

2 台の同型カメラを併用した HDR 動画像の生成法

HDR Video Generation Using Two Identical Cameras

小田仁己¹⁾ 吉田麻子²⁾ 柴田史久¹⁾ 木村朝子¹⁾ 田村秀行¹⁾

Masaki Oda, Asako Yoshida, Fumihisa Shibata, Asako Kimura, and Hideyuki Tamura

1) 立命館大学大学院 情報理工学研究科 (〒525-8577 滋賀県草津市野路東 1-1-1)

2) 立命館大学大学院 理工学研究科 (〒525-8577 滋賀県草津市野路東 1-1-1)

Abstract: This paper proposes a method for High Dynamic Range (HDR) video generation using two identical cameras. When using a traditional image sensor or an Analog-Digital (AD) converter, the brightness range of the produced image is limited. Thus, people often manipulate the exposure level of the camera to generate HDR images. However, this technique cannot be applied when generating HDR videos because the camera and the object are not static. The method in this paper utilizes two identical cameras but with different exposure levels and combines the obtained images to generate HDR videos. This paper also describes techniques used to solve problems when combining the obtained images.

Key Words: High Dynamic Range Image, Identical Cameras, Tone Mapping

1. はじめに

近年, 携帯電話, デジタルカメラ, パソコンなどの普及により, デジタル画像を手軽に扱えるようになった. さらに最近では, 通常よりも広いダイナミックレンジを持った High Dynamic Range (HDR) 画像が扱われるようになり, さまざまな分野に応用されている. デジタル画像でのダイナミックレンジとは表現可能な最小輝度, 最大輝度の比率, さらには画像が持つビット数を意味しており, ダイナミックレンジが広いほど明るさの分解能が高く, 現実シーンにおける幅広い輝度情報を保持できる. これにより, HDR 画像は, より人間の視覚に近い明るさや色の変化を表現可能であり, 通常の画像で輝度の高低が激しすぎる際に発生する白とびや黒つぶれを軽減できる.

通常, デジタルカメラで撮影される画像において, 各画素が持つ情報量は RGB 各 8 ビット (計 24 ビット) であり, HDR 画像はこれよりも大きなビット数を持つ. このようなビット数の大きな画像を取得する際には, 明るさの分解能が高い画像センサや高性能な AD 変換器を使用する必要がある. しかし, 現在一般に使用されているそれらの機器の性能では, 高ビット数の画像を取得は困難である.

そのため, 通常の画像, Low Dynamic Range (LDR) 画像から HDR 画像を生成する手法が Debevec ら[1]により考案された. Debevec らが考案した手法は, 同じカメラで露出を変更して撮影した複数枚の LDR 画像を合成し, ラディアンスマップを生成するというものである. ラディアンスマップとは, 実際のシーンの輝度そのものを画素値として保持している画像で, これが HDR 画像となる. ただ

し, この場合, 保持している画素値が通常のディスプレイで表示可能なビット数を大きく超えるため表示できない. そこで, 図 1 (b) に示すように擬似的にカラーをつけるか, 図 1 (c) のように, トーンマッピングにより画素値が RGB 各 8 ビットに収まるように変換するか, もしくは, 通常のディスプレイよりも大きなビット数の画像が表示可能な HDR ディスプレイを使用する必要がある. なお, トーンマッピングについては 3 章で詳しく述べる.

また, [1]の手法では, 複数枚の LDR 画像から HDR 画像を生成するため, カメラの位置が固定され, かつ対象物体が静止していること, つまり静止画像であることが前提条件となる. よって, 動画像を HDR 化したい場合には, この手法で実現するのは困難となる.

そのため本稿では, 2 台の同型カメラを併用した HDR 動画像の生成を実現する手法を提案する. さらに, 生成する際の諸問題についても取り組み, 実際に生成した HDR 動画像について考察する.

