

# 全方位時空間画像からの超解像度化

長原一 八木康史 谷内田正彦  
大阪大学大学院基礎工学研究科

あらまし： 全方位視覚センサHyperOmniVision は周囲360度が観測でき、かつ容易に透視投影画像に変換できる利点をもつ。しかしその反面、その構造上から一般のカメラに比べ、同一画角で比較した場合、解像度が低いという問題をもつ。本報告では、HyperOmniVisionを回転させて得られる時空間画像を用いることにより、超解像を実現する方法について報告する。また、光学シミュレーションからその最適化について論じる。

## Super-resolution from An Omnidirectional Image Sequence

Hajime NAGAHARA, Yasushi YAGI and Masahiko YACHIDA  
Graduate School of Engineering Science, OSAKA University

Abstract: The omnidirectional image sensor, we call HyperOmniVision, can observe a 360 degree field of view and can have its images transformed to perspective images easily. However, it has an intrinsic problem where the resolution is lower than an ordinary video camera in the same field of view, because of its mechanism. We propose the super-resolution method which fuses the image sequence within the turning motion of the sensor. And we discuss the optimization of the method by simulation result of its optics.

### 1. はじめに

全方位視覚センサは、水平360度全方位が撮像できる広角視覚センサであり、主にロボットの視覚誘導の目的として数々研究されてきた。ここで扱うHyperOmniVisionを含め、回転カメラを用いる方法<sup>[1]</sup>や魚眼レンズを用いる方法<sup>[2]</sup>、複数のカメラ画像を合成する方法<sup>[3]</sup>などが提案されている。一方近年、テレビ会議システムやテレグジスタンスなどヴァーチャルリアリティやロボティクスの世界において、人に提示される画像の視野角が臨場感や作業性に影響すると報告されている<sup>[4]</sup>。そこで、視野角の狭い一般的なカメラに代わりにこれら全方位視覚センサを人間への提示のための広視野角撮像センサとして適用することで臨場感や作業性が向上すると期待できる。先に挙げた全方位視覚センサにおいて、回転カメラを用いる方法は、全方位撮像にカメラの回転を必要とするため、実時間計測は不可能である。また魚眼レンズを用いる

方法は、その複雑な光学系により画像の歪み補正や画像変換の計算コストが大きく実時間提示が困難である。複数カメラ画像を合成する方法は、同様にカメラの歪み補正の計算コストやセンサシステムの複雑化、大型化が問題となる。それに対して、HyperOmniVision<sup>[5]</sup>は実時間で全方位が撮像でき、双曲面鏡を用いたその光学特性から人が直感的に理解しやすい通常のカメラ画像(平面射影変換)やパノラマ画像(円筒面射影変換)に容易に変換出来る利点を持つ。これら人による実時間監視目的への応用は当然ながら実時間撮像、変換及び提示が必須とされるため、HyperOmniVisionが最も適していると言える。しかしながら、HyperOmniVisionは一枚のCCDで全方位を撮像するというその撮像系の構造上、撮像素子の分解能が制限となり変換画像は一般のビデオカメラ画像に比べて画角分解能が低い。そのため、見た目の解像度がオペレータの印象や作業効率を大きく左右する対象に対

しては、その低解像度の問題が人に提示するそれら応用の障害となっている。解像度を向上させる直接的な方法として、ハイビジョンカメラなどの高解像度カメラを用いる方法も考えられるが、カメラや周辺機材のコストを考えると現状ではまだまだ困難である。また、これら機材のコストが下がったとしても、通常画角のカメラと同様の比較において同じ問題をはらむ。従来、カメラの動きにより画素ずれを仮定した画像列を用いて仮想的に CCD 分解能を越える情報、すなわち超解像度を画像処理により実現する手法が提案されてきた。Pelg と Keren<sup>[6]</sup> は初期高解像度画像を仮定し、それが撮像される光学過程をシミュレートして得られた画像と実際撮像された画像列の誤差を最小にするような高解像度画像を反復推定する方法を提案した。また、Irani と Pelg<sup>[7]</sup> は、トモグラフィに似たバックプロジェクション手法を提案した。しかし、これら手法は、通常のテレビカメラを対象としたものであり、全方位センサを対象とした超解像度化ではない。

そこで本報告では、全方位画像センサ HyperOmniVision の光学特性を考慮し、本センサに適した超解像化手法を提案する。

## 2. HyperOmniVision の撮像原理

HyperOmniVision は図1,2に示すように、鉛直下向きの双曲面鏡と上向きのカメラから構成される。双曲面鏡は式 1,2 のように表され 2 つの焦点を持ち、カメラは双曲面の中心軸と同軸に主点が双曲面鏡のその一方の焦点にくるよう配置されている。双曲面鏡に写った像をカメラで下方から撮像することにより、実時間でセンサの水平方向 360° を一度に撮像することができる(図3)。また、双曲面鏡のミラー焦点に向かって入ってくる光は、双曲面の特性によりもう一方の焦点に入射し、撮像される。画像上での任意の写像点  $p(x,y)$  からミラー焦点に対する俯角  $\alpha$ 、方位角  $\beta$  は式 3-5 に示すように一意に求まる。そのため、任意の射影変換を行うためには、この俯角  $\alpha$  と方位角  $\beta$  の直線と変換画像面の交

点を求めればよい(図3)。すなわち、ミラー焦点を視点とした一般のカメラ撮像画像(図4-b)やカメラを回転して得られるパノラマ画像(図4-c)に容易に変換できる。このような原理により、HyperOmniVision センサは実時間で全方位を撮像でき、さらに平面や円筒面といった任意の出力画像に容易に変換できるという特性を持つ。

$$\frac{X^2 - Y^2}{a^2} - \frac{Z^2}{b^2} = -1 \quad (Z > 0) \quad (1)$$

$$c = \sqrt{a^2 + b^2} \quad (2)$$

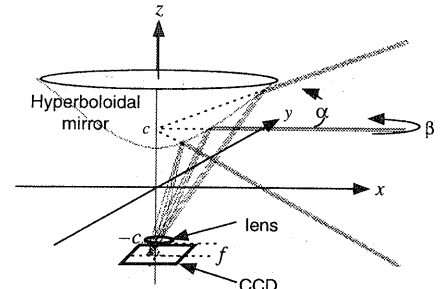


Fig.1: Structure of HyperOmniVision

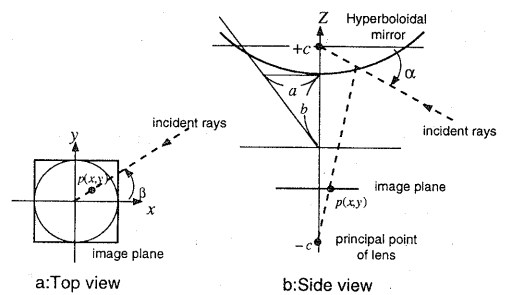


Fig.2: Optics of HyperOmniVision

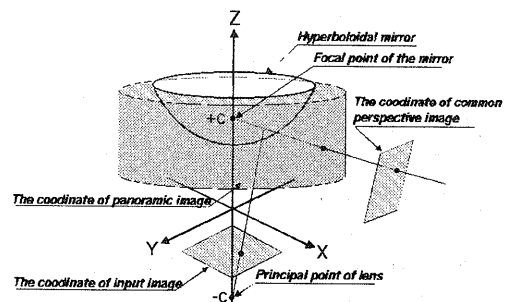


Fig.3: Geometrical relation between the input and transformed image coordinate

$$\beta = \tan^{-1}(y/x) \quad (3)$$

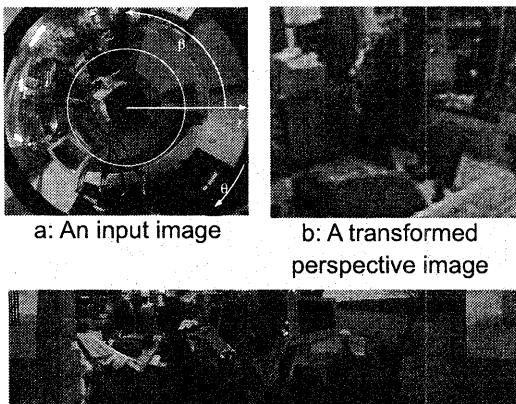
$$\alpha = \tan^{-1} \frac{(b^2 + c^2)\sin\gamma - 2bc}{(b^2 + c^2)\cos\gamma} \quad (4)$$

$$\gamma = \tan^{-1} \frac{f}{\sqrt{x^2 + y^2}} \quad (5)$$

### 3. HyperOmniVision に適した超解像度化

#### 3.1 超解像度化原理

CCDは離散的に配置されたフォトセンサの集合であり、その画像分解能はその数により決まる。そこで、カメラを能動的に運動させながら撮像した時系列画像により、フォトセンサの相対位置関係をずらすことにより疑似的にサンプリング点を増やすことで超解像度化を図る。図1に示す様にHyperOmniVisionは回転対象の光学系である。すなわち、センサのz軸回転運動においては、光学的な変化はCCDの回転のみであり視差や視点の移動は起こらない。そのため、カメラパラメータの変化はCCDの回転角 $\theta$ のみなので、時系列画像間の位置合わせが容易であるという利点を持つ。また、入力画像は矩形座標系のため、各画素は回転により回転中心からの距離が変化する。つまり、回転により画素のずれが出力画像の水平方向（入力座標系円周方向）だけでなく、垂直方向（入力座標系放射方



c: A transformed panoramic image  
Fig.4: An example of omnidirectional image and transformed images

向)においても起こるため、鉛直軸回りの動作により両方向の超解像度化が行える(図5)。

以上のHyperOmniVisionセンサに適した特徴から、超解像度化のための画素ずれを生むカメラの動きとしてこのz軸回転を採用した。図5にz軸回転による画素ずれ画像列とその変換画像、

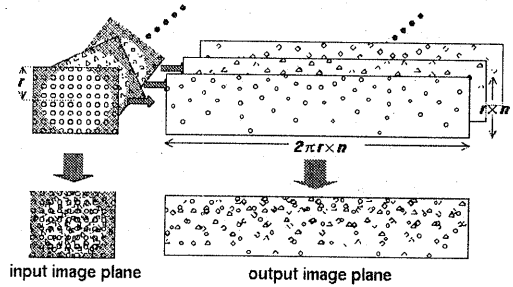


Fig.5: Relation of sampling point between input image sequence and fused output image

出力画像の関係を示す。以後出力画像サイズは図中の係数 $n$ によって示す。

#### 3.2. 光学特性と超解像度化

ここで述べる超解像度化手法は、CCDの画素限界をカメラを運動させることによりサンプリング点を疑似的に増やすことにより行われる。

最近傍内挿の様な、直接的超解像度化手法は、撮像過程による劣化の復元は考慮していない。しかし、カメラには集光限界があり、それを越えるといくら画素ずれを細かくしても解像度の向上は望めない。また、反復推定やバックプロジェクションにおいても光学特性の把握の正確さが、結果に大きく影響を与える。つまり超解像度化を考える際、センサシステム個々の光学特性について十分に考慮する必要がある。

そこで今回実験で用いたセンサを例にして各撮像過程の物理モデルを立て、シミュレーションから各特性を求め、超解像度限界と最適化について論じる。

表1に評価に用いたセンサ系のパラメータを、図6に撮像変換モデルを示す。途中のアナログ



Fig.6: The image formation process

系を含む部分は、無視している。HyperOmniVisionの双曲面鏡とカメラの集光レンズの特性をレイトレーシングによりシミュレートし、各焦点位置におけるMTFを求めた。MTFの例を図8に示す。CCDの受光面は実際には点ではなく、一定の感度を保つため有限の大きさを持つ。そのため、得られた値はこの窓内の平均輝度となるためローパスフィルタとして振る舞い、その帯域は制限される。図7にCCDの周波数通過特性を示す。図4-cに示したようなパノラマ画像変換を仮定した場合、出力画像の点は入力画像の円周方向において中心に近づくほど密になる。このことは、出力画像で同じ空

Table1: Mirror and Lens parameter

Hyperboloidal mirror	a	68.8mm
	b	88.4mm
CCD Camera	Lens	paraxial lens
	F	1.4
	CCD sell pitch	6.35 μm
	CCD sell size	3.175 μm

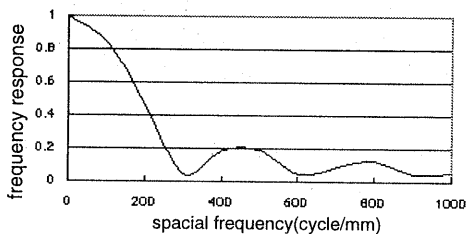


Fig.7: Frequency response of CCD

間周波数であったとしても、俯角により入力での周波数要求に違いがあることを示している。図9,10に出力における最大要求周波数の入力換算を周囲方向,放射方向についてそれぞれ示す。本手法における超解像度の再現性は、以上3つの特性の合成によって決められる。従来のHyperOmniVisionの設計は画素数の多い外側(俯角の小さい)や、または視線中心の水平線上でMTF特性が最大となるように設計されてきた。

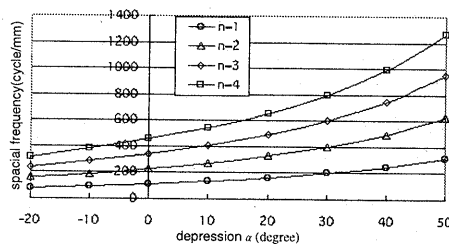


Fig.9: Frequency requirement for each output images about circumferential direction

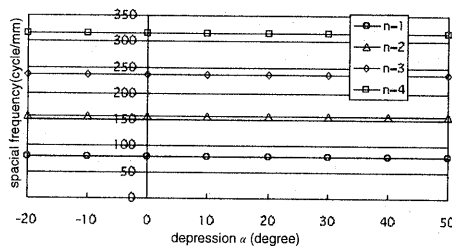


Fig.10: Frequency requirement for each output images about radial direction

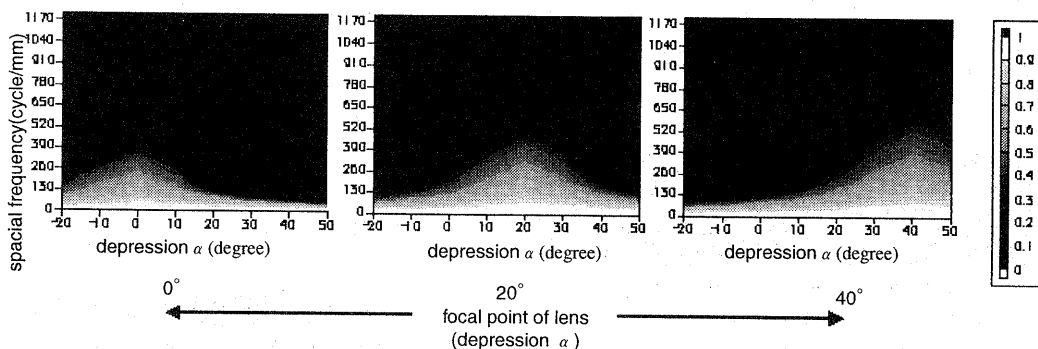


Fig.8: MTF of circumferential direction on CCD plane

超解像度化を考慮した場合、全体的なMTF特性を考慮に入れることはもちろんのこと、出力画像の解像度からの入力必要周波数をも考慮に入れなくてはならないことが分かる。図11,12にレンズの焦点位置を俯角 $0^\circ$ ,  $20^\circ$ ,  $40^\circ$  にとつた際の撮像系全体の空間周波数特性をx軸, y軸方向についてそれぞれ示す。これより入出力座標系の違いからx軸方位の周波数限界がy軸に比べて特に低いことがわかる。また、目的出力要求解像度により最適なレンズの焦点位置や全体の超解像度再現性に違いが見られる。これは、解像度が上がるにつれ俯角の大きい方から限界に近づくためである。直接的な情報量の向上は、x,y軸2倍あたりであることがわかる。また、図11,12で示す特性はあくまで画像全体の平均特性を表すものである。そのため、平面射影変換のように、画像の一部のみを最適化するにはMTFピークを目的視野に合わせるのがよい。

#### 4. アルゴリズム

画素ずれを仮定した時空間画像列の情報を統

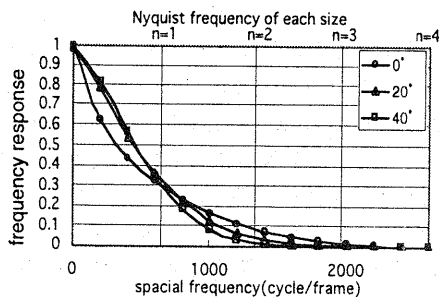


Fig.11: Frequency response of output (x direction)

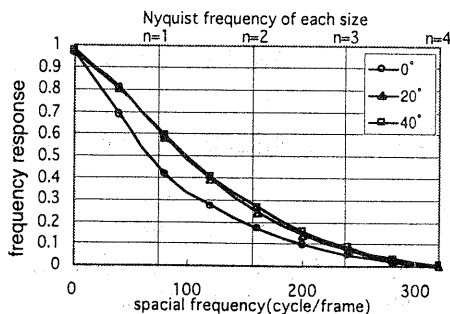


Fig.12: Frequency response of output (y direction)

合し超解像度化を行うためのアルゴリズムについて述べる。

#### 4.1. 最近傍内挿による超解像度手法

現在までに超解像度化手法として提案されている反復推定やバックプロジェクションによる間接的な方法は、精度は高いものの解が不安定であり、また計算コストが大きく実時間処理には向かない。これでは、テレビ会議システムやトレイグジスタンスなどの実時間監視を目的とする応用は不可能である。そこで、時空間画像列から直接的に最近傍内挿補間を行うことにより超解像度化を行う。

出力画像空間から入力を見た場合、図5の様に入力画素は回転による画素ずれにより相対位置が変化する。撮像対象が静止していると仮定した場合、出力画素に対して最近傍の入力画素の値がより正確な情報であると言える。そこで、時系列画像中において出力画素の最近傍の画素値を出力値として採用する。この方法は反復推定やバックプロジェクションのように反復演算が必要なく、また逐次的に処理できるため計算コストが軽く、実時間処理が可能であると考えられる。

#### 4.2. バックプロジェクション法

バックプロジェクションによる超解像度化<sup>(7)</sup>は、間接的に再帰的に計算することにより行われる。最近傍内挿等によって作られた初期画像から、光学過程をシミュレートして得られた推定入力画像列を生成する。その推定入力画像列と実際に撮像された画像列との差をとり、その誤差をバックプロジェクションし、目的画像を更新することにより、超解像度化を行う。式6にこの過程を式で示す。この方法は、超解像度化と同時にぼけ復元を行えるという利点を持つ。

$$f^{n+1}(x) = f^n(x) +$$

$$\sum_{y \in \cup_k Y_{kx}} (g_k(y) - g^n(y)) \frac{(h_{xy}^{BP})^2}{c \sum_{y \in \cup_k Y_{kx}} h_{xy}^{BP}} \quad (6)$$

ここで、

$f^n(x)$ : 超解像度推定画像

- $g_k(y)$ : 撮像入力画像列
- $g_k^d(y)$ : 推定入力画像列
- $x$ : 超解像度画像の画素
- $y$ : 入力画像の画素
- $Y_{k,x}$ :  $x$  に影響を与える  $y$
- $h_{xy}^{BP}$ : バックプロジェクションカーネル
- $c$ : 正規化定数

## 5. 実験

### 5.1. シミュレーション

HyperOmniVision の光学過程をシミュレートして作成した入力画像 (図 13) を用いて超解像度化を視覚的, 数値的に評価した. 円筒面射影を対象変換画像とし, 時空間画像を用いた最近傍内挿とバックプロジェクションの2つの方法による結果と比較のために単一画像を用いた通常の最近傍内挿による出力結果を図13の点線部について図14-16に示す. また, 超解像度処理に用いた画像枚数に対する復元精度を理想画像とのPSNR(Peak Signal to Noise Ratio)を用いて図17-aに示す. その際の使用枚数と計算コストの関係を図17-bに示した. PSNR が大きいほど理想画像との差が小さいことになり, 両手法とも超解像度化の有効性が表れていることが分かる. 超解像効果の点からはバックプロジェクションの方が勝っている. また, 撮像枚数よりも演算の繰り返し回数の方が解像度向上に貢献していることが分かる. しかしながら, 最近傍内挿に

よる手法は, ぼけ過程の復元を考慮せず, また計算コストがケタ違いに軽いわりには良い結果と言える. また, 撮像と計算を同時に逐次的に処理することができるため, 実時間処理の可能性を持つ.

すなわち, 処理時間においては最近傍内挿, 精度ではバックプロジェクションによる超解像度化とそれぞれ長所短所をもち, 対象に応じて適応するのが得策と言える.

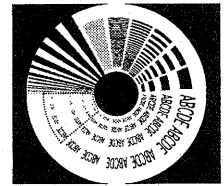


Fig.13: Simulated input image

### 5.2. 実画像実験

実際に超解像度撮像システムを作成し実画像について実験を行った. 本システムは表1に示したパラメータの双曲面鏡とカメラ(EVI-310 Sony), 回転テーブル (DR-160 シグマ光機) により構成される.

図 18-20 に実画像を用いて各サイズにおける結果を示す. この結果からもシミュレーション同様, 実画像においても超解像度化の効果が実験的に確認できた. 特に俯角の大きい情報量の少ない部位について著しい向上が見られる. このことより HyperOmniVision の従来の問題点である俯角に対する解像度差の問題が改善されている. 図 19,20 を比較すると, バックプロジェクションの結果の方がコントラストが高い. しかし, 最近傍内挿による超解像度化の結果の方が俯角の大きい領域において, ぼけにより滑らか

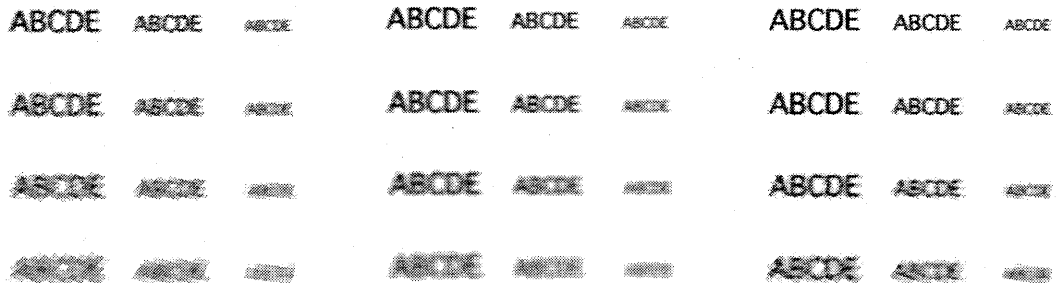


Fig.14: Nearest neighbor interpolation from one image (1 input image,  $n=2$ , PSNR 19.72)

Fig.15: Super-resolution by nearest neighbor interpolation (25 input images,  $n=2$ , PSNR 21.62)

Fig.16: Super-resolution by back projection (25 input images,  $n=2$ , 50 iterations, PSNR 22.62)

な画像となっており、現実のシーンにおいては視覚的に自然に見える。

## 6. まとめ

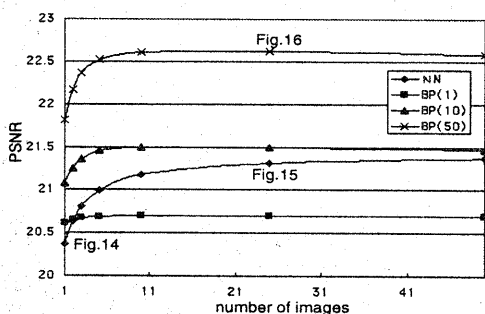
本研究において、全方位視覚センサ HyperOmniVision の超解像度化手法を提案した。カメラの動きによる画素ずれを仮定した時系列画像の情報を統合することにより実現される。カメラの動きとしてz軸同軸回転を用いることにより、画像のマッチングの容易さと垂直解像度の違いという従来の問題を改善するという意味で、HyperOmniVision の光学特性に適応しているといえる。また、物理モデルをたて検証することにより、本センサにおける超解像度化の限界と最適化について述べた。シミュレーション画像、実画像の両方の結果により超解像度化を確認し有効性を示した。従来主流とされてきたバックプロジェクションによる超解像度化の実験結果はよりよい性能を示しているが、計算コストから実時間実装は不可能である。そのため、実時間提示が必要でない応用に対しては、当然ながらバックプロジェクション、実時間提示が必要な応用に対しては本文で提案した、最近傍内挿による超解像度化手法と使い分けるのが妥当であると考えられる。

一般的に時系列画像を用いた画像処理は撮像する環境中の対象物が静止していることを仮定

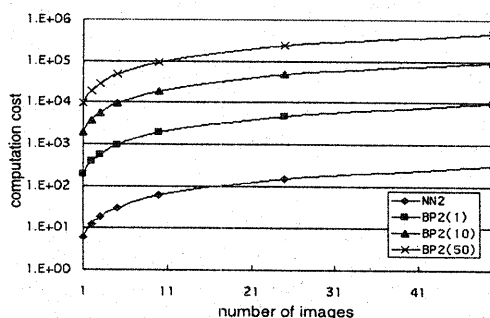
している。本手法も例外ではなく、動領域に対しては適用できることはおろか、逆に画像のモーションブラーを生む。しかし、実際の撮像システムを構築する際には、動領域を背景差分等により検出し、その領域をマスクするなどすれば、静止環境の超解像度化並びに HyperOmniVision の特徴の一つである全方位実時間計測を同時に可能にする。

## 参考文献

- [1] J.Y.Zheng and S.Tuji: "Parametric representation of scenes for route understanding", Proc. of the 10th Int. Conf. on Pattern Recognition, pp.161-167 (1990).
- [2] Z. L. Cao, S. J. Oh and E. L. Hall: "Dynamic omnidirectional vision for mobile robots", J. Robotic Systems, 3, 1, pp.5-17(1986).
- [3] T. Kawanish, et al.: "Generation of high-resolution stereo panoramic images omnidirectional imaging sensor using hexagonal pyramidal mirrors", ICPR, pp.485-489,1998.
- [4] Darwin G, et al.: "Sensory Requirements and Performance Assessment of Tele-Presence Controlled Robots", Proc. 1996 IEEE Int. Conf. on Robotics and Automation, pp.1375-1380, 1996.
- [5] K. Yamazawa, Y. Yagi and M. Yachida: "New real-time omnidirectional image sensor with hyper-



a: PSNR against the number of images used for superresolution



b: Computation cost of each method against the number of images used for superresolution

Fig.17: Comparison of each method  
 NN: Super-resolution by nearest neighbor interpolation  
 BP(i): Back projection  
 (i: number of iterations, n=2)

boloidal mirror” , Proc. 8th Scandinavian Conf. on Image Analysis, Vol.2, 1381-1387,1995.

[6] D. Keren, S. Peig, and R. Brada: “Image sequence enhancement using sub-pixel displacements” , CVPR, pp.742-746, Jun, 1998.

[7] M.Irani and S. Peig: “Improving resolution by image registration” , CVGIP:Graphical models and image precessing, Vol.53, pp.231-239, May, 1991.



Fig.18: Nearest neighbor interpolation from one image  
(1 input image,  $n=2$ )



Fig.19: Super-resolution by nearest neighbor interpolation  
(25 input images,  $n=2$ )



Fig.20: Super-resolution by back projection  
(25 input images,  $n=2$ , 50 iterations)