

ボリュームデータの3次元DCT圧縮

森下広之 大野義夫

mori@on.cs.keio.ac.jp ohno@on.cs.keio.ac.jp

慶應義塾大学理工学研究科

概要

ボリュームデータは3次元のデータであり、様々な利点を持つがデータ量が大きという欠点をもつ。そのため、本論文では3次元離散コサイン変換を用いてボリュームデータを圧縮し、評価を行ったのでその結果を報告する。

この手法は、現在画像圧縮の国際標準であるJPEGに用いられている離散コサイン変換を3次元に拡張してボリュームデータを圧縮するもので、評価用のテストデータとしてはCTスキャナから得たデータと、CGによって作成した人工的な画像を用いた。

3次元の離散コサイン変換を用いて圧縮の実験を行った結果、一定の画質を保ちながら、原画像の1/10程度までの圧縮を行えることが確認できた。

Volume Data Compression
Using Three-Dimensional Discrete Cosine TransformHiroyuki Morishita and Yoshio Ohno
Faculty of Science and Technology, Keio University

Abstract

Volume data is widely used for the expression of 3D objects. It has various advantages, but its handling is still inconvenient because of huge size. This paper discusses a 3D extension of DCT (Discrete Cosine Transform) for the compression of volume data. 2D DCT is used in JPEG for the image data compression.

In this paper, we used two sets of volume data for the evaluation of the 3D DCT: one is obtained through a CT scanner and the other is artificial one made by computer graphics. It turned out that 3D DCT can compress image data to 1/10 sizes without dropping of image quality.

1 はじめに

近年のコンピュータのハードウェア性能の向上により、大量のデータを扱うことが容易になってきている。そのため、従来は困難であった空間的な3次元のデータであるボリュームデータが様々な分野で活用されるようになった。ここでボリュームデータとは、3次元空間上に設定した等間隔の3次元格子点での実数値集合として表されるデータをさすものとする。ボリュームデータはCT(コンピュータ断层撮影)画像や、MRI(核磁気共鳴法)画像などを積み重ねることによって作成できる。また、コンピュータシミュレーションの分野においても用いられており、これからのコンピュータの進歩、応用分野の広がりなどを考えると、ボリュームデータの活用の幅はさらに広がるものと考えられる。

ボリュームデータ圧縮としてMPEGを用いることも考えられるが、その場合時間軸として取った軸方向以外の展開が極めて不便となる。そこでわれわれはJPEGに使われる離散コサイン変換を3次元に拡張することによりボリュームデータを圧縮することを試みた。この方法ではブロックごとの境界であればどの方向からでも切り出すことができ、部分的な展開も可能である。

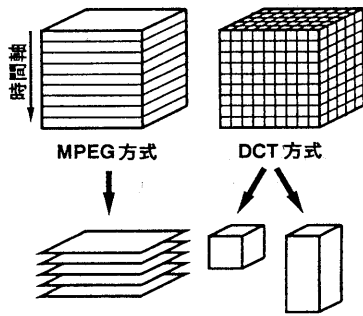


図1: 画像展開

2 本研究の手法

ボリュームデータは、物体の内部構造を表現できるなどの利点を持つが、データ量が膨大なため、保持、蓄積に非常にコストがかかるのが欠点である。

本論文では2次元離散コサイン変換[1], [2]を3次

元に拡張し、それをボリュームデータに適用して圧縮し、評価を行う。圧縮の手順は以下のステップからなっている。

1. ボリュームデータを $N \times N \times N$ のブロックに分割する
2. 各ブロックに3次元離散コサイン変換を適用する
3. 量子化を行い変換係数の並べ変えを行う
4. 変換係数のエントロピー符号化を行う

2.1 離散コサイン変換

ボリュームデータから切り出したブロックに離散コサイン変換を適用する。3次元離散コサイン変換による変換係数は以下の式(1), (2)で与えられる。

正順3次元離散コサイン変換:

$$F(u, v, w) = \frac{8C(u)C(v)C(w)}{N^3} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} f(j, k, l) \cos \frac{(2j+1)u\pi}{2N} \cos \frac{(2k+1)v\pi}{2N} \cos \frac{(2l+1)w\pi}{2N} \quad (1)$$

