

USB 接続型トランスコーダの開発

山下 剛*・高島 稔 弘・楠本 高 裕

Development of USB-Connected Transcoder — by Tsuyoshi Yamashita, Toshihiro Takashima and Takahiro Kusumoto — Due to the spread of mobile devices and the evolution of mobile broadband networks, a larger number of mobile device users watch video contents on their own devices. The authors have developed a USB-connected transcoder to convert high-quality contents into suitable formats and qualities for mobile devices. This development allows existing recorders and broadcast receivers to make MP4 files and to send streaming video data to video devices via IP networks.

Keywords: transcoder, mobile, video, USB

1. 緒 言

携帯電話やスマートフォンなどのモバイル端末の普及と高機能化、加えてモバイルブロードバンドの拡大に伴い、モバイル端末上で映像コンテンツ^{*1}を利用する形態が一般化してきている。しかし、コンテンツ自体が据置型機器の大型ディスプレイの普及に合わせて高品質化が進んでおり、ディスプレイサイズや処理性能の制限のあるモバイル機器でそのまま視聴するのは難しい。そこで、高品質な映像コンテンツをモバイル機器で再生可能な形式に変換するトランスコード技術に注目が集まっている。今回、既存の機器に外付けしてトランスコード^{*2}機能を提供する、USB接続型トランスコーダを開発した。本稿では、その概要と評価結果について報告する。

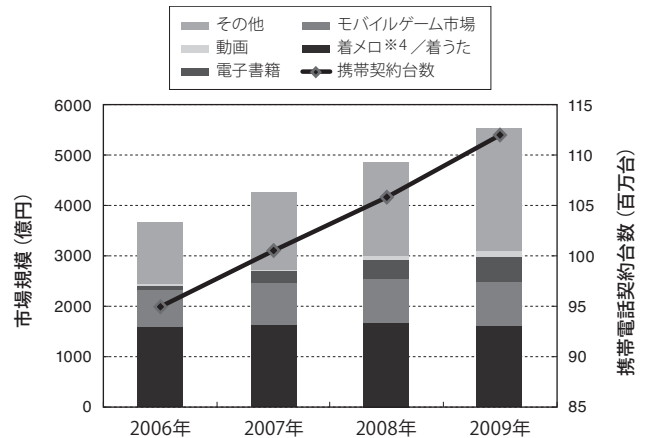


図1 モバイルコンテンツ配信市場の拡大

2. モバイル向け映像配信

2-1 モバイルコンテンツ市場の拡大 固定系ブロードバンドにおいてはマルチメディアコンテンツ配信が一般化して久しいが、携帯電話などのモバイルの分野においても、高性能な端末の普及と、無線ネットワークの大容量化（高速化）が進むことで、いわゆる「着うた^{*3}」に代表される音楽コンテンツの配信が可能となり、さらに近年では、「EZチャンネル」、「BeeTV」等の携帯電話専用の動画配信サービスも広まっている。携帯電話以外の、スマートフォンやゲーム機、携帯音楽プレイヤー、あるいはデジタルフォトフレームなどでも動画視聴機能のサポートが一般化してきており、2009年度に5000億円を突破したモバイルコンテンツ配信市場（図1）のさらなる拡大につながるものと期待されている。

2-2 モバイル機器におけるコンテンツ利用方法 モバイル機器で動画を視聴するスタイルの一般化に伴い、

ユーザ自身が録り溜めたコンテンツや、テレビなどの据置型機器向けのコンテンツをモバイル端末上で利用したいというニーズが高まっている。しかし現状では、一般の放送などで利用されているコンテンツをユーザの携帯電話などでそのまま利用することは難しい。理由としては、高性能化したとはいえ、据置型機器との性能差は大きく、家庭用のテレビに映し出すことを前提とした高品質な映像を扱うだけのスペックがモバイル機器に不足していること、また対応動画フォーマットが異なることなどが挙げられる。さらにモバイル端末同士でも、ディスプレイサイズや対応フォーマットには大きな差がある（表1）。このため、異種端末間で動画コンテンツを共有するためには、各端末に適した形式に変換してやる必要がある。

