

高鮮鋭・高品質な映像拡大技術の開発

携帯端末には、より高品質で見やすい映像の視聴が求められている。動画共有サイトで閲覧可能な符号化歪みの多い映像でも、より高品質な映像視聴を可能にするために、高画質に映像（画像）を拡大する技術を開発した。

先進技術研究所 藤林 暁

ブン チュンセン

1. まえがき

近年、通信のブロードバンド化が進み、動画共有サイトやビデオ・オン・デマンドなどの映像サービスが人気を集めている。モバイル環境下においても、移動端末の高性能化、HSDPA（High Speed Downlink Packet Access）^{*1}の高速通信、パケット定額制の浸透により、映像サービスやワンセグ放送など、移動端末で気軽に映像を楽しめる環境が整いつつある。

2008年5月の調査[1]によれば、移動端末で動画を見る頻度や時間が1年前よりも増加し、特に動画共有サイトの閲覧が増加していて、移動端末で映像を楽しむことが一般的になりつつあるといえる。

また他の調査[2]によると、ユーザは動画共有サイトに対して、「画質」や「画面サイズ」など、より高品質で画面サイズの大きな映像視聴に対する要望が大きい（表1）。こうした要望は、ワンセグ放送に対して

も同様である（表2）[3]。

映像の見やすさにかかわる移動端末機能の1つとして、高解像度ディスプレイの搭載が挙げられる。FOMA 905i以降のハイスペックシリーズでは、VGA（Video Graphics Array）^{*2}以上の表示解像度をもつディスプレイが標準搭載となり、サイズの大きな画像の表示が可能となった。さらに、各メーカーが独自に大型液晶テレビなどで培った高画質技術を移動端末にも搭載し、移動端末上での映像表示はより高精細なものとなった。

一方で、現状の映像サービスの品質は十分に高いとはいえない。動画共有サイトで視聴可能な映像の品質も高いとはいえず、また画像サイズにかかわる映像の解像度も、現在の

映像サービスの主流解像度は移動端末の表示可能解像度に比べて低く、約1/4程度にとどまっている。

そのため、低解像度の映像でも移動端末の高解像度ディスプレイ性能を活かし、画像サイズが大きく見やすい映像を提供する技術を開発する必要がある。

そこで筆者らは、ユーザがより高品質で高精細な映像を楽しむことができるように、ドコモUSA研究所、移動機開発部、株式会社モルフォと連携して低解像度の映像を高画質な映像（画像）に拡大する技術の開発に取り組んだ。

本稿では、映像の見やすさの観点から映像の高解像度化に注目し、映像サービスの現状と課題を整理する。そして、映像を高解像度化する

表1 動画共有サイトへ期待する点

期待する改善点	回答(%)
画質	55.4
画面サイズ	37.8
回線速度	38.2

表2 ワンセグ放送視聴において重視する点

重視する点	回答(%)
映像がきれい	46.6
受信感度が良い	40.8
画面が大きい	34.1

*1 HSDPA：W-CDMA方式に基づくダウンリンクの高速パケット伝送方式。3GPP（3rd Generation Partnership Project）規格上の下り伝送速度は、最大約14Mbit/sである。移動端末の電波受信状況に応じて、変調方式と符号化率を最適化する。

*2 VGA：表示される画像の画素数が640×480dotであること。

映像拡大技術の従来手法の特徴や課題と、その課題を解決することができる開発技術の概要と技術概念およびその性能について解説する。

2. 映像サービスの現状

モバイル環境下の映像サービスでは、伝送帯域が限られているため多大な情報量をもつ映像の圧縮は必須である。また、YouTube[®]*3やニコニコ動画[®]*4に代表される動画共有サイトで視聴可能な映像も、サーバ運用の観点で効率よく映像を圧縮する必要がある。

