

研究速報

ダウンサンプリングされた光線空間データのレートひずみ特性についての一考察

高橋 桂太^{†a)} (正員)

A Study on Rate-Distortion Performance of Down-sampled Light Field Data

Keita TAKAHASHI^{†a)}, Member[†] 名古屋大学大学院工学研究科, 名古屋市

Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya-shi, 464-8603 Japan

a) E-mail: keita.takahashi@ieee.org

あらまし 光線空間データの伝送において, 送信側で視点や画素をダウンサンプリングした上で符号化し, 受信側でアップサンプリングする方法が考えられる. 本論文では, このような符号化シナリオのレートひずみ特性を直感的に把握するための理論的考察を示す.

キーワード 光線空間データ, レートひずみ理論

1. まえがき

三次元映像や自由視点映像を実現するための入力として, 同一の対象を非常に多数の視点から撮影した画像群を用いることが有効である. このような画像群を光線空間データと呼ばれる. 光線空間データは膨大な情報となるため, 蓄積・伝送のためには, 符号化が重要となる. 本論文では, 符号化における, 光線空間データのサブサンプリングの有効性を, 理論的に考察する.

まず, 問題設定を述べる. 元の光線空間データが N 視点, $W \times H$ 画素の画像で構成されるとする. ダウンサンプリングされた光線空間データを, N' 視点, $W' \times H'$ 画素とする. ここで, $N < N'$ であれば視点を間引いた (または, より少ない視点でのみ撮影した) ことに, $W' < W, H' < H$ であれば画素を間引いた (または, より低解像度のカメラを用いた) ことに相当する. ダウンサンプリングを伴う場合には, より少ないサンプル数のデータを符号化して伝送するが, 受信側でアップサンプリングを行い, 元のサンプル密度を復元するものとする. 視点を間引いた場合には近接視点の画像を用いた視点補間処理が, 画素を間引いた場合には複数フレームを用いた超解像処理が行われる. ここで興味があるのは, 元の光線空間データを直接符号化する場合と, ダウンサンプリングを伴う場合を比較して, レートひずみ特性の観点でどちらが優れているのか, という点である. ただし, レートひずみ特性は, 元のデータのサンプル数, すなわち, N 視点 \times $W \times H$ 画素において比較する. また, 単純化のた

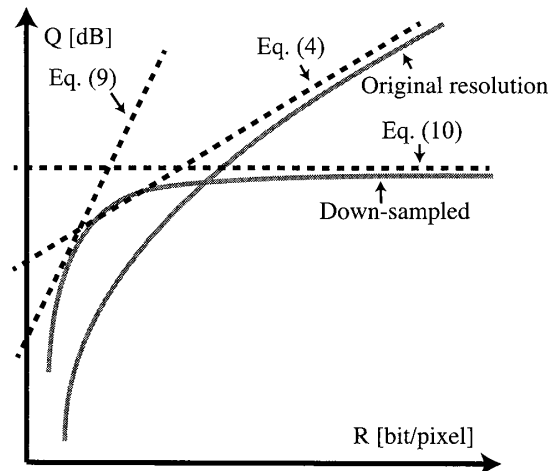


図1 サブサンプリングを伴う符号化のレートひずみ特性
Fig. 1 R-D performance of down-sampled light fields.

め, 符号化において視点間予測は用いず, 画像を個別にイントラ符号化すると仮定する.

先行研究において, 視点を間引いたケース [1], 及び低解像度のカメラを用いたケース [2] について, 個別に詳細に解析した. これらを比較すると, 両方のケースは類似した特性を示すことが分かっている. 図1に示すように, ダウンサンプリングを伴う場合 (実線: Down-sampled) は, 元のデータを直接符号化した場合 (実線: Original resolution) と比較して, 低レートでは優位性を示す可能性があるが, 高レートでは品質が飽和するため優位とはならない. 本論文では, これらを包括的に扱う枠組みを示す. この枠組みは, 光線空間データのダウンサンプリングを伴う符号化に関して, より統一かつ直感的な解釈を可能にする.

2. 画像信号のレートひずみ特性

画像信号が, 広義の定常ガウス過程によって生成されると仮定し, その電力スペクトル密度関数を $\Phi(u, v)$ とすると, そのレートひずみ関数は以下となる [3], [4].

$$D(\theta) = \frac{1}{4\pi^2} \iint_{[-\pi, \pi]^2} \min[\theta, \Phi(u, v)] dudv \quad (1)$$

$$R(\theta) = \frac{1}{8\pi^2} \iint_{[-\pi, \pi]^2} \max\left[0, \log_2 \frac{\Phi(u, v)}{\theta}\right] dudv \quad (2)$$

ここで, ひずみ D は平均二乗誤差, レート R は bit/pixel で表される. θ はレートとひずみのトレードオフを制御するパラメータであり, θ が大きいほど, レートは小さくひずみは大きくなる. また, デシベルで表したひずみの尺度を $Q(\theta) = -10 \log_{10} D(\theta)$ と表記する.

