

# 次世代画像圧縮技術研究の方向性

八島 由幸 千葉工業大学 情報科学部

## AVC/H.264 は圧縮技術の完成形？

画像圧縮符号化とは、画像信号の持つ膨大な情報量を削減するデータ圧縮技術である。静止画としてはデジタルカメラや Web での閲覧画像、動画としては、地上デジタル放送、ワンセグ放送、DVD、デジタルビデオカメラ、IPTV、Web 動画投稿など、画像圧縮技術は私たちの身の回りのいたるところで使われるようになっていく。たとえば、HDTV 信号の場合、非圧縮の状態だと毎秒 1Gbit の情報量となる。これを、ネットワーク帯域、端末の性能、蓄積媒体の容量等に応じて、何分の 1 かに減らす。圧縮するときのポイントがいくつかあり、できるだけ圧縮率を高くする「高圧縮性」、画質をできるだけ維持する「高品質性」、できるだけ計算処理を軽くする「低演算性」などが重要である。

図-1 に画像圧縮符号化に関する簡単な年表を示す。画像を圧縮符号化して送るという考え方は、古くは 1960 年代から研究があり、1970 年代に入って、フレーム内予測符号化やフレーム間予測符号化の考え方とともに実際の装置化が進んだ。1980 年代に入る

と、動きに追従したフレーム間予測により動きの冗長度を削減する「動き補償フレーム間予測 (MC, Motion Compensation)」と、画像の空間周波数成分で表現することで空間的冗長度を削減する「離散コサイン変換 (DCT, Discrete Cosine Transform)」を組み合わせた方法が主流となった。そして 1991 年にこの「MC+DCT」の考え方に基づいた国際標準規格 H.261 が制定される。H.261 は元々テレビ電話およびテレビ会議を目的にして策定された標準であるが、H.261 がその後に与えた影響は大きく、MPEG-2、H.263、MPEG-4 Visual、MPEG-4 AVC/H.264 (以降 AVC/H.264)、VC-1 など、放送や通信分野で現在幅広く用いられている動画国際標準のほとんどが MC+DCT をベースにして改良を重ねられてきたものである。

MC+DCT をベースとして 1980 年代後半から積み上げられた技術蓄積により、最新の国際標準 AVC/H.264<sup>1)</sup> では大幅な圧縮効率向上が達成されている。1920 × 1080 のフル HDTV に対して画質劣化をほとんど生じさせることなく 10Mbps 以下に圧縮できる。すなわち圧縮率は 1/100 を超えるということになる。

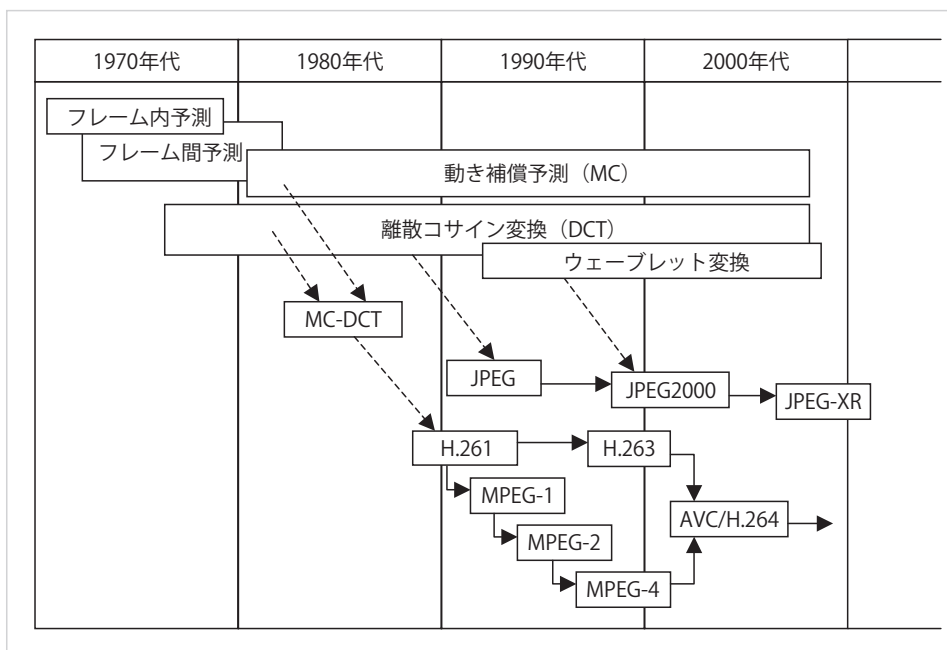


図-1 画像符号化技術と国際標準化の流れ

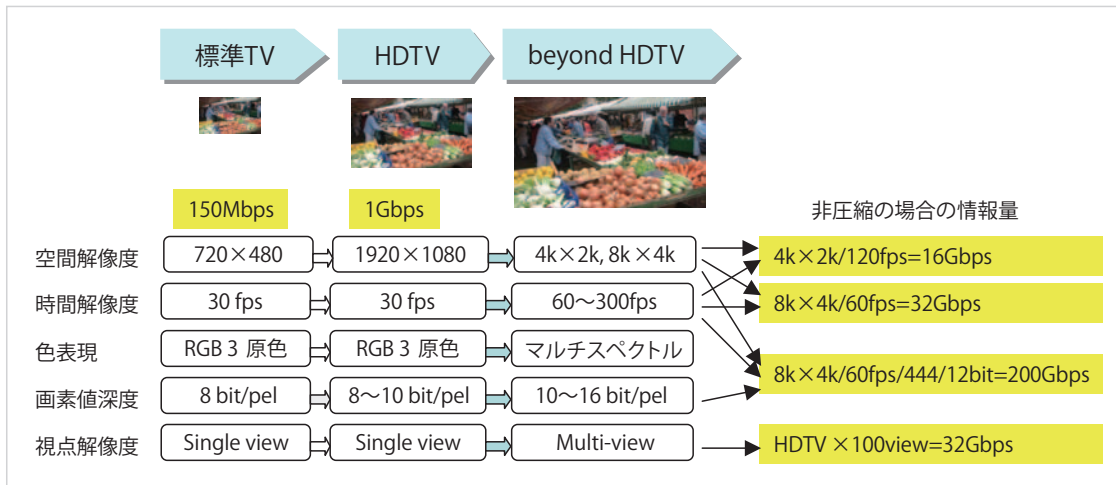


図-2 画像表現形態の拡張と情報量の増加

AVC/H.264 は Blu-ray などの次世代 DVD, IPTV, デジタル放送の IP 再送信, 携帯端末向けワンセグ放送をはじめ産業界への普及も順調に進んでいる。MPEG-2 の 2 倍という高い圧縮効率を達成できる AVC/H.264 があるにもかかわらず今後は語ろうというのが本稿の目的であるから, この上, 何が不足で, 何が求められているのかを見極めなければならない。

