Motion JPEG2000と MJ2 File Format

福原 隆浩†

キーワード：JPEG2000, Motion JPEG2000, MJ2, MP4, ISO国際標準化

1. ま え が き

昨今のイメージセンサの高解像度・高ビット深度化、メモリーからの画像データの読出し・書き込み速度の高速化、メモリー・バンド幅の高帯域化、画像コーデックの高性能化等に伴い、デジタルスチルカメラ(DSC)で静止画のみならず高品質な動画記録が可能になっている。一方、カムコーダも本来の動画記録のほかに、高解像度のイメージセンサを用いた高品質な静止画記録にも対応している。このように静止画と動画の記録を1台で対応する場合、画像コーデックを共通化すればハードウェアコストの面でメリットが大きい。初代の動画デジカメは、静止画用にJPEG、動画用にMotion JPEGを採用しており、コーデックの共通化を実現していた。その後、記録メディア(フラッシュメモリー等)の大容量化と並行して長時間記録への要望が高まり、JPEGの約2倍の圧縮性能を持つJPEG2000¹⁾⁶⁾の動画版であるMotion JPEG2000の標準化への期待が大きくなった。JPEGでは、規格外の8ビット超の高ビット深度画像をJPEG2000が対応可能としていることも大きなメリットの一つであった。

2. Motion JPEG2000標準化の背景

1998年7月に開催されたISO/IEC/JTC1/SC29/WG1バンクーバ会合において、筆者はJNB (Japan National Body) 提案として、Motion JPEG2000を標準化する必要性について、以下のように説明した。JNB案とは、WG1国内委員会のメンバー間で合意した案であり、個々の機関からの独自提案よりも強い影響力がある。したがって、国際会合においても、日本からの提案として尊重される傾向が強い。

(1) 当時存在していたJPEGの動画版であるMotion JPEGは、国際標準規格がないため、いくつかのデファクトのフォーマットが乱立している。そのため、

機器間で互換性が取れないケースがある。

(2) JPEG2000は、JPEGの約2倍の圧縮性能のほか、スケーラビリティ、強いエラー耐性等のJPEGにはない豊富な機能を備えており、これらの機能を利用したJPEG2000の動画規格が期待されている。

(3) Motion JPEG2000の国際標準化を行わない場合、いくつかのデファクト規格が乱立し、DSCをはじめとする機器やデバイス間での互換性が確保できなくなるリスクがある。

(4) 動画と音声との同期手段の確立が必要になるが、JPEG2000には、MPEGシステムのような同期手段を実現する規格が存在していないので、新たに標準化で定義する必要がある。

その後、バンクーバ会合の会合期間内にWG1委員会のメンバーが議論を行い、JPEG2000 Part3として、Motion JPEG2000を標準化する提案が合意された。当時発行されたドラフトに記載されたMotion JPEG2000のスコープは、以下の通りである(原文)²⁾。

1. Still and Video codec equipped on the DSC and Camcorder
 2. Built on the baseline JPEG2000 (Part1) coding
 3. Easy editing of encoded bitstream (frame by frame)
 4. Synchronization of video (Motion JPEG2000) and audio via a file format
 5. Very flexible choice of audio (compressed or uncompressed)
 6. Providing fully interoperability for consumers
- Motion JPEG2000のターゲットアプリケーションに関しては、以下の記載がある。

- * DSC and Camcorder with Still and Video Codec
- * Remote Surveillance Systems
- * Digital Video Recording Systems
- * Video Capture Card

したがって、この時点では、その後大きなアプリケーション

† ソニー株式会社

"Motion JPEG2000 and MJ2 File Format" by Takahiro Fukuhara (Sony Corporation, Tokyo)

