

円形カメラ配置光線空間情報を利用した圧縮符号化

木村 真治[†] 圓道 知博^{††} 藤井 俊彰^{††} 谷本 正幸^{††}

[†]名古屋大学大学院 工学研究科 〒464-8603 愛知県名古屋市千種区不老町

E-mail: †kimura@tanimoto.nuee.nagoya-u.ac.jp, ††{yendo,fujii,tanimoto}@nuee.nagoya-u.ac.jp

あらまし 我々の研究室では Image Based Rendering の一種である光線空間法を用いて任意視点画像生成に関する研究を行っている。任意視点画像生成には非常に密に配置されたカメラによって撮影された画像を用いるため、非常に多くの原画像データが必要となる。そこで、情報圧縮が必要となる。本稿では、カメラが円形に配置された場合の情報量削減手法として、光線空間上では点の軌跡がサイン波を描くという性質に着目した新たな圧縮符号化方式を提案する。この圧縮符号化を用いることにより、約 100 分の 1 の圧縮率で平均 31.5dB の PSNR を得ることができたので、これを報告する。

キーワード 光線空間法, 多視点画像圧縮, 自由視点画像合成, 水平断面画像, 円形カメラ配置

Image Compression using Ray-Space Information in a Circular Camera Arrangement

Shinji KIMURAI[†], Tomohiro YENDO^{††}, Toshiaki FUJII^{††}, and Masayuki TANIMOTO^{††}

[†] Graduate school of Engineering, Nagoya University Furo-cho, Chikusa-ku, Nagoya, 464-8603 Japan

E-mail: †kimura@tanimoto.nuee.nagoya-u.ac.jp, ††{yendo,fujii,tanimoto}@nuee.nagoya-u.ac.jp

Abstract We propose to generate arbitrary viewpoint images using Ray-Space method, one of the Image Based Rendering methods. In order to generate a new viewpoint image, we need large amount of original images captured by densely-arranged cameras. Therefore, we need compression of the data. In this paper, we propose a new coding scheme taking advantage of property that each point in real world has a sinusoidal trajectory in ray space in condition that one object is captured by cameras arranged in a circle. As a result, we indicate the relation between PSNR and bit rate.

Key words Ray-Space method, Multi-view image coding, Free viewpoint image synthesis, Horizontal section image, Circular camera arrangement

1. はじめに

計算機やネットワークの能力向上に伴い、エンターテインメントやアーカイブ、教育など多くの分野で 3 次元空間情報を扱う機会が増えている。従来では 3 次元空間情報をモデルベースで考えていたのに対し、現在は実際に撮影された画像を元により自然な形で 3 次元空間を処理することが可能になってきた。実画像を元にした処理技術は、モデルベースでの処理に比べ、我々にとってより自然に 3 次元物体を表現できるため、歴史遺産のデジタル保存などの分野で用いられることが多い。このように 2 次元の実画像を用いて 3 次元空間のシーンを再構成する手法は Image Based Rendering (IBR) と呼ばれ、多くの研究が活発に行われている [1], [2]。

我々の研究室では IBR の一種である光線空間法 [3] を用いて

自由視点テレビ (Free Viewpoint Television: FTV) [4] を提案している。自由視点テレビは、多数のカメラによって撮影された画像を元に視聴者が見たい視点の画像を作り出すことのできるテレビである。このテレビは現在のような放送局からの一方通行の映像配信でなく、視聴者も視点動かせるということで、全く新しい形での双方向性を実現するテレビである。自由視点テレビを実現するために必要となる技術としては、撮影画像を補正する技術 [5] や、補間で自由視点画像を作成する技術 [6], [7]、効果的に情報を圧縮する技術などが必要となる。

撮影対象の形状モデルデータを必要としない光線空間法は 3 次元シーンを再構成するのに非常に有用であるが、再構成のためにはシーン全体の光線をすべてサンプリングしておかなければならないため、モデルベースの手法に比べ扱うデータ量が膨大になる問題がある。つまり、任意視点画像を作るためには非