### AVC/H.264 の 2~10 倍の圧縮効率が必要

そこで, まずは画像を取り巻く最近の情勢を俯瞰してみよう。地上デジタル放送への完全移行が 2 年後に迫るとともに, 次世代ネットワーク (NGN) により IP ネットワークを利用した画像配信の拡大が予想され, いよいよ HDTV が家庭向けの画像関連産業においても主流となる時代に来たといえる。

一方で図-2 に示す 5 つの観点から HDTV を超える画像表現が可能な技術が研究開発されてきており, すでに一部は市場に登場してきている。5 つの観点とは, 空間解像度, 時間解像度, 色表現方法, 画素値深度, 視点解像度である<sup>2)</sup>。

空間解像度の観点からは, デジタルシネマやパブリックビューイングなどで 4k×2k の画像が利用され, すでに家庭をターゲットにできると思われる 60 インチを超える液晶ディスプレイも実用段階を迎えている。また, NHK では次世代放送に向けてスーパーハイビジョン (SHV) の研究開発が進められ, カメラやディスプレイの開発のほか, コーデックによる圧縮伝送実験も行われている。一方, ディスプレイにおいては 1 秒に 30 フレームという通常の表示を超えて, 2 倍速, 4 倍速で表示をして滑らかな動きを演出する機能がすでに市販品の多くに搭載されており, 時間解像度の観点からも拡張が進んでいる。さらに, より正確な色表現を可能とする色空

間として, RGB/4:4:4 形式, 色空間拡張 xvYCC とともに, マルチスペクトルによって 3 原色表現を超え, 実物と同じ色を表現できるシステムも検討されている。画素値深度では, 従来は 1 画素あたり RGB 各 8 ビット (256 階調) で表現することが多かったが, デジタルシネマや医療・芸術の領域ではこれを拡張し 10~12bit で表現をする画像が扱われるようになってきた。「視点」という観点から考えると, これまで視聴者は画像の作り手依存で決められたアングルに従い, 送られてくる画像をただ見ることしかできなかった。しかしながら, 視聴者が自ら視点を変えて見たい部分を見るという「自由視点テレビ (FTV, Free Viewpoint Television) の概念が提唱されており, この環境ではカメラを複数台 (たとえば 100 視点ならカメラ 100 台) 用意し, そこから出力される複数のアングルに対応した膨大な画像を処理することになる。

以上のような観点から, 画像圧縮符号化への要求条件をまとめると次の 3 点に集約される。

- (1) 情報量の爆発  
質感や臨場感を保持しつつ AVC/H.264 を大幅に上回る超高圧縮符号化が必要となる。
- (2) 新表現形態への対応  
超高精細・高フレームレート・多視点・3 次元など新しい画像表現形態に適した手法が必要となる。
- (3) 画像視聴環境や個人嗜好の多様化  
多様なネットワーク/多様な端末でもストレスなく視聴が可能なスケーラビリティとコーデックアーキテクチャが必要となる。

図-2 には, 前述したような画像表現拡張が進むと情報量が具体的にどうなるかも示している。HDTV が 1Gbps であることを考えれば, 膨大に膨れ上がることは一目瞭然であり, たとえば 8k×4k×60fps で

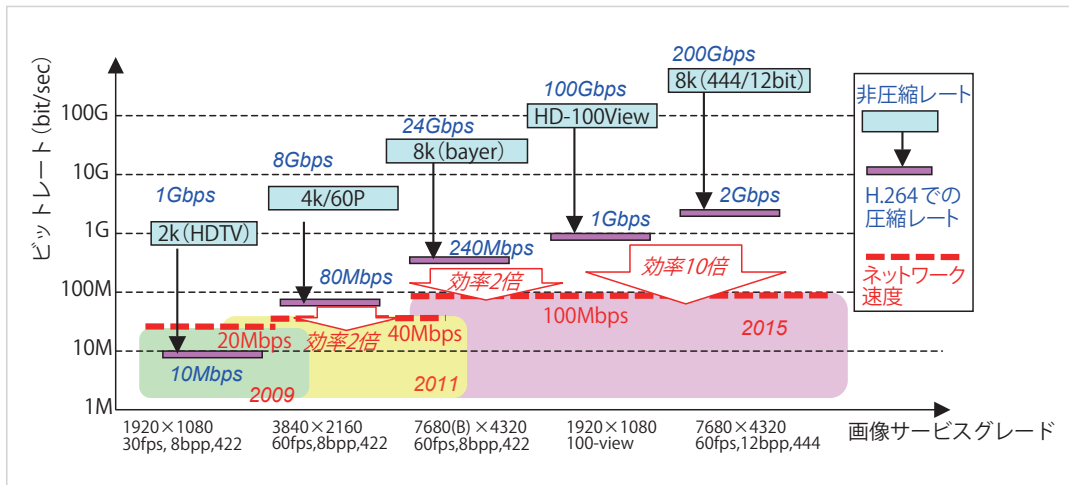


図-3 AVC/H.264 の性能限界

32Gbps, 4k × 2k × 120fps で 16Gbps, HDTV の 100 視点画像で 100Gbps となる。

拡張表現されたこれらの画像を最新の AVC/H.264 で圧縮するとどのくらいまで圧縮できるだろうか。AVC/H.264 では 1Gbps の 4:2:2 非圧縮 HDTV 信号を高画質を保ったまま 10Mbps 程度まで圧縮できることを先に述べた。いま、画像サイズによって圧縮率が変わらないと仮定すると、拡張表現された画像の AVC/H.264 による圧縮データは図-3 に示すようになる。4k × 2k × 60fps (非圧縮 8Gbps) が 80Mbps, 現状のスーパーハイビジョン 8k × 4k (Bayer sampling) × 60fps (非圧縮 24Gbps) が 240Mbps, もしこれが拡張されて 444/12bit の SHV になると非圧縮 200Gbps になるので圧縮後は 2Gbps となる。また、多視点の場合には、32 視点 HDTV (非圧縮 32Gbps) が 320Mbps, 100 視点 HDTV (非圧縮 100Gbps) だと 1Gbps なる。

