

HDTV用リアルタイムトーンマッピング処理システムの検討 Development of real-time tone mapping system for HDTV

大久保 英彦†
Hidehiko Okubo

三ツ峰 秀樹†
Hideki Mitsumine

深谷 崇史†
Takashi Fukaya

山内 結子†
Yuko Yamanouchi

1. はじめに

CG制作では、環境マッピングやイメージベースドライビングのための画像として高ダイナミックレンジ画像(HDRI)を使用することがリアリティー向上に非常に効果的であると言われて久しい。実世界のシーンからHDRI取得のために、撮影用デジタルカメラの応答曲線の計測と、露光時間の異なる複数撮影画像を利用したHDRI生成手法が確立され、静止画像であればツールを利用して比較的簡単にHDRIを作成することができるようになってきている。

その一方でテレビ番組制作としても、映像加工時の自由度の大きさから、特殊な映像素材として高ダイナミックレンジ画像を取得することへの強い要求がある。

現在本研究と共に、番組制作用の高ダイナミックレンジHDTV動画取得のためのカメラシステムの研究を行っている。収録とあわせて、そこで取得される12bitを超える輝度レンジを持つ画像を表示させる必要がある。しかし実際に番組制作において撮影現場で利用することを考慮すると、表示デバイスとしては現在一般的に使用されている映像モニターを使用し、かつ撮影時のモニタリング環境としても幅広い輝度にわたるディティール確認のための理想的な環境が準備できないことが予想される。そこで、撮影された高ダイナミックレンジ画像を適切な手法で圧縮するためのトーンマッピングを行うことが、モニタリングのための現実的かつ有効な手法となる。

本研究では、RGB各色16bit float (Half float)のHDTV解像度の映像に対してリアルタイムにピクセル輝度値の対数平均を算出させてトーンマッピングを行うシステムをPCベースで試作し、そのパフォーマンスを評価した。

2. トーンマッピング手法

トーンマッピングについては多くの手法が提案されている。その基になるものが、シーン全体の大域的な光の変化と主に反射による局所的な照明効果を分別し、大域的な変化のみを圧縮させてレンジ内に納めようとする考え方である。トーンマッピングオペレータの種別としては、画像全体にわたって同一のオペレータが作用するか、局所的に変化するかによってグローバルオペレータ・ローカルオペレータに分類され、また時間依存性などの特徴によっても分類がなされている。

本研究においては、オペレータが比較的シンプルでリアルタイム実装に親和性が高いことと、さまざまな画に対する適用結果に不自然なアーティファクトが少ないという理由から、Automatic dodging-and-burningと呼ばれているReinhardの手法[1]をベースに実装することとした。

dodging-and-burningとは、写真をプリントする際に明るすぎる部分のライトを抑えることで減感し、暗すぎる部分に

はよりライトを加え増感することにより、全体を一様にプリントした場合に比べてコントラストを大きくする伝統的な写真技法である。そのアルゴリズムについて、以下に概要をまとめる。

$L_w(x, y)$ を画像中のピクセル (x, y) に対応する実世界の輝度であるとする、その輝度の対数平均 \tilde{L}_w は、以下で与えられる。

$$\tilde{L}_w = \exp\left(\frac{1}{N} \sum_{x,y} \log(\delta + L_w(x, y))\right) \quad (1)$$

ここで N は画像のピクセル数を、そして δ は輝度値が0の時に起こる計算上のアンダーフローを防ぐための微小値である。 \tilde{L}_w が写真光学のゾーンシステムにおける中間値(middle-gray)ゾーンにマップされる時、ピクセル輝度のスケールリングとして以下が適用される。

$$L(x, y) = \frac{a}{\tilde{L}_w} L_w(x, y) \quad (2)$$

ここで a はキー値とよばれ、この画像が主観的に明るい(ハイキー)、暗い(ローキー)ということを決める値になる。通常の場合、キー値は $a=0.18$ が用いられる。もっともシンプルなグローバルトーンマッピングオペレータは、ディスプレイ輝度が L_d で与えるとする、

$$L_d(x, y) = \frac{L(x, y)}{1 + L(x, y)} \quad (3)$$

で与えることができる。これによって、低いコントラストエリアにおけるディティールを保持しながら、輝度全体を0から1の表示可能域に納めることが可能になる。

次に、周辺領域の輝度値と大きく異なるピクセルのコントラストを強調させるようにローカルオペレータを作用させ、dodging-and-burningを実現させる。まず、複数の空間スケール s_i で定義された円形のコンボリューションカーネルで画像をたたみ込む。これによって得られた応答を $V(x, y, s_i)$ とすると、空間スケール s_i に応じたローカルなコントラストを以下の式で与えることができる。

$$Activity(x, y, s_i) = \frac{V(x, y, s_i) - V(x, y, s_{i+1})}{2^\varphi a s_i^{-2} + V(x, y, s_i)} \quad (4)$$

ここで φ はエッジ強調をコントロールする先鋭度パラメータになる。Reinhardは s_i について8つのスケールレベルを考慮しており、最小の s_1 は0.35で、そこから1.6倍づつ大きくしている。そして、各ピクセルに対して最初に $|Activity(x, y, s_m)| > \epsilon$ を満たす s_m を見つける。この s_m は、そのピクセル周辺部で急激な輝度変化をしない最大の領域を与えることになる。このようにして得られた $V(x, y, s_m)$ を利用して、以下の式からローカルトーンマッピング処理後の輝度を得ることになる。

† 日本放送協会 放送技術研究所 (人間・情報科学)

$$L_d(x, y) = \frac{L(x, y)}{1 + V(x, y, s_m)} \quad (5)$$

3. 実装

本研究ではPCワークステーション上への実装を行い、特にパフォーマンス面での実用性を評価するために、上記のアルゴリズムのうちグローバルオペレータ部分に対する実装のみを行った。

3.1. 処理の流れ

入力される高ダイナミックレンジ画像データはRGB各チャンネルが16bit floatのHD解像度(1920x1080)で、ワークステーションのメインメモリに展開されるものとした。その画像データをGPU(Graphics Processing Unit)によってトーンマッピング処理し、結果をモニタに表示させた。インプリメントは標準的なCG描画APIであるOpenGLとプログラマブルシェーダー言語CgFXを利用した。図1にその処理の流れを示す。

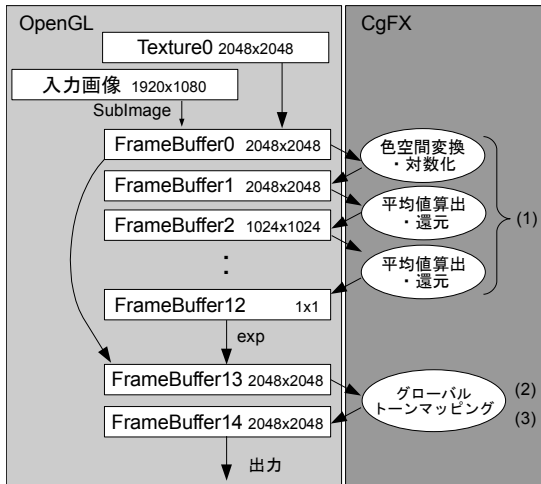


図1. トーンマッピング処理の流れ (右端の数は処理が行われる式の番号を示す)

3.2. Non-power-of-two 画像の扱い

一般的にシェーダー言語で画像のデータを処理する場合、画像全体から一つの量を算出するための直接的な手法はあてられない。しかし上記の処理では、画像全体の輝度の対数平均値を求める処理が必要になる。そこで広く用いられる手法が並列還元処理(parallel reduction)であるが、その際に処理対象となる画像の縦横サイズが2の累乗でないためにそれを考慮した処理が必要となってくる。本研究では、図1の還元処理内でアルファチャンネルを処理対象ピクセルのピクセル占有率として使用することで、2048x2048のテクスチャメモリのうち処理対象とする1920x1080の領域に関するピクセル平均値を算出させた。

4. 実験

実験は以下の図2で示すシステム上でプログラムを実行させてHDRI画像のトーンマッピングを行った。

Workstation: HP xw9400
 AMD Opteron 2220 (2.8GHz) x 2
 main memory 8GB
 Graphics Card:
 NVidia QuadroFX 5600 (1.5GB)
 driver version 180.29

図2. 実験で使用したシステム

本実験での入力画像としては、同じ静止画像を繰り返し入力させることでフレームレート測定をおこなった。

結果として得られた画像の例を以下の図3に示す。これはDurand [2]が公開しているRadiance MapをHDTV解像度の16bit float OpenEXR画像に変換を行って入力画像としたものである。このサンプルの室内外部分のように輝度の違う領域が均等に大きく分かれている場合は、グローバルオペレータでも十分良好な結果を得ることができた。トーンマッピング処理部のフレームレートは約180fpsであった。



図3. 入力画像(上2枚)と処理結果画像(下1枚)。2枚の入力画像はそれぞれ室外と室内に露出を合わせたものを示した。

5. 考察

処理速度で考えると、グローバルオペレータであれば実験システムでHDTV解像度のトーンマッピング処理は十分現実的であることが確認された。ローカルオペレータに関しては、HDTVより小さな画像を利用したGPUによるリアルタイム実装を行った研究がある[3]。これによれば、512x512の画像の場合にグローバルオペレータで約250fpsの処理速度であるシステムにおいて、空間スケールで2つのレベルを考慮したローカルオペレータを作用させて30fpsを超える結果をだしている。それを踏まえると、今後進めていくローカルオペレータ実装後のパフォーマンスでもリアルタイム処理の可能性が十分あると考えられる。

6. まとめ

本研究では、HDTV向けのリアルタイムトーンマッピングシステムを検討し、グローバルオペレータであればPCベースで十分処理が可能であることを示した。実験ではスタティックな画像でのオペレータに対するパフォーマンス評価のみを行ったが、今後は照明環境や輝度分布が異なる素材を対象に処理画像の評価を行い、カメラシステムと併せて運用できるシステムの構築を目指していく。

参考文献

[1] Reinhard E. et.al., "Photographic tone reproduction for digital images.", ACM Transactions on Graphics, Proceedings Siggraph 2002
 [2] <http://people.csail.mit.edu/fredo/PUBLI/Siggraph2002/>
 [3] Goodnight N. et.al., "Interactive Time-Dependent Tone Mapping Using Programmable Graphics Hardware", Proceedings of the 14th Eurographics Workshop on Rendering Techniques, Leuven, Belgium, June 25-27, 2003