映像の圧縮には、TV電話・iモーションではMPEG-4 Visual (Moving Picture Experts Group phase 4 Visual)^{*5}、ワンセグではH.264/AVC (Advanced Video Coding)^{*6}などの技術が利用され、動画共有サイトの映像の多くは、FLV (Flash[®]*7 Video)形式の映像コーデックであるSorenson Spark^{*8}やOn2 VP6^{*9}などの圧縮技術が利用される。このような圧縮技術は、映像を少ない情報量に抑えることができるため、伝送帯域の有効活用やサーバ容量軽減の観点で有効な手段といえる。

一方で圧縮技術を利用すると、映像の本来あるべき情報の一部が失われ、映像内にはブロック歪み^{*10}、輪郭付近の映像のちらつき(モスキート歪み)やぼやけなどの符号化歪みと呼ばれるさまざまな歪みが現れる。

符号化歪みの例を図1に示す。符号化後の画像は画面全体がぼやけ、多くのブロック歪みやモスキート歪

みが確認できる。特に動画共有サイトで視聴可能な映像には、このような歪みが非常に多く含まれ、高品質な映像とはいえない場合が数多く存在する。

また、現状の映像サービスの主流解像度は、帯域制限やサーバ管理などサービス運用の観点から、VGA以下にとどまっている。ワンセグやiモーションでは320×240pixであるQVGA (Quarter Video Graphics Array)、TV電話では176×144pixであるQCIF (Quarter Common Intermediate Format)を採用している。動画共有サイトで視聴可能な映像はQVGA程度が一般的となっており、移動端末に搭載されるディスプレイの表示可能な解像度に比べて低く、映像自体の画像サイズが小さい。

そのため、そのままディスプレイに表示すると画像サイズが小さく表示されたり、全画面表示機能を利用した際も、現状では簡単な拡大処理であるため映像がぼやけてしまったりするなど、高品質で画像サイズが大きな映像を十分に楽しめるとはい

えない。

このようなことから、現状の映像サービスは、符号化歪みによる映像品質の低下・低解像度という点で課題があり、移動端末の高解像度ディスプレイ性能に合った高品質で見やすい映像を提供する技術を開発する必要がある。

3. 移動端末での映像の高解像度化

映像サービスの品質・解像度を向上させるためには、伝送帯域向上やシステム変更、コンテンツの用意などの改善を必要とするため、すぐに実施することは難しい。

そこで、より高品質で画像サイズの大きな映像を視聴することができるように、移動端末側で映像を高品質に拡大表示する技術を開発した。

受信した映像を移動端末上でディスプレイの性能に合わせて高画質に拡大(高解像度化)することができれば、さまざまな制限により画質・解像度が低い映像でも、高品質で見やすい映像として視聴することがで



図1 符号化歪みの例

*3 YouTube[®]: Google, Inc.の登録商標。
 *4 ニコニコ動画[®]: 株式会社ドワンゴの登録商標。
 *5 MPEG-4 Visual: ISO (International Organization for Standardization)で規格化した低遅延・高画質を目指した動画配信を目的とした動画フォーマットであるMPEG-4規格の一部で、動画像符号化の方式を定める部分。
 *6 H.264/AVC: ISO/IEC (International

Electrotechnical Commission)のJTC1 (Joint Technical Committee 1)のMPEGとの合同チームであるJVT (Joint Video Team)で規格化された動画像符号化方式の1つ。MPEG-2などの約2倍の圧縮効率を実現し、ワンセグ放送などで標準動画形式として採用されている。
 *7 Flash[®]: Adobe Systems Inc.の米国およびその他の国における商標または登録

商標。
 *8 Sorenson Spark: Adobe社がFLV形式の映像コーデックとして採用している方式の1つ。
 *9 On2 VP6: Flash Player 8から採用されたOn2 Technologies Inc.独自開発のコーデック。

きるようになる (図2)。また、移動端末に取得した映像コンテンツを外部の大型ディスプレイなどに出力するような応用も可能となる。

3.1 従来の映像拡大技術

従来の映像を拡大する技術には、画素補間技術や超解像度技術などが存在する (図3)。

画素補間技術とは、映像を拡大する際に、図3に示す三角印のように本来存在しない画素値を、周辺の画素値を利用して補間する技術である。画素補間技術には代表的なものとしてバイリニア法、ニアレストネイバー法^{*11}、バイキュービック法^{*12}などの技術が存在する。