ここで、これら圧縮後のビットレートと、それらの画像が普及しはじめる時期とを対比させたときにどう捉えるかが重要となる。4k × 2k 画像に関しては、撮像機器やディスプレイはもはや利用可能な状態と言ってよく、ここ 2～3 年 (2011 年ごろ) で普及が始まる可能性を秘めている。SHV に関してはまずはネットワークを利用したパブリックビューイングなどが中心であるが、2015 年以降は家庭への配信も視野に入れた検討が進むものと考えられる。一方で、家庭で自由に利用できるネットワーク帯域としては、現在 HDTV が 2 チャンネル程度 (20Mbps) として<sup>3)</sup>、仮に 2011 年で 2 倍 (40Mbps), 2015 年で 5 倍 (100Mbps) としよう。図-3 において、右上がりの直線はこの利用可能なネットワーク帯域を示したものである。すると、2011 年には 4k × 2k × 60fps の画像を 40Mbps まで下げなくてはならない。すなわち AVC/H.264 の 2 倍の効率を持つ符号化方式が必要である。一方、2015 年において SHV や複数視点

HDTV を扱う際には、AVC/H.264 の 2～10 倍の効率向上が必要だということになる。

## 実現形態から見た課題

圧縮率の要求条件が分かったところで、今度は演算量について考えてみたい。画像圧縮において高圧縮率かつ高画質を実現しようとした場合、レート歪最適化処理 (RDO, Rate-Distortion Optimization) という処理が適用される。RDO とは、簡単に言えば、画像の小領域ごとに、あらかじめ用意されている圧縮ツール (たとえば動き補償のブロックの分け方や、DCT を行う際のブロックの大きさなど) の組合せを片っ端からすべて試して、その領域に一番効率の良いツールの組合せで圧縮処理を実行するというものである。AVC/H.264 は、MPEG-2 等に比べてきわめて多くの符号化ツールを搭載しているため、ユーザはそれらを自由に選択して最適化を図り高圧縮にすることが可能な構成になっている反面、RDO 処理にかかる演算量が膨大なものになる。それでも、最適化の工夫と処理デバイス群の性能進展によって、最近では、Dual CPU 搭載の PC 上で HDTV のリアルタイムソフトウェア符号化処理がなんとか可能な状態にまでなってきた。また、ハードウェアに関しては高性能チップの開発によりはがき大の HDTV 符号化ボードが開発されてきているので、これを 4 枚搭載することで、1U ラック (幅 19 インチ×高さ約 4cm) のコーデック装置で 4k × 2k/30P の画像を処理できると考えてよい。

さて、画像の精細度が HDTV を超えてくるとどうなるかを考えてみる。今、圧縮処理にかかる演算が画素数に比例して増えると仮定しよう (動き検索範囲の拡大などが伴う場合には画素数みあいよりも演算量は増えることに注意が必要)。図-4 (a) の横線は、今後の画像表現拡張に伴って HDTV に比較してどれほどの演算量を必

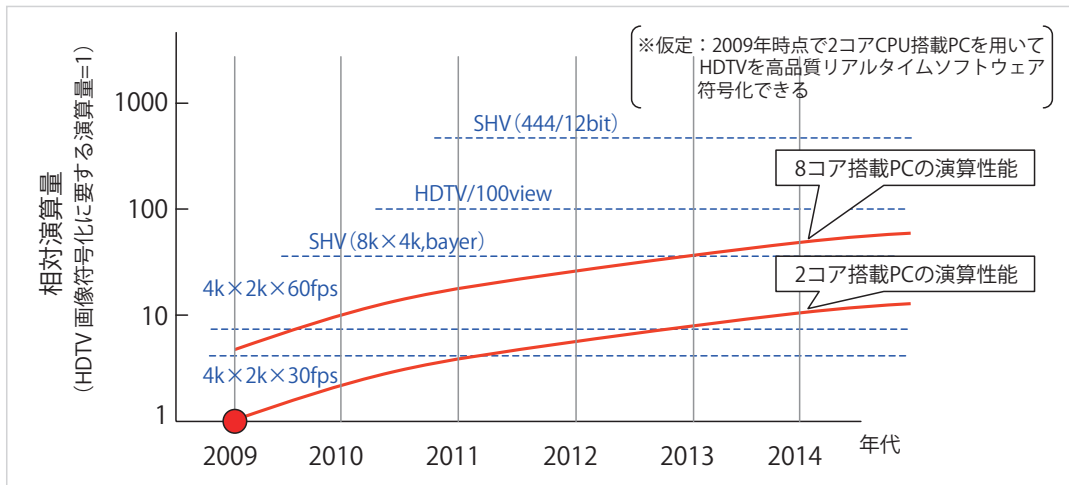


図-4 (a) 画像符号化演算量の見極め (ソフトウェア)

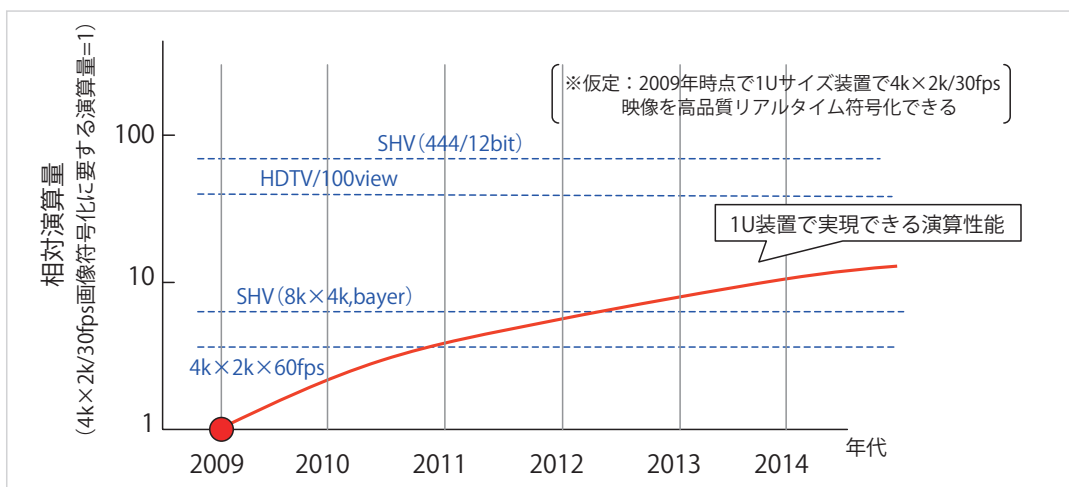


図-4 (b) 画像符号化演算量の見極め (ハードウェア装置)

