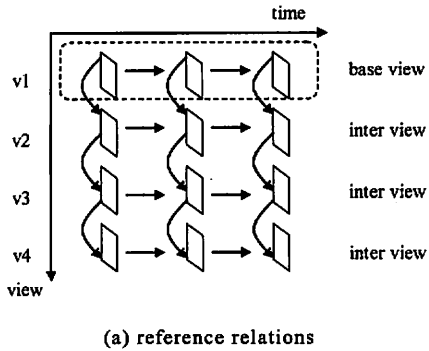


Fig. 6: MVC reference relations and bitstream structure

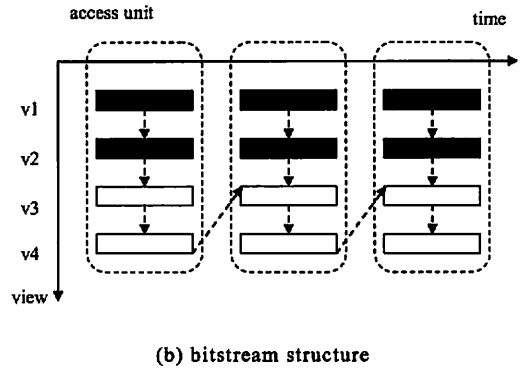
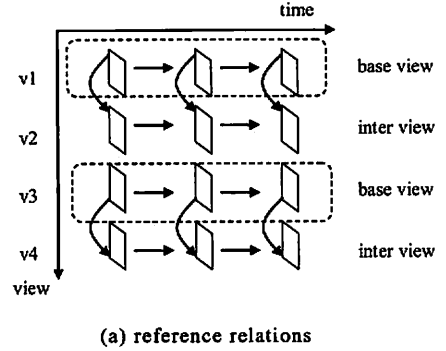


Fig. 8: MVC reference relations and bitstream structure for two base views

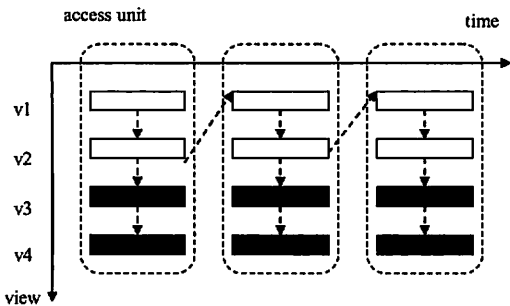


Fig. 7: MVC bitstream structure for view scalability

並列処理をするためにフレームの復号順序について、視点映像(view)の中で同一時刻のフレームを連続して復号するように制限を加えている。例えば図 6(a)に示すような 4view がある場合に、符号化データは図 6(b)のような構成になる。ここで図 6(a)の方の実線矢印はフレーム間の参照関係を表し、図 6(b)の方の破線矢印はフレームの復号順序を表す。Base view は同一 view の復号画像のみを参照する、すなわち時間方向の予測のみする view を表し、inter view は更に他の view の復号画像も参照する view を表す。Inter view では視差補償が適用可能となり、base view よりも符号化効率

が高い。しかしながら inter view のフレームを復号するために base view の復号画像が必要である。すなわち inter view の復号には依存する base view が必要である。図 6(a)で view3 と 4 を復号せずに view1 と view2 のみを復号することが可能である。この場合には図 7 に示すように、view3 と view4 の符号化データをスキップして復号できる。このように MVC は view の scalability を備える。更に図 8(a)に示すように、base view を複数設定することができる。この場合には view3 と view4 のみ復号することが可能である。図 8(b)に示すように view1 と view2 の符号化データをスキップして復号することができる。

MVC は H.264/AVC の拡張方式であるため、同じく拡張方式である SVC (Scalable Video Coding)とも組み合わせることが可能である。SVC では H.264/AVC をベースレイヤにして、空間スケーラブル機能や SNR スケーラブル機能を実現する。MVC と SVC を組み合わせる場合には、ある view について空間スケーラブル機能を有するように符号化することが出来る。この場合には、その view の符号化データ中に空間スケーラブル用の符号化データを更に加える(図 9)。この場合には view 間の依存関係だけではなく layer 間の依存関係も発生