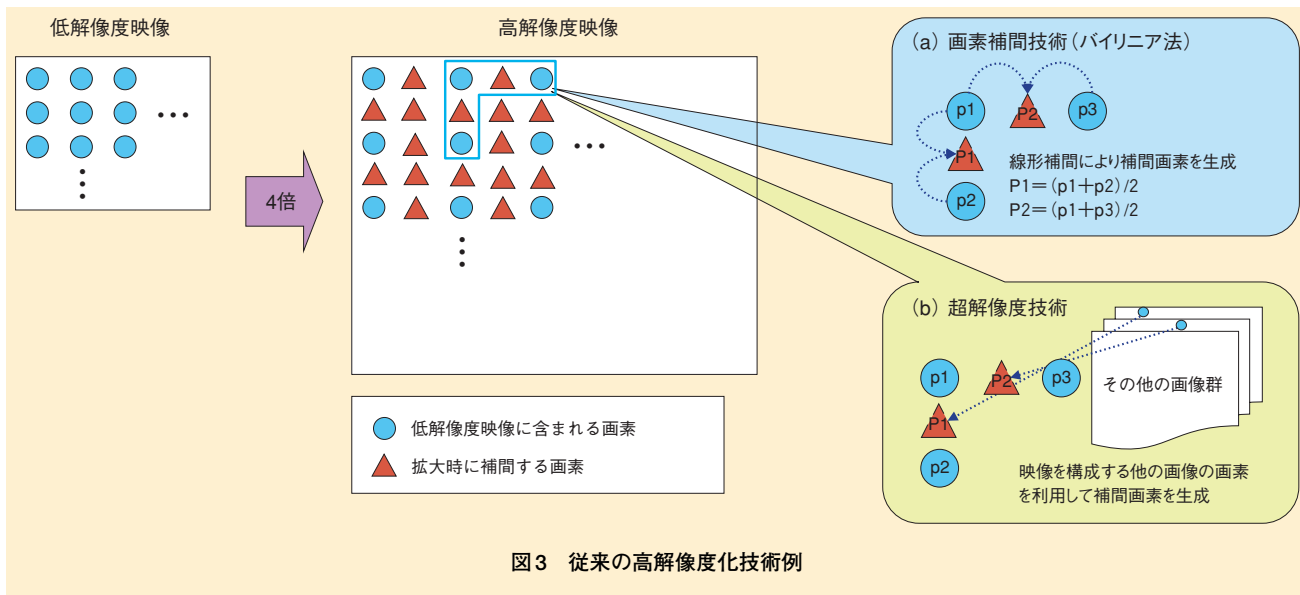
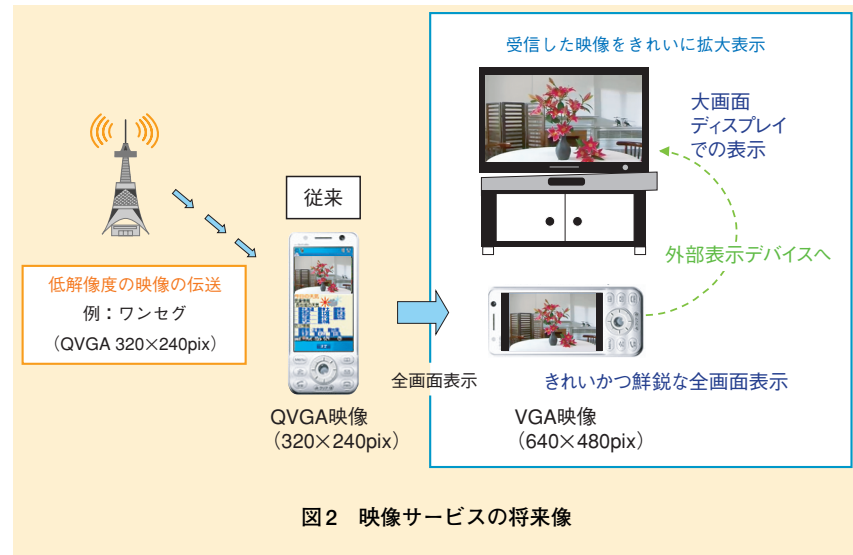
例えばバイリニア法の場合 (図3(a))、隣接する画素の線形補間により、画素と画素の間を補間する。一般に画素補間技術は、演算量が少なく高速な処理が可能であるが、画素

値が滑らかに変化するように画素値を補間するため、鮮鋭度向上に弱く、拡大された画像がぼやけてしまうという課題が存在する。

一方、映像をより鮮鋭に拡大する技術として、超解像度技術[4][5]と呼ばれる技術が存在する。超解像度技術は、一般的に高度な演算を用い

て、映像を構成する複数の画像から本来存在しない画素値を生成し、非常に鮮明な拡大映像を生成する (図3(b))。

しかし、移動端末のような演算能力の限られた機器で、高度な演算処理を行うことは難しい。また、これらの従来技術では、映像を拡大する



* 10 **ブロック歪み**：画面の切り替えや動きの大きなシーンなどで特に見られる、モザイクのようにブロック状の境界が見えてしまう歪み。
 * 11 **ニアレストネイバー法**：画素位置がもっとも近い画素値を補間画素として複製し利用する画素補間手法の1つ。
 * 12 **バイキュービック法**：周辺16画素を用いた多項式演算により補間画素を生成する画素補間手法の1つ。

過程で符号化歪みが強調されたり、映像サイズが小さい時には意識されなかったような歪みが、拡大されて目立つようになっていたりするなど、拡大映像が高品質に見えない場合が存在する。

3.2 映像拡大技術の独自技術

映像に含まれる符号化歪みを軽減でき、画素補間によって発生するエイリアシングノイズ^{*13}やぼやけを抑えることで、高鮮鋭で高品質に映像を拡大する技術を開発した。本技術は、1枚の画像で高鮮鋭化処理を行うことが可能であるため、複数枚の画像を利用する超解像度技術と比較して、低演算量での処理が可能となる[6][7]。

3.3 独自技術の概要

本技術は、「歪み軽減」、「画素補間」、「画像鮮鋭化」の処理を組み合わせ、高画質な拡大映像を生成する(図4)。

本技術には、2つのプロフィールが存在する。主に動画共有サイトに投稿されるような、歪みの多い映像の高画質化を目的とした低演算プロフィールと、技術性能を最大限に発揮し、より高鮮鋭な映像を生成するフルプロフィールである。低演算プロフィールでは、移動端末上のソフトウェアでワンセグ放送をリアルタイム処理することが可能である。

3つの処理のうち、ドコモ独自技術である歪み軽減処理と画像鮮鋭化

処理は、両方とも Over complete transform と呼ばれる手法に基づく処理を行っている。

3.4 Over complete transform

Over complete transform は、画像を構成する画素値が局所的に似ている性質(類似性)を利用した技術である(図5)。図5(a)の赤枠の部位を拡大すると、図5(b)のようにA(茶

色)とB(肌色)と輪郭を示す画素領域が含まれることが分かる。ある画素P(図中「×」)は本来領域Bに属する他の画素と近い値で構成されるが、図5(a)右図緑枠の部位のように符号化歪みが加わると、画素Pが領域Bに属する他の画素と大きく異なる値になることがある。

そこで、さまざまな小領域を考えて局所的類似性が高い部位を特定

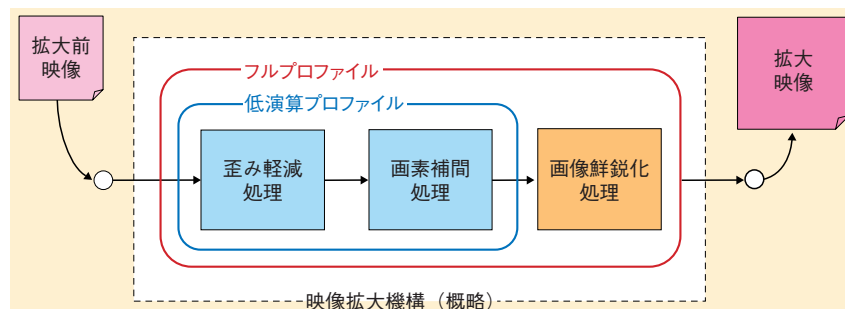
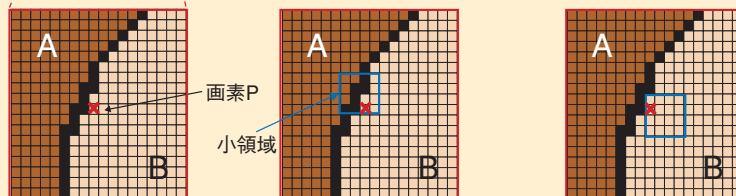


図4 開発技術の構成



(a) 符号化前と符号化後の画像の様子



(b) 赤枠画素領域部の拡大図 (c) 画素の類似性が低い小領域 (d) 画素の類似性が高い小領域

図5 画素の局所的類似性

*13 エイリアシングノイズ：画像に含まれる線や輪郭のキザツキや画像のむら。

し、画素Pと類似性の高い小領域内の値に画素Pの値を近づけることで、歪みの軽減を行う。

例えば、図5(b)について図5(c)と図5(d)のような画素Pを含む異なる小領域を考える。図5(c)の青枠の小領域には画素値の大きく異なる領域AとBの画素が含まれるため、小領域内の画素どうしは類似性が高いとはいえない。一方、図5(d)の場合、小領域内の画素どうしは、同じ領域Bに属する画素で構成されるため類似性が高いといえる。

このとき類似性は、周波数変換時の変換係数の様子で判断する(図6)。図6において、横軸は空間周波数、縦軸は各変換係数の大きさを示す。図5(d)のように画素の類似性が高い場合、その領域を周波数変換すると、大きな値をもつ変換係数の数は図6(a)のように少なく、変換係数 $X(k)$ で表現される。

この小領域に歪みが含まれると、図6(b)のように $X(k)$ は歪みによる影響により小さな変換係数を含む $c(k) (=X(k) + \text{noise})$ となる。そこで、

図6(c)のように $c(k)$ のうち値の小さな変換係数を除去することで、歪みが軽減され符号化歪みのない $X(k)$ に近い $c'(k) (\cong X(k))$ が得られ、 $c(k)$ の類似性を高めることになる。

このように、画素の類似性を高める周波数領域での処理を、各画素につき複数の小領域で行い、画素の類似性の高さに合わせて処理結果に重み付け加算をする。

この処理により、拡大前画像に含まれる符号化歪み、画素補間後の拡大画像に含まれる画素補間に伴うエイリアシングノイズや信号の広がりが軽減される。

そのため、本技術で拡大された映像は、主観的に歪みが軽減され、輪郭のはっきりした映像として見えるようになる。

3.5 独自技術の性能

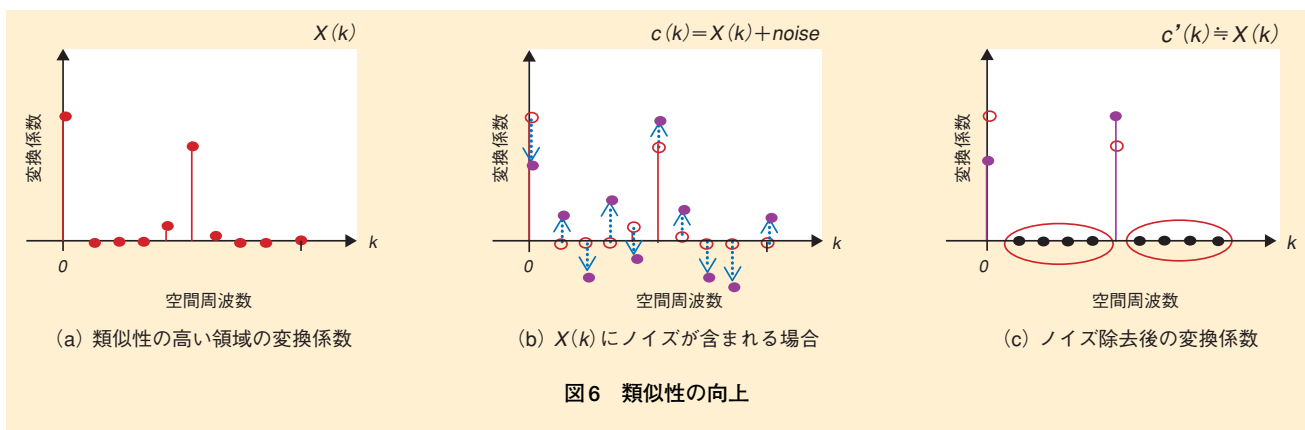
従来技術と独自技術による拡大映像の画質の差を図7と図8に示す。両図とも、(a)は拡大前画像、(b)は従来技術としてのバイリニア法による拡大画像、(c)は独自技術を用い

て縦横2倍に拡大した画像である。両図から従来技術では、符号化歪みが画像拡大されることで見えやすくなり、高品質に見えないことが確認できる。このような歪みは、映像として見る場合にはちらつきとして知覚され、映像品質の低下を招く。

一方、独自技術では人の周りや文字周辺の歪みが軽減され、かつ本来画像に存在する輪郭線などがはっきりとしており、よりメリハリの利いた映像として視聴することができる。

4. あとがき

本稿では、映像サービスの現状とその課題を解説し、移動端末のディスプレイ性能を活かした、より高品質で画像サイズの大きな映像提供を目的とした映像拡大技術の概要と、その利用技術および性能を解説した。本技術では、符号化歪みやぼやけを軽減するメカニズムを組み込むことにより、動画共有サイトに投稿される映像のようにあまり画質が高くない低解像度の映像でも、より鮮鋭で高品質に映像を拡大することが





できる。現在、独自技術の2009年冬以降に販売する移動端末への搭載に向け、各社メーカーやベンダにて搭載検討を実施している。

今後の課題として、ベンダなどとの協力体制の構築や、フルプロファイルも含め、本技術の商用展開に向けた各種アプリケーション（映像資産の有効活用やDVDプレーヤー、メディアプレーヤーでの高画質拡大表示など）での映像の拡大性能検証や改善などを進めていく。

文 献

- [1] モバイルリサーチ：“ケータイの動画コンテンツ視聴に関する調査。”
- [2] 株式会社アイシェア：“動画共有サイトとテレビ連携に関する調査。”
- [3] ネットリサーチDIMSDRIVEアンケート：“ワンセグ機器に関するアンケート。”
- [4] Y. Altunbasak, A. Patti and R. Mersereau：“Super-Resolution Still and Video Reconstruction From Mpeg-Coded Video,” IEEE Trans. Circuits Systems for Video Tech., Vol 12, 2002.
- [5] S. Farsiu, D. Robinson, M. Elad and P. Milanfar：“Fast and robust multi-frame super-resolution,” IEEE Trans. on Image Processing, Oct. 2004.
- [6] S. Kanumuri, O. Guleryuz and M. R. Civanlar：“Sparse directional transforms for super resolution reconstruction of mobile video,” in Proc. SPIE, San Diego, Aug. 2007.
- [7] S. Kanumuri, O. Guleryuz, M. R. Civanlar, A. Fujibayashi and C. S. Boon：“Temporal Flicker Reduction and Denoising in Video using Sparse Directional Transforms,” in Proc. SPIE, San Diego, Feb. 2008.