要とするかを表したものである。また右上がりの曲線は、ムーアの法則に従って演算処理プロセッサの速度が1.5年で2倍になるという向上曲線である。この図から、たとえば、2012年に4k×2k/30P（空間解像度4k×2k、フレームレート30frame/秒）の画像が現状の2コアPCと同じ感覚でソフトウェアエンコード可能になり、2014年には4k×2k/60Pまで拡張しても2コアPCでエンコード処理に多少の余裕が出ることになる。また、8コアCPUを利用すれば2014年にはSHVのエンコード処理も可能で、演算量に余裕が出るが、2コアでは処理量過多で間に合わない。一方、100視点HDTVや444/12bitのSHVに対しては、8コアクラスでも処理が追いつかない。

また、図-4 (b) は、同様に4k×2k/30P画像に比較してどれほどの演算量を必要とするかを表したものである。右上がりの曲線は、図-4 (a)と同様、ムーアの法則に従う演算処理プロセッサの向上曲線である。図-4 (b)ではこの向上曲線をLSIの集積度と読み替えてみると、2011年には4k×2k/60Pが1Uラックサイズで実現可能、2012～13年にはSHVが1Uラックサイズで

実現可能ということになる。

以上、図-4 (a)、図-4 (b)から言えることは、AVC/H.264の演算量は膨大ではあるが、現状のコーデック消費電力や処理PC規模が受け入れられるとするならば、少なくとも2～3年後の4k×2kや5年後のSHVの圧縮符号化をターゲットとした場合には、演算量を削減する必要は必ずしもないということを意味している。さらに言えば、PCの演算時間やチップへの搭載機能には若干の余裕ができ、AVC/H.264に加えて高圧縮化を可能とする新しい符号化処理技術を搭載できるということの意味している。これは今後の新しいコーデック方式を設計していく際には重要な知見である。前章における圧縮効率との議論と絡ませると、「この5年間で新たな圧縮ツールを搭載して符号化効率を2～10倍にする」という設計指針が得られよう。もちろん携帯電話など現状のAVC/H.264では演算量的に重たいという状況も指摘されているので、すべての適用領域に上記が当てはまるわけではないことに注意を要したい。

## AVC/H.264 を拡張する取り組み

画像圧縮符号化の現状と将来についてはここまで述べたような大局的分析ができるのだが、現実はどうのように動いているのかを本章で触れておきたい。MPEG-2 や AVC/H.264 はすでに産業界で広く用いられていることはすでに述べた。現在、国際標準化においては AVC/H.264 の機能拡張が議論されている。図-5 に ISO および ITU で検討されている項目を示す。大きくは、AVC/H.264 の機能的拡張と、AVC/H.264 を超える圧縮率向上の2つに分けられる。

機能拡張においては、以下の4つが代表的なものがある。

### (1) スケーラブル符号化 SVC

AVC/H.264 をベースとした階層符号化。新しい階層間予測などの仕組みを取り入れ、空間解像度、時間解像度、画質の3要素のスケラビリティを高効率に実現する。規格化済み。

### (2) 多視点画像符号化 MVC/FTV

AVC/H.264 をベースにして複数カメラからの画像をカメラ間予測などを使って効率よく圧縮する MVC は規格化済み。FTV ではさらに奥行き情報などを利用してカメラが実際に存在しない仮想カメラ位置の画像合成も可能とするもので、要求条件や基本技術議論が行われている。

### (3) 再構成型画像符号化 RVC

MPEG-2 や AVC/H.264 を構成する圧縮ツールを定義し、それらの組合せによりデコーダを動的に記述できる新しい符号化アーキテクチャ。圧縮データとともにデコーダ記述情報も再生側に送る。ツールの組合せで圧縮方式を決められるため、コーデック設計に自由度が増すという利点がある。

### (4) 高画質画像対応 (xyVCC, 444 画像)

広域色空間への対応や、色信号を空間的に間引かずに輝度と同じ解像度を保ちつつ圧縮符号化するための規格。

これらの詳細は紙面の都合上紹介できないので、文献4)も参考にしていきたい。

一方、高圧縮を狙う方向性としては、MPEG における HVC (High-performance Video Coding)、および ITU-T/VCEG における NGVC (New Generation Video Coding) の2つのプロジェクトがある。前者は、2009年2月の会合にて Call for Evidence (標準化を進めるのに十分な技術の検証を求める呼びかけ) のドラフトが発行され<sup>5)</sup>、2009年度中にもベースとなる方式提

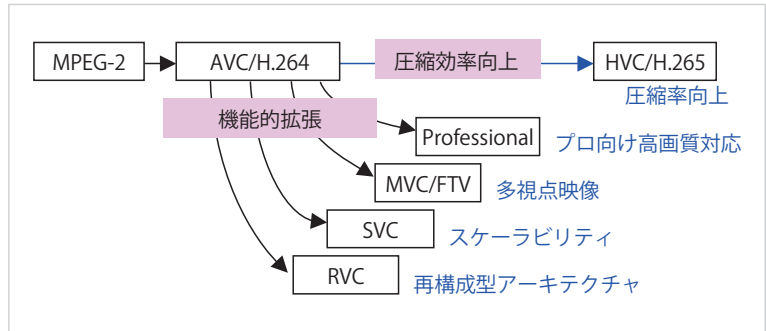


図-5 国際標準化の動向

案が募集される可能性が高い。一方後者は KTA (Key Technology Area) とよばれる参照ソフトウェアをベースにして新しい要素技術組込みによる性能評価が継続されている。いずれも、要求条件や具体的アプリケーションの議論は続けられているところであるが、メインターゲットは「新技術を追加して圧縮率を向上すること」、および「圧縮率は同じ程度でも演算量を削減すること」の2つに集約されている。また、取り扱う画像に関しても、4k × 2k や WVGA (Wide VGA) が明記されるなど、前述した高精細化の方向性を明確化しているのも特徴である。

## 画像圧縮率向上の法則 0.85<sup>n</sup>

このように圧縮率向上は次世代符号化のメインターゲットの1つとなっているが、一方で、AVC/H.264 が多彩な符号化ツールを搭載しており、ある意味で究極の方式といわれる中、この先いったいどこまで圧縮率を向上させることが可能なのかという疑問がわくかもしれない。

