

## ライトフィールド符号化の国際標準化動向

石川孝明†

キーワード：ライトフィールド画像、画像符号化、JPEG Pleno、イマーシブメディアの符号化表現、国際標準化

## 1. まえがき

ISO/IEC JTC 1/SC 29/WG 1 (JPEG) および WG11 (MPEG) は、それぞれ静止画像符号化方式と動画画像符号化方式に関する国際規格を審議している。現在各WGでは、新たな画像表現であるライトフィールド画像の符号化方式の規格化に着手しており、静止画像符号化については、JPEG Pleno (ISO/IEC 21794 シリーズ) のパート2として審議中であり、動画画像符号化方式については、MPEG-I (ISO/IEC 23090 シリーズ) の枠組みにおいて技術探索を進めている。本稿では、各WGで審議中のライトフィールド符号化の最新動向について述べる。なお、執筆時点での最新の審議情報を含むため、本稿の内容は最終的に出版される規格と異なる可能性があることに留意していただきたい。

## 2. ライトフィールド画像

## 2.1 新しい画像表現

デジタルカメラやスマートフォンで撮影される画像は、水平垂直に画素値が並ぶ2次元平面を前提としており、撮像時にフォーカスを確定することで2次元画像またはその連続シーケンスを信号として記録することが一般的である。これは、撮像素子に到達する光学情報を記録することに他ならないが、この際、光がどの方向から到達したのかについての情報は失われてしまう。これに対してライトフィールドカメラでは、マイクロレンズアレイを用いることで、撮像素子に到達する光の進行方向を光線情報として記録することができる。そのため、ライトフィールド画像は、撮影後にフォーカスを変更可能であり、また、従来は手前の物体に隠されて表示できなかった情報も視点を変えることで再生可能になる。このような特性から、製品検査、映像制作、監視カメラ、顕微鏡などの分野で採用され始めており、今後さらに注目されていくと予想される。その一

方で、従来は扱われていない光学情報を記録するため、ライトフィールド画像の情報量は必然的に増加する。これは、複数の単眼カメラを格子状に並べて構築するカメラアレイにおいて顕著であり、ライトフィールド画像のデータ圧縮の必要性が理解される。

## 2.2 ライトフィールドの記録

ライトフィールド画像が広く認知されるきっかけとなったのは、Ren Ngらが2005年に発表した4次元ライトフィールド画像を取得可能なライトフィールドカメラといえる<sup>1)</sup>。ライトフィールドに関する基礎理論では、光線情報を記録するための理想的な光学表現として7次元の情報  $(P(\theta, \phi, \lambda, t, V_x, V_y, V_z))$  があればよいことが知られている。各パラメータは、方向  $(\theta, \phi)$ 、波長  $\lambda$ 、時間  $t$ 、そして、 $V_x, V_y, V_z$  が3次元空間の任意の位置を表している。これに対し、Ren Ngらは、Marc Levoyらが提案した位置と方向の4次元に縮退した表現形式を用いてライトフィールドカメラを実装した。静止画像符号化の観点からは、従来の水平垂直方向にサンプル点を有する画像が上下左右の空間位置に拡張されていると解釈でき、4次元のサンプリング情報を効率よく符号化することが課題となる。また、動画画像符号化の場合には時系列画像が対象となるため、フレーム間の冗長性を削減する高効率な符号化方式の開発が求められる。

## 2.3 レンズレット画像とマルチアパーチャ画像群

ライトフィールドカメラから得られるレンズレット画像の一例を図1に示す。ライトフィールドカメラで得られる図1のようなレンズレット画像にデモザイキング等の前処理を施すと、図2に示す複数のビュー(サブアパーチャ画像群)

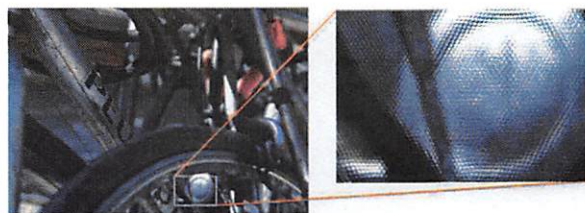


図1 レンズレット画像 (Bikes: 7728×5368 [画素])

†早稲田大学 国際情報通信研究センター  
"Report on International Standardization of Light Field Image Coding"  
by Takaaki Ishikawa (Global Information and Telecommunication  
Institute, Waseda University, Tokyo)