画像信号のレートひずみ特性は, それぞれの画像の

性質に依存する。平坦な画像は、低レートでも高品質に符号化できるが、変動の激しい画像は、品質を保つためにより大きなレートを必要とする。しかし、次に述べる性質は、個々の画像信号に依存せず、普遍的に成り立つ。低レート (θ が大きい) では、レートの増加に伴って画質が急しゅんに向上する。一方、高レート (θ が小さい) では、画質の向上の度合いが徐々に飽和し、約 6 dB/bit に漸近する。この性質は、以下のようにして確かめられる。

$$\begin{aligned} \frac{dQ}{dR} &= \frac{dQ/d\theta}{dR/d\theta} = 20 \log_{10} 2 \cdot \frac{\theta}{D(\theta)} \\ &\geq 20 \log_{10} 2 \simeq 6 \text{ [dB/bit]} \end{aligned} \quad (3)$$

ここで、等号付き不等号は、式 (1) より $D(\theta) \geq \theta$ であることから導かれる。等号が成立するのは、全ての周波数に対して $\theta \leq \Phi(u, v)$ となる場合であり、 $\theta \rightarrow 0$ の極限 (高レートの極限) ではこれが成立する。極端なケースとして、符号化対象が白色雑音の場合には、傾きは θ に依存せず常に約 6 dB/bit で一定となる。ここで述べた性質を図 2 に示す。

上記に基づき、本論文では、画像信号のレートひずみ関数の上限を決める漸近線

$$Q_{up}(\theta) = g_{const} R(\theta) + b \quad (4)$$

を導入する。添字の up は、これがひずみ尺度 Q の上限または上界であることを表す。傾き $g_{const} = 20 \log_{10} 2$ は定数となるが、切片 b は画像の性質 (テキストチャの複雑さ) に依存し、複雑な画像ほど b は小さくなる。図 2 では、それぞれのレートひずみ曲線に対する漸近線を点線で表している。本論文では、この漸近線を用いて、レートひずみ特性を解析する。

3. ダウンサンプリングを伴うレートひずみ特性

元の光線空間を直接符号化伝送する場合、各画像を個別にイントラ符号化すると仮定しているため、レートひずみ曲線の上限は式 (4) で与えられる。ダウンサンプリングされた光線空間データに対しても同様に、

$$Q'_{up}(\theta) = g_{const} R'(\theta) + b' \quad (5)$$

を上限を決める漸近線とする。ここで、 $Q'(\theta) = -10 \log_{10} D'(\theta)$ である。視点のみを間引いた場合には、各画像の性質は元の光線空間データの画像と変わらないと考えられるため、 $b' = b$ である。しかし、画素を間引いた場合には、画素ピッチを基準に考えると、同一の信号がより高い周波数をもつように見えるため、

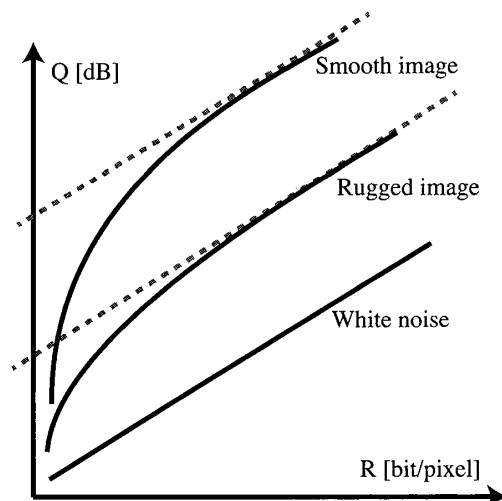


図 2 画像信号のレートひずみ関数
Fig. 2 R-D function of image signals.

一般に $b' < b$ となる。

受信側で、ダウンサンプリングされたデータを、元のサンプリング密度にアップサンプリングした後のレート \hat{R} (bit/pixel) 及びひずみ \hat{D} (平均二乗誤差) は、以下のように表せる。