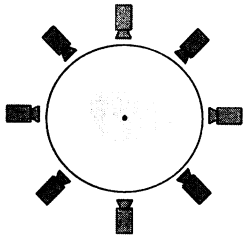


図1 Camera Arrangement

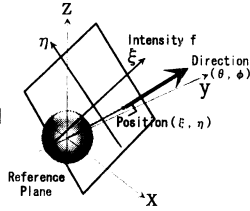


図2 Ray Parameter

常に密に配置されたカメラによって撮影された画像群を必要とする。そのため、情報量の削減が求められる。このように多数のカメラから撮影されたデータを圧縮する方法としては、空間方向と時間方向の冗長性を利用した手法 [8], [9] や、自由視点画像合成に適し、JPEG と互換性を持つ手法 [10] などが提案されている。

本稿ではカメラを円形に多数配置した状態で、対象物の撮影を行い、光線空間を構築し、その光線空間情報の圧縮を行った。実空間における撮影対象の各点の軌跡が光線空間上ではサイン波の構造を示すという性質に着目することで、円形カメラ配置における多眼画像の新たな情報圧縮法を提案し、実験、検証を行う。

本稿の構成は以下の通りである。2章で円形にカメラを配置した場合の光線空間について説明を行う。3章で提案法について説明し、4章で提案法を用いた実験を行い、結果を示す。最後に5章でむすびとする。

2. 光線空間法 -円形カメラ配置-

光線空間法とは、3次元空間の情報を空間中を伝搬する光線の情報を用いて記述、処理する方法である。図1のような円形カメラ配置、または球面カメラ配置において、光線のパラメータのとり方は図2の通りである。光線が通過し、また、その光線の向きに垂直な面を基準面に取り、3次元実空間上での光線の向き (θ, ϕ) と基準面上での光線が通過する位置 (ξ, η) とで光強度を表現する。つまり、光強度 f を $f(\xi, \eta, \theta, \phi)$ のように表現することができる。本稿では多数のカメラを球面ではなく円形に配置しているので縦方向の視差を無視して考え、光強度 f を $f(\xi, \theta)$ と表現する。

次に、光線空間の構築について説明する。まず、円形に多数配置されたカメラから対象物体を撮影する。次に、撮影された画

像を図3の左図のように1,2,3,4,5...と並べる。この作業によってできた直方体が光線空間である。この直方体の光線空間をある一定の高さで切ると図3の右図のような一枚の画像が得られる。本稿では以下、このようにして得られた画像を「水平断面画像」と呼ぶ。

3. 提案手法

3.1 概要

光線空間情報を効率的に圧縮するために、光線空間の性質を利用した圧縮手法を提案する。まず、水平断面画像に着目してほしい。図3の水平断面画像から見てもわかるように、カメラを円形に配置した場合の光線空間においては「実空間中の1点の軌跡が水平断面画像上ではサイン波の構造を示す」という非常に重要な性質がある。このため、水平断面画像は撮影対象表面の多数の点の軌跡から成るサイン波の集まりで成り立っていると考えることができる。分かりやすい例を図4に示した。これは光線空間をある高さ（線が引かれたところ）で切ることで得られた水平断面画像を示しており、原画像群の帽子部分の点の軌跡を表している。

次に、各々のサイン波上で考える。各サイン波は実空間中の1点の軌跡であるので、そのサイン波上には撮影対象物上の同じ点を違った角度からカメラで捉えた画素値が並んでいる。同じ点を撮影した画素値であるので、照明による影響などをすべて無視した場合、各サイン波上では同じ画素値が並ぶはずである。しかし実際には、光の反射特性やカメラノイズ、照明などの影響を受けるため、画素値は変化する。しかしながら、鏡などの反射特性が大きい物体をのぞけば、各サイン波上では画素値の変化は小さいと考えることができる。つまり各サイン波上での画素値列の冗長性は高く、空間周波数は低周波領域に集まると考えることができる。そこで、JPEGなどで一般的に行われているブロック単位での離散コサイン変換 (Discrete Cosine Transform: DCT) ではなく、同じサイン波上に位置する画素値列に対する DCT を行い、DCT 係数の量子化を行うことで、光線空間情報の情報量を削減できると考えた。

符号側ではサイン波の振幅情報と、そのサイン波に対する DCT 係数を求め、それを復号側に渡す。復号側では符号側から得られた情報を元に、一本一本のサイン波を復元することができる。水平断面画像を復元することができる。光線空間のすべての高さにおいて水平断面画像を復元すれば、原画像も復元することができる。

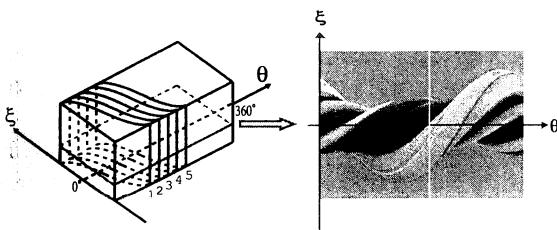


図3 Ray Space

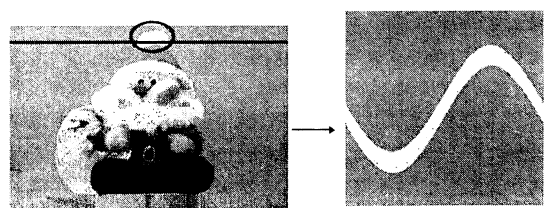


図4 An Example of Ray Trajectory

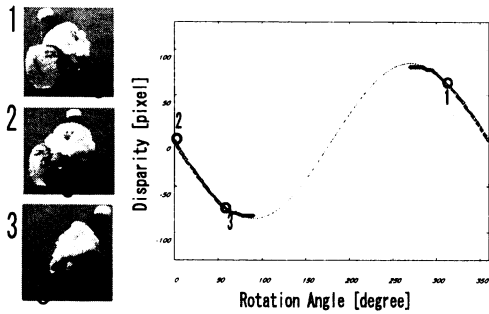


図 5 Sin Fitting

3.2 振幅抽出

提案法実装のために、まずは各サイン波の振幅を求める。サイン波の振幅を求めるということは、カメラが配置されている円の中心から撮影対象物上の点までの距離を求めることに相当する。各ピクセル(実空間上では各点)の振幅を求めるために、まずはある1ピクセル(原画像上の横方向中心点)を探索開始ピクセルとして指定し、そのピクセルの動き情報を5×5のブロックマッチングを用い、±90°の範囲で求める。ピクセルの動きは高さ方向には変化しないとし、水平方向の動きベクトルのみを求める。求められた動きの軌跡は3.1の概要で述べた光線空間の性質より、理想的にはサイン波を示すはずである。しかしながら、ブロックマッチングの誤対応などにより得られた軌跡は理想的なサイン波を描かない。そこで、最小二乗法を用いて、実際に得られた軌跡情報に最も適したサイン波の振幅を求める。その様子を図5に示す。右図は左図2番の画像の○で囲んだピクセルを探索開始点として動きの軌跡を求めた結果と、その軌跡に最適なサイン波をフィッティングした結果である。このようにブロックマッチングの誤差が少なければ実際の軌跡においてもほぼサイン波になる軌跡が得られる。

また、撮影対象物は本質的には無数の点で構成されており、画像のサンプリング間隔を考えない場合は水平断面画像も無数のサイン波によって構成されているはずである。しかしながら、撮影画像枚数を N (以下、 N は撮影画像枚数を表す) とすると、本稿では水平断面画像1枚は N 本のサイン波によって復元できると仮定した。つまり、探索開始ピクセルは N 点であり、光線空間全体(光線空間上のすべての高さでの水平断面画像)を復元するには(撮影画像の高さ)× N 本のサイン波の振幅と、そのサイン波上の画素値が必要になる。また、各々のサイン波の位相間隔を n (以下、 n は撮影画像の位相間隔を表す) とすると、 $n = \frac{360^\circ}{N}$ である。

サイン波一本一本に対して上記の作業を行い、振幅を求めるが、ブロックマッチングの誤対応などにより、実際の振幅とは大幅に異なる振幅が得られる場合がある。提案手法では撮影対象物の正確なモデル情報は必要とせず、たとえ違っていてもDCT係数の高次の項を送ることで正確な画像復元は可能である。しかしながら、サイン波の振幅が実際の振幅と大きくずれている場合は、サイン波上の画素値列の相関が小さくなってしまいうので圧縮効率が下がってしまう。そこで、振幅の誤差を小さくする

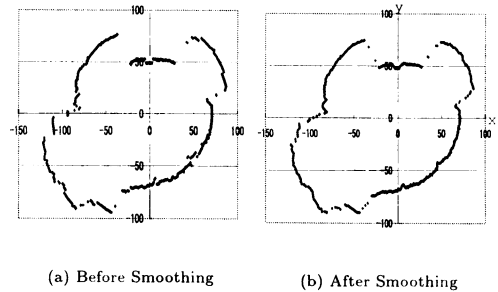


図 6 Amplitude Smoothing

ために振幅のスムージングを行った。本来、物質というものは基本的には表面がなだらかであり、急激な振幅の変化はないと考えられる。そこで、同じ水平断面画像上の両隣の位相の振幅と、カットする高さを上下1ピクセル変化させて得られた水平断面画像の同じ位相での振幅を用いることで、スムージングを行った。

h は光線空間から水平断面画像を得るときにカットした高さを表し、 p は位相を表す。また、 $A(h, p)$ は高さ h 、位相 p の時の振幅を表す。スムージングされた振幅を $A_{smoothed}$ とし、スムージングの式を以下に示す。

$$A_{smoothed}(h, p) = \frac{1}{5} \{A(h, p) + A(h-1, p) + A(h+1, p) + A(h, p+n) + A(h, p-n)\} \quad (1)$$

図6(a)に実際の振幅-位相関係を曲座標系で示す。図は振幅を A 、位相を θ としたときの、 $x = A \cos \theta$ を横軸に $y = A \sin \theta$ を縦軸にとったものである。図6(a)の振幅をスムージングした結果を図6(b)に示す。

3.3 サイン波データ列

水平断面画像1枚に対しサイン波を N 本描くとするのだが、 N が多ければ多いほど同じ画素を通るサイン波は増えてしまう。つまり、1つの画素に対し、対応するサイン波が何本もあるのである。本来、1つの画素には1つのサイン波が対応するも

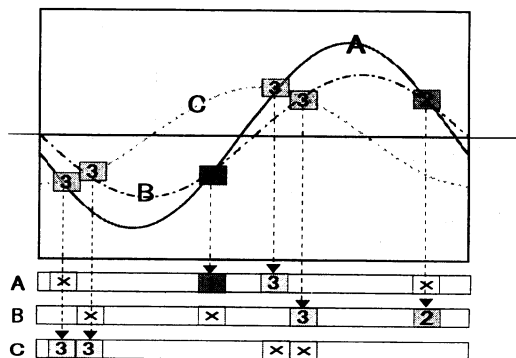


図 7 Correspondence of Pixel to Each 1-D Data Array of Sinusoidal Curve

のであり、このままではサイン波一本一本に対し、画素値を割り振ることができない。そこで、以下の法則に従い、1つの画素に対し1つのサイン波を対応させる。

- (1) カメラの撮像面に対し平行でカメラが配置された円の中心を通る面よりカメラ側に対象物がある場合は振幅が大きいサイン波を優先
- (2) 1の逆の場合は振幅が小さいほうのサイン波を優先
- (3) 2の場合と1の場合のサイン波両方が通る場合は振幅の大きさに関わらず、1の場合を優先

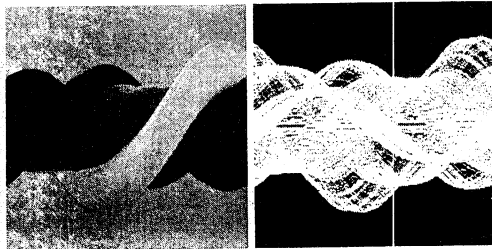
これらの法則により、各画素において対応するサイン波が一意に求まる。この様子を図7に示す。図中の数字は上記の法則の中で、どの法則によって判定されるかを表しており、これによりA,B,Cのサイン波が交わる場合、各画素がどのサイン波に対応するかが求まる。どのサイン波に対応するか、という情報を持つ水平断面画像と同じ大きさの2次元の配列を本稿では「インデックス配列」と呼ぶ。インデックス配列はN個の振幅を与えれば上記の法則より一意に求まる。

また、図7中の「X」のように画素が他のサイン波に対応し、画素値が求まらない画素は、同じサイン波の画素値列の周りの画素から線形補間を行い、画素値の空白をなくす。この作業をすべてのサイン波について行くと、一枚の水平断面画像に対し、N個の1次元画素値列を得ることができる。また、各画素値配列のサイズは $1 \times N$ となる。

