

FTVを実現する3Dシーン表現

3D Scene Representation for FTV



谷本 正幸 (Masayuki TANIMOTO, Ph.D.)

名古屋産業科学研究所 上席研究員
(Senior Research Fellow, Nagoya Industrial Science Research Institute)

映像情報メディア学会 電子情報通信学会 IEEE
表彰：文部科学大臣表彰科学技術賞 電子情報通信学会業績賞 映像情報メディア学会功績賞
研究専門分野：画像情報圧縮 画像処 3次元映像、ITS

あらまし 無限の視点を持ち3次元空間情報の全てを伝達する自由視点テレビFTV (Free-viewpoint Television)は、ユーザが視点を自由に移動させて臨場感あふれる3次元シーンを見ることができる没入型3次元(3D)メディアであり、世界の映像メディアの頂点に立つ。現在のFTVは、MVD (Multi-view plus depth)と呼ばれるデータ形式で3次元シーンを表現し、depth情報を利用してview情報から自由視点画像を生成している。本研究では、新しい3次元シーン表現としてMVS (Multi-view plus surface)を提案し、表面情報を利用して画像生成を行うFTVの新しいフレームワークを構築した。これにより、高精度な奥行情報の取得と高品質な画像生成が行える。さらに、物体の反射特性や透明度も付加データとして持つことができるため、鏡面反射や半透明のシーンも表現でき、照明条件の変更なども可能となる。

1. 研究の目的

自由視点テレビFTV (Free-viewpoint Television) [1]-[4]は、あたかもその場にいるかのように、自由に視点を変えて遠隔地の情景を見ることができる映像メディアである。無限の視点を持ち、好きな位置に視点

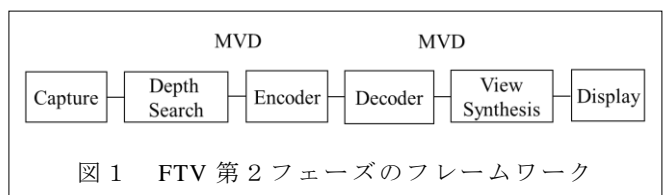
を置くことができるFTVは、臨場感あふれる3次元シーンを見ることができる没入型3次元(3D)メディアであり、世界の映像メディアの頂点に立つ。

FTVを実現するには、3D空間のあらゆる地点の光線情報が必要であるが、あらゆる地点にカメラを設置することはできない。このため、FTVでは限られた地点にカメラを設置し、そのカメラが取得した光線情報からその他の地点の光線情報を補間して自由視点画像を生成している。この画像生成は3Dシーンの表現方式と密接に関係している。本研究では、新しい3Dシーン表現方式を研究し、それに基づくFTVの新しいフレームワークを構築することを目的とする。

2. 研究の背景

研究代表者は2001年にFTVを動画像の国際標準を定めるMPEG (Motion Picture Experts Group)に提案し、フェーズ毎にその国際標準化を行ってきた。FTVの第1フェーズでは、多数のカメラで撮影した多視点映像の圧縮符号化MVC (Multiview Video Coding)[5]の標準化を実現した。第2フェーズでは、多視点3D表示のため、奥行き情報を利用して少数の多視点映像から視点合成を行う仕組みを導入し、多視点映像と奥行き情報の統合圧縮符号化技術3DV (3D Video) [6][7]の標準化を行った。2013年8月には、3Dシーン鑑賞を革新するMPEG-FTVと名付けたFTVの第3フェーズ標準化[8]を開始した。MPEG-FTVでは、超多視点3Dとフリーナビゲーション、及び全方位3Dの実現を目指している[9]。

FTVの第2フェーズで採用されたFTVのフレームワークを図1に示す。3DシーンをMVD (Multi-view plus depth)[6]と呼ばれるデータ形式で表現し、奥行き情報を利用してview情報を3D warping[10]することにより、自由視点画像を生成する。このとき、奥行き推定参照ソフトウェアDERS (Depth Estimation Reference Software)[11]と画像合成参照ソフトウェアVSRS (View Synthesis Reference Software)[12]が用いられる。