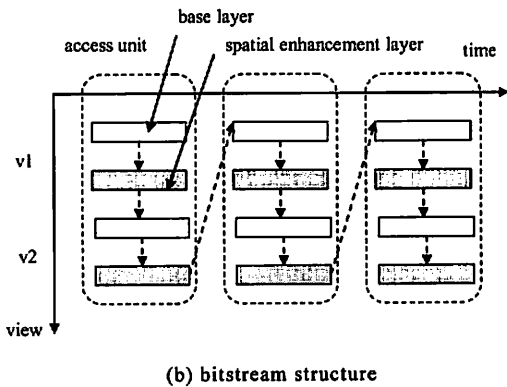
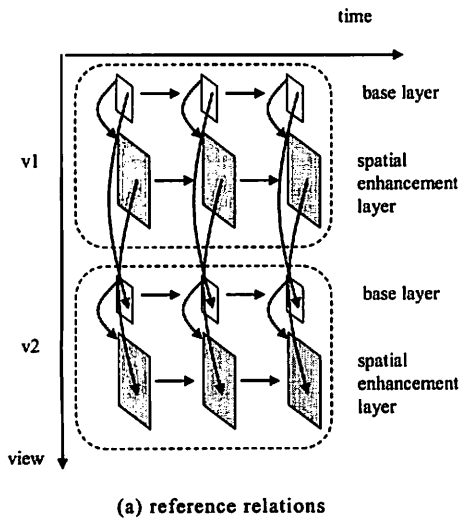


Fig. 9: MVC and SVC combined case

する。なお空間解像度については、MVCの符号化データ中に異なる空間解像度の view を含むようにしても良い。

View 間や layer 間の依存関係は SPS (Sequence Parameter Set)で記述される。この SPS は GOP 単位に変更することが可能である。スライスで通信の packets を構成し、そのヘッダには view や layer を識別する ID がふられる。本稿で想定する通信システムでは、サーバはこの ID に従って packets の送信を制御すれば良い。

3.2. 分割映像用の多視点映像符号化の提案

本稿では多画面映像システムにおいて、サブ画面に ROI した映像を復号して表示するために、図 3 に示し

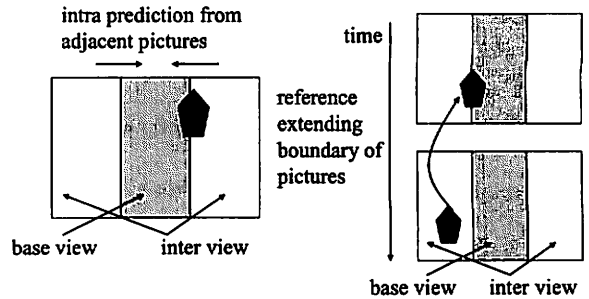


Fig. 10: Prediction from boundaries

たように分割映像を view とみなして符号化する場合の符号化方式を提案する。分割映像を個別に符号化すると符号化効率が低下する問題がある。そこで分割映像間に予測符号化を適用する符号化方式を提案する。提案方式では MVC を拡張して分割映像を符号化する。

提案方式の手順の例として、3つの分割映像がある場合の符号化方法を示す。基本的には MVC と同様に base view の画像情報を符号化した後に、inter view を符号化する。Inter view を符号化するには base view の復号画像信号を利用する。Base view の画像信号の利用方法により適用される符号化方法が異なる。まず先頭フレームでは、Base view の復号画像信号はイントラ予測とフレーム間予測に利用することが出来る。イントラ予測については base view と inter view の幾何情報を利用して、base view の外周画像情報から予測画像を作りイントラ予測をする(図 10(a))。フレーム間予測については単純に base view の復号画像を参照画像として動き補償を適用する。次に先頭フレーム以外については、base view と inter view の過去のフレームを参照画像にして動き補償をする。過去のフレームの復号画像であるため、base view と inter view では並列に復号することができる。Base view は base view の復号画像のみを使い動き補償をする。Inter view は inter view と base view の復号画像を使い動き補償をする。Base view と inter view の過去のフレームを単純に異なる参照画像として設定しても良いが、base view と inter view の幾何情報を利用して領域をまたがるような動き補償をすることもできる(図 10(b))。ここで、現在の MVC では他の view の復号画像のうち、同一時刻の画像のみ予測符号化に利用できる。従って、本稿で提案する方式では、同一時刻だけではなく、過去のフレームも参照できるように MVC を拡張する。

Table 1: MB mode of MVC

Code number	Mode
0	SkipMB
1	Inter 16x16
2	Inter 16x8
3	Inter 8x16
4	Inter 8x8
5	Intra 4x4
6...	Intra 16x16