逆順3次元離散コサイン変換:

$$f(j, k, l) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{w=0}^{N-1} C(u)C(v)C(w)F(u, v, w) \cos \frac{(2j+1)u\pi}{2N} \cos \frac{(2k+1)v\pi}{2N} \cos \frac{(2l+1)w\pi}{2N} \quad (2)$$

ここで,

$$C(\omega) = \begin{cases} \frac{1}{\sqrt{2}} & \omega = 0 \\ 1 & \omega = 1, 2, \dots, N-1 \end{cases} \quad (3)$$

$f(j, k, l)$: 格子点 (j, k, l) におけるデータのボクセル値

$F(u, v, w)$: 離散コサイン変換後の係数

である。

この変換により、1つの波としてとらえた画像を空間周波数別に分解し、それぞれの波の係数を求めることができる。

2.2 量子化

今回、離散コサイン変換の評価に重点をおいたため量子化に用いる量子化テーブルは1つの式で与えられる簡単なものを用い、下に示すように量子化テーブルの座標を表す、 u, v, w の和が等しくなる平面ごとに同じ値を与えた。ここで、 $\text{quant_table}[u][v][w]$ は座標 u, v, w に対応する量子化テーブルの値、 m は量子化のレベルを設定するパラメータで、これを変えることにより量子化の粗さ、すなわち、圧縮率をコントロールした。

$u + v + w = 0$ に対して、

$$\text{quant_table}[u][v][w] = 16$$

$1 \leq u + v + w < 4$ に対して、

$$\text{quant_table}[u][v][w] = u + v + w + 10$$

$4 \leq u + v + w < 11$ に対して、

$$\text{quant_table}[u][v][w] = u + v + w + 20$$

$11 \leq u + v + w < 16$ に対して、

$$\text{quant_table}[u][v][w] = u + v + w + 10m$$

$16 \leq u + v + w < 21$ に対して、

$$\text{quant_table}[u][v][w] = u + v + w + 20m$$

2.3 符号化

2.3.1 変換係数の並べ替え

符号化を行う際に、符号化の効率をあげるため、離散コサイン変換により得られた変換係数を並べ替えてやる必要がある。2次元画像の場合には、高周波成分をつなげ、値の連続性を保つ方法としてジグザグスキャンが用いられている [1], [4]。本研究では、これを3次元に拡張する。すなわち図2のように座標 u, v, w の和が等しくなる平面でカットし、その平面ごとにジグザグスキャンを行った。また、比較のための2次元圧縮ではスライスごとにジグザグスキャンを行った。

2.3.2 エントロピー符号化

エントロピー符号化については、2次元離散コサイン変換圧縮と同じ処理を行う。本研究の符号化はGNUのgzip (ver1.2.4)を用いて行った。gzip[3]は、LZSS法(スライド辞書法)とハフマン法を組み合わせたものである。本論文で用いた手法がもとのデータを完全には

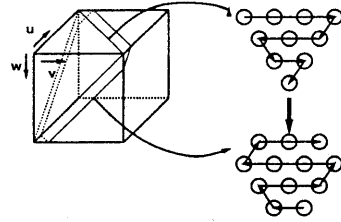


図2: 変換係数の並べ替え

復元できない非可逆圧縮であるのに対し、gzipは可逆圧縮である。今回用いたデータをgzipのみで圧縮した場合、圧縮率は19.8%であった。

3 結果

CTスキャナのスライスデータを積み上げて作ったボリュームデータと、CGによる人工的な画像の例として作成したボリュームデータ(図3)に、前節のアルゴリズムを適用し圧縮、評価を行った。