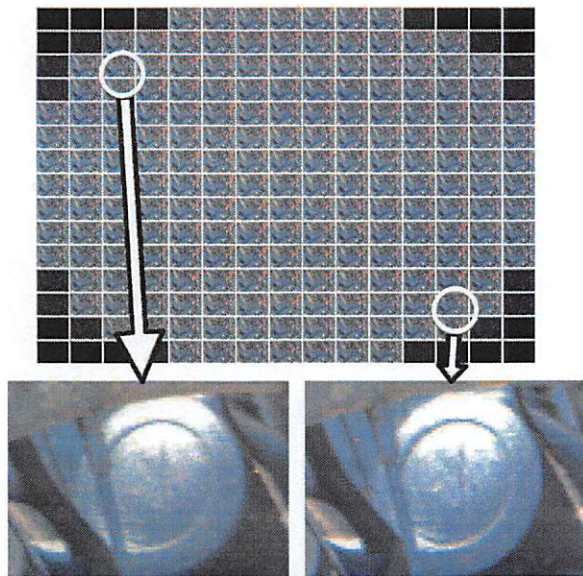


図2 サバパーチャ画像群 (Bikes: 15x15[ビュー])

が得られる。図2の丸で囲んだ箇所を拡大した画像を観察すると、ビュー間の視差を確認できる。なお、図1は、JPEG Plenoの審議で使用されている試験画像の一つである。

### 3. 最新の標準化動向

#### 3.1 データ変換と圧縮

マイクロレンズアレイによるライトフィールドカメラや複数台のカメラを組み合わせたカメラアレイにて撮影される画像は、図1に示すレンズレット画像または図2に示すマルチアパーチャ画像や視差の異なる複数画像群として記録される。両者は交換可能な情報であるため、画像入力から表示までの一連のフローでは、両者の形式変換と圧縮の対象となるデータについて、図3のような複数の処理パターンが想定される。図3の中心の点線で囲った部分はデータ圧縮処理部であり、圧縮の対象となるデータには形式変換後の時系列レンズレット画像群、時系列マルチビュー画像群、マルチフレーム画像群の3種類が想定される。マルチフレーム画像群は、任意時刻におけるすべてのビューを所定の順番で走査し、擬似的な時系列画像群を構築する方式である。なお、第3.2節で述べる静止画像符号化では、レンズレット画像をマルチアパーチャ画像群に変換して複数のビューとして扱うため、図3におけるマルチフレーム画像群の処理に分類できる。ただし、静止画像符号化では時間方向のフレーム群は考慮されない。

#### 3.2 静止画像符号化

##### 3.2.1 規格化の経緯と審議状況

2次元静止画像の次元数を上回る高次元の画像表現を対象として、WG1ではライトフィールド、ポイントクラウド、ホログラフィなどの新しい画像表現に対応する規格の審議

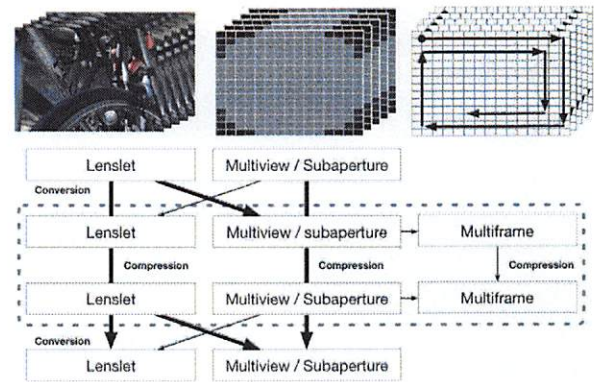


図3 ライトフィールド画像の処理フロー

に着手している。2014年10月の第67回会合から欧州の専門家を中心とする調査活動が開始され、ワークショップの開催やWG 11との合同会議などを経て、JPEG Pleno Plenoptic image coding system (ISO/IEC 21794シリーズ)として規格審議が開始された。Plenoはラテン語で「完全な・すべて」を表す言葉である。