ョンとなるデジタルシネマ⁵⁾については、具体的には想定していなかったことがわかる。

上記のWG1委員会での合意を受けて、1999年12月13日～15日に米国ハワイ州マウイ島で開催されたISO/IEC/JTC1/SC29プレナリ会合(議長:元NTT渡邊裕氏)において、Motion JPEG2000の標準化開始に関する採決が行われ、賛成国13、反対国0、棄権1の賛成多数で可決された。会合結果のResolutions³⁾には、Motion JPEG2000について以下の記述があり、MPEGに対する配慮がなされた表現になっている。これは元々MPEG側から動画像標準化に対する反対が根強かったことと、MPEGとMotion JPEG2000の想定するアプリケーションの相違点を明らかにして、市場に混乱を起こさないように注意を払ったためである。

SC 29 encourages WG 1 to develop its proposed Motion JPEG2000. This project should be developed having in mind the benefits of a solution that integrates with other SC 29 standards. Departure from this goal should be duly motivated.

3. Motion JPEG2000のエンコード技術

図1は、通常のJPEG2000のエンコーダのブロック図である。JPEGのDCTの代わりに、Wavelet Transform (ウェーブレット変換)を用い、エントロピー符号化にEBCOT (Embedded Block Coding with Optimized Truncation)を採用していることが、JPEG2000の大きな特徴である。

一方、Motion JPEG2000は、JPEG2000画像の時系列的連続であるためコーデックを共通化しているが、以下がMotion JPEG2000用に新たに定義された項目である。

- (1) フレームベース・フィールドベース符号化
- (2) Chromaフォーマットの定義
- (3) 動画用のVisual Weighting Tableの定義
- (4) 独自のファイルフォーマット(MJ2)

Motion JPEG2000の場合、フレームベース符号化用には、図2に示すように、First Field/Second Fieldを統

図2 フレームベース符号化時のWavelet Transform

図3 フィールドベース符号化時のWavelet Transform

合したまま(フレーム構造)、ウェーブレット変換する。一方、インタレース画像をフィールドベース符号化する場合は、図3のように、Even FieldとOdd Fieldを別々にウェーブレット変換する。これは、被写体の動きの再現性向上に効果がある。

4. Motion JPEG2000のファイルフォーマット

Motion JPEG2000ファイルフォーマット(MJ2)は、主に以下の特徴を持っている。

- (1) QuickTimeフォーマットをベースに構築
- (2) MPEG-4ファイルフォーマット(MP4)との共通部分が多く、共通のBox構造を多数有する。
- (3) Video部分をMotion JPEG2000, MPEG-4 Audio以外の多くのAudioフォーマットをサポート可能で、編集用途向けのアプリケーションのために非圧縮のAudioもサポート可能。
- (4) VideoとAudioとの同期はファイルフォーマットでサポートされる(60分の1秒単位で可能)。

図4は、JPEG2000ファイルフォーマット(JP2)とMJ2を併記した図である。MJ2の大きな特徴はmoov, moof, mdatから成るbox構造であり、これらはQuickTimeフォーマットの構成要素である。MJ2にQuickTimeフォーマットを採用した背景としては、すでにMP4がQuickTimeフォーマットベースで標準化されており、これをベースにすることが作業的にも技術的にも簡潔であったことが理由として挙げられる。図4において実線は必須のboxで、点線はオプションを意味している。したがって、JP2の場合には、最低1ピクチャのCodestreamを含むが、後続に任意のピクチャ数分のCodestreamを包含することもできる。また、MJ2では

図1 JPEG2000エンコーダのブロック図

オプションでJPEG2000のCodestreamを格納することができる。これは例えば、mdat内に格納されているJPEG2000動画の中で最も特徴的な静止画1枚を抽出して、それを図4のJPEG2000 Codestream部に格納することで、見出し的な使い方ができる。簡単のために、先頭のフレームの画像を選択してもよい。これによって、MJ2ファイルにどのような動画像が入っているのかを、1ピクチャのCodestreamをデコードするだけで実現できるので便利である。その後、Apple社のQuickTimeエンジニアのDr. David SingerがMP4、MJ2双方のEditorになり、筆者がMJ2のCo-editorになって標準化作業が進められた。