デジタル映像・音声を別の符号化方式やフォーマットに変換したり、解像度やビットレートなどの動画パラメータ

表1 動画視聴端末仕様一覧

	テレビ	携帯電話 (*1)	A社製携帯型 メディアプレイヤー	S社製携帯型 ゲーム機
ディスプレイ 解像度	HD (1920 × 1080)	QVGA (320 × 240) VGA (640 × 480) WVGA (800 × 480) FWVGA (854 × 480)	480 × 320	720 × 480
コーデック	MPEG2 (MP)	MPEG4 Video (SP) H.264/AVC (BP) VC1	MPEG4 Video (SP) H.264/AVC (BP)	MPEG4 Video (SP) H.264/AVC (BP/MP)
フォーマット	MPEG2-TS	MP4 WMV ワンセグ (TS)	MP4	MP4 avi

*1 携帯電話の仕様は端末依存
SP = Simple Profile, BP = Baseline Profile, MP = Main Profile, HP = High Profile

を変更する機能は、一般的にトランスコード機能と呼ばれ、モバイル機器と動画をつなぐキーテクノロジーとして注目を集めている。

3. USB トランスコーダの開発

3-1 トランスコーダデバイスの概要 今回の開発では、トランスコード機能を有していない既存の機器に対して、外付けでトランスコード機能を追加するための、USB^{*5}ドングルタイプのトランスコーダデバイスの試作を行った。本USB接続型のトランスコーダの機能を利用することにより、録画したコンテンツを変換して携帯電話に持ち出したり、据置型機器で視聴中の放送をリアルタイムに変換してネットワーク配信し、別の部屋の機器や、モバイル端末上で同時に視聴する機能などの実現が可能となる。表2に本USBトランスコーダの入力仕様と用途別の出力仕様の一覧を示す。

3-2 ハードウェア仕様 表3に試作したUSBトランスコーダのハードウェア仕様を、図2にハードウェアブロック構成を示す。機器のデザインは既存の機器のUSBポートに直接接続できるドングルタイプを採用し、サイズは一般的なUSBメモリデバイスより2回り程度大きなサイズに抑えた(図3)。また、ユーザビリティを向上させるため、通常はUSBバスパワー^{*6}で動作する設計としている(ただし、MicroSD^{*7}カードインターフェイスを使用する場合はACアダプタからの外部給電が必要)。メモリは256MBのDDR3 SDRAM^{*8}と1MBのPROM^{*9}を搭載した。大容量のROM^{*10}を搭載するとUSBバスパワー駆動が不可能になるため、ROM上にはブートローダ^{*11}のみを格納し、ブートローダ起動後にUSBホスト側からファームウェア^{*12}をダウンロードしてRAM^{*13}上に展開して起動する方式を採用した。

表2 USB トランスコーダ入出力仕様一覧

入力仕様	
項目	内容
フォーマット	TTS
ビデオコーデック	H.264/AVC MPEG2
オーディオコーデック	MPEG2 AAC MPEG1 Layer II
ビデオビットレート	20Mbps以下
オーディオビットレート	350kbps以下
解像度	1920 × 1080 1440 × 1080 1280 × 720 720 × 480
出力仕様 (動画持ち出し)	
項目	内容
フォーマット	MP4
ビデオコーデック	H.264/AVC
ビデオビットレート	64 ~ 384kbps
オーディオコーデック	MPEG4 AAC
オーディオビットレート	32 ~ 128kbps
解像度	QVGA (320 × 240) QCIF (176 × 144)
フレームレート	30 fps, 15fps
出力仕様 (リアルタイムストリーミング)	
項目	内容
フォーマット	TTS
ビデオコーデック	H.264/AVC
ビデオビットレート	1 ~ 3Mbps
オーディオコーデック	パススルー
オーディオビットレート	パススルー
解像度	VGA (640 × 480) QVGA (320 × 240)
フレームレート	30 fps

表3 USBトランスコーダ・ハードウェア仕様

項目	内容
メモリ	256 MB (DDR3 × 2個)
PROM	1 MB
消費電力	800 mA 以下
筐体サイズ	117 mm × 33 mm × 13.6 mm (図3参照) (突起部含まず)
外部 I/F 等	LED × 1 USB2.0 I/F (プラグタイプ A) × 1 MicroSD カードスロット × 1 5V 外部電源 × 1 *外部電源はMicroSD カード使用時のみ必要

