バイラテラルフィルタによるコントラスト強調画像融合

Contrast Enhancing Image Fusion Using Bilateral Filters

学生会員 吉 武 謙[†],浦 浜 喜 一[†]Yuzuru Yoshitake[†] and Kiichi Urahama[†]

Abstract We propose a method for fusing multiple images into a single one. Our method is an extension of a technique for enhancing the contrast of a single image to multiple ones. First, we calculate the logarithm of pixel values in component images and smooth them using a bilateral filter. The smoothed image is the basic component and the difference between the original image and its basic component is the detailed component. We then compress the basic component and add it to the detailed component. Finally we recover the pixel values with an exponential function to obtain contrast-enhanced images. When adding detailed components together, we weigh both components according to their magnitude to preserve salient information in both component images. We tested this method with flash shot, multiple exposure, multiple focus, multiple light and multimodal medical images.

キーワード：画像融合，コントラスト強調，バイラテラルフィルタ，softmax 融合

1. ま え が き

多重画像を1枚の画像に融合する例として，マルチバンド衛星写真の融合や，CTとPETなどの医用画像の融合などがある．その際，鮮明に写っている物体が成分画像ごとに異なることが多いので，各物体の有用な情報が失われないように1枚の画像のなかに重ねる必要がある．融合するに先立って各成分画像の位置合わせ (registration) を行う必要があるが，本論文では位置合わせ終了後に行う融合処理法を提案するので，以下の実験ではすべて位置合わせ済みの画像を扱う．

画像融合法としては，単純に平均画像を表示する場合も多いが，主成分分析¹⁾や射影追跡法²⁾³⁾，ニューラルネット⁴⁾，ウェーブレット⁵⁾など，古くから多くの手法が提案されている．Grundland⁶⁾は，クロスディゾルブのような異なるシーンの重ね合わせ法として，フィルタ処理などを用いずに画素毎の非線形平均による簡便な融合法を提案している．

画像融合の評価量としては相互情報量や相互相関などが用いられる⁷⁾が，本論文では局所コントラストに基づく方法を考える．局所コントラストに基づく画像融合法も数多く提案されているが，非線形最適化による方法³⁾⁸⁾は反復解法の計算量が多い．

本論文では，最適化問題の反復解法によらず，単一画像のコントラスト強調処理を多重画像に拡張した画像融合アルゴリズムを提案する．単一画像のコントラスト強調法も種々あるが，ここではRetinex⁹⁾等の視覚順応を模倣した手法を採用し，エッジ付近でハローが生じないようにバイラテラルフィルタによる平滑化を用いる．バイラテラルフィルタによる単一画像のコントラスト強調では，入力画像を平滑化した基調成分を圧縮して，入力画像から基調成分を引いた詳細成分を相対的に伸長する．本論文では，この手法を多重画像に拡張し，詳細成分をコントラストに応じて加重して，多重画像を単一画像に融合する．

Fattalら¹⁰⁾もバイラテラルフィルタを用いた画像融合法を提案しているが，多重スケール平滑化を要し，多方向照明画像に特化しているので融合処理も複雑である．本論文で提案する手法は1回の平滑化だけを用い，対象画像を限定しない，汎用的で基礎的な手法である．

なお，以下の実験の計算環境はPentium(R) 4 CPU 2.8GHz, 1GB RAM, C言語である．

2. 単一画像のコントラスト強調

提案法を述べる前に，その基礎となる単一画像のコントラスト強調法を概説しておく．

視覚順応を模したコントラスト強調法として，モードフィルタによる多重スケール平滑化によるコントラスト強調法が提案されている¹¹⁾が，ここでは計算量を減らすためにモードフィルタを簡略化したバイラテラルフィルタを用い，また単一スケールの平滑化だけにする．

2007年12月20日受付，2008年3月21日再受付，2008年4月5日採録
[†]九州大学 大学院芸術工学研究院 視覚情報部門

(〒815-8540 福岡市南区塩原 4-9-1, TEL 092-553-4512)