ここで、図-6を見ていただきたい。これは、2006年に北京で開催された Picture Coding Symposium において、ハノーバー大学の Mussman 教授が講演で示したスライドをベースに筆者が一部追記したものである<sup>6)</sup>。横軸は年代、縦軸は標準テレビ信号を圧縮符号化した場合のビットレートを対数表示している。その時代の代表的なコーデックプロダクトである、ヨーロッパの国際共同開発プロジェクト COST211 (主な技術はフレーム間予測と DCT)、MPEG-2 (主な技術は MC+DCT)、AVC/H.264 (MC+DCT の最適化) をプロットすると、見事に一直線上に乗る (ただし、これらの画質が一定であるという条件を課さなければならないが、厳密にそれを言うことは難しい)。これまで、MPEG-2 が規格化されたとき、あるいは MPEG-4 が規格化されたとき、いろいろな学会で「画像符号化研究はもうおしまいか？」を議論するようなシンポジウム企画が何度かされてきたが、おしまいどころか MPEG-2 以降も図-6 のようにきれ

いな直線で性能が伸びてきているのである。かなり大雑把ではあるがこの直線を式で近似してみると、1年で15%の符号化効率向上となる。すなわち、ある年にR (bps)まで圧縮できていれば、n年後のビットレートは、

$$R_n = R \times 0.85^n$$

になることを示している。これが現在以降も成り立つとすれば、4年後には $R_n = 0.52$ 、10年後に $R_n = 0.20$ となり、それぞれAVC/H.264の半分、および2割まで圧縮が可能となる。たとえばHDTVを例にとれば、現在AVC/H.264で8Mbps程度で符号化されているものが、4Mbps、1.6Mbpsで同じ画質の保持が可能になるであろうことを意味する。

図-6をベースにもう少しミクロに表現したものが図-7である。1980年を過ぎると「動き補償予測とDCTを組み合わせたいわゆる「MC+DCT」の研究がしだいに注目され、これが第1の大きなブレイクスルーであると位置づけられる。その後、H.261、MPEG-2、H.263、MPEG-4 Visual、AVC/H.264という流れは、MC+DCT方式に多くの改良を施しつつ進展してきている。今、AVC/H.264の次の技術を論じるとき、図-7に示したように次の2つの方向性を考えることができる。

#### (1) Middle term approach

基本的フレームワークとしては、従来のMC+DCTの枠組みを踏襲し、部分的な要素技術を積み重ねて効率向上を狙う。

#### (2) Long term approach

従来のMC+DCTの枠組みにとらわれず新たな観点からの圧縮アルゴリズムを模索する。MC+DCTにかわるブレイクスルーと位置づけられる。

画像圧縮という技術の本質をとらえた場合、これら2つの方向性が具体的にどのような位置づけにあるのかを次章で述べる。

### コンピュータ／ネットワークの発展と画像符号化の新しい体系化

画像圧縮の本質をつきつめると、限られたデータ量の中でいかに伝えたい画像内容を表現するかということに帰着する。1980年代後半に原島らによって提唱された

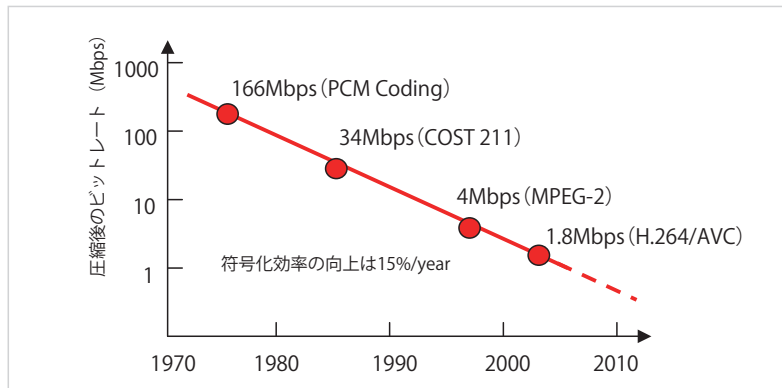


図-6 標準テレビ信号に対する圧縮効率の向上 (Mussman@PCS2006でのプレゼンテーション資料, 「DATA RATE OF VIDEO SIGNALS IN BROADCAST QUALITY」に一部追加)

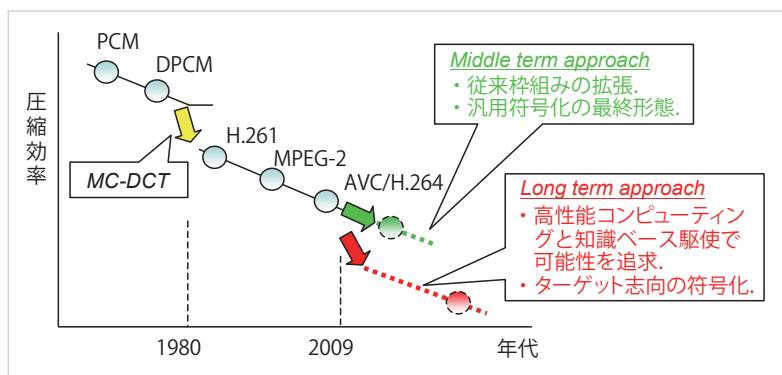


図-7 高圧縮画像符号化研究：2つのアプローチ

「知的符号化」は、たとえば、エンコーダ／デコーダ側に顔などのモデルを持たせ、エンコーダ側で検出した表情情報に従ってデコーダ側でモデルを動かすというコンセプトであり、画像圧縮の本質を突いたものといえる。一方、近年のコンピュータやネットワークの発展、および画像処理技術の進展を鑑みると、図-8に示すような画像圧縮技術の体系化が可能である。

図-8では、「PCM」、「ロスレス符号化 (Lossless Coding)」、「ロッキー符号化 (Lossy Coding)」、「擬似表現符号化 (Pseudo-representation Coding)」、「セマンティック符号化 (Semantic Coding)」の5つのレイヤとして符号化をとらえ、それぞれのレイヤにおける現段階での主要技術と、符号化画像の評価基準も同時に示してある。私たちが目で見ているアナログデータの画像をPCMとしてデジタルデータにしたものを符号化の第一歩として位置づける。PCM符号化における評価基準はデジタル化歪となる。次の「ロスレス符号化」レイヤは、その言葉通り、圧縮データから圧縮前の画像が完全に再構成できる可逆性を持つものである。ロスレス符号化では元画像が完全に再構成されるので画質は評価基準には入らず、圧縮データの符号量だけを競う情報理論的評価となる。そして、次の「ロッキー符号化」のレイヤが、今