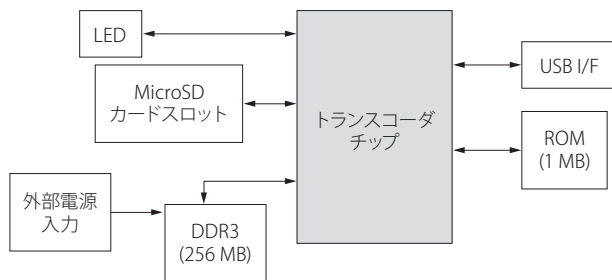


図2 ハードウェアブロック図

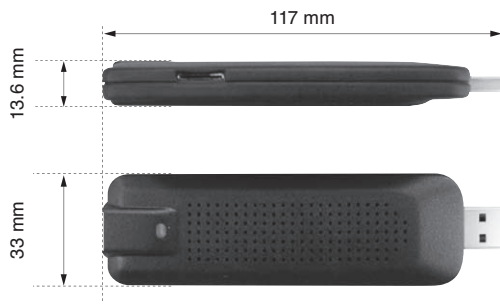


図3 USBトランスコーダ外観

3-3 ソフトウェア及びシステム構成 本USBトランスコーダは、USBホストとIP^{※14}通信を行う。USBホスト側からのUSBトランスコーダの設定、動画ストリーム転送などを全てIP通信で行うため、容易にホスト側のアプリケーションを構築することができる。なお、今回は試作であるためホストとUSBトランスコーダ間のデータの暗号化は行っていない。

続いて、トランスコーダを利用する場合のシステム構成を示す。まず、モバイル端末向けに蓄積コンテンツを変換する場合のシステム構成例を図4に示す。ホストとなる録画装置側では、HDDなどのストレージ^{※15}デバイスに蓄積したコンテンツを読み出して、USBトランスコーダに入力

する。トランスコーダは、入力されたコンテンツをMP4^{※16}形式のファイルにトランスコードしてホスト側に返す。ホスト側ではトランスコーダから受け取ったデータをストレージに書き出す。コンテンツの変換が完了後、ストレージに書き出したファイルを、別途MicroSDカードなどの記録メディアにコピーすることによって、携帯電話などで視聴が可能となる。

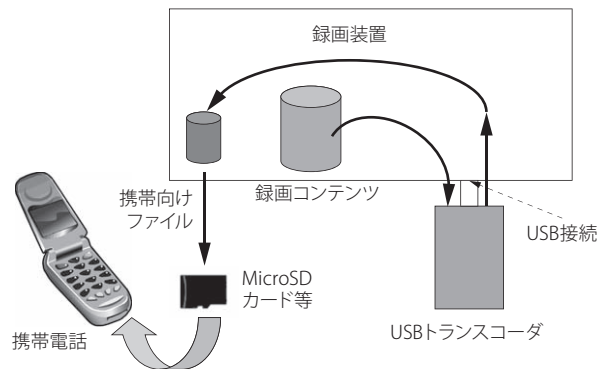


図4 蓄積コンテンツ変換時システム構成例

次に、リアルタイムトランスコードを行う場合のシステム構成例を図5に示す。デジタル放送の受信機などがホスト側となり、視聴中のストリームのコピーをUSBトランスコーダに入力する。トランスコーダ側は入力されたストリームをTTS^{※17}形式にリアルタイムにトランスコードしてホスト側へ返す。ホスト側では、トランスコーダから受信したストリームを、さらに外部のIPネットワークに対して転送することにより、放送波を直接受信することができないモバイル端末上でも、コンテンツの視聴が可能となる。