FTVを実現する3Dシーン表現

3D Scene Representation for FTV

FTVの第3フェーズで行った探索実験[13]の結果、鏡面反射や半透明のシーンを含むアプリケーションにはMVD方式に基づくフレームワークでは十分に対処できないことが明らかとなり、新たな3Dシーン表現方式が必要となった。

3. 3Dシーン表現方式の比較検討

FTVをどのようにして実現するかは、3Dシーンの表現方式と密接に結びついている。これまでの3Dシーン表現方式には、イメージベースの光線空間からモデルベースまでの様々なものがある。

3.1. 光線空間

光線空間法では、3D実空間の1本の光線を、それを表すパラメータを座標とするパラメータ空間の1点で表す。このパラメータ空間が光線空間である。光線空間として直交光線空間と極座標光線空間を定義した。直交光線空間は3Dシーンを一方向から見るときに、極座標光線空間は3Dシーンを様々な方向から取り囲むようにして見るときに用いる。

光線空間は、多くの視点から撮影した画像を集めることによって作られる。垂直方向の視差を無視した場合、直線配置カメラ画像によって直交光線空間、円形配置カメラ画像によって極座標光線空間が得られる。

直交光線空間の構造を図2に示す。垂直断面が2次元画像で、水平断面が直線構造となることが特徴である。図3に示すように、直交光線空間を垂直に切るとその断面が自由視点画像となる。視点を指定すると切断する位置が定まる。異なる視点での自由視点画像生成は切断面を変えることで容易に実現できる。多眼式3次元ディスプレイに表示するには、同時に眼数分の平行した断面を作る。

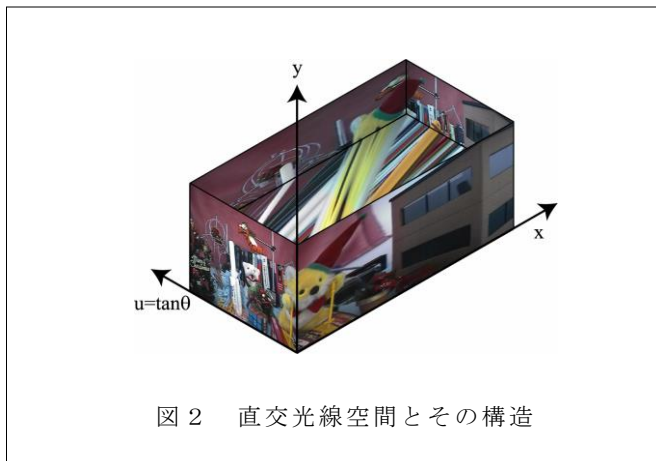


図2 直交光線空間とその構造

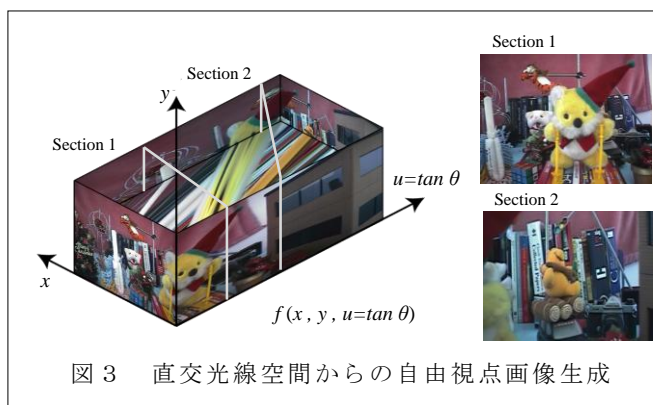


図3 直交光線空間からの自由視点画像生成

極座標光線空間の構造を図4に示す。極座標光線空間の水平断面は正弦波構造となる。図5に示すように、極座標光線空間から自由視点画像を生成するには、正弦波曲面で切断する。