Table 2: MB mode of the proposed method

Code number	Mode
1	Intra 4x4
2...	Intra 16x16
26	SkipMB
27	Inter 16x16
28	Inter 16x8
29	Inter 8x16
30	Inter 8x8

また分割映像を拡大した画像を符号化するために、SVCの空間スケーラブル符号化を利用する。

3.3. 分割映像のフレーム間予測符号化

ここで先頭フレームの符号化において、base viewの復号画像をinter view符号化におけるフレーム間予測に利用する方法について説明する。本稿で想定する図3に示すような分割映像では、視差補償の特徴を踏まえると、分割映像間でオーバーラップする空間の領域が無い場合、分割映像間では画像信号の相関は低いと考えられる。他方で、静止画像におけるフラクタル符号化にもあるように、画像中には自己相似性があるとすると、ブロックレベルでの相関もあると考えられる。ここで提案するフレーム間予測で扱う画像信号の相関は後者の方にむしろ属する。本稿で提案する方式では、base viewの復号画像をinter viewの参照画像に設定して通常の動き補償を行う。従って、フラクタル符号化でも提案されるような平行移動は動き補償で代用される。

次に提案するsyntaxを説明する。MVCを拡張したsyntaxを提案する。具体的にはマクロブロックモード情報を次のように拡張する。提案方式では動き補償を行うため、Pスライス(Pピクチャ)としてinter viewを符号化する。MVCのPスライスにおけるマクロブロックモードを表1に示す。このうちスキップMBは他のマクロブロックモード情報とは別に符号化され、CAVLC(可変長符号化)の場合にはランレングス符号化される。提案方式におけるマクロブロックモード情報を表2に示す。スキップMBは別途符号化せずに、他のマクロブロックモード情報と同じsyntaxで符号化する。提案方式ではIntra予測の方に短い符号長を割り当てる。

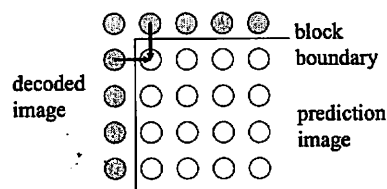


Fig. 11: Prediction in RWIP

スキップMBの時には動きベクトル情報や残差信号情報も符号化しない。MVCでは動きベクトル情報を周囲のブロックから予測して生成し、それを使って予測画像を作成して、予測画像を復号画像とする。スキップMBはフレーム間の相関が高い場合には有効な符号化方法と考えられる。しかしながら、本提案方式ではフレーム間で高い相関を望めない。そこで、スキップMBではイントラ予測を利用する方式を提案する。ここでMVCでは、H.264/AVCと同じく、周囲のブロックの復号画像を利用した方向性予測や平均値予測を行う。しかしながら、方向性予測や平均値予測では周囲のブロックとの連続性が一部で保たれない問題がある。このため、予測残差信号を符号化しないとブロック歪が大きくなる問題がある。そこで、本稿では[6]で提案されたRWIP (Recursively Weighted Intra Prediction)を利用する。図11に示すように、首位のブロックの復号画像と、予測画像を畳み込みながら新たに予測画像を作る。

4. 分割映像をROIする通信方式

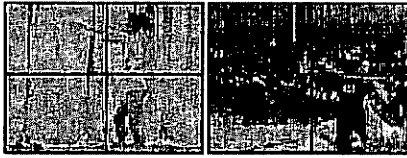
多画面映像通信システムでは、MVCを拡張した方式で符号化された分割映像の符号化データから所望の領域を得るための符号化データを取り出して復号する。この時、どの分割映像の符号化データが必要かについては、MVCにあるSPSの仕組みを利用する。図4に示すシステムの構成では、受信側は所望の分割映像をサーバに要求し、サーバはSPSを解析して、必要な分割映像符号化データを配信する。更に、分割領域を拡大したい場合には更に空間スケーラブル符号化データを要求する。

5. 評価実験

提案する符号化方式の有効性を評価する。本稿ではまず第3.3節で提案した分割映像のフレーム間予測符号化について評価した。符号化条件を表3に示す。分割映像は元々の画像を均等に縦横4分割して作成した(図12)。このうち左上の分割映像をbase viewとし、その他をinter viewとして符号化した。図13に各QPにおける全ての分割映像を個別に符号化した場合からの符号量削減率を示す。このようにQPの値の大きい低

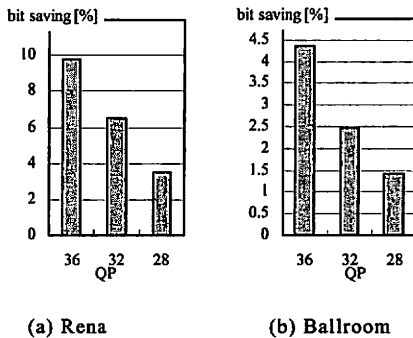
Table 3: Simulation conditions

Base codec	MVC (as extension of AVC)
QPs	36, 32, 28
Mode decision	Rate distortion optimization with Lagrange manipulator



(a) Rena (b) Ballroom

Fig. 12: Test sequences



(a) Rena (b) Ballroom
Fig. 13: Bit saving by proposed method

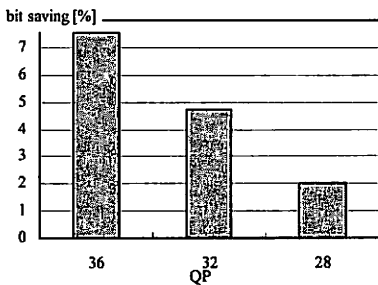
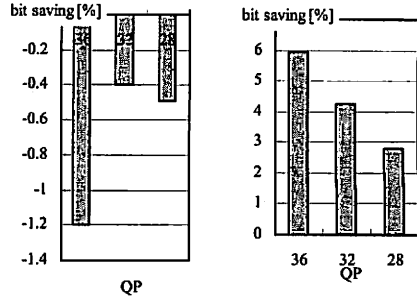


Fig. 14: Bit saving comparing non-divided images

レート付近では約 10%程度の符号量削減率が達成できることが分かる. 図 14 に分割しない場合からの符号量削減率を示す. 提案方式により分割することによる符号量の増加分を超える符号量の削減を達成できることが分かる.

また図 15 に MVC の P スライスのマクロブロックモードを使った場合と, マクロブロックモードのうちスキップ MB 以外を提案方式にしてスキップ MB を符号化しない場合について, base view として符号化した場合からの符号量削減率(符号量増加率)を示す. マイナスの値になっているのは符号量が増加していることを



(a) MVC P-slice (b) No Skip MB

Fig. 15: Bit saving by reference method

示す. 図 15(a) から P スライスのマクロブロックモードをそのまま使うと符号量が増加してしまうことが分かる. 図 15(b) から提案方式でマクロブロックモードを Intra 予測の方が符号量を短くすることで符号量を削減できることが分かる. 図 13 と比較すると RWIP を使うスキップ MB を取り入れることによって更に符号量の削減が達成できることが分かる.

6. まとめ

本稿では多画面映像通信システムを想定して, 有効な映像符号化方式を提案した. 提案方式では画面を分割した分割映像を単位に, 領域を view とみなして多視点映像符号化 MVC を拡張して符号化する. 提案方式のうち, 特にフレーム間予測符号化を適用する部分につき評価実験を行い, 分割しない場合よりも高い符号化効率を実現できることを示した. 今後は, 提案方式のうち動き補償部分の符号化方式とイントラ予測部分の符号化方式を検討する.

文献

- [1] http://newsroom.cisco.com/dlls/2006/prod_102306b.html.
- [2] “超大画面映像システムの高度な適用に関する調査研究報告書,” 機械システム振興協会, Mar., 2003.
- [3] “Report on 3DAV Exploration,” N5878 MPEG document, Jul., 2003.
- [4] 奏泉寺久美, 上倉一人, 岩津茂太郎, 八島由幸, “MPEG-4 ASP/FGS を用いたユーザ主導型 ROI 機能の実現,” D-11-1 電子情報通信学会総合大会, 2005.
- [5] 木全英明, 北原正樹, 上倉一人, 八島由幸, “自由視点映像通信のためのオープンループ・スケーラブル映像符号化に関する一検討,” 信学技報 IE2004-231, pp.49-54, 2005.
- [6] H. Kimata, M. Kitahara, and Y. Yashima, “RECURSIVELY WEIGHTING PIXEL DOMAIN INTRA PREDICTION ON H.264,” SPIE Visual Communication and Image Processing 2003 (VCIP), vol.5150, pp.2035-2042, 2001.