MJ2のbox構造は、表1に示したように、ベースとなったQuickTimeフォーマットと同様に6階層から構成されている(表中、左の階層boxは右の階層boxを内包)。この中でMJ2にとって重要なboxについて以下説明する。

(1) moov

MJ2ファイル中1個だけ存在し、画像と音声の同期情報、データへの参照情報、時間管理情報を持つ。

(2) mdat

MJ2ファイル中1個だけ存在し、実際のデータ、例えばJPEG2000のCodestreamやMP3 Audio圧縮データ等が、このmdatに記録される。

(3) moof

MJ2ファイル中複数存在することができ、moovと類似した機能を持ち、moofとmdatをペアにしながらデータをMJ2ファイルに記録することができる。本来のQuickTimeフォーマットには存在していなかったが、MJ2独自のboxとして採用された。

5. 標準化活動での体験談

筆者がJPEG2000の標準化に初めて参加したのが1997年11月のシドニー会合である。この会合でJPEG2000の主観評価テストが行われ、アリゾナ大学提案のウェーブレット変換+TCQ量子化+算術符号化方式が最高の評価を獲得した。この会合の後、DCT派とウェーブ

表1 Box types, structure, cross-reference

JP					
ftyp					
moov					
	mvhd				
	trak				
		tkhd			
		tref			
		edts			
			elst		
			mdia		
			mdhd		
			hdlr		
			minf		
				vmhd	
				smhd	
				hmhd	
				dinf	
					dref
				stbl	
					stsd
					stts
					stsc
					stsz
					stco
	mvex				
		trex			
moof					
	mfhd				
	traf				
		tfhd			
		trun			
mdat					
free					
skip					
	udta				

レット派とに分かれて多くの実験と議論が行われたが、結果的には、JPEGとの互換性よりも新機能や高圧縮を重視すべきとの意見が優勢となり、変換方式はウェーブレット変換に決定した。

その後、JPEG2000はパラメータが豊富に設定できるため、この選択枝の多さがハードウェア(LSI含む)を作る際のハードルになることが懸念された。そこで筆者がPart1 Amendment-1の議長兼Editorに就任して、パラメータの制約条件を定めたProfile-0、Profile-1を標準化した。ここで一番問題になったのは、JPEG2000のエンコード・デコード単位であるコードブロックのサイズである。ハードウェア化する際にメモリーサイズを小さくする場合には、32×32が有利である。他方、コードブロックサイズが小さいとコードブロック数が増大する。その結果、メモリーアクセス数が増えるため、コードブロックサイズの大きい64×64を推奨する機関との対立が起こった。そこで妥協案として、32×32、64×64の両方のサイズをPart1デコーダは必須と定めた。これは、その後のLSI開発の進歩の早さを見込んでの決定であった。

Part1の標準化が順調に進むのに並行して、Part3の標準化を開始することになった。先に述べた1999年12月

のマウイ会合での、WG1、WG11の両コンビーナを含めたキー・メンバの会議は、WG11からの抵抗・反発が予想以上に大きかった。しかし筆者は、JPEG2000の標準化に参画する前にはMPEG-4の標準化会合に出席しており、私事で恐縮であるが、1995年11月のMPEG-4主観評価テストで、自らの提案方式が、最低ビットレートのクラスで世界No.1の評価を得ている。その後、MPEG-4 Video Sub-Groupのチェアをしていた時期もあり、すでにWG11委員会メンバに多くの知人がいた。これらの人的コネクションが役立ち、MPEG Systemグループ・議長のDr. Olivie Avaroをはじめ、多くのオピニオンリーダーがMotion JPEG2000標準化推進に助力して頂いた。この場を借りて、諸氏に感謝の意を表したい。

最後に、Part3の標準化で残念に思うこととしては、Part3の標準化がスタートする時点で、Part1のコーデックのコア技術の多くがすでに確定していたことである。もしもPart1とPart3の同時並行の時間があったならば、動画像にとって効果的な技術のPart1への入れ込みも可能だったかもしれない。例えば、Part1のウェーブレット変換フィルタは、Lossy符号化用には9×7フィルタしか規定されていないが、5×3フィルタもPart1に入れるべきであった。5×3フィルタはLosslessだけでなく、Lossy符号化でも中低ビットレートでは9×7フィルタと画質差もほとんどないうえ、フィルタリングの際に必要なメモリー量が大幅に減らせるメリットがある。これはリアルタイム性を要求される動画像符号化では非常に大きなアドバンテージになる。

この他にも、画質のデテイル向上に効果があるDeadzoneなしの量子化手法や、Lazyモード拡張法等がある。

6. 国際標準化に従事する技術者に向けて

国際標準化活動に従事するエンジニアは、どのようにモチベーションを持てば良いのだろうか？ 国際標準化策定には通常2～3年の年月を要する。その期間は、企業特にメーカーのエンジニアにとっては、すぐに利益に結び付かない業務に従事しているわけで、最先端の技術の標準を作るという強い意志と誇りを持つことが大切である。その後、苦

勞して世界の英知を集めて策定した国際標準は、企業にとって利益の源になる製品の中に組込まれて、われわれの苦勞は報われるのである。

具体的に言えば、ISO/IEC/JTC1/SC29/WG1委員会の策定したJPEGのお陰で、デジカメメーカーは利潤を上げることができ、WG11委員会の策定したMPEG-2やH.264の恩恵で、カムコーダやBluRayレコーダで大きな売り上げを立てているのであり、これらの事実に対して、標準化に従事したエンジニアはもっと誇りを持つべきなのである。これから国際標準化に携わるエンジニアの諸氏に対しては、国際標準化は、利益追求のための道具という論理ではなく、己の持つ技術やアイデアを通じて世の中の発展に貢献するという、もっと高い志を抱いて活動して頂きたい。それが結果的には、企業にとっても己にとっても成功の近道になると考える。

(2011年6月28日受付)

〔文 献〕

- 1) ISO/IEC 15444-1: JPEG2000 Part-1 Second Edition (2002)
- 2) ISO/IEC/JTC1/SC29/WG1 N1390: "Proposal of subdivision ISO/IEC 15444-3: Standardization of Motion JPEG2000"
- 3) ISO/IEC/JTC1/SC29/WG1 N1625: "13th SC29 Maui Plenary Meeting, Resolutions"
- 4) ISO/IEC/JTC1/SC29/WG1 N2705: "Motion JPEG2000 FDIS Draft (including COR1)", T. Fukuhara and D. Singer
- 5) "Digital Cinema System Specification V1.0" (July 20 (2005))
- 6) 福原隆浩, 板倉英三郎: "JPEG2000詳細解説", CQ出版社 (Sep. (2004))
- 7) 福原隆浩: "次世代画像符号化方式Motion-JPEG2000", トリケブス (2002)

福原 隆浩 1986年、早稲田大学理工学部電子通信学科卒業。1988年、早稲田大学理工学研究科修了。同年、三菱電機(株)入社。1997年、ソニー(株)入社。2009年、首都大学東京システムデザイン学部博士課程後期修了。ISO/IEC/JTC1/SC29/WG1委員会メンバ。JPEG2000 Part1 AMD-1, AMD-2のEditorの他、Part3 (Motion JPEG2000) を世界で最初に提唱してTechnical Editorを務める。Digital Cinema Initiative (DCI) の規格化にも従事する。JPEG2000や低遅延動画像通信・映像検索・デジタルシネマ等に関する研究開発に従事。現在、ソニー(株)PDSG PSG本部。1992年電子情報通信学会学術奨励賞受賞。2007年情報処理学会規格調査委員会より標準化貢献賞受賞。博士(工学)。