$$\hat{R}(\theta) = \alpha R'(\theta) \quad (0 < \alpha < 1) \quad (6)$$

$$\hat{D}(\theta) = E_{res} + f(\theta) \quad (7)$$

まず、式 (6) については、アップサンプリングによりサンプル点数が増えるため、 $\hat{R}(\theta)$ は $R'(\theta)$ よりも定数倍小さくなる。例えば、サンプル点数が 4 倍になる場合には、 $\alpha = 1/4$ である。一方、式 (7) については、やや複雑である。まず、第 1 項の E_{res} は、 $\theta \rightarrow 0$ の極限、すなわち符号化ひずみが全くない場合に生じる、アップサンプリングに伴う誤差を表すとする。視点補間を行う場合には、奥行推定の誤差や、オクルージョン領域等が E_{res} に寄与する。超解像を行う場合には、フレーム同士の位置合せ誤差や、点広がり関数の形状に依存する誤差等が E_{res} に含まれる。一方、第 2 項の $f(\theta)$ は、符号化のひずみに伴って増大する成分を表しており、定義から $\lim_{\theta \rightarrow 0} f(\theta) = 0$ とする。関数 f の符号化ひずみ $D'(\theta)$ への直接的な依存性は、必ずしも単純な形式にはならないため、ここでは f の下界を

$$f(\theta) \geq \beta D'(\theta) \quad (8)$$

として一次式で表現する。 β は、上記の条件を満たすべく大きな正の数とする。 $f(\theta)$ も $D'(\theta)$ も平均二乗誤差であり、式 (8) の両辺の次元は一致する。

ここまでの定式化に基づき、ダウンサンプリングを伴うケースのレートひずみ曲線の上界を導出する。まず、 θ が大きい（低レートの）ときには、 $f(\theta)$ は E_{res} に対して相対的に大きくなり、式 (7) は $\hat{D}(\theta) \simeq f(\theta)$ と近似される。この近似式を式 (6)(8) とともに式 (5) に代入することにより、

$$\hat{Q}_{up}(\theta) = \frac{g_{const}}{\alpha} \hat{R}(\theta) + b' - 10 \log_{10} \beta \quad (9)$$

としてレートひずみ曲線の上界を決める直線が導出される^(注1)。一方、 θ が小さい（高レートの）ときには、 $\hat{D}(\theta) \simeq E_{res}$ となり、品質の上限は以下で定められる。

$$\hat{Q}_{up}(\theta) = -10 \log_{10} E_{res} \quad (10)$$

結論として、ダウンサンプリングを伴うケースのレートひずみ曲線は、式 (9)(10) で表される 2 本の直線を上界とする。この様子を図 1 に示す。 $0 < \alpha < 1$ より、式 (9) の直線の傾きは、式 (4) で与えられる元の光線空間データの漸近線の傾きよりも大きい。この傾きの大きさが、低レートにおいて、ダウンサンプリングを伴う符号化が優位になる可能性がある現象と結び付いている。ただし、式 (9) の切片は β に依存し、 β が小さいほど切片は大きくなることに、注意が必要である。正の数 β を任意に小さくとれば、式 (8) を容易に満たせることは自明である。しかし、よりタイトな上界（達成可能な上限により近い上界）を得るためには、式 (8) を満足する、なるべく大きな β を求める必要がある。一方、式 (10) は、アップサンプリングの性能によって定められる品質の上限を表しており、

E_{res} が大きいほど上限が小さくなる。この直線によって、ダウンサンプリングを伴う符号化の高レートでの性能が頭打ちになる。したがって、 β または E_{res} の値が大きい場合には、式 (9) が式 (4) を上回る前に、式 (10) の上限に到達する可能性がある。この場合、ダウンサンプリングを伴う符号化は、元のデータを直接符号化する場合と比べて優位にはなりにくい。逆に、 β 及び E_{res} が十分に小さければ、ダウンサンプリングを伴う符号化が、低レートで優位となりやすい。ここに述べた性質は、先行研究 [1], [2] と一致するが、本論文の解析は、わずかなパラメータで簡潔に表現されるため、より直感的に把握できる。

4. む す び

本論文では、ダウンサンプリングを伴う光線空間データの符号化の問題を理論的に考察し、レートひずみ曲線の上界を簡潔な形で導いた。符号化において視点間の予測を伴う場合の解析は今後の課題である。

謝辞 本研究は JSPS 科研費 24700162 の助成を受けた。

文 献

- [1] M. Ishii, K. Takahashi, and T. Naemura, "Rate-distortion performance of multi-view image coding with subsampling of viewpoints," Proc. IEEE Int. Conf. on Image Processing, pp.2428-2431, 2008.
 - [2] K. Takahashi, T. Naemura, and M. Tanaka, "Rate-distortion analysis of super-resolution image/video coding," Proc. IEEE Int. Conf. on Image Processing, pp.1665-1668, 2011.
 - [3] T. Berger, Rate Distortion Theory, Prentice-Hall Electrical Engineering Series, Prentice-Hall, 1971.
 - [4] 酒井善則, 吉田俊之, 映像情報符号化, ヒューマンコミュニケーション工学シリーズ, オーム社, 2001.
- (平成 24 年 12 月 10 日受付, 25 年 2 月 18 日再受付)

(注1) : 式 (8) より、この直線は、厳密な上限ではなく、上界の一つの表現に過ぎないことに注意を要する。