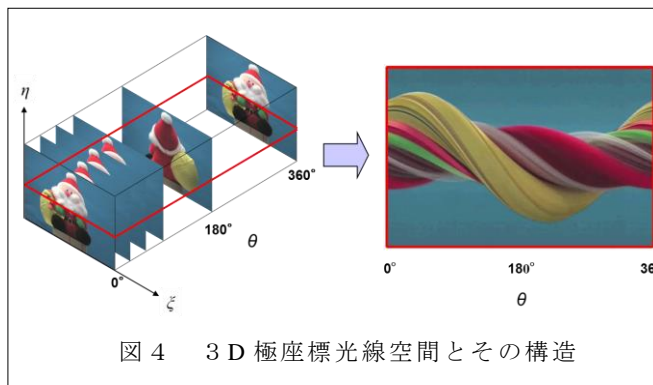


図4 3D極座標光線空間とその構造

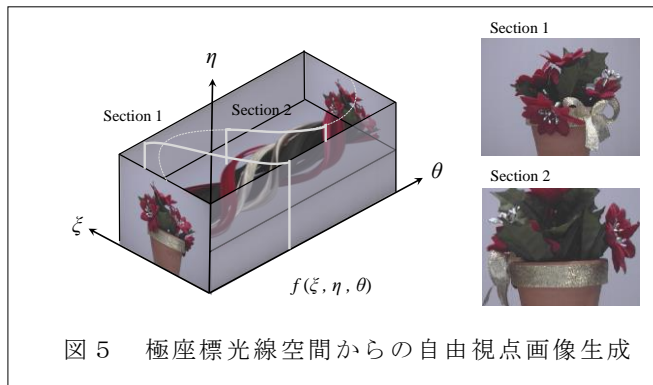


図5 極座標光線空間からの自由視点画像生成

光線空間法では、自由視点画像生成は容易であるが、データ量が膨大となる。

3.2. MV (Multi-view)

図6に示すMV (Multi-view)は3Dシーンを様々な位置で撮影した多視点映像である。多数のカメラを密に配置して撮影すれば光線空間が得られる。カメラ配置が粗な場合には、カメラ間を通過する光線は取得されないため、近接したviewの光線から補間して生成する。

FTVを実現する3Dシーン表現

3D Scene Representation for FTV



図6 多視点映像 MV (Multi-view)

MV はデータがビデオであり、従来のビデオ技術の蓄積を利用できるが、自由視点画像生成に奥行き探索を伴う補間処理が必要である。

3.3. MVD (Multi-view plus depth)

MV に奥行きマップを加えたものが図7に示す MVD (Multi-view plus depth)である。奥行きマップはデプスカメラで撮影するか、DERS (Depth Estimation Reference Software)[11]を用いて MV から生成する。MVD では、depth 情報を利用して view 情報を 3D warping[10]する VSRS (View Synthesis Reference Software)[12]により、自由視点画像を生成する。



図7 MVD (Multi-view plus depth)

MVD では、view も depth もビデオの形式であるため、ビデオ技術の蓄積を利用できる上、depth 情報を利用できるので、自由視点画像生成が容易である。

3.4. ポイントクラウド

ポイントクラウドは、3次元位置情報と色情報を持った点の集まり（点群）によって3Dシーンを表現するものである。点群を3D warpingして画像を生成する。点群が粗な場合には生成画像に隙間ができるので、点を膨らませる splattingなどで隙間を埋める。

ポイントクラウドは、距離測定で求められるため実時間生成が可能であるが、自由視点画像生成には点をつなげて面にする処理が必要である。

3.5. 3D モデル

3Dモデルの代表例は3Dメッシュである。物体表面の点を3角形の頂点として結び、物体を3角形パッチによって表現する。テクスチャは3角形パッチに貼り付けられる。自由視点画像生成は容易であるが、自然シーンの3Dモデルを作るのが難しい。

4. 新しい3Dシーン表現方式とFTVフレームワーク

4.1. MVDの問題点

MVDでは、depth探索のblock matchingにおいて、物体表面が傾いていると、matching blockと物体表面が平行にならず、図8に示すようにview Aとview Bで異なる範囲を比較するため正しいマッチングが行えない。

また、3D warpingによって画像生成を行うとき、図9に示すように物体表面が傾きを持つ場合には、生成画像の画素が粗になる。

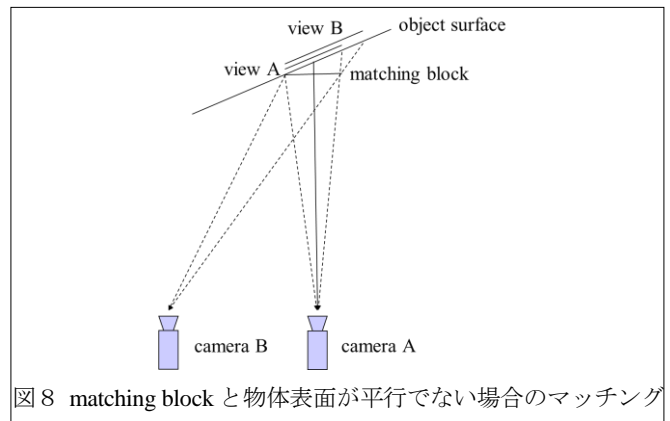


図8 matching block と物体表面が平行でない場合のマッチング

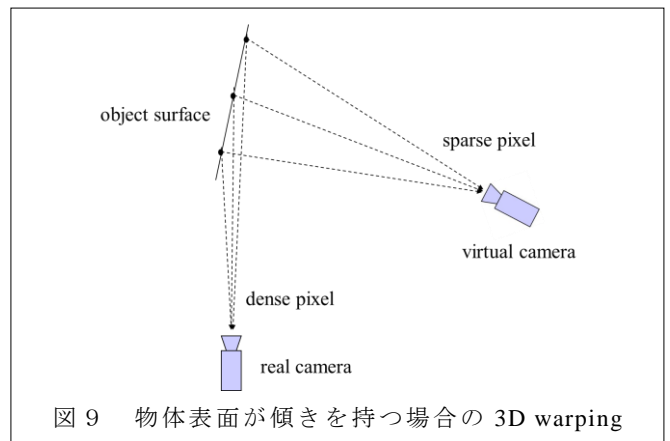


図9 物体表面が傾きを持つ場合の3D warping

4.2. MVS (Multi-view plus surface)の提案

新しい3Dシーン表現として、図10に示すMVS (Multi-view plus surface)を提案した。MVSはmulti-viewにシーンのsurface情報を付加したものである。図11に示すように、surface情報はmulti-viewの各画素の奥行き d とその点に接する平面の法線方向 (θ, ϕ) から成る。すなわち、MVDが3D情報を点情報として持つのに対して、MVSは面情報として持つ。

viewはR,G,BまたはY,U,Vの3成分から成る。例えば、viewをY,U,V成分で表せば、MVSの成分は図12に示すようになり、surface情報はカラー情報と類似のデータ構造を持つことが分かる。

FTVを実現する3Dシーン表現

3D Scene Representation for FTV

MVS は表面情報を持っているので、物体の反射特性や透明度も付加データとして与えることができる。このため、鏡面反射や半透明のシーンも表現でき、照明条件の変更なども可能となる。

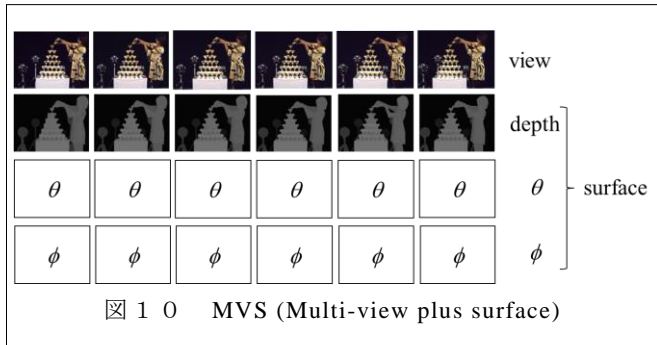


図 1.0 MVS (Multi-view plus surface)

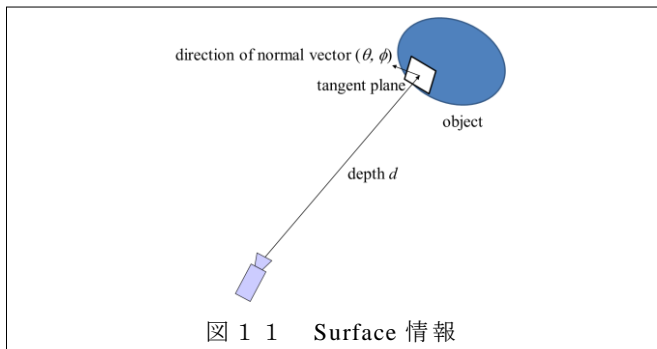


図 1.1 Surface 情報

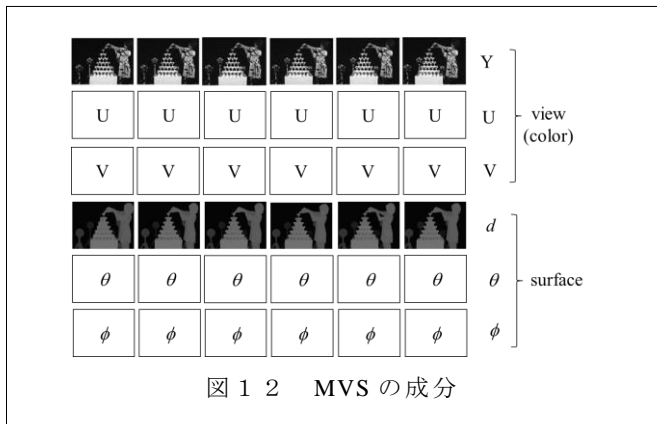


図 1.2 MVS の成分

MVS は図 1.3 に示すように、3Dシーンを点で表すMVDと、構造として記述する3Dモデルの間に位置付けられる。

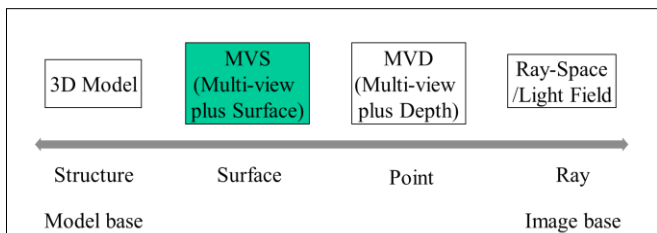


図 1.3 3D シーン表現方式における MVS の位置付け

4.3. 新しい FTV のフレームワーク

図 1.4 に示す MVS を 3D シーン表現方式とする新しい FTV のフレームワーク[14]を構築し、MPEG に提案した。

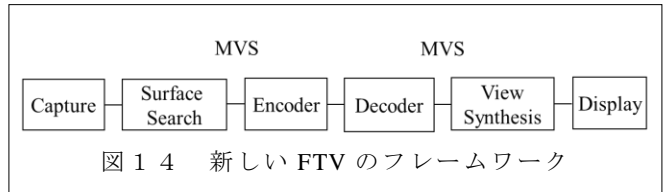


図 1.4 新しい FTV のフレームワーク

surface探索では、図 1.5 に示すように、 d とともに、面の方向 θ, ϕ も変えながら探索を行い、最も整合する d, θ, ϕ を求める。matching blockと物体表面を平行にできるので、図 1.6 に示すように、同じ範囲のview Aとview Bが比較でき、正しいマッチングが行える。これにより、高精度な奥行情報の取得と高品質な画像生成が可能となる。

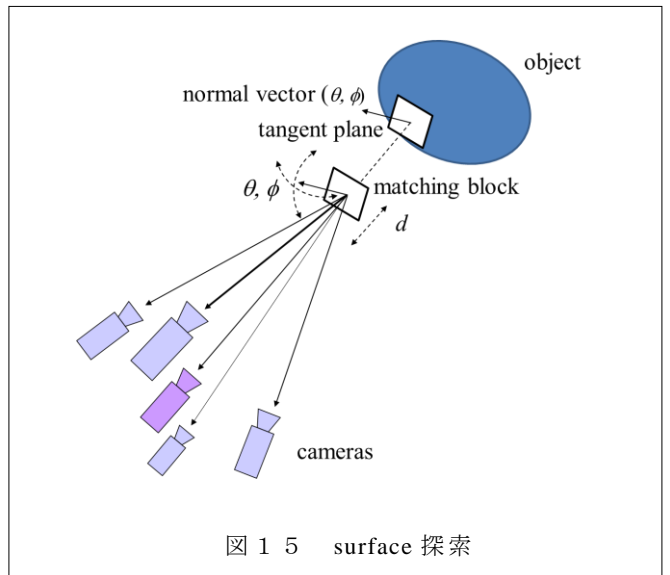


図 1.5 surface 探索

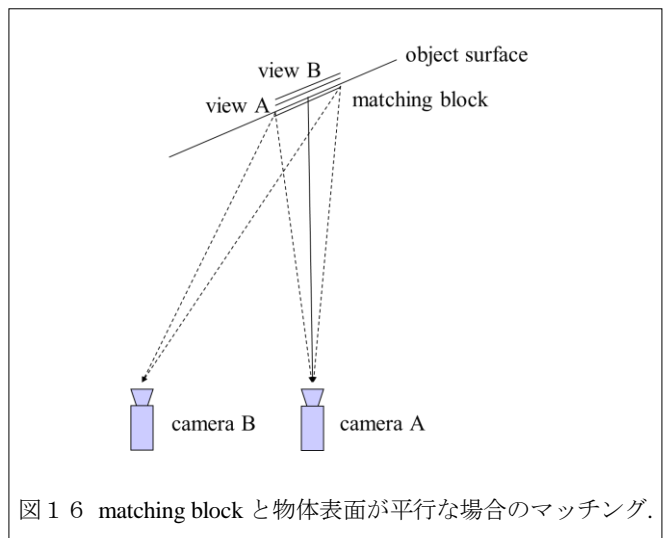
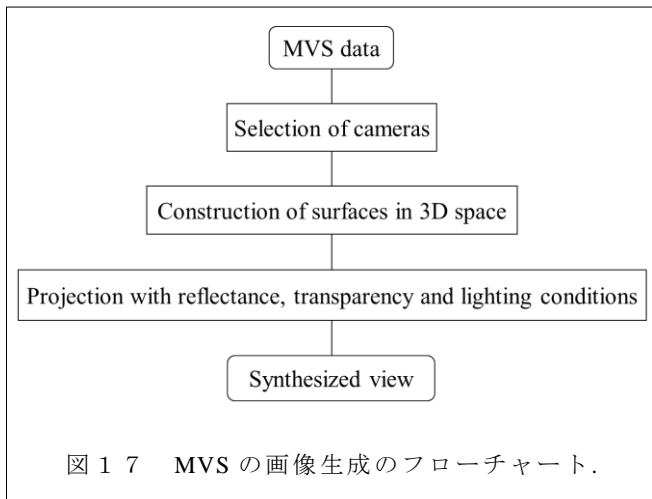


図 1.6 matching block と物体表面が平行な場合のマッチング.

FTVを実現する3Dシーン表現

3D Scene Representation for FTV

画像生成のフローチャートを図 1 7 に示す。まず、画像を生成する仮想カメラ位置に近いカメラ群を選択する。選択したカメラ群の view の画素をその奥行情報を利用して 3 D 空間に逆投影し、3 D 空間の点を得る。これらの点を距離、面の方向、色情報等を用いてグループ化し、surface を構成する。これを仮想カメラ位置に投影して、画像を生成する。



5. 将来展望

新しい 3 D シーン表現として、MVS (Multi-view plus surface) を提案し、表面情報を用いて画像生成を行う FTV の新しいフレームワークを構築した。MVS は従来の MVD の自然な拡張となっており、これにより高精度な奥行情報の取得と高品質な画像生成が行える。さらに、物体の反射特性や透明度も付加データとして持つことができるため、鏡面反射や半透明のシーンも表現でき、照明条件の変更なども可能となる。MVS は、オリエンテーションや透明度を持つポイントクラウド表現と類似点があるので、両方式の統合方式に発展する可能性もある。