評価の基準は以下の3つである。

- 3次元圧縮による圧縮率と画質の劣化との関係
- ブロックサイズの変更による処理速度と誤差の変化
- アニメーションによる評価

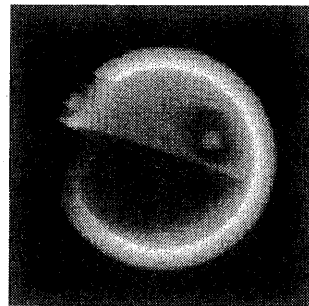


図3: CGによる人工的なボリュームデータの例

3.1 3次元圧縮による圧縮率と画質の劣化

CTデータ(原画像を図9に示す)については、圧縮率が11.5%, 7.2%(図10), 5.4%(図11), 5.1%の4通

りについて評価を行った。また、比較のため2次元離散コサイン変換を行い、同程度の圧縮率が得られた画像を示した。

また、画像の劣化の数値的な評価基準として、CTデータの原画像と圧縮画像とのボクセル値の差ごとにボクセル数をカウントしたものを図4に示し、CTデータ、CGデータの両方においてそれぞれの圧縮率での誤差の分布の幅を表1に示した。さらに、圧縮率ごとの画像の平均二乗誤差を図5に示した。

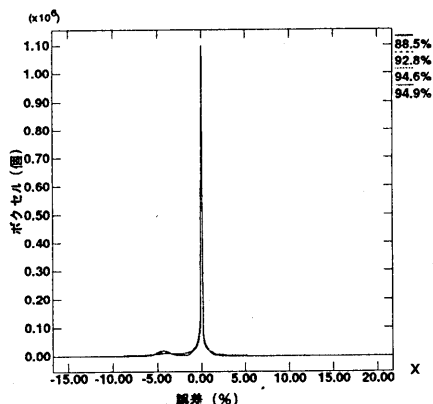


図 4: ボクセル値の誤差の分散 (CT データ)

表 1: 誤差の分布幅

データ	圧縮率 (%)	分布の範囲
CT データ	11.5	-10.7% ~ 6.3%
	7.2	-14.3% ~ 10.7%
	5.4	-12.2% ~ 17.8%
	5.1	-7.8% ~ 18.3%
CG データ	10.4	-9.0% ~ 2.75%
	7.2	-17.0% ~ 10.3%
	5.8	-25.0% ~ 22.3%
	5.1	-29.8% ~ 27.8%

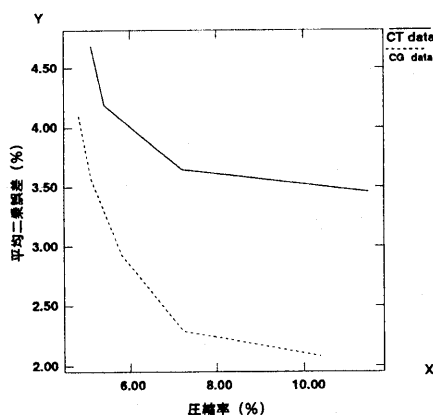


図 5: 圧縮率と平均二乗誤差

3.2 ブロックサイズによる処理速度と誤差の変化

離散コサイン変換を行う際のブロックの大きさを変え、それによる処理速度の変化とそれぞれのブロックサイズでの平均二乗誤差を測定した。ブロックサイズは2, 4, 6, 8の4通りとし、評価にはCTデータを用いた。処理速度の変化を図6に、平均二乗誤差を表2、図7に示した。離散コサイン変換の特性を比較するため、この測定に限り量子化は行わなかった。また、低周波成分1/8だけを使った場合の平均二乗誤差も測定した。これも図7に示した。この時に生成された画像をブロックサイズ2, 4, 8について図15, 16, 17に示した。計算時間の測定はSUN SPARCstation IPX (倍速チップ搭載)で行った。

表 2: ブロックサイズと計算時間, 誤差

サイズ	圧縮時間 (秒)	展開時間 (秒)	誤差 (%)
2	52.3	18.9	1.3442
4	82.3	57.2	1.3449
6	244.0	234.2	1.3442
8	518.6	833.9	1.3349

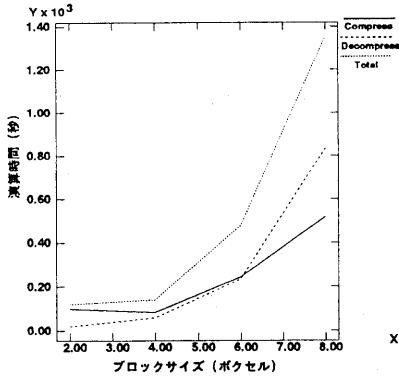


図 6: ブロックサイズと計算時間

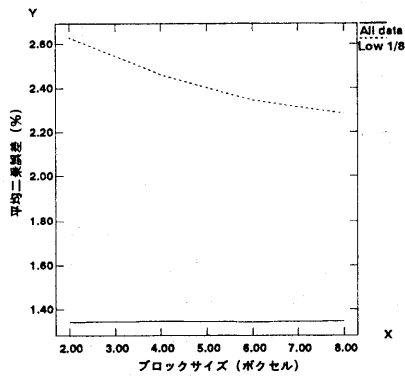


図 7: ブロックサイズと誤差

4 評価

4.1 画質の評価

図 9～図 11を見て分かるように、圧縮率 5.4% の画像では劣化が目立つ。また、これ以上圧縮率をあげると急激に画質が悪化した。劣化の主な原因はブロック歪みである。離散コサイン変換は 1 つ 1 つのブロックで完結してしまい、ブロック間の境界条件を持たない。そのため圧縮率が上がると情報量が少なくなり、ブロックとブロックの間に格差が目立ってしまう。

4.2 2次元圧縮画像との比較

同程度の圧縮率で比較すると図 10 と図 11 の間で格差が目立つ。2次元圧縮ではスライスごとに圧縮が行わ

れるので縦方向にはブロック歪みは現れない。その分やや劣化が目立たないとも取れるが、圧縮率 5.3% になると横方向のブロック歪みがひどくなる。このように、圧縮率が高いときにはブロック歪みは 3次元圧縮の方がやや低いことが分かる。3次元の場合は、立方体のブロックに対して離散コサイン変換が施されるため、同じだけの情報を削っても 2次元の場合に比べ画質は良くなる。これはブロックサイズを大きくした場合と同じ効果があるためである。

4.3 誤差の分布

誤差とその誤差を持つボクセル数をカウントした図 4 をみると、ほとんどのボクセルが誤差 $\pm 1\%$ の範囲に集まっている。この傾向は圧縮率が変わってもほぼ同じである。また、CG データと CT データの誤差の分布幅 (表 1) を比較すると、圧縮率が低い場合はあまり違いは見られないが、圧縮率が上がると CG データの誤差の幅が急激に大きくなった。これは、CG などによって作られた画像が急激な画素値の変化を持つためと考えられる。図 8 に示すように、急激な変化を持つ部分は、周波数分解が困難なためである。

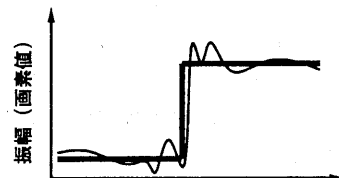


図 8: エッジでの波形

4.4 圧縮率と誤差

圧縮率とボクセルの誤差の関係を示す図 5 では、圧縮率が 6% を超えたあたりで誤差の増加傾向に大きな変化が見られる。実際の画像でもこのあたりで劣化が目立ち始めた。圧縮率は、量子化テーブルの設定による影響も大きいため、一概にこの圧縮率が限界であるとは言えないが、一定の限界が存在することは明らかである。われわれの実装においては、この限界が 10%～8% 位にあ

るものと考えられる。また、前節で述べたように、この圧縮法には不向きのはずのCGデータの方が誤差の値が小さくなっているのは、これがCTデータと比較して非常に冗長さの大きな画像であったためと考えられる。

4.5 ブロックサイズと計算時間

ブロックサイズと計算時間の関係を示した図6を見ると、ブロックサイズの増加に伴って計算時間も増加の割合を増しながら増大している。これは主に離散コサイン変換の演算規模の増大によるものである。式(1)、(2)から分かるように、1ブロックに対する計算量は、ブロックの大きさを N としたとき、 $O(N^3)$ となる。

4.6 ブロックサイズと誤差

ブロックサイズの変化によるブロック値の平均2乗誤差の変化を表した図7を見ると、全ての変換係数を保存して圧縮展開を行ったデータについては、ブロックサイズによらずほぼ一定の誤差の値を保っている。しかし、変換係数の内、低周波成分1/8のみを用いたデータについては、ブロックサイズによって誤差の値に違いが見られる。値としては図5で得られたものよりも小さいが、このデータは量子化を行っておらず、量子化を行うと誤差は増大する。図15、16、17を見ると、同じデータ量(全体の1/8)にも関わらず、画質の違いがはっきりとでている。このようなことからブロックサイズは可能な限り大きい方が、少ないデータ量でも高品質の画像が得られることが分かる。

5 評価

5.1 3次元離散コサイン変換について

われわれの実装では、実用的な圧縮率は8%~10%位までであった。あまり圧縮率が上がらなかった要因としては量子化テーブルを固定したことが考えられる。また、同程度の量子化を行った場合、圧縮率が低い間は縦方向のブロック歪みの見られない分2次元離散コサイン変換による圧縮の方が画質がやや良かったが、高圧縮率においては2次元圧縮よりも3次元圧縮の方がブロック歪みの程度は小さかった。このことよりボリュームデータに高圧縮を行う場合には、3次元圧縮の方が適してい

るといえる。しかし、同時にある程度以下の情報量になると急激に誤差が増加したことから、限界もはっきり存在する。さらに、これは3次元、2次元の両方に共通した問題であるが、画素値が急激に変化するCGで作成されたような図形に関しては、圧縮率を上げた場合に誤差の分布の範囲が急激に広がった。この実験も離散コサイン変換が自然画像(CTデータも含む)向きの圧縮法であることを裏付けている。

5.2 実用化について

実用化において最も問題となるのは、圧縮にかかる時間である。ブロックサイズ8の場合では、圧縮に10分前後、展開に15分前後を要した。処理速度を上げる最も容易な方法は、離散コサイン変換のブロックサイズを小さくすることである。しかし、画質を維持し、高圧縮率を得るためには、ハードウェア化が不可欠であろう。

6 結論

今回、量子化テーブルを固定するなど、比較的単純な実験であったにもかかわらず、あまり画質の劣化も起こらずに原画像の10%程度に圧縮することができた。2次元圧縮との比較では、圧縮率についてはやや有利な結果が得られたが、圧縮時間などの面でははるかに劣っている。また、縦方向のブロック歪みが目立ってしまうのも問題である。しかし、高い圧縮率においては2次元圧縮に比べてブロック歪みが小さいなどの利点も見られた。

今後の課題としては、圧縮率があまり上がらなかった主な要因である量子化テーブルの改良が考えられる。今回は1つの式によって与える簡単なものを用いたが、画像ブロックごとに異なった量子化テーブルを与えることにより、画質の劣化を抑えて圧縮率を上げることができる。これは、離散コサイン変換を行った後の変換係数の交流成分の分布によって画像の精密度を定義し、これによってブロックをいくつかのレベルに分解し、量子化テーブルを適応的に割り当てるといったものである。

このように、まだまだ改良すべき点は多いものの、ボリュームデータを高い圧縮率で圧縮する手法として、3次元DCT圧縮は有効な手段であるといえる。

謝辞

本研究にあたって有益な助言をして下さった、慶應義塾大学工学部 小沢慎治教授、慶應義塾大学工学部 野寺隆専任講師にお礼を申し上げます。

参考文献

- [1] N. Ahmed, T. Natarajan, K. R. Rao, "Discrete Cosine Transform," *IEEE Transactions on Computers*, January 1974, pp. 90-93, 1974.
- [2] K. R. Rao, P. Yip (安田 浩, 藤原 洋 訳) "画像符号化技術 DCT とその国際標準," オーム社, 1992.
- [3] 奥村 晴彦, "ファイル圧縮技術 (上)," bit, 共立出版, 1994 年, 12 月号, pp. 4-13.
- [4] テレビジョン学会 編, 原島 博 監修, "画像情報圧縮," オーム社, 1991.

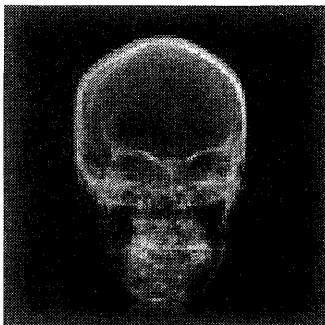


図 9: オリジナルの CT データ

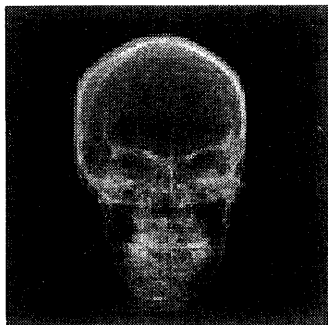


図 10: CT データ 7.2% (3D)

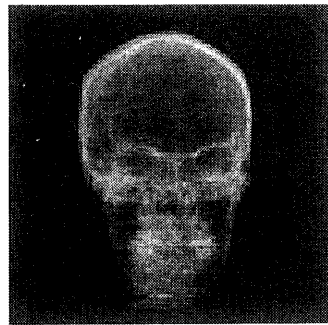


図 11: CT データ 5.4% (3D)

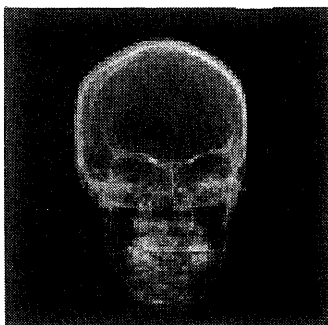


図 12: CT データ 11.6% (2D)

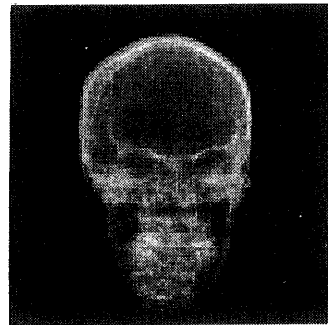


図 13: CT データ 7.2% (2D)

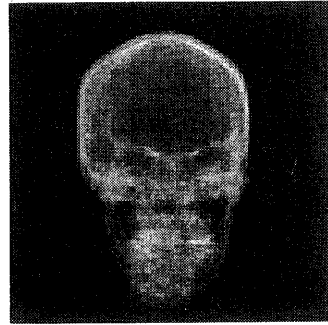


図 14: CT データ 5.3% (2D)

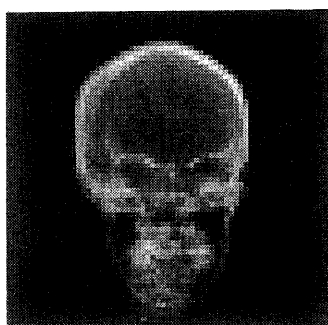


図 15: ブロックサイズ 2

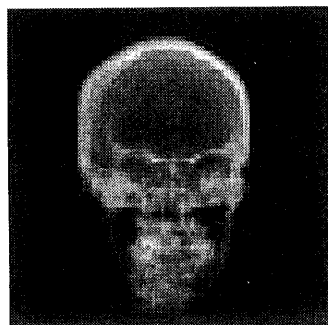


図 16: ブロックサイズ 4

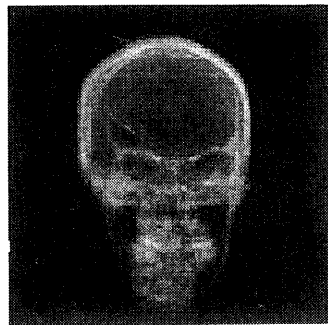


図 17: ブロックサイズ 8