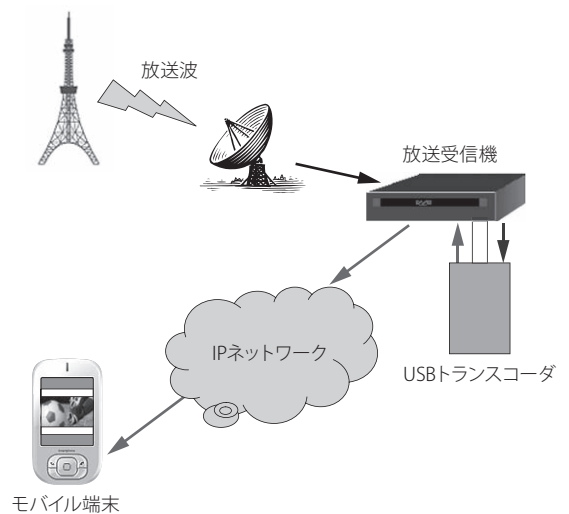


図5 リアルタイムトランスコード時のシステム構成例

4. USB トランスコードの評価

本章では、開発したUSB トランスコードの評価内容と結果について報告する。

4-1 トランスコード組み合わせ評価 まず各種の放送サービス等で用いられている代表的な映像／音声パラメータのストリームのトランスコードが可能かどうかを確認するために、入力ストリームと出力パラメータの組み合わせ試験を行った。表4に入力ストリームのパラメータ一覧を、表5に出力パラメータ一覧を示す。加えて、安定性や放送固有の特殊な挙動への対応評価のために、エージング試験（約3時間）と、マルチ編成^{*18}番組の各チャンネルの選択トランスコード試験、マルチ編成番組と単一編成番組間の切り替え試験などを行った。

組み合わせトランスコード試験に関してはいずれも問題はなかったが、特殊ストリームのトランスコードにおいて、マルチ編成番組と単一編成番組の切り替え時に発生する、コーデック種別の切り替えに対応できないことが分かった（表6）。本機能は今後の対応が必要な課題と考えられる。

表4 テストストリームパラメータ一覧

ビデオコーデック	解像度	オーディオコーデック	オーディオタイプ
MPEG2	1920×1080i	AAC-LC	ステレオ
MPEG2	1440×1080i	AAC-LC	ステレオ5.1ch
MPEG2	720×480i	AAC-LC	ステレオモノラル
MPEG2	720×480i	MPEG1 Layer2	ステレオ
H.264/AVC	1920×1080i	AAC-LC	ステレオ
H.264/AVC	1440×1080i	AAC-LC	ステレオ5.1ch
H.264/AVC	720×480i	AAC-LC	ステレオモノラル

表5 出力パラメータ一覧

動作モード	フォーマット	ビデオパラメータ	オーディオパラメータ
携帯持出・高品質	MP4	H.264/AVC, 320×240, 30 fps, 384 kbps	AAC-LC/2ch, 96kbps
携帯持出・中品質	MP4	H.264/AVC, 320×240, 15 fps, 192 kbps	AAC-LC/2ch, 64kbps
携帯持出・低品質	MP4	H.264/AVC, 176×144, 15 fps, 64 kbps	AAC-LC/1ch, 32kbps
リアルタイムトランスコード・高品質	TTS	H.264/AVC, 640×480, 30 fps, 3Mbps	パススルー
リアルタイムトランスコード・低品質	TTS	H.264/AVC, 320×240, 30 fps, 1Mbps	パススルー

表6 特殊ストリームのトランスコード結果

入力ストリーム	結果
複数の音声ESを含むストリームのトランスコード（オーディオパススルー設定時）	OK
複数の音声ESを含むストリームのトランスコード（オーディオトランスコード設定時）	OK
マルチ編成番組・チャンネル選択トランスコード	OK
マルチ編成番組・単一編成番組切り替え試験 切り替え時にコーデック変化なし	OK
マルチ編成番組・単一編成番組切り替え試験 切り替え時にコーデック変化あり (MPEG2⇔H.264/AVC)	NG

4-2 トランスコード速度評価 次に、トランスコードに要する処理時間の評価を行った。表7にトランスコード速度の測定結果を示す。入力コンテンツの実時間でトランスコードされた場合を1倍速として、各組み合わせパターンにおいて最大何倍速で変換可能であるかを評価した。その結果、オーディオのトランスコードを伴う場合は、ほぼ1倍速にとどまるが、オーディオトランスコードを行わずにパススルーする場合は、2倍速以上の速度が得られることが分かった。対応機器を広げるためにはオーディオも任意の形式に変換できることが望ましく、一方ユーザビリティの観点からは変換速度が速いほど望ましいため、オーディオのトランスコードしながらいかに高速化を図るかが、速度面での今後の課題と考えられる。

表7 トランスコード速度評価結果

入力ストリーム	出力パラメータ	トランスコード速度(倍速)
H.264/AVC, 1440×1080i AAC-LC ステレオ	ビデオ：H.264/AVC, 320×240, 30 fps, 384 kbps オーディオ：AAC-LC/2ch, 96kbps	1.10
	ビデオ：H.264/AVC, 320×240, 30 fps, 384 kbps オーディオ：パススルー	2.21
	ビデオ：H.264/AVC, 640×480, 30 fps, 3Mbps オーディオ：パススルー	2.25
MPEG2, 1440×1080i AAC-LC ステレオ	ビデオ：MPEG2, 1440×1080i, 30 fps, 6Mbps オーディオ：パススルー	3.39
	ビデオ：H.264/AVC, 1440×1080i, 30 fps, 6Mbps オーディオ：パススルー	2.46
	ビデオ：H.264/AVC, 320×240, 30 fps, 384 kbps オーディオ：AAC-LC/2ch, 96kbps	1.03

*トランスコード速度(倍速) = (コンテンツの実時間) / トランスコードに要した時間

5. モバイル機器固有の課題への対応

今回開発したトランスコードの評価の過程で、モバイル機器固有の課題がいくつか判明した。本章では、それらの

課題と対策内容について説明する。

5-1 リップシンク問題への対応 放送波のストリームは、映像ストリームと音声ストリームがトランスポートストリーム (TS) 形式で多重化されて伝送される (図6)。TS形式では映像と音声それぞれに、表示タイミングを示すタイムスタンプ (PTS) が含まれており、受信側でタイムスタンプに従って映像と音声の表示タイミングを調整することで、映像と音声の同期 (リップシンク) を維持する仕組みとなっている。

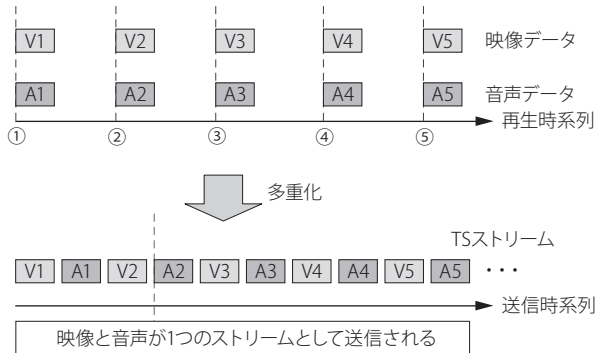


図6 トランスポートストリームによる多重化

一方、携帯電話などで採用されているMP4ファイルフォーマットの場合は、各映像と音声の表示タイミングは、先頭のデータからの相対時刻で記述されている。放送波のストリームをある時点から取り出してトランスコードすると、映像と音声異なる時刻から開始される場合があり、これを単純にMP4ファイルフォーマットに格納すると、映像と音声の開始時刻が一致していないため、先頭からの相対時刻による表示タイミング管理では、リップシンクが維持できなくなってしまう (図7)。

本問題に対応するために、MP4ファイルフォーマットの規格には「Edit List」と呼ばれる仕組みが存在し、同期調整のために映像もしくは音声の開始時刻にオフセットを与えることが可能となっている。今回の試作でも、当初はこの「Edit List」の機能を利用する予定であったが、実際に評価してみると、市販の携帯電話などでは本機能に対応していない場合があることが判明した。その結果、映像と音声の同期に最大数百 msec 程度のずれが発生した。根本的には携帯電話側で規格に準拠できていないことが原因ではあるが、トランスコードとしてはできる限り再生端末仕様に依存しない調整方法を取るべきであると考え、「Edit List」による調整に代えて、ストリーム中の映像と音声のPTSを考慮してMP4ファイルフォーマットへの映像・音声の格納方法を調整する方式を導入することにした。新しい方式では、「Edit List」非対応の端末でも、視聴中に違和感がないレベルでリップシンクを維持することが可能となった。

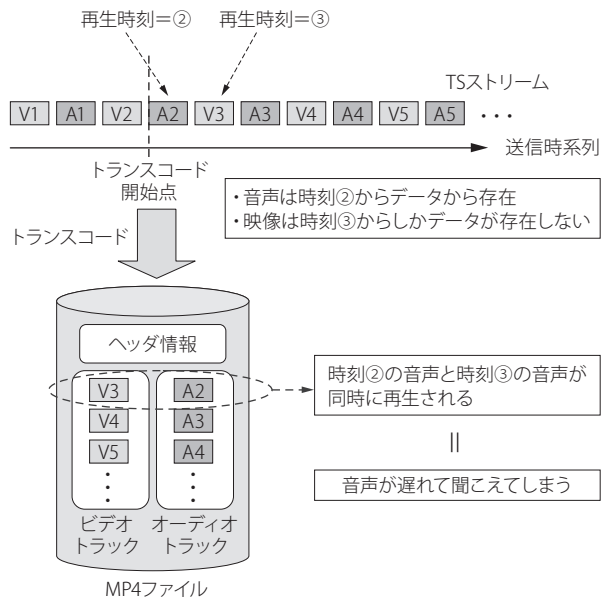


図7 MP4ファイル化によるリップシンクのずれ

5-2 表示アスペクト比問題への対応 映像の各画素の縦横比を示す情報はSAR^{*19}と呼ばれ、通常は映像コーデック部分にこの情報が含まれている。SARは映像の表示端末側で処理されており、例えばSARが4:3の映像の場合は、垂直方向に対して3分の4倍に引き伸ばして表示することにより、適切な縦横比で表示させることができる。異なる解像度へトランスコードを行う場合には、このSARを適切に書き換える必要がある。例えば、図8に示すように、SARが1:1で、縦横比16:9の映像ソースを、320×240の画素数にトランスコードする場合、SARは4:3に書き換える。これにより、トランスコードされた各画素には、水平方向に圧縮された映像が格納されているものの、表示端末側で映像のSARを参照して、画素の水平方向を垂直方向に対して3分の4倍に引き伸ばすことで、正しいアスペクト比で表示することができる。ところが、評価に用いた携帯電話側でSARに対応しておらず常に1:1

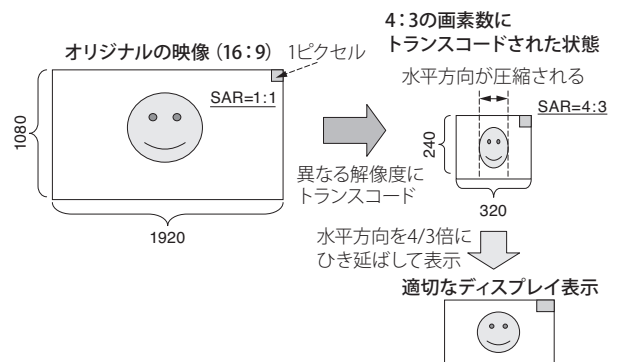


図8 異なる解像度にトランスコードされた映像の表示

で表示されるため、入力映像のSARと出力解像度の縦横比が一致しない場合、歪んだ状態で表示されてしまうことが分かった。

そこで今回の試作では、トランスコード時に入力映像のSARと出力解像度の縦横比を比較して、一致しない場合には図9のように上下に黒帯を挿入したレターボックス^{※20}化することにした。これにより、携帯電話上でも適切なアスペクト比で表示することが可能となった。

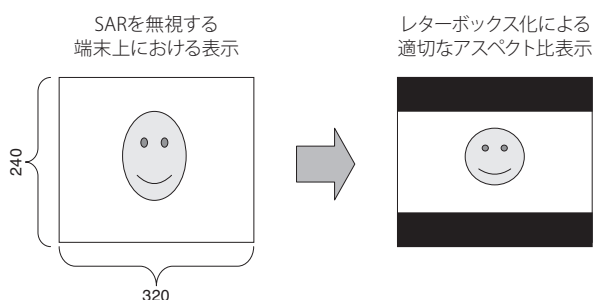


図9 レターボックス化による表示アスペクト比の調整

6. 結 言

今回、トランスコードチップを内蔵したUSB Dongle型トランスコーダを開発した。本デバイスにより、既存の録画装置や受信装置に外付けでトランスコード機能を追加し、蓄積コンテンツのMP4ファイルフォーマットへ変換してモバイル端末に持ち出す機能、並びに視聴中の映像をリアルタイムにトランスコードしてモバイル端末で視聴する機能を実現した。

試作の評価では、コーデックの切り替えを含む特殊ストリームへの対応や、変換速度などに課題があることが判明した。データの暗号化や、モバイル機器に持ち出した場合の著作権保護機能など、サービスを提供するために必須の機能の開発と合わせて、今後対応が必要な課題であると考えられる。

7. 謝 辞

本開発にご協力いただいた住友電工ネットワークス(株)の関係者に感謝の意を表す。

用語集

※1 コンテンツ

Contents：特に、メディアによって提供される、ニュースなどの情報や音楽・映画・漫画・アニメ・ゲームなど各種の創作物を指す。書籍、ウェブページにおいても同様である。

※2 トランスコード／トランスコーダ

Transcode/Transcoder：ある特定のフォーマット・符号化パラメータを有する動画／音声データを、別のフォーマットや符号化パラメータに変換する技術及び変換装置。

※3 着うた

ちゃくうた：携帯電話の着信音をMP3やAACなどのフォーマットで符号化された30秒程度の長さの楽曲にするサービス。

※4 着メロ

ちゃくめろ：携帯電話の着信音として使用される、電子音で作られた楽曲・メロディーのこと。

※5 USB

Universal Serial Bus：主にパソコンと周辺機器を接続する規格として策定され、その後様々な機器同士の接続に使用されるようになった汎用バス規格。

※6 USBバスパワー

USB Bus Power：USBホストからUSBデバイスに給電する仕組み。USB規格で制定されており、これにより対応機器はACアダプタ等による外部給電なしで動作可能となる。

※7 MicroSD

Micro Secure Digital：SDカードアソシエーションにより規定された不揮発性記憶媒体の規格の一種。

※8 DDR3 SDRAM

Double-Data-Rate3 Synchronous Dynamic Random Access Memory：揮発性の半導体メモリであるDRAMの規格の一種。

※9 PROM

Programmable ROM：ROMの一種で、通常は読み込み専用だが、特定の手順でアクセスすることで書き込みが可能になるデバイス。

※10 ROM

Read-Only Memory：コンピュータ用の読み込み専用の不揮発性記憶媒体。

※11 ブートローダ

Boot Loader：システム開始時に呼び出される小さなプログラム。オペレーティングシステムなどの大きなプログラムをメモリ上にロードして開始するための機能を有する。

※12 ファームウェア

Firmware：ハードウェアの制御を行うために機器に組み込まれたソフトウェアのことを指す。一般的なソフトウェアよりもハードウェアに近い位置づけのソフトウェア。

※13 RAM

Random Access Memory：データの読み書きが随時可能な揮発性の半導体メモリ。

※14 IP

Internet Protocol：インターネット上でパケット伝送を行うための規格。

※15 ストレージ

Storage：ハードディスク等、コンピュータ内でデータを記憶する装置。一般的には電気を供給しなくとも記録が消えず（不揮発性）、かつRAMと比べて記憶容量の大きなものを指す。

※16 MP4

ISO/IEC 14496-14で規定された、動画データを格納するためのファイル形式の一種。

※17 TTS

Timestamped Transport Stream：MPEG2システムのTSパケットの先頭に4バイトのタイムスタンプ情報を付加したフォーマット。

※18 マルチ編成

Multi Channel：地上波デジタル放送やBS放送で採用されている、同一周波数で複数のチャンネル（番組）を伝送する仕組み。単一編成時にHD（ハイビジョン）映像を流している帯域に、最大3チャンネルのSD（標準画質）番組を流すことができる。

※19 SAR

Sample Aspect Ratio：デジタル映像の1画素あたりの縦横比。

※20 レターボックス

Letter Box：画面の縦横比と映像の縦横比が異なる場合に、映像の上下に黒帯を追加して表示サイズを合わせこむ方式。

- ・着うたは、(株)ソニー・ミュージックエンタテインメントの登録商標です。
- ・EZチャンネルは、KDDI(株)の登録商標です。
- ・BeeTVは、エイベックス・エンタテインメント(株)の登録商標です。
- ・着メロは、(株)ビジュアルアーツの登録商標です。
- ・MicroSDは、米国SD-3C, LLCの米国及びその他の国における商標または登録商標です。

参 考 文 献

- (1) ISO/IEC 14496-12 (2004-02-01), "Information technology – Coding of audio-visual objects – Part 12", ISO base media file format
- (2) ISO/IEC 14496-14 (2003-11-15), "Information technology – Coding of audio-visual objects – Part 14", MP4 file format
- (3) ISO/IEC 14496-10 (2004-10-01), "Information technology – Coding of audio-visual objects – Part 10", Advanced Video Coding

執 筆 者

山下 剛*：情報通信研究所
映像伝送機器の開発に従事



高島 稔弘：情報通信研究所 グループ長

楠本 高裕：住友電工ネットワークス(株) 映像機器部 担当課長

*主執筆者