[†] Faculty of Design, Kyushu University

(4-9-1, Shiobaru, Minami-ku, Fukuoka-shi, 815-8540 Japan)

以下、モノクロ画像で説明する。カラー画像の場合は、RGB を YUV に変換して明度 Y だけに以下の処理を施し、UV はそのまま RGB に戻す。

画素 (i, j) の値を d_{ij} とする。 $d_{ij} \in [0, 255]$ である。まず画素値の対数をとって $f_{ij} = \ln(d_{ij} + \epsilon)$ とする (ϵ は $\ln 0$ を避けるための微小な正定数であり、以下の実験では $\epsilon = 1$ とした)。この対数変換は視覚の感度特性に対応している。次に、この f_{ij} をバイラテラルフィルタで平滑化する。すなわち

$$g_{ij} = \frac{\sum_{k=-p}^p \sum_{l=-p}^p e^{-\alpha(k^2+l^2)-\beta(f_{ij}-f_{i+k,j+l})^2} f_{i+k,j+l}}{\sum_{k=-p}^p \sum_{l=-p}^p e^{-\alpha(k^2+l^2)-\beta(f_{ij}-f_{i+k,j+l})^2}} \quad (1)$$

を画素毎に計算する。

f_{ij} は概略変化を表す基調成分 g_{ij} と詳細成分 $s_{ij} = f_{ij} - g_{ij}$ に分解される。コントラストを強調するには、詳細成分を相対的に拡大すればよいから、基調成分に $0 < \delta < 1$ をかけて減衰させ、 $h_{ij} = \delta g_{ij} + s_{ij}$ としてから $d'_{ij} = e^{h_{ij}}$ で対数を元に戻し、値の範囲を $[0, 255]$ に規格化して出力する。すなわち、 d'_{ij} の最大値を d_{max} 、最小値を d_{min} とすると、 $d''_{ij} = 255(d'_{ij} - d_{min}) / (d_{max} - d_{min})$ を画素値とする画像を出力する。

なお、この基調成分の減衰処理はアンシャープマスキング (USM) と等価である。なぜなら USM では $h'_{ij} = f_{ij} + \delta' s_{ij}$ とするが、これは $\delta' = 1/\delta - 1$ ならば $h_{ij} = \delta g_{ij} + s_{ij}$ と一致するからである (ただし、両方法とも値を $[0, 255]$ に正規化するものとする)。よって、アンシャープマスキングと本方法との違いは対数変換があるかないかである。

以上の方法を図 1 に示す 2 枚のカラー写真 (サイズ 311×218) にそれぞれ別々に適用した結果を図 2 に示す。図 1(a) はフラッシュをたいた写真であり、フラッシュが後ろの絵に反射して白くなっている。図 1(b) はフラッシュなしの通常撮影であるが、顔が暗くなっている。図 2(a) は図 1(a) のコントラスト強調結果で、当然ながら反射の白いところは改善されていない。図 2(b) は図 1(b) の強調結果で、顔が改善されているが、まだ充分ではない。 $p = 5, \alpha = 0.001, \beta = 0.001, \delta = 0.5$ (以下の実験でもすべて同じ値) とした。

3. 多重画像のコントラスト強調融合

以上の単一画像のコントラスト強調法を多重画像の融合に応用する。画像が n 枚あるとし、第 k 枚めの画像の画素 (i, j) の値を d_{ijk} とする。以下、前節と同じくモノクロ画像で説明する。カラー画像の場合は、画素値 $R_{ijk}, G_{ijk}, B_{ijk}$ を $Y_{ijk}, U_{ijk}, V_{ijk}$ に変換し、明度 Y_{ijk} を d_{ijk} として以下の処理を適用して融合画像の明度 Y'_{ij} を求め、U と V は平均

$$\begin{aligned} U'_{ij} &= \sum_{k=1}^n U_{ijk} / n, \\ V'_{ij} &= \sum_{k=1}^n V_{ijk} / n \end{aligned} \quad (2)$$

として、 $Y'_{ij}, U'_{ij}, V'_{ij}$ を RGB に戻し、RGB それぞれを $[0, 255]$ に規格化して融合画像の画素値とする。