ライトフィールド画像の符号化技術については、2017年4月に提案募集が発行され、同年6月に募集が締め切られた<sup>2)</sup>。計6件(企業2件、大学4件)の応答があり、2017年7月に開催された第76回会合にて、客観および主観評価実験の結果が報告された。各方式の比較評価の結果、テストシーケンスとビットレートの違いにより最良方式が異なることが示された。そのため、引き続き、符号化ツールごとの優位性を明らかにするためのコア実験を継続することで合意した。

2019年3月開催の第83回会合では7度目のコア実験が設定されるなど、現在も活発な審議が継続されている。これまでの審議の結果を反映し、JPEG Plenoパート2 (Light Field Coding)には、4D Transform Modeと4D Prediction modeの2種類のモードが規定されている。前者は4次元DCTを用いる方式であり、後者は参照ビューと視差マップを用いて中間ビューを予測する方式である。これらの方式の違いは、第4章で概説する。

##### 3.2.2 JPEG Pleno今後の作業計画

表1に、JPEG Plenoの作業計画を示す。規格文書は、作業原案(WD)、委員会原案(CD)、国際規格案(DIS)、最終

表1 JPEG Pleno (ISO/IEC 21794) の作業計画

Part	Title	WD	CD	DIS	IS
1	Framework	18/01	19/03	19/07	20/01
2	Light Field Coding	18/04	19/01	19/07	20/01
3	Conformance Testing	19/10	20/01	20/04	20/10
4	Reference Software	19/10	20/01	20/04	20/10



国際規格案 (FDIS), 国際規格 (IS) とステータスが進むが, ライトフィールド符号化を規定するパート2は, 現在DIS投票に向けた議論とコア実験を実施している状況である. またパート1は, プロジェクト全体で共通するフレームワークとファイルフォーマットを定めており, パート2と連動して審議されている. その他の画像表現であるポイントクラウドやホログラフィについては, 提案募集の準備段階であり, アドホックグループにおいて試験画像, 評価方式等の調査および性能評価を実施している. なお, 第83回会合ではホログラフィに関するワークショップが開催されており, 提案募集の発行に向けた議論が加速している<sup>3)</sup>.

### 3.3 動画像符号化

WG 11では, 動画像符号化方式を中心とするさまざまな規格化を審議している. 同WGではこれまでも, 視差情報を動き補償フレーム間予測にて圧縮するMVC (Multi-view Video Coding) や, 多視点映像から奥行き情報を抽出して復号時に中間視点を合成する方式3DV (3-Dimensional Video) などを長年検討している.

2018年1月に開催された第121回MPEG会合において, WG 11が今後注力していく領域が示され, イマーシブメディア (MPEG-I) の符号化に関するロードマップが示された. MPEG-I (ISO/IEC 23090シリーズ) の審議におけるフェーズ1では, ポイントクラウドに関する動画像符号化方式としてISO/IEC 23090パート5がIS発行済である. そして, フェーズ2に組み込まれている6DoFに対応する映像表現やライトフィールド映像の符号化は, 2022年をIS発行の大まかな目標に定める探索段階にある<sup>4)</sup>. なお, 2017年7月の第119回会合にてライトフィールド映像のテストマテリアルの提供がすでに呼びかけられている. MPEG-Iに関する標準化動向は, 青木らの解説記事が詳しい<sup>5)</sup>.

図3は, 現在検討中のライトフィールド映像の符号化フローである. 時系列の連続シーケンスを扱うため, JPEG Plenoパート2よりも構造が複雑であり, 検討すべき項目も多い. そのため現在は, 複数の種類が想定される処理フローの中でも, 図3の太線で示したフローに該当するレンズレット映像からマルチビュー映像への変換 (EE\_conversion), レンズレット映像の圧縮 (EE\_LL), マルチビュー映像の圧縮 (EE\_MV) を優先度の高い調査実験と位置づけている<sup>6)</sup>. EE\_conversionは, 同変換に起因する品質の変化について明らかにする実験であり, EE\_LLおよびEE\_MVは, それぞれの処理フローにおける圧縮性能を測る実験である.

## 4. ライトフィールド画像の符号化

DIS投票に向けて継続審議中のJPEG Plenoパート2 (ISO/IEC 21794-2) では, 各サブアパーチャ画像を効率よく符号化するために, 現時点で二つの符号化モードが規定されている. また検討段階ではあるものの, サブアパー



図4 Lenslet試験画像 (Bikes, Fountain, Stone, Danger)

チャ画像間の相関を利用して階層的に画素を予測するロスレス符号化や高いランダムアクセス性とスケーラビリティを実現するための符号化ツールも審議されている.

### 4.1 共通実験条件

ライトフィールド画像を符号化する各方式を同じ実験条件で比較する目的から, 共通実験条件 (CTC: Common Test Conditions) が定められている<sup>7)</sup>. 本稿執筆時点の最新版はバージョン3.2である. 符号化ツールの性能改善やユースケースの対象拡大などの審議状況に合わせて, CTCは今後も随時更新される可能性がある. CTCには, 符号化性能の評価基準, 試験画像と特性, 符号化性能を議論するための基準となる方式 (アンカー) などが記載されている. 提案募集の発行段階ではマイクロレンズアレイを用いるライトフィールドカメラを用いた試験画像 (Lensletと呼ぶ) と複数台のカメラを格子状に並べて撮影するカメラアレイを用いた試験画像 (HDCA: High Density Camera Arrayと呼ぶ) について, それぞれ4種および5種類の画像群が規定された<sup>2)</sup>. その後, 審議の過程において, 奥行き情報や視差情報を正確に与えることが可能なコンピュータ上で合成された2種類の試験画像 (Syntheticと呼ぶ) が追加された. また, HDCAについても2種類が追加されている. 図4に4種類のLenslet試験画像を示す. CTCに定められた各試験画像は, JPEGウェブサイトにて公開されている<sup>8)</sup>. また同サイトでは, その他の試験画像やポイントクラウドとホログラフィのデータも入手可能である.

JPEG Plenoパート2では, 図2のサブアパーチャ画像群を符号化対象とし, カメラごとに異なるキャリブレーションに関する情報も符号化対象である. カメラアレイから得られる画像群もサブアパーチャ画像群と同様に複数のビューとして扱えるため, LensletとHDCAは同一のフレームワークで処理できる. ただし, Lensletについてはライトフィールドカメラの特性上, 外周に位置するビューの輝度レベルが低いため, 符号化実験ではそれらを除外した13×13のビューを処理対象としている. 各ビューは, 626×434 [画素] の解像度であり, 各画素のビット深度は10ビットである.

### 4.2 4D Transform mode

4D Transformモードは, 4次元DCTを用いる符号化方式である. 図5(a)に示すように, サブアパーチャ画像上の任意の画素は, 画像内の水平垂直の2軸 (spatial) と視差が生じる上下左右方向の2軸 (view) からなる4次元空間で表現される. これは, 2.2節で述べたライトフィールド画像

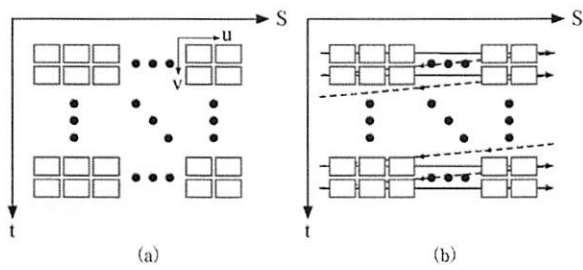


図5 サバパーチャ画像群の4次元構造

を表現するために必要な次元数と一致している。さらに、それぞれの空間軸では画素間に高い相関性が存在するため、JPEG標準で使われる2次元DCTを拡張する方式を採用するのは自然な発想である。

4D Transformモードでは、まず図5(a)の画素群をRate-Distortion最適化に基づいてブロック分割し、各ブロックを独立に4次元DCTで直交変換する。得られた係数群を量子化した後、非有意係数を効率的に表現するために16分木によるビットプレーン分割で表現し、最後に算術符号化を施してビットストリームを構築する。これらをカラーコンポーネントごとに独立して処理する。

4D Transformモードによる圧縮は、カメラアレイのように隣接するビュー間の視差が大きくなるにつれ、後述の4D Predictionモードと比較して符号化効率の面では不利になる。しかし、視差情報や奥行き情報を用いる複雑な予測処理を必要としないメリットがあり、特に低ビットレートでの符号化効率について、優位性が確認されている。

#### 4.3 4D Prediction mode

4D Predictionモードは、参照ビューと視差マップを用いて中間ビューを予測するモードである。各ビューは図5(b)に示すようなラスタスキャンで順番に処理される。図5(b)に示したスキャン方法以外にも中心から外周に向けて回転するようにスキャンする方法も試されたが、同程度の符号化効率であり有意差が示されていない。また、ランダムアクセス性を高めるために、ラスタスキャンではなく、所定のビューをグループ化して階層的に扱う手法も規定されている。

ビットストリームを構成する参照ビュー、視差マップ、予測パラメータ、予測差分データ等のすべての要素が、ボックス形式のJPEG Plenoファイルフォーマットに従って記述される。復号側では、それらの情報を用いて、ワーピング、合成、予測誤差の訂正などの処理を施して、すべてのビューを再構築する。ワーピングと合成により再構成される中間ビューまたは視差マップには、オクルージョンに起因する画素の欠落が生じる場合があるため、そのような情報は、復号器にてメディアンフィルタによるインペインティング法を用いて補間する。

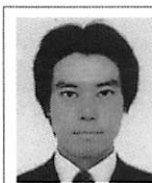
参照ビュー、視差マップ、予測誤差データの圧縮方式は、JPEGファミリー標準のうちJPEG 2000, JPEG, JPEG LS, JPEG XT, JPEG XR, JPEG XSから選ぶことができる。使用可能な圧縮方式は、拡張可能なTCODECフィールドで復号器に伝達される。なお、デフォルトで使用されるJPEG 2000は、視差マップのような不連続性を有する信号を圧縮するために最適化されておらず、復号画像のオブジェクト境界付近で画質が低下する場合がある。この問題を解消するために、不連続メディアを対象とする拡張符号化方式がJPEG 2000パート17として議論されており、コア実験において画質の改善効果が確認されている。

## 5. むすび

本稿では、静止画像符号化の国際規格を審議するISO/IEC JTC 1/SC 29/WG 1 (JPEG)とWG11 (MPEG)の最新標準化動向について述べた。撮像デバイスとしても進化を続けるスマートフォンには、近年複数のレンズを搭載する機種も増えており、多眼化と高解像度化がさらに進めば、ライトフィールド符号化の必要性が高まると予想される。また表示デバイスにおいてもライトフィールドを活かしたヘッドマウントディスプレイや立体ディスプレイなどが開発されており、VR/AR/MRの各領域やホログラフィでの応用に対する要求も強い。今後、これらの領域においてJPEG Plenoまたは現在MPEGで検討中の最新規格が広く普及することを期待したい。(2019年6月24日受付)

## 〔文 献〕

- 1) R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, P. Hanrahan: "Light field photography with a hand-held plenoptic camera", Stanford University Computer Science (2005)
- 2) T. Ebrahimi, P. Schelkens, F. Pereira, S. Foessel, T. Richter, D. McNally: "JPEG Pleno Call for Proposals on Light Field Coding", ISO/IEC JTC 1/SC 29/WG 1 (Jan. 2017)
- 3) JPEG Pleno Holography Workshop Proceedings. [https://jpeg.org/items/20190329\\_pleno\\_holography\\_workshop\\_geneva\\_proceedings.html](https://jpeg.org/items/20190329_pleno_holography_workshop_geneva_proceedings.html) (2019)
- 4) ISO/IEC JTC 1/SC 29/WG 11: "MPEG-121 Version of MPEG Standardisation RoadMap", N17332 (Jan. 2018)
- 5) 青木秀一: "次世代メディアの探求! MPEGにおけるイマーシブメディアの標準化", 情報処理, 60, no.7 (2019)
- 6) ISO/IEC JTC 1/SC 29/WG 11: "Exploration Experiments and Common Test Conditions for Dense Light Fields", N18446 (Apr. 2019)
- 7) F. Pereira, C. Pagliari, E.d. Silva, I. Tabus, H. Amirpour, M. Bernardo, A. Pinheiro: "JPEG Pleno Light Field Coding Common Test Conditions V3.2", ISO/IEC JTC 1/SC 29/WG 1 (Mar. 2019)
- 8) JPEG Pleno Database, <https://jpeg.org/plenodb/> (2019)



石川 孝明 2005年、早稲田大学大学院国際情報通信研究科修士課程修了。2007年、同大学国際情報通信研究センター助手。2010年、同センター招聘研究員となり、現在に至る。画像符号化の研究と応用開発に従事。2014年より、ISO/IEC JTC 1/SC 29/WG 1国内小委員会主査。