F T V は、空間情報の全てを伝達する没入型 3 D メディアであり、映像メディアの頂点に立つ。無限の眼を持つ F T V は極めて高いセンシングや表現の能力を持つため、産業や生活、社会、学術、教育、文化、スポーツ、アミューズメント等、多くの分野において大きな貢献が期待される。

参考文献

[1] Masayuki Tanimoto, Mehrdad Panahpour Tehrani, Toshiaki Fujii, Tomohiro Yendo, “Free-Viewpoint TV”, IEEE Signal Processing Magazine, vol.28, no.1, pp.67-76, January 2011.

[2] Masayuki Tanimoto, Mehrdad Panahpour Tehrani, Toshiaki Fujii, Tomohiro Yendo, “FTV for 3-D Spatial Communication”, Proceedings of the IEEE, Vol. 100, No. 4, pp. 905-917 (April 2012). (invited paper)

[3] Masayuki Tanimoto, “FTV: Free-viewpoint Television”, Signal Processing : Image Communication, Vol. 27, Issue 6, pp. 555-570 (June 2012). doi:10.1016/j.image.2012.02.016 (invited paper)

[4] Masayuki Tanimoto, “FTV (Free-viewpoint Television)”, APSIPA Transactions on Signal and Information Processing, Vol. 1, Issue 1, e4 (14 pages) (August 2012). doi: 10.1017/ATSIP.2012.5 (invited paper)

[5] “Introduction to Multi-view Video Coding”, ISO/IEC JTC1/SC29/WG11 N7328, July 2005.

[6] “Introduction to 3D Video”, ISO/IEC JTC1/SC29/WG11 MPEG2008/N9784, Archamps, France, May 2008.

[7] “Report on Experimental Framework in 3D Video Coding”, ISO/IEC JTC1/SC29/WG11 N11631, October 2010.

[8] M. Tanimoto, T. Senoh, S. Naito, S. Shimizu, H. Horimai, M. Domański, A. Vetro, M. Preda and K. Mueller, “Proposal on a New Activity for the Third Phase of FTV,” ISO/IEC JTC1/SC29/WG11 MPEG2013/M30229, Vienna, Austria, July 2013.

[9] “Use Cases and Requirements on Free-viewpoint Television (FTV) v.3,” ISO/IEC JTC1/SC29/WG11 MPEG2016/N16130, San Diego, USA, February 2016.

[10] Y. Mori, N. Fukushima, T. Yendo, T. Fujii, M. Tanimoto, “View generation with 3D warping using depth information for FTV,” Signal Processing: Image Communication, Vol. 24, pp. 65-72, January 2009.

[11] Masayuki Tanimoto, Toshiaki Fujii, Mehrdad Panahpour Tehrani and Menno Wildeboer, “Depth Estimation Reference Software (DERS) 4.0,” ISO/IEC JTC1/SC29/WG11 MPEG2009/M16605, London, UK, June 2009.

[12] Masayuki Tanimoto, Toshiaki Fujii and Kazuyoshi Suzuki, “View synthesis algorithm in View Synthesis Reference Software 2.0 (VSRS2.0),” ISO/IEC JTC1/SC29/WG11 MPEG2009/M16090, Lausanne, Switzerland, February 2009.

[13] “Description of Exploration Experiments on Free-viewpoint Television (FTV),” ISO/IEC JTC1/SC29/WG11 MPEG2013/N14105, Geneva, Switzerland, October 2013.

[14] Masayuki Tanimoto, “New Framework of FTV,” ISO/IEC JTC1/SC29/WG11 MPEG2016/M38547, Geneva, Switzerland, May 2016.

この研究は、平成 2 5 年度 S C A T 研究助成の対象として採用され、平成 2 6 ~ 2 8 年度に実施されたものです。