3.4 穴画素問題

図8(a)にある高さの水平断面画像を示す。図8(b)は図8(a)の各画素に対し、N本のサイン波どれかに対応した画素のみを白くした画像である。本手法では一枚の水平断面画像がN本のサイン波によって成り立つと仮定しているが、本質的には無限本のサイン波によって撮影対象物は構成されているため、また、求められた振幅の実際の値との誤差のために図8(b)に示すように、背景部分以外の本来画素値を持つべきところでもサイン波が対応しない画素がある。本稿ではこのような画素を「穴画素」と呼ぶ(図9中の「H」は穴画素を示す)

また、穴画素以外の問題として、本来Dのサイン波上に位置するべき図9中の「2」と書かれた画素がA,B,Cのサイン波だけではCのサイン波に対応してしまうという問題がある。誤対応した画素2は本来AとBの間にあるDのサイン波に対応す



(a) Original

(b) Described by Sin

図8 Example of Hole Pixels

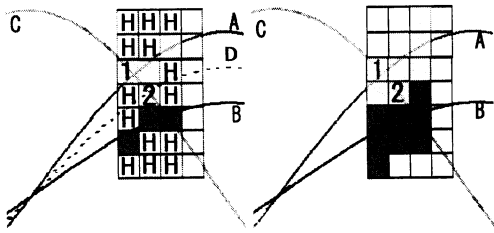


図9 Hole Problem

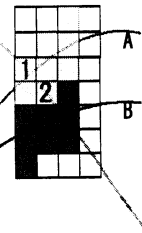
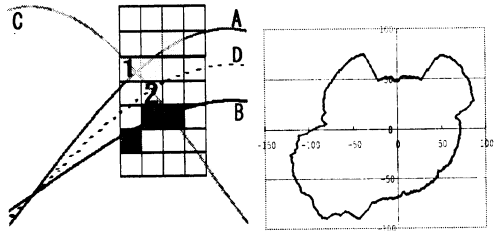


図10 Technique 1



(a) Correspondence of Pixel to Sinusoidal Curve

(b) Interpolated Sin Curves

図11 Technique 2

るものであるため、AとBの画素値列とは相関が高い画素値を持つと考えられるが、Cのサイン波の画素値列の中では他の画素値との相関が低いと考えられる。DCT係数から画素値への復元の際に高次の項を持っていけば適切な値が復元可能だが、圧縮率をあげるために高次の項をカットすると、多くの場合ノイズとなって画像品質はさがってしまう。そこで、穴画素をなくし、画素とサイン波画素値列との誤対応を減らすための手法を2つ提案する。

手法1は、図10に示したように、画素の穴を減らすためにサイン波それぞれに幅を持たせる手法である。(図10ではサイン波が通るピクセル+上下1ピクセルの幅を持たせて各サイン波に対応させている) 穴画素は少なくなり、画素2がCではなく、Aのサイン波に含まれるようになるので、復元時のノイズを少なくすることができる。手法1ではインデックス配列にはN本のサイン波のうち、どのサイン波に対応するか、の情報は無い。

手法2は、図11(a)のように、その画素2を本来通るはずであろうサイン波Dに対応させるため、サイン波の補間を行う手法である。振幅の補間は図6(b)のような振幅-位相関係のグラフにおいて、図11(b)のように離散な点を位相間で結ぶことによって位相間の振幅を補間する。この補間を行うことにより、画素のサイン波への誤対応を防ぐことができるため、より冗長性の高い1次元画素値配列が得られる。サイン波の補間は位相を n° 刻み $\rightarrow \frac{n}{2}^\circ$ 刻み $\rightarrow \frac{n}{4}^\circ$ 刻み $\rightarrow \frac{n}{8}^\circ$ 刻み $\rightarrow \dots$ と同様に行い、その都度、各画素を通過するサイン波を3.3の法則を用いて決定する。この繰り返しによって穴画素をなくすことができ、各画素に対応するサイン波が一意に決定する。手法2では繰り返し回

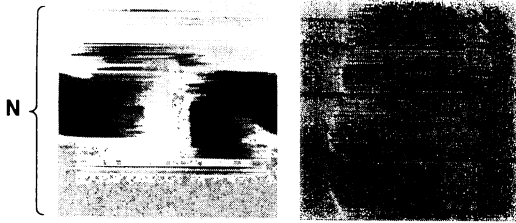


図 12 Examples of Color Variation

数を増やすことで、穴画素を完全になくすことができる。手法 2 のインデックス配列には補間したサイン波に対応するという情報も入る。

3.5 データ圧縮

上記の作業により得られた、 N 本の各サイン波の画素値配列 ($1 \times N$) を縦に並べたものの例を図 12 に 2 つ示す。縦軸はサイン波の移り変わりを表し、横軸は各々のサイン波の画素値列を表している。こうして得られた N 本のサイン波の画素値列に対し、それぞれ 1 次元の離散コサイン変換 (DCT) を行う。次に、各配列で得られた DCT 係数を量子化することにより DCT 係数を減らし、冗長データの削減を行う。このため、サイン波を復元するために必要なデータは振幅、量子化された DCT 係数の直流分・交流分となる。

本稿では自然画像を扱っているため、 N 本のサイン波の振幅はなだらかに変化すると考えることができるので、振幅の値は隣り合う値との相関性が高いと考えることができる。そこで、隣の振幅から予測符号化を行い、その残差をハフマン符号化してデータの圧縮を行う。また、同様にして N 本のサイン波の DCT 係数直流成分は相関性が高いので予測符号化+ハフマン符号化をしてデータ圧縮を行う。DCT 係数の交流成分は相関性が低いため、ハフマン符号化のみを行いデータ圧縮を図る。

また、隣り合う位相のサイン波は振幅の大きさの相関性が高く、それに伴い、隣り合うサイン波の画素値列も相関が高いと考えられる。図 12 の例からもその様子が分かる。よって、圧縮率をより高めるため、各サイン波毎に DCT を行うのではなく、サイン波の画素値列を何本かをまとめて 2 次元の画素値列とし、その画素値列に対し 2 次元の DCT をかける手法も用いる。

3.6 画像復元

水平断面画像復元は 3.4 で求めたインデックス配列を元に行う。まず、 N 本のサイン波に対し、符号側から送られてきた量子化済みの DCT 係数を逆 DCT することにより画素値が求まる。これは図 12 の画像に対し DCT をかけ、復元したものに相当する。この画像からインデックス配列を元にして、水平断面画像上の対応する画素に画素値をコピーする。

次に N 本のサイン波では復元されない部分 (穴画素) を、 N 本のサイン波によって復元された画素値を元にして補間する。補間は水平断面画像上で同じ列にあり、 N 本のサイン波の復元によって求められた画素を元にして行う。補間を行う穴画素に最も近い二つの復元された画素値を線形補間することで穴画素の

画素値を決定する。この作業を行うことで、穴画素のない水平断面画像を得ることができ、それから撮影画像を復元することができる。

4. 実験

4.1 実験条件

- 原画像サイズ：横 360×縦 240 ピクセル (カラー画像)
- 原画像枚数：360 枚 ($n = 1^\circ$)
- 撮影方法：対象物を回転ステージの上に置き、回転させながら 1 つのカメラで撮影
- 撮影対象：人形
- 背景：一様

手法 1 ではサイン波が通るピクセルに対し、上下に 1 ピクセルずつの幅をもたせ、実験を行った。また、手法 1 は 1 次元の DCT をかけ、手法 2 は 1,2 次元の DCT をかけ、圧縮実験を行った。また、2 次元 DCT を用いた実験ではサイン波を 8 本ずつまとめたものに対し情報量削減を行った。つまり、一枚の水平断面画像 ($N \times N$) を圧縮するのに、 $\frac{N}{8} = 45$ 個の二次元配列 ($N \times 8$) に 2 次元 DCT をかけて情報量の削減を行った。

4.2 実験結果

図 13 に実験の結果をビットレート (bpp) と PSNR (dB) の指標で表した。PSNR は生成画像 360 枚と原画像 360 枚の輝度値を比較した値の平均をとっている。また、PSNR の導出は背景部分を除き、対象物範囲 (復元画像範囲) のみの比較で行っている。

また、図 14 に今回の実験で使用した画像を示す。図 15 に手法 2 の 1 次元 DCT を用いて復元した画像を示す。図 15 はビットレート 0.464bpp (圧縮率 0.0194) で、平均 PSNR 32.6dB の画像を表す。図 15 の黒くなっている部分はサイン波によっては復元されなかった部分で、一様な背景の部分とみなす。

4.3 考察

1 次元 DCT において、手法 2 は手法 1 の手法にくらべ PSNR の向上をはかることができた。これは位相間の振幅を補間することにより、手法 2 の各サイン波の画素値列の誤対応が減ったため、画素値列の冗長性が高まり、同圧縮率における PSNR 向上につながったと考えられる。

また、手法 2 において、各サイン波の画素値列に 1 次元の DCT をかけるのではなく、8 本のサイン波をまとめて 2 次元の DCT をかけることで同程度の圧縮率に対し、約 1dB の PSNR の向上

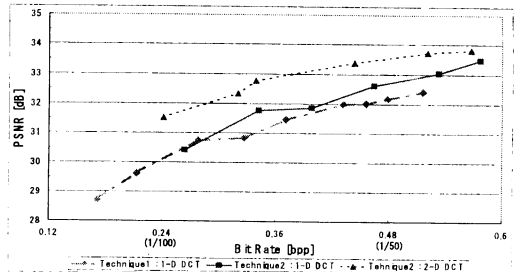


図 13 Compression Results



図 14 Original Images

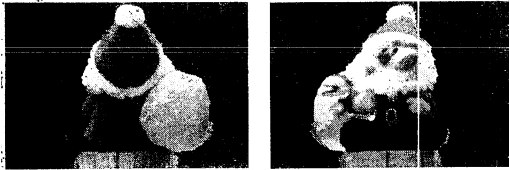


図 15 Decompressed Images (0.46bpp 32.6dB)

を図ることができた。これにより、図 12 にも示すように、サイン波の画素値列が隣り合う位相と相関性があることが明らかになった。

5. む す び

本稿では、光線空間上での点の軌跡の性質に着目した新たな円形多視点カメラ情報圧縮法を提案した。実験結果より、本手法が円形にカメラを配置した場合の多視点画像の情報圧縮に非常に有用であることが示された。

今後は、本稿では水平断面画像 1 枚 1 枚に対しサイン波上の画素値列に 1 次元、ないし 2 次元の DCT を行っていたものを拡張し、水平断面画像間での相関性も使った 3 次元画素値列に対し DCT 行うことで、さらなる情報量の圧縮を目指す。対象物は縦方向に急激に変化することは少ないと考えることができ、つまり水平断面画像間での相関性も高いと考えることができるので 3 次元の DCT を用いることで圧縮率の向上が可能であると考えられる。また、各サイン波の振幅情報をより正確に得ることで圧縮率をさらに向上させることができるのは明らかであるので、より実際の値に近いサイン波の振幅の導出を目指す。

文 献

- [1] Marc Levoy and Pat Hanrahan, "Light Field Rendering", ACM SIGGRAPH96, pp. 31-42, 1996
- [2] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski and Michael F. Cohen, "The lumigraph", ACM SIGGRAPH96, pp. 43-54, 1996
- [3] 藤井, 金子, 原島: "光線群による 3 次元空間情報の表現とその応用", テレビジョン学会誌, Vol.50, No.9, pp.1312-1318, 1996.
- [4] Toshiaki Fujii, Masayuki Tanimoto, "Free-Viewpoint TV System Based on Ray-Space Representation", SPIE ITcom, Vol. 4864-22, pp. 175- 189, Aug. 2002.
- [5] Kenji Yamamoto, Toshiaki Fujii and Masayuki Tanimoto, "Projective Transformation to Align Multi-Camera Images for Efficient FTV Compression", IWAIT2005, pp. 1-6, 2005
- [6] 山中, 藤井, 谷本, "円形カメラ配置における間引き・補間を用いた光線空間情報圧縮", PCSJ2004, pp.93-94, 2004
- [7] ヤーセル, ドローゼ, 藤井, 谷本, "視差平均化と後置適応フィル

タを用いた光線空間補間", IMPS2004, pp.35-36, 2004

- [8] Shinya Oka, Toshiaki Fujii and Masayuki Tanimoto, "Dynamic Ray-Space Coding Using Inter-view Prediction", IWAIT2005, pp. 19-24, 2005
- [9] 木全, 北原, 志水, 上倉, 八島, "自由視点映像通信のための多視点符号化の一検討", FIT2004, pp.225-226, 2004
- [10] 田口, 苗村 "光線空間符号化のための自由視点画像合成に関する検討", 3 次元画像コンファレンス'05, pp.9-12, 2005.,