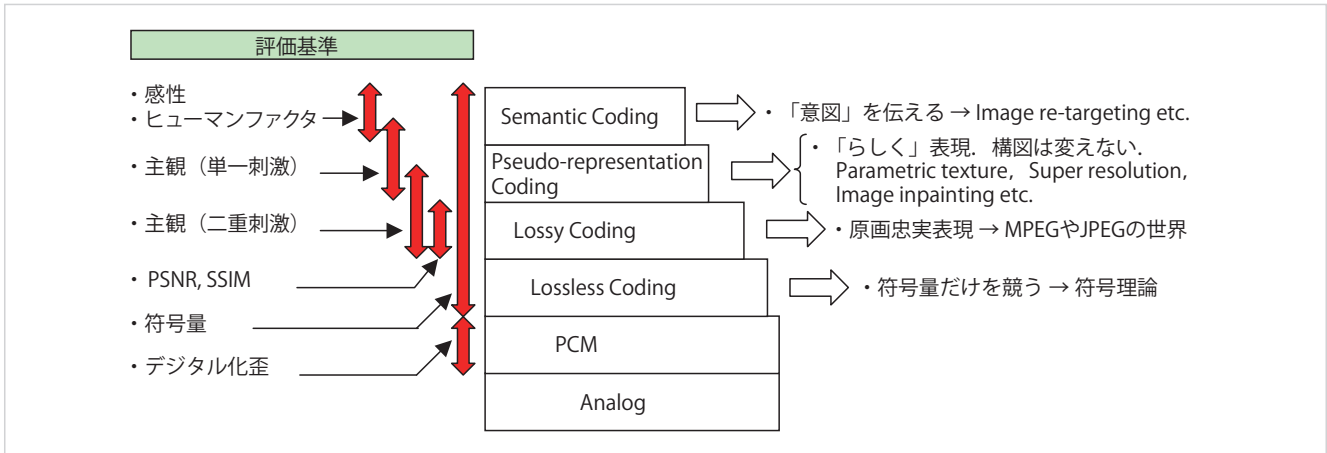


図-8 高圧縮画像符号化ハイアラキー

世の中で使われている MPEG-2 や AVC/H.264 の世界である。このレイヤでは画質の劣化を許容するが、エンコーダ設計指針としては原画にできるだけ忠実に表現することが目的となる。ロッキー符号化における、符号化画像の評価基準として端的なものは、原画像と符号化画像の差分、すなわち SN 比である。MPEG などの国際標準化においては種々の符号化レートでの圧縮性能を総合的に判断するために  $BD^7)$  という指標が用いられているが、これも基本的には SN 比の域を出るものではない。一方、MPEG-2 や AVC/H.264 のエンコーダにおいて、実際には、人間の視覚特性を利用した情報配分が行われる場合もあり、SN 比よりも見た目を重視する符号化制御も採用されている。次に述べるセマンティックなレベルの情報までは利用しないが、視覚特性を利用した周波数変換や量子化方法にはまだ改善の余地は多分に残されており、前章で述べた Middle term approach の技術的な狙いどころではないだろうか。この場合、符号化画像の評価には、SSIM など視覚特性と画素構造の類似性を考慮した画質評価<sup>8)</sup>が適していると考えられる。

さて、次階層の「擬似表現符号化」が、MPEG-2 や AVC/H.264 をはじめとするこれまでの圧縮のフレームワークを大きく変える新しい考え方の1つである。セマンティック符号化では、評価基準として、「似ていればよい」・「らしく表現されていればよい」という主観的側面が重要視される。たとえば、「波」の画像があったとして、必ずしも1つ1つのしぶきを忠実に再現する必要はないとすれば、原画を忠実に再生することを狙うのではなく、なるべく「波らしく表現する」ことで圧縮技術としての評価が高くなるようなものである。

擬似表現符号化が、画像内の個々の対象物をターゲットとしてそれが何であるかを表現して伝える、という位置づけとすると、最終階層の「セマンティック符号化 (Semantic Coding)」は対象物の時空間的位置関係まで

踏み込んで、「画像の意図」を伝えるという考え方である。圧縮によって画像の構図までも変わる可能性があるため、「画質劣化」という概念はもはや通じず、圧縮画像の評価手段は難しくなる。視認性や画像の価値などヒューマンファクタを考慮した評価指標が必要となろう<sup>9)</sup>。

擬似表現符号化、セマンティック符号化の2つを前章の Long term approach と位置づけることができる。

## 新しい観点からの圧縮手法例

本章では、前章で述べた Long term approach の画像圧縮に利用できそうな技術例をいくつか紹介する。(1)は擬似表現符号化、(4)はセマンティック符号化、(3)はその中間に位置する技術となる。

### (1) パラメトリックテクスチャ

画像の中で細かい時空間テクスチャを有する領域、たとえば、芝生、土、空、花畑、水、波、煙、風によるゆらぎなどの領域は、現状の MPEG-2 や AVC/H.264 のブロックベース符号化は得意ではない。ところが、このようなテクスチャ領域は、全体がまとまってそれらしく再生されていれば、画像全体を見たときに違和感がないという特徴を持つ。要するに、再生画の画素値1つ1つが原画の上で対応する画素値と大きく異なっても、ある大きさの領域の中でのいくつかの統計量が似ていれば「それらしく」見える。そこで、図-9に示すように、この領域のテクスチャを表現できるパラメータをエンコーダ側で抽出して、そのパラメータセットを符号化して伝送、デコーダ側では送られてきたパラメータセットからテクスチャを生成することで再生画を得るという手法が考えられる。たとえばグラフカットを用いたテクスチャ合成は該手法の1つと位置づけられる<sup>10)</sup>。