3.1 融合の計算手順

まず、画素値を対数変換し、 $f_{ijk} = \ln(d_{ijk} + \epsilon)$ とする。次に、 f_{ijk} を画像 k ごとに式 (1) のバイラテラルフィルタで平滑化し、それぞれの基調成分 g_{ijk} を計算する。

このようにして各画像を基調成分 g_{ijk} と詳細成分 $s_{ijk} = f_{ijk} - g_{ijk}$ に分解してから、それらを統合する。まず基調成分は単純に全画像の平均値にする。すなわち

$$g_{ij} = \sum_{k=1}^n g_{ijk} / n \quad (3)$$

を融合画像の基調成分とする。次に、融合画像の詳細成分

図 1 フラッシュ撮影 (a) と通常撮影 (b)
Photographs with (a) and without (b) flash.

図 2 コントラスト強調結果
Result of contrast enhancement.

は各詳細成分の加重和

$$s_{ij} = \sum_{k=1}^n w_{ijk} s_{ijk} / \sum_{k=1}^n w_{ijk} \quad (4)$$

とする。重み w_{ijk} については後述する。引き続き、単一画像のときと同じく $h_{ij} = \delta g_{ij} + s_{ij}$ を計算して、 $d'_{ij} = e^{h_{ij}}$ を $[0, 255]$ に規格化した値を画素値とする画像を融合画像として出力する。

3.2 重み w_{ijk} の設定

式 (4) の重み w_{ijk} の設定法として、以下の3通りを考える。

1) 平均

すべての w_{ijk} を 1 にすると、式 (4) は式 (3) と同じになり、詳細成分も平均化される。このような、対数をとった後の相加平均は、例えば、 $e^{(f_1+f_2)/2} = e^{(\ln d_1 + \ln d_2)/2} = \sqrt{d_1 d_2}$ となることから、元の画素値の相乗平均になる。今の場合にはフィルタによる成分分離があるので単純な相乗平均とはならないが、基調成分と詳細成分の相加平均は元画像の非線形な融合になる。単純な相加平均による融合では分散が増えるのでコントラストが低下する⁹⁾が、このような非線形融合法ではコントラストが向上する可能性がある。

図1の2枚の画像の平均による融合結果を図3(計算時間は1.46秒)に示す。図1(a)のフラッシュ反射もなくなり、顔も改善されており、図1(a)と(b)を個別にコントラスト強調した図2の2枚よりも良好な結果が得られている。

2) 最大値選択

単純に平均するよりも、絶対値が最大の詳細成分だけを選択したほうがコントラストが高くなると思われる。すなわち、各画素において

$$w_{ijk} = \begin{cases} 1, & \text{if } k = \arg \max_l \{|s_{ijl}|\} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

とすると、各画素において最も変化が激しい成分画像の詳細成分が選ばれる。この方法の統合詳細成分 s_{ij} は、上記の平均の詳細成分よりも大きくなり、コントラストが高くなる。

図1の2枚の画像の最大値選択による融合結果を図4(計算時間1.51秒)に示す。図3よりも更に、全体的にコントラストが強調されている。しかし、図4の1部を図5に拡大して示すように、絵の額縁のところにノイズ状の極端な

濃淡変化が現れたり、額と髪の境界などに擬似輪郭が現れたりして、不連続な濃淡変化の多い粗い画像になっている。

これは、一般に式 (4) の w_{ijk} が画素ごとに異なる値をとることが原因である。図1の画像での w_{ijk} の値を図6に示す。図6の白画素は $w_{ij1} = 1, w_{ij2} = 0$ の画素であり、黒画素は逆に $w_{ij1} = 0, w_{ij2} = 1$ の画素である。このように w_{ijk} は画素ごとに乱雑に変化しており、これが図4のノイズの原因である。

3) softmax による融合

上記二つの平均と最大値選択は両極端な設定法であり、これらを統一する一般的な設定法は

$$w_{ijk} = e^{\lambda |s_{ijk}|} / