

RI-002

非線形特性を応用した画像の超解像度化 A Novel Super Resolution Method with Non-Linear Function

合志 清一[†] 寺川 雅嗣[†] 三上 浩[†] 今井 繁規[†]
Seiichi GOHSHI[†], Masatsugu TERAGAWA[†] Hiroshi MIKAMI[†] and Shigeki IMAI[†]

1. まえがき

低解像度画像から高解像度画像を作成する試みは数多く行われ様々な手法が提案されてきた。この分野は Image Restoration と呼ばれ、精力的に研究がおこなわれている。近年、Super Resolution と呼ばれる新たな分野が形成され新たな注目を集めている[1]。Super Resolution の和訳である「超解像」は、テレビ産業でも使用され一般的な言葉となりつつある[2]。Super Resolution に関する研究は幅広く行われ、数多くの提案がおこなわれてきた。多くの提案は実験に用いる画像にある仮定を設け、その仮定の基に議論を進めている。これらの研究は成果を上げているものの、制限された画像や静止画に絞った実験結果の報告が多い。また、フィードバックループを使用した巡回処理（イテレーション）を必要とするものも多い[3][8]。

放送や映画は我々に身近なメディアであり高解像度化の要求は高く様々な研究が行われ、エンハンサー等実用化されたものも多い。しかしながら、一部を除き Super Resolution の研究成果を放送や映画に十分生かすことができていない状況にある。放送は1秒間に60または50フレーム、映画でも1秒間に24フレームのリアルタイム処理が必要である。遅延や負荷の問題からイテレーションを必要とする処理は導入困難である。また、放送、映画ともに映像の性質も多岐にわたる。このため、映像に一定の性質を仮定する手法の導入も困難である。一部の受信機には最新の研究成果が導入されているが、ノウハウの固まりであり公開された文献から詳細を知ることは困難である[2][4]。

著者らは絶対値回路を用いた簡単な非線形処理でナイキスト周波数を超える周波数成分を再生し、画像の高解像度が可能であることを示した[5]。

今回、非線形回路に3乗特性を用いた新たな手法を考案した。考案した手法を、絶対値を用いた処理で画質改善効果が十分ではなかった画像に適用し、良好な結果を得たので報告する。以下、2章では従来技術とその課題について述べ、3章では提案技術について述べる。4章では実験結果を示し、5章では実験結果について周波数領域で考察を行い、6章で結論を述べる。本論文ではテレビ、映画等の動画をリアルタイムで高解像度化するアルゴリズムに議論を集中し、イテレーションを必要とするノンリアルタイムアルゴリズムは検討範囲外とする。

2. 従来技術とその課題

精細感を増す手法として広く用いられている手法にエンハンサーがある(図1)。エンハンサーはリアルタイムで動画像を処理することが可能であり、テレビカメラ、テレビ受信機、画像処理ソフトウェア等に幅広く利用されている。

しかしながら、その構成から明らかなように、ナイキスト周波数を超える周波数成分の再生は困難であり、ナイキスト周波数近傍にエネルギーがない画像に対しては精細感を向上させることができない。SDTV をアップコンバートして HDTV ディスプレイに表示する場合、変換後の信号は HDTV のナイキスト周波数近傍の信号を保有していない。このため、アップコンバートした信号に上記エンハンサーを施しても、解像度はほとんど向上しない。近年注目されている 4K ディスプレイ[6]に HDTV 信号をアップコンバートして表示する場合も同様の課題がある。

$$\begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{pmatrix} \dots (1)$$

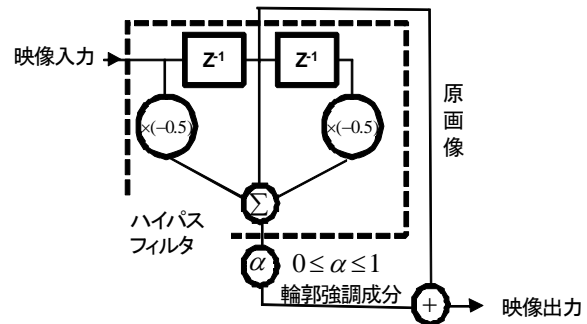


図1 従来技術のエンハンサー

図2 実験に用いた原画像



具体例を用いてこの課題を検証する。図2に示す画像にゼロを挿入して縦横それぞれ2倍に拡大し、式(1)を用いて補間を行う。図3はその一部を切り取った画像図である。

[†]シャープ株式会社 研究開発本部, Corporate Research and Development Group, Sharp Corporation



図3 実験に用いた画像(図2を縦横2倍に拡大し切取)

図3の画像を図1のエンハンサーにより処理した結果を図4に示す。処理は水平方向と垂直方向にそれぞれ1次元処理として行った。式(1)により作成した図3はナイキスト周波数近傍の信号は保有しない。前述のように図1のエンハンス処理はナイキスト周波数近傍の帯域を強調するので、該当する周波数成分を有しない画像に適用しても解像度感の向上は期待できない。実際、図3と図4を比較しても画質差は認められない。

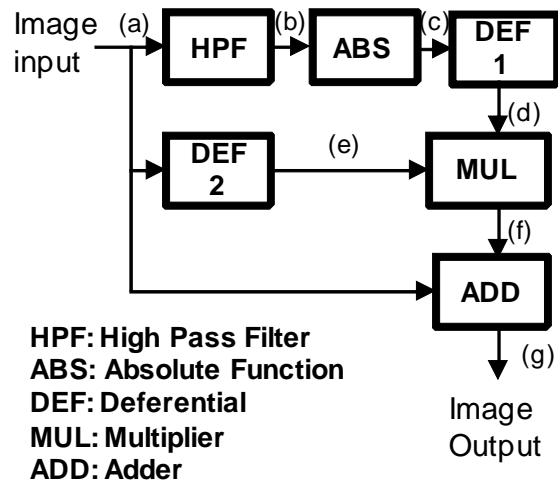


図5 ノンリニアエンハンサー



図4 図1の従来技術のエンハンサーによる処理結果

紙面の都合により図2は縮小して掲載しているが、画質を判定するために図3,4および後掲の図6,10,11は縮小せずに掲載している。



図6 図5の処理による高解像度化処理結果

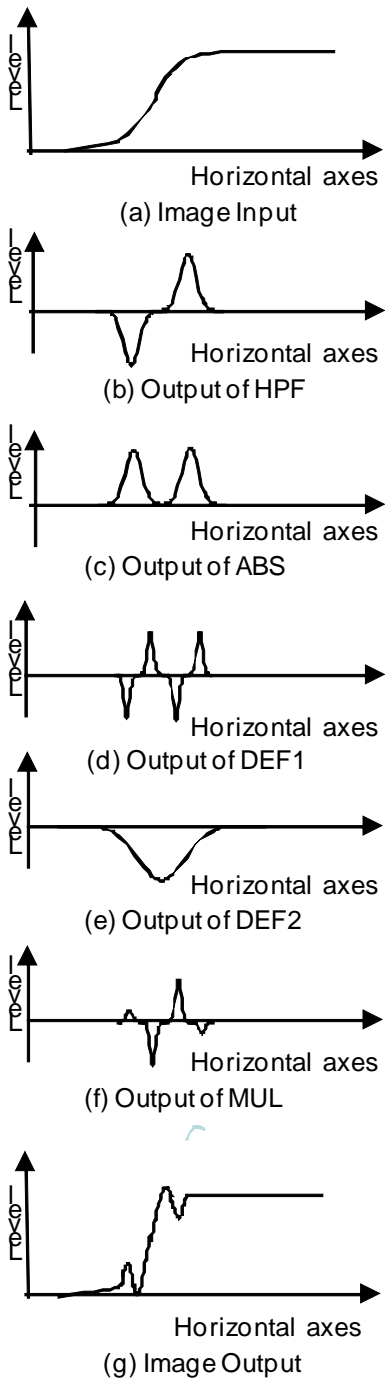


図7 図5の(a)から(g)の信号

エンハンサーはリアルタイムで画像の解像度感を改善する手法として広く用いられているが、ナイキスト周波数周辺に成分を有しない画像には効果がない。SDTV⇒NTSCアップコンバート、HDTV⇒4K 画像は動画像であり、ナイキスト周波数近傍の成分を有しない。アップコンバートした画像の解像度をリアルタイムで改善する有効な手法はこれまで存在しなかった。

ナイキスト周波数を超える周波数成分をリアルタイムで生成する手法として著者らが提案した手法を図5に示す。

図5(a)に図7(a)に示すなまの信号が入力された時の図5(b)~(g)の波形を図7(b)~(g)に示す。処理の詳細は参考文献[5]を参照いただきたい。図5のアルゴリズムでは非線形処理であるため、信号の周波数、レベルにより特性が異なる。このため、線形フィルタのように単一の図で周波数特性を示すことはできない。

図3の画像を図5により処理した結果を図6に示す。処理は水平方向と垂直方向にそれぞれ1次元処理として行った。図6を図3,4と比較すると、あごの下部分のスクリーンや右上の椅子の背もたれの解像度が確認できる。しかしながら、目や鼻等の改善はやや物足りない。

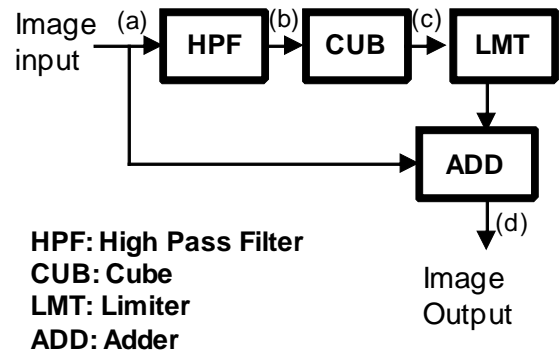


図8 本論文で提案する信号処理

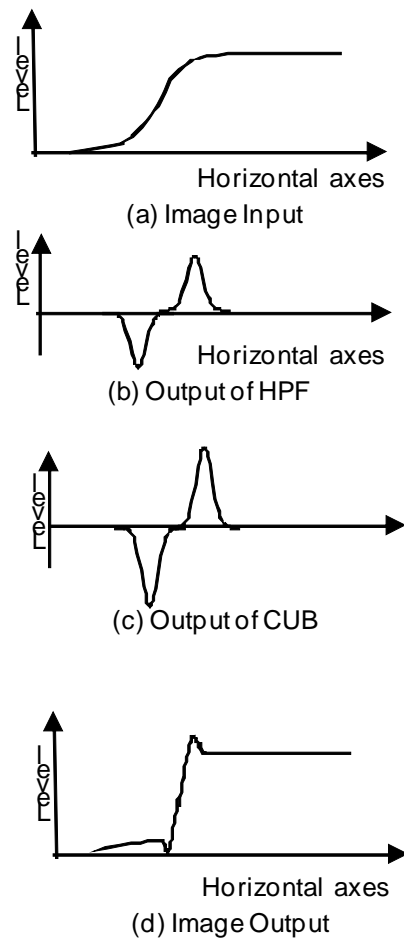


図9 図8の(a)から(d)の信号

人間の顔は様々な画像に現れるので、画質評価には重要な素材である。スカーフや椅子の背もたれは図 3 でもはっきりしたエッジとなっているが、目鼻は比較的低いレベルのエッジである。微少レベルのエッジは総合的な画質評価に影響するので、何らかの手法による改善が求められる。

3. 提案技術

新たに提案する高解像度化信号処理のブロックダイアグラムを図 8 に示す。同図の各部の働きを説明するために、図 9(a)に示すエッジ信号を例に取る。図 9(a)~(d)各信号の横軸は画面の水平または垂直方向、縦軸は信号レベルである。尚、図 9(b),(c)は微小信号であり、縦軸のスケールは図 9(a),(d)とは異なる。

図 9(a)の緩やかなエッジを有する信号が、図 8 左上の画像入力に供給されたとする。信号は 2 系統に分岐し、一方はハイパスフィルタで処理され、図 9(b)に示す波形となる。ハイパスフィルタ出力を CUB ブロックで 3 乗し、図 9(c)を得る。CUB ブロックの出力は LMT ブロックで α ($0 < \alpha < 1$) と乗算し強度を調整する。更に、一定値以下の値となるように波形をクリップする。クリップの条件の一例としては ± 32 が考えられる。CUB ブロック入りに α を乗じた結果が、クリップレベルに達しない場合クリップ処理は行わず、ADD ブロックに信号を供給する。ADD ブロックでは入力信号図 9(a)と図 9(c)に LMT 処理を行った信号が加算され、エッジが急峻化された図 9(d)を得る。

図 8(d)はエッジ部分が幅の狭い輪郭強調信号で高解像度化されている。このエッジは原信号のナイキスト周波数よりも高い周波数成分を持っている(APPENDIX A 参照)。



図 10 図 8 に示す提案手法による高画質化処理結果



図 11 図 5 および 8 による高画質化処理結果

本提案は図 1 に示すエンハンサーのように精細感を増す技術ではなく、原信号が保有しないナイキスト周波数を超える精細化成分を再現し、画像の精細化を図る手法である。

4. 実験結果

図 3 の画像を用いて図 8 の信号処理を行った結果を図 10 に示す。図 8 の処理を水平方向と垂直方向に一回ずつ行い、図 10 を得た。

図 5 の処理による図 6 と図 10 とを比較すると、スカーフやいすの背もたれの改善は図 6 が勝るが、目、唇等の画像信号の振幅が少ない部分は図 10 が勝っている。

他の画像でも実験を行ったが、振幅が小さな信号は図 8 に示す信号処理が良い結果を残し、振幅が大きな信号は図 5 に示す信号処理が良い結果を示している。

画像信号が 8 ビットで $0 \sim 255$ の場合、ハイパスフィルタ出力は $-255 \sim 255$ となる。図 8 と図 5 では何れもハイパスフィルタ出力に対して非線形処理を行っている。図 8 の非線形処理は 3 乗であり、図 5 のそれは絶対値処理である。これらの非線形処理への入力が $-255 \sim 255$ であれば、3 乗処理は絶対値処理よりも大きな値を出力する。このため微少振幅は 3 乗処理の方が、絶対値処理よりも良い結果を与えるものと考えられる。

画像全体の解像度を向上させるためには、絶対値処理と 3 乗処理を共に行う方が良い結果を与えるものと考えられる。

図 11 は図 5 と図 8 の両者の処理を行った結果である。いずれの処理も水平、垂直に 1 次元処理として各 1 回行った。図 11 は微細な部分まで精細度が向上しており、図 6,10,11 のなかで、最も良い結果を示している。

図5の処理と図8の処理は非線形処理であるため、処理の順序を変更すると若干異なった結果を与える。ここでは図8の処理により微細な部分を精細化した後、図5の処理のより振幅が大きな部分の精細化を行った。

逆順の処理も行って見たが、目鼻等の細かい部分の精細化が十分ではなく、一部の輪郭が強調されたアニメーションの様な画像となる。主観画質は上記順序の方が勝る結果を与えた。

図5、図8のアルゴリズムはともに非線形処理であるため、処理の順序によって処理結果が変化することは予想される。図5のアルゴリズムは複数の周波数を発生するが、図8のアルゴリズムは3倍の周波数のみを生成する。少ないリソースを有効に活用してリアルタイム処理装置を開発するためには更なる検証が必要である。

5. 考察

2章で述べた様に図1に示す従来型のエンハンサーは、解像度感は向上させるが、解像度は向上しない。一方、本提案を画像に適用すると解像度が向上する。この点を周波数領域で考察する。

図3に2次元FFTを施した結果を図12に示す。図3は図2を縦横それぞれ2倍に拡大した画像であるため水平方向は $-\pi/2$ から $\pi/2$ の領域にのみスペクトラムが存在する。垂直方向は $\pm\pi/2$ 近傍に折り返して生じたと考えられるスペクトラムがわずかに存在するが、ほぼ $-\pi/2$ から $\pi/2$ の制限された帯域のみにスペクトラムが存在すると言える。この状況は2章で述べたことと一致する。

図3に図1のハンハンス処理を施した図4の2次元FFTの計算結果を図13に示す。図12と図13とを比較すると、図13では垂直方向 $\pm\pi/2$ 近傍のスペクトラムがやや強調されているが、両図に大きな差はない。

図1のエンハンサーは $\pm\pi$ 近傍を中心に強調する。図12から明らかなように、図3の $\pm\pi$ 近傍には強調すべき成分が存在しない。よって、図1のエンハンサーを図3に適用しても大きな効果は得られない。上記の議論により図1のエンハンサーが縦横2倍に拡大した画像に効果のないことを周波数領域で示した。

図11に2次元FFTを施した結果を図14に示す。図14から明らかなように、スペクトラムは全領域に幅広く拡大している。スペクトラムの拡大は画像が精細化されたことを示しており、図3から図11へ改善結果とも整合する。

ここで注目すべき点は提案方式が原画像である図3が保有しないスペクトラムを作り出している事である。図3,11,12,14より、提案方式は原画像が保有しない成分を再生し高画質化していることを周波数領域で示した。

リアルタイムで動画像の解像度を増す簡単なアルゴリズムはほぼエンハンサーに限られ、テレビやビデオカメラに幅広く搭載されてきた。しかしながら、エンハンサーは解像度感を向上させることは可能であるが、解像度を向上させることは不可能であった。提案するアルゴリズムは解像度を向上させることが可能であり、ハードウェア化・リアルタイム動作が可能である。ハードウェア化の実現により、テレビやビデオカメラの他SDTV⇒HDTVコンバータ、4K表示装置等、幅広い応用が可能となる。これまでいてレーシ

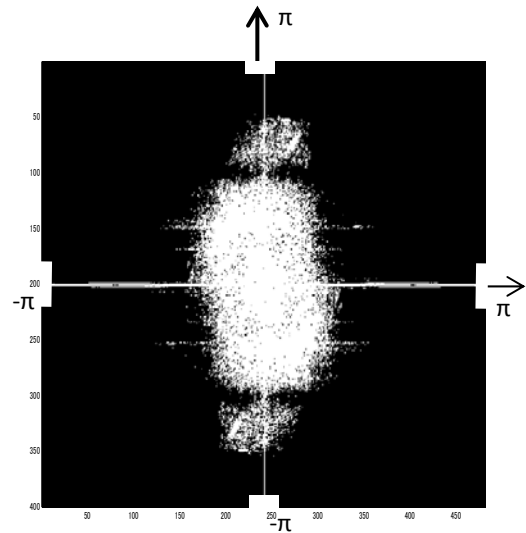


図12 図3の振幅特性

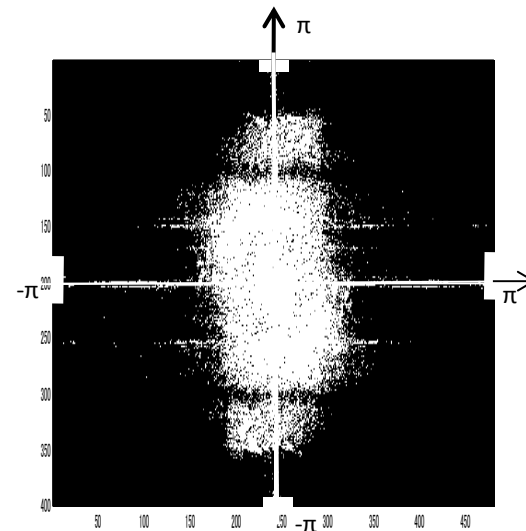


図13 図4の振幅特性

ョンを用いた高解像度技術も提案されているが、本研究と

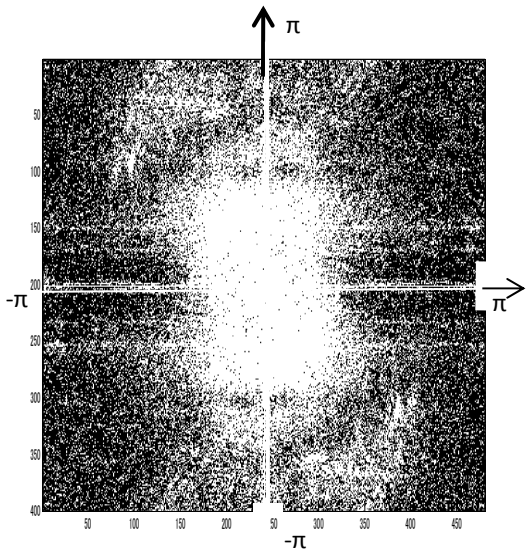


図14 図11の振幅特性

は目的・応用が異なること確認しておきたい[3][8]。

6. むすび

画像の解像度改善手法として、ハイパスフィルタと3乗特性を組み合わせた信号処理を提案した。提案した信号処理は微小振幅に対して有効に作用し、解像度を改善する。本提案と先に提案した絶対値を用いた信号処理を組み合わせることで、更に良好な結果が得られることも示した。提案手法はナイキスト周波数を越えた周波数成分を再生することを周波数領域で示した。

提案方式はTV受信機の高画質化や静止画の高精細化等幅広い分野に応用可能であると考えている。これまで困難であったナイキスト周波数を越えた周波数成分を再生するので、SDTVからHDTVへのアップコンバート、あるいはHDTVから4K、8Kシステムへのアップコンバートにも応用可能である。

今後の課題としてはハードウェア化があげられる。本方式はコンパクトなハードウェアで実現可能であるため、実用化も比較的容易であると考えている。

APPENDIX A

画像信号においてエッジは重要な情報である。画像が何らかの処理でボケを生じた場合、エッジ部分の高周波成分が失われたと考えることができる。高解像度化はエッジをいかに復元するかが重要な課題である。エッジは空間的に局在化した信号であり、エッジの連続性を利用した高解像度化が行われている[8]。

本論文では空間的に存在するエッジは、周波数領域でも局在化することを仮定している。即ち、エッジが周波数 ω_0 を持つならば、 $2\omega_0, 3\omega_0, \dots$ も保有するとする。即ち、エッジがもつ周波数の整数倍に信号が局在すると仮定する。この仮定に立てば、画像からエッジ成分を検出して、その高調波を加算することにより、高解像度化が可能と考えられる。

一次元信号を例として説明する。画像信号が

$$f(x) = \sum_{n=-N}^N a_n \cos n\omega x + b_n \sin n\omega x$$

$$= \sum_{n=-N}^{-1} a_n \cos n\omega x + a_0 + \sum_{n=1}^N a_n \cos n\omega x + \sum_{n=-N}^N b_n \sin n\omega x \dots (2)$$

にフーリエ級数展開されたものと仮定する。式(2)における最高周波数は $\sin(\pm N\omega x)$ および $\cos(\pm N\omega x)$ により表される。

ここで

$$g(x) = \sum_{n=-N}^{-1} a_n \cos n\omega x + \sum_{n=1}^N a_n \cos n\omega x + \sum_{n=-N}^N b_n \sin n\omega x \dots (3)$$

とすると、

$$f(x) = g(x) + a_0 \dots (4)$$

となる。

式(4)において a_0 は画像の直流成分、 $g(x)$ は交流成分となる。

本論文で提案した図8の信号処理をにおいて $f(x)$ を入力とすると $g(x)$ はハイパスフィルタ出力となり、図7の信号(b)に相当する。

図7ではこの信号を3乗している。これは $\{g(x)\}^3$ の処理を行うことと等価である。三角関数には以下の式が成立する。

$$(\sin \theta)^3 = \frac{3}{4} \sin \theta - \frac{1}{4} \sin 3\theta \dots (4)$$

$$(\cos \theta)^3 = \cos \theta + \frac{1}{2} \cos 3\theta \dots (5)$$

この式は3乗により3倍の周波数成分が現れることを示している。よって 原信号が保有しないナイキスト周波数を越える周波数成分を有することになり、以下の形に級数展開される。式(3)と式(6)の Σ に注目いただきたい。

$$(g(x))^3 = \sum_{n=-3N}^{-1} c_n \cos n\omega x + \sum_{n=1}^{3N} c_n \cos n\omega x$$

$$+ \sum_{n=-N}^N d_n \sin n\omega x \dots (6)$$

式(3)は-NからNの加算となっているが、式(6)では-3Nから3Nとなっており、原信号の3倍の周波数成分が発生したことを示している。この成分を原信号に加算することで精細化した画像信号を得ることが可能となる。

3乗を例にして説明したが、エッジが空間的かつ周波数領域でも局在すれば、他の非線形特性でも解像度を改善できる可能性があり、今後の研究が必要である。

参考文献

[1]Subhausis Chaudhuri et al. "Motion Free Super Resolution, isbn0-387-25890-6
 [2] 松本信幸, 井田孝, "フレーム内再構成型超解像の領域適応処理による高画質化の検討", SIP2008-6 IE2008-6, pp.31-36, 2008
 [3]Luminita A. Vese and Stanley J.Osher, "Modeling Textures with Total variation Minimization and Oscillating Patterns in Image Processing", Journal of Scientific Computing, Vol.19, Nos.1-3, December 2003
 [4]http://www.sony.co.jp/SonyInfo/News/Press_Archive/200408/04-042/
 「5」合志、寺川、三上、今井、"画像の超解像度化に関する検討 "、映メ学技報 IEEE-BT, Vol.33, No.25, pp9~13, BCT2009-57 (June.2009)
 [6] "美に挑む規格外テレビ", 日経エレクトロニクス, 2009, 2-9
 [7]S.C.Park, et al, "Super Resolution Image Reconstruction :Technical Overview", IEEE Signal Processing magazine, pp21-36, May, 2003
 [8] 寛幸、小川、長谷山、"エッジの連続性を考慮した IFS による画像の拡大", 信学論誌, Vol.J92-D, No.3, pp.382~392, 2009