バリエーションとしては、大きなテクスチャ領域 S が

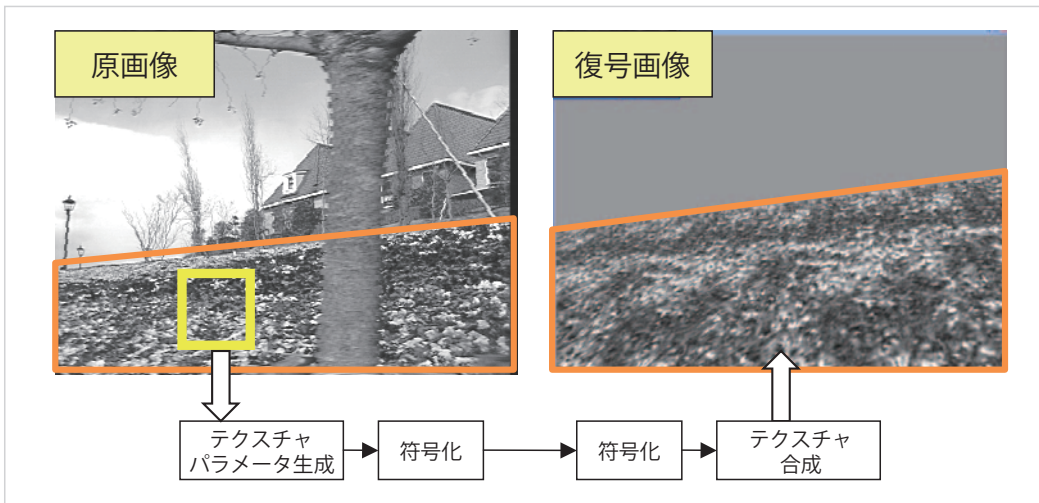


図 -9 Parametric texture coding

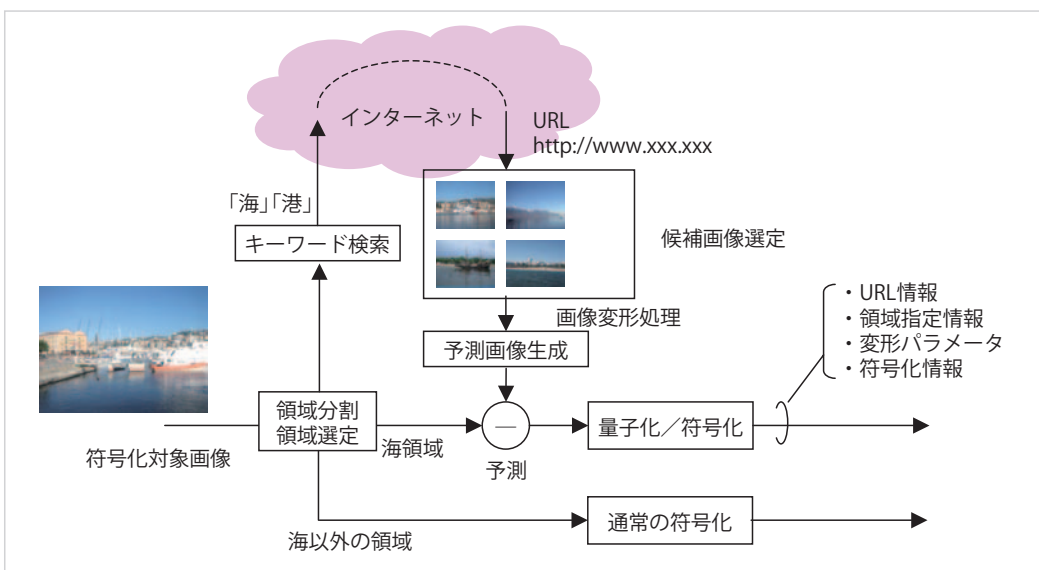


図 -10 Web 画像参照符号化

あるとき、その一部の小さな領域  $s \subset S$  を通常の AVC/H.264 で符号化しておき、デコード側では一部領域  $s$  を再生するとともに、再生された画素情報を時空間にわたる周辺領域にパラメトリックに拡散することで全体領域  $S$  を再構築する手法が考えられる (Texture expansion coding)。

いずれも、テクスチャパラメータ伝送あるいはごく小さな領域の符号化情報だけを伝送すればよいので、画像によっては AVC/H.264 に比べて相当高い圧縮率の達成が見込める。

## (2) 超解像

超解像とは、解像度の低い画像から、本来そこには含まれていない周波数成分を推定して、解像度の高い画像を作り出すものである<sup>11)</sup>。近年では民生のテレビジョンにも搭載されるなど技術的にかなり成熟してきているものもある。超解像を、無から有を作り出す技術というように広義に解釈すれば、一部の領域が欠けていると

きに時空間的に周囲の情報を駆使してその領域を再生する inpainting<sup>12)</sup>や、白黒画像からカラー画像を再生する colorization など含めることができる。画像符号化への応用としては、あらかじめ解像度を落としてから符号化することで符号量を削減し、デコード側では、復号された低解像度画像から超解像技術により高解像度画像を再生すればよい。Inpainting では一部の領域データを完全に送らずデコード側で穴埋めをする、Colorization では色信号情報を最小限(たとえば数十~数百点の画素値)しか送らないようにしてデコード側で画像全体に色をつけることで同様の効果が得られる<sup>13)</sup>。

## (3) Web 画像参照予測

これは、ネットワークの発展と検索技術が可能にした方式と言える。図-10 に考え方を示す。通常、ある対象画像を予測符号化しようとするときに、その対象画像の中でこれから符号化しようとする部分を予測する参照領域は、その対象画像の既符号化部分である。しかしな





図 -11 Seam Carving 適用例



図 -12 Seam Carving 適用例 (表示形の縦横比が原画像と異なる場合)

がら、今や世の中には画像データがあふれてきており、Web で検索すれば、これから符号化しようとする画像に対しての類似画像が一瞬にして多数リストアップされる。それらの類似画像を参照画像として、予測してやろうというのが Web 画像参照予測の考え方である<sup>14)</sup>。さらに一歩進めて、たとえば、画像の中の一部分をまったく符号化せず、検索された類似画像中の一領域で完全に置き換えてしまう<sup>15)</sup> ことでさらに高い圧縮率実現が可能かもしれない。Web 画像参照予測は、手元に蓄積しておく圧縮データ量は少なくできるものの、復号の際にネットワーク経由で参照画像をダウンロードしてこなければならない、この情報量をどのようにとらえるかが課題になる。

#### (4) Seam Carving

動画の中で、そのコンテンツを表現するのに重要な時間部分に対応するフレームだけを抜き出してくる技術はすでにある程度確立されている。シーンチェンジ検出を基本としたいわゆる「画像の速見」を可能にする技術である。コンテンツの意図を伝えるのに不要あるいは貢献度

のフレームを符号化しないことでデータ圧縮が図れる。

Seam Carving<sup>16), 17)</sup> はこれを空間方向に拡張したものであると位置づけられる。図-11 に示す従来例のように、画像を縮小するとき全体を同じ比率で縮小すると、その画像の中で重要な部分(たとえば図-11 では船)も同じ割合で小さくなってしまいますので、携帯電話のような小画面に画像を提示する場合に好ましくない。Seam Carving は、画像の中の重要な部分は間引かずに、たとえば背景などその画像の意図を表現するのに貢献度の少ない領域だけを間引いていくものである。画像の意図を的確に表現しつつデータ量を減らすことができ、画像圧縮の1つの手段として利用が可能だ。さらに図-12 のように画像表示の際に縦横比が変わっても柔軟に対処できる。「重要な領域」「貢献度の少ない領域」をどう定義するかが課題であり、いろいろなアプローチが模索されている。

## まとめ

画像圧縮符号化は今後の画像表現拡張に伴ってますます新しい技術が必要とされる分野である。本稿では、最近のコンピュータとネットワークの進展を背景として今後の符号化技術の方向性をアルゴリズムの観点から論じてみた。ターゲットとする時期を今後3～4年とするかあるいは10年先とするかによりアプローチも異なってくるし、どちらを優先すべきかという議論は不毛であると思う。研究としては両者並立で進められるべきであり、AVC/H.264ベースの高画質化技術もきわめて重要(本文中で述べた Long term approach がすべての画像/すべての対象物に使えるかどうかは疑問の余地があるだろう)、新しい観点から将来の芽を探求することもこれまた軽視できない。なお、本文で取り上げた次世代技術は、筆者の独断も多分に含むことをご承知おきいただきたい。これら以外にも斬新な手法あるいは研究の芽が多数提案されているはずであり、筆者も気付かないだけである。

なお、画像圧縮符号化はアルゴリズムだけでなく、ソフトウェアやハードウェアといった実現技術もきわめて重要である。コーデックの世界もクラウドコンピューティングを駆使し、ローカルに符号化装置やソフトウェアを持つのではなく、Networkの中にコーデック機能あるいは圧縮要素技術を分散させて持ち、それらの協調作業で高圧縮を実現する CaaS (Codec as a Service) という考え方があってもよい。このあたりの詳細については別の機会に触れたいと思う。

**謝辞** 執筆にあたり処理画像の作成にご協力いただいた NTT サイバースペース研究所画像メディア通信プロジェクト画像符号化技術グループの皆様に感謝いたします。

### 参考文献

- 1) Recommendation ITU-T H.264, ISO/IEC 14496-10 AVC, Advanced Video Coding for Generic Audiovisual Services (2003).
- 2) 八島由幸: 高臨場感通信を実現する次世代映像符号化技術, 電子情報通信学会通信ソサイエティマガジン No.2, 秋号, 解説論文, pp.141-166 (Sep. 2007).

- 3) インプレス R&D: インターネット白書 2008, pp.111-114 (2008).
- 4) 八島由幸: SC29/WG11におけるビデオ圧縮符号化最前線, 情報処理学会情報規格調査会, 情報技術標準 NEWSLETTER, No.79 (Sep. 2008).
- 5) Draft Call for Evidence on High-Performance Video Coding (HVC), MPEG Document N10363 (2009).
- 6) Mussman, H. G.: PCS2006 Keynote, Beijing (Apr. 2006).
- 7) Bjontegaard, G.: Calculation of Average PSNR Differences between RD-curves, VCEG-M33, ITU-T. Q.6/SG16 VCEG (Apr. 2001).
- 8) Wang, Z. and Bovik, A. C.: Mean Squared Error: Love it or Leave it? - A New Look at Signal Fidelity Measures, IEEE Signal Processing Magazine, Vol.26, No.1, pp.98-117 (Jan. 2009).
- 9) Image Aesthetics, Mood and Emotion, ICIP2008 Special Session, MA-L1 (Oct. 2008).
- 10) Kwatra, V., Schödl, A., Essa, I., Turk, G. and Bobick, A.: Graphcut Textures: Image and Video Synthesis Using Graph Cuts, ACM Transactions on Graphics, Vol.22, Issue 3, Proceedings of ACM SIGGRAPH 2003, pp.277-286 (July 2003).
- 11) 田中正行: 超解像度映像処理技術, ハイクオリティメディアを追求する映像システム技術の最新動向, Advanced Image Seminar (2008), 画像電子学会講習会 (Apr. 2008).
- 12) Shih, T. K.: Digital Inpainting, IEEE International Conference on Multimedia and Expo, Tutorial T1.3, ICME2006 (2006).
- 13) 宮田高道, 小宮山裕樹, 稲積泰宏, 八島由幸, 酒井善則: Colorization を用いた画像圧縮符号化のための色指定情報抽出法, 映像情報メディア学会論文誌, Vol.62, No.11, pp.1810-1818 (Nov. 2008).
- 14) 遠間正真, 近藤敏志, ハニアアミルシャヒ, 青木孝文: 画像共有サービスを利用した静止画符号化方式, 信学論 (D), Vol.J91-D, No.8, pp.1959-1962 (Apr. 2008).
- 15) Hays, J. and Efros, A. A.: Scene Completion Using Millions of Photographs, ACM Transactions on Graphics (SIGGRAPH 2007), Vol.26, No.3 (Aug. 2007).
- 16) Avidan, S. and Shamir, A.: Seam Carving for Content-Aware Image Resizing, ACM SIGGRAPH 2007, ACM Transactions on Graphics, Vol.26, No.3, Article 10 (July 2007).
- 17) Rubinstein, M., Shamir, A. and Avidan, S.: Improved Seam Carving for Video Retargeting, ACM SIGGRAPH 2008, ACM Transactions on Graphics, Vol.27, No.3, Article 16 (Aug. 2008).

(平成 21 年 4 月 27 日受付)

**八島由幸** (正会員)  
yashima@net.it-chiba.ac.jp

1981 年名大・工・電子工学卒業, 1983 年同大学院工学研究科電子工学専攻修士課程修了。同年日本電信電話公社 (現 NTT) 入社。以来, 主として画像信号の高効率符号化アルゴリズム, HDTV 信号符号化装置, 動画画像信号処理, MPEG 関連システムの研究開発に従事。2003～07 年まで東京工業大学大学院理工学研究科連携教授。2009 年より, 千葉工業大学情報科学部情報ネットワーク学科教授。工博。2004 年画像符号化シンポジウムフロンティア賞, 2005 年度高柳記念電子科学技術振興財団高柳記念奨励賞, 2008 年テレコムシステム技術賞奨励賞, 2008 年 FIT2008 船井ベストペーパー賞を受賞。IEEE シニアメンバ, 電子情報通信学会, 映像情報メディア学会各会員。