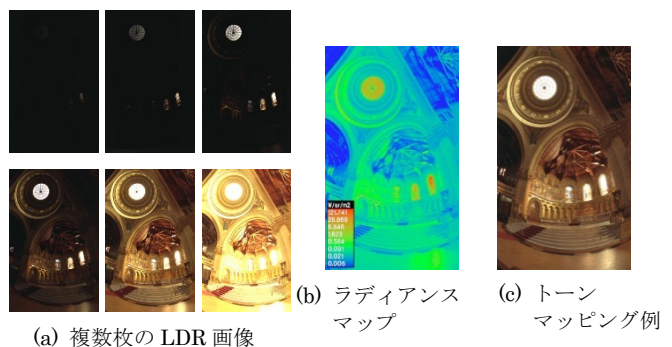
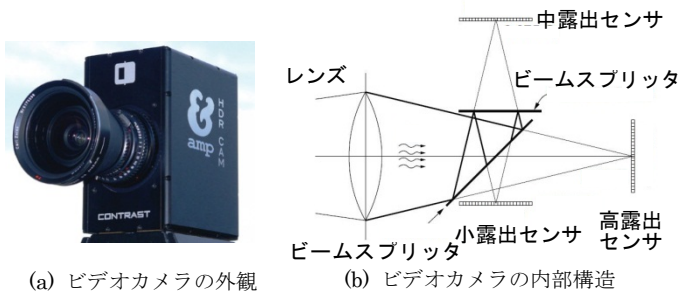


図 1 Debevec ら[1]により考案された HDR 画像の生成



(a) ビデオカメラの外観 (b) ビデオカメラの内部構造
図2 Tocciら[2]が開発したHDRビデオカメラ

2. 関連研究

HDR動画を生成するために、Tocciら[2]は図2に示すようなビデオカメラを開発した。このカメラはレンズを通過した光がカメラ内部に配置されたビームスプリッターによって3つに分けられるようになっている。分けられた光はそれぞれ露出の異なる撮像素子によって記録され、露出の異なる3つの画像を取得することができる。これらを合成することでHDR動画を生成している。また、Kangら[3]は、露出の高低を交互に切り替えながら動画を撮影し、それらの前後フレームを用いてHDR動画を生成している。この手法では、あるフレームと前後のフレームの3フレームを用いている。これらのフレームは全て時刻が異なる画像であるため、前後のフレームを中間のフレームに射影させながら画像間の対応関係を取得している。

これらの研究のように、特殊な内部機構を備えたカメラを用いればHDR動画を生成可能である。また、通常のRGB各8ビットで駆動するディスプレイよりもダイナミックレンジの広いディスプレイを用いることで、HDR動画の表示も可能である。しかし、そのような内部機構を備えたカメラやディスプレイは極めて特殊である。本研究では、映画撮影などで一般的に使用されるカメラを用いたHDR動画の生成を目指す。また、生成したHDR動画の表示に関しても一般的なディスプレイでの表示を目指す。

本研究で提案する手法では、特殊な内部機構を備えたカメラを使用する代わりに、一般的な同型カメラを2台用いて同じ光景を異なる露出で撮影する。ここで撮影には、3D映像の撮影で使用するビームスプリッターリグを利用する。ビームスプリッターリグは2台のカメラの光軸を合わせて同時に撮影することが可能である。しかし、2台の同型カメラを用いたとしても、カメラには個体差が存在し、画像センサの感度やレンズの歪み方などに違いがある。また、カメラの取り付け方によっては、それぞれのカメラが撮影した画像間で画像の位置ずれが発生する。よって、2台のカメラで露出以外が完全に同じ画像を取得するのは困難であるため、本研究ではHDR動画の生成に加えて、2台のカメラで取得した動画間における画質ずれや位置ずれの問題についても取り組む。さらに、生成したHDR動画を通常のディスプレイで表示するために、トーンマッピング処理も行う。

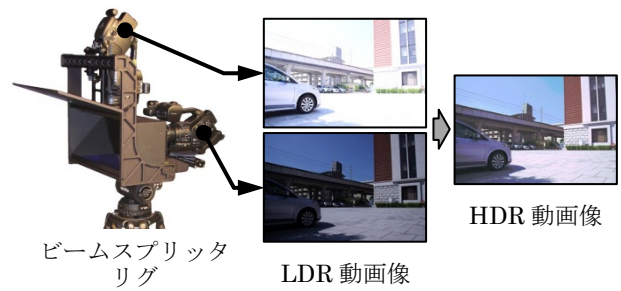


図3 本手法のイメージ

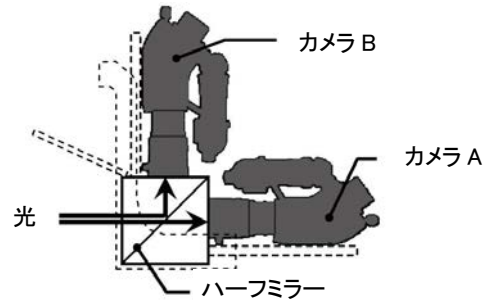


図4 ビームスプリッターリグのサイドビュー

3. 提案手法

3.1 概要

我々が提案する手法ではビームスプリッターリグを用いてHDR動画を生成する(図3)。この機材は図4に示すように2台のカメラとハーフミラーで構成されており、各カメラの設定を変更することで、同じシーンで露出の異なる2枚の画像を取得できる。本研究では、全てのフレームにおける2枚のLDR画像対からHDR動画を生成する。以下に本手法の大きな流れを説明する。

- (1) 2台のカメラの画角、ズーム値などの設定を全て同じにし、Genlockを同期する
- (2) 露出の異なる2台のカメラで撮影したLDR画像をそれぞれ連番画像として出力する
- (3) 画質ずれ、位置ずれの補正を行う
- (4) 各キーフレームのキー値から中間フレームの明るさを補間し、フレーム毎にトーンマッピングを施したHDR画像を生成する
- (5) 生成した連番画像を統合し、動画として出力する

3.2 HDR動画の生成

RGB各8ビットの情報量しか持たないLDR画像対から実際のシーンにおける輝度を推定し、その値を画素として持つ画像を生成する。以下の3.2.1から3.2.2項にて主な処理に関して詳しく述べる。

3.2.1 画質ずれと位置ずれの軽減

後述するHDR画像の生成には「露出時間 Δt をある値だけ変更したとき、画像上の画素値 Z がどのように変化するか」という関係を利用して実際の輝度値 E を算出している。つまり、複数枚のLDR画像における画素値の変化は露出のみの変更でなければならないが、2台のカメラで撮影した画像の間には画質ずれや位置ずれが生じてしまう。そのため、これらを解消する必要がある。

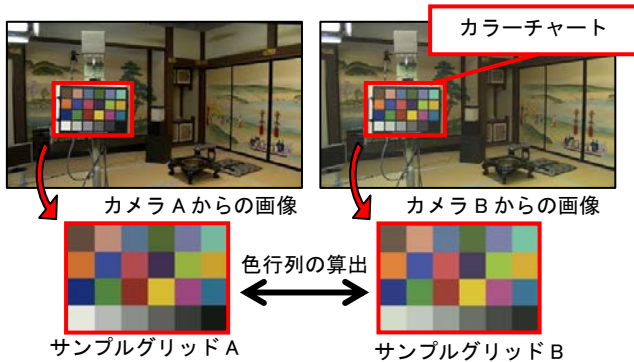


図5 カラーチャートを用いたキャリブレーション

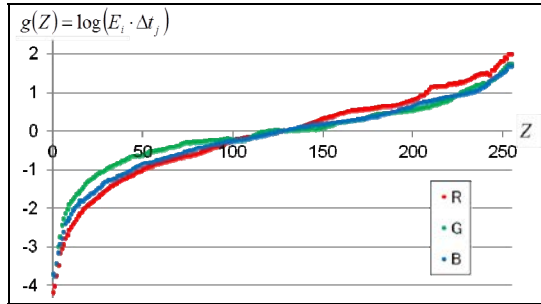


図6 レスポンスカーブ

画質ずれを軽減するために図5のようにカラーチャートを用いてキャリブレーションを行う。2つのカメラから取得した画像中のサンプルグリッドから色行列を算出し、片方のLDR画像にかけることで画質ずれの軽減を図る。この色行列とは式(3.1)のようにRGBで表現されている色を別の色に変換する 3×3 行列のことである。この行列を片方の画像に掛け合わせることで色相を統一させる。

$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (3.1)$$

位置ずれは使用しているカメラのレンズ歪みの相違とビームスプリッターにカメラを取り付ける際のわずかな位置や角度のずれによって生じる。そのため、レンズの歪みをステレオキャリブレーションによって補正しレンズの歪み係数や光学中心などのカメラ固有の内部パラメータを算出する。その後、画像のずれ量の計測を行い、位置ずれの解消を行う。ずれ量の計測にはLucasKanade法を用いることで、各画素の移動量である回転行列や並進行列を算出することが可能である。これらの行列を利用し、カメラBのLDR画像へ幾何学的な変換を行い、画素をカメラAのものと同じと一致させることで位置ずれを軽減する。

3.2.2 トーンマッピングの処理

実際の輝度値をLDR画像から算出することは困難であるため、RGB各チャンネルにおいて定義されているレスポンスカーブを用いる(図6)。レスポンスカーブの算出には露出時間におけるレンズを通過した光の総量、エクスポージャー値 $g(Z)$ を利用する。そして、レスポンスカーブと式(3.2)を用いて、実際の輝度 E をRGB各チャンネルについて算出し適当な明るさを持つHDR画像を生成する。

$$\ln E_i = \frac{\sum_{j=1}^P w(Z_{ij}) (g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^P w(Z_{ij})} \quad (3.2)$$

$$w(Z) = \begin{cases} Z - Z_{\min} & (Z \leq \frac{1}{2}(Z_{\min} + Z_{\max})) \\ Z_{\max} - Z & (Z > \frac{1}{2}(Z_{\min} + Z_{\max})) \end{cases} \quad (3.3)$$

ここで、 P は使用するLDR画像の枚数、 Z はLDR画像における画素値、 Δt は露出時間、 $g(Z)$ はエクスポージャー値、 $w(Z)$ は重みづけ関数をそれぞれ表している。 $g(Z)$ はレスポンスカーブにより求められ、 $w(Z)$ は式(3.3)により求められる。式(3.2)の重み $w(Z)$ は、計測された画素値の中間値に相当する部分で最も大きくなり、最大・最小となる部分で最も小さくなる。本手法では2枚のLDR画像を用いるため $P=2$ とする。

その後、通常のディスプレイで表示するために各チャンネルが8ビットに収まるように変換する。この処理をトーンマッピング[4]と呼ぶ(式(3.4-3.6))。

$$\bar{L}_w = \frac{1}{N} \exp \left(\sum_{x,y} \log(\varepsilon + L_w(x,y)) \right) \quad (3.4)$$

$$L(x,y) = \frac{a}{L_w} L_w(x,y) \quad (3.5)$$

$$L_d(x,y) = \frac{L(x,y) \left(1 + \frac{L(x,y)}{L_{white}^2} \right)}{1 + L(x,y)} \quad (3.6)$$

$L_w(x,y)$ は各画素における輝度を表し、 ε はlogが発散しないための定数である。平均輝度 \bar{L}_w は、各画素の輝度に対するスケージングに用いられる。 a はキー値と呼ばれ、明るさの調整で用いられる。 L_{white} は、最も輝度が高い部分の輝度を表している。

式(3.4)はシーン内の基準となる明るさを決定し、式(3.5)でトーンマッピング後の画像の明るさを決定する。

トーンマッピングの最も単純な方法は平均輝度を中心に線形に変換することであるが、HDR画像は非常に幅広い明るさの情報を持っているため、線形的な調整では暗い部分が黒くつぶれてしまう。式(3.6)は非線形の変換関数であり、明るい部分の情報を間引いて暗い部分の明るさを持ち上げることで、より人間の視覚に近い画像を生成できる。

このトーンマッピングの処理をカメラから取得した連番画像に適用することでHDR動画画像を生成する。このとき、明るさを調整するキー値の設定は非常に困難であるため、GUIを用いて手動で調整している。ただし、1フレームずつ行っているのは効率が悪いので、キー値を変更する代表となるフレーム(キーフレーム)を選択し、そのキーフレームのキー値を決定する。その後、その他のフレームのキー値をキーフレームのキー値を基にスプライン補間によって算出し、一意に決定する。



カメラ A の撮影画像 カメラ B の撮影画像

図 7 入力画像



図 8 画質ずれの補正結果 (カメラ B)



図 9 カメラ A,B のキャリブレーション結果

4. 実験と考察

4.1 実験内容

3章で述べた方法で実際に HDR 動画の生成を行う。

本実験では、ダイナミックレンジが広がっていることを確認しやすいように、明るい部分と暗い部分の輝度の差が激しいシーンで撮影を行う。図 7 に本実験の入力画像を示す。ビームスプリッタリグに取り付ける 2 台のカメラは SONY PMW-EX3 を使用する。

本実験では、使用する 2 台のカメラを「カメラ A」「カメラ B」とし、それぞれのカメラの設定を表 1 に示す。

表 1 撮影時の各カメラの設定

	F 値	ゲイン	シャッター スピード
カメラ A	1.9 F	0 dB	1 / 250 秒
カメラ B	1.9 F	0 dB	1 / 60 秒

4.2 実験結果と考察

図 7 の入力画像に対して画質ずれと位置ずれを軽減した結果を図 8, 9, 10 に示す。

画質ずれの軽減では、片方の入力画像の画質に近づければ良いため、算出した色行列を用いてカメラ B の撮影画像を補正した。位置ずれに関しては図 9 のカメラ A, B の画像を重ねたものを図 10 (a) に示す。(b) は (a) の重ね合わせる画像を補正したものである。(a) と (b) を比較すると補正後の方ではブレが軽減されていることが分かる。画質ずれと位置ずれを軽減した後、ラディアンスマップを生成し、トーンマッピングにて HDR 画像を生成した (図 11)。処理の際のキー値は GUI のスライダーを用いて手動で設定している。キー値は 0 から 1 までの間を小数点第三位刻みで変更できるようになっており、また、同様に任意のフレームを選択可能である。

図 11 (a) は適正露出で撮影した LDR 画像、(b) は図 7



(a) 位置ずれの補正前 (b) 位置ずれの補正後

図 10 位置ずれの補正結果 (画像の一部)



(a) 比較画像 (LDR 画像) (b) 結果画像 (HDR 画像)

図 11 適正露出で撮影した LDR 画像と結果画像の比較

を入力画像としたときの HDR 画像である。図 11 の (a), (b) を比較すると、(a) では畳などでは白飛びが生じているが、(b) では軽減できている。このことから 1 台のカメラから取得した際の画像よりもダイナミックレンジの広い画像が生成できたことがわかる。

5. むすび

本稿では、2 台の同型カメラとハーフミラーから構成されるビームスプリッタリグを用いて HDR 動画を生成する方法について述べた。動画の HDR 化は、各フレームの画像を HDR 画像化させることであり、露出の異なるシーンの複数枚の LDR 画像から生成することが可能である。しかし、そのためには、複数枚の LDR 画像の画質ずれや位置ずれを軽減する必要があり、それらを合成することでより正確なラディアンスマップを算出できる。その後、通常のディスプレイに画像を表示させるために生成した画像に対してトーンマッピングを施した。

本手法による HDR 動画の生成結果から、同じタイプのカメラ 1 台で撮影した場合よりもダイナミックレンジが広がっていることが確認できた。

参考文献

- [1] P. E. Debevec and J. Malik : "Recovering high dynamic range radiance maps from photograph," Proc. SIGGRAPH 1997, pp. 369 - 378, 1997.
- [2] M. D. Tocci, C. Kiser, N. Tocci, and P. Sen : "A versatile HDR video production system," Proc. SIGGRAPH 2011, Vol. 30, 2011.
- [3] S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski : "High dynamic range video," Proc. SIGGRAPH 2003, Vol. 22, pp. 319 - 325, 2003.
- [4] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda : "Photographic tone reproduction for digital images," Proc. SIGGRAPH 2002, Vol. 21, pp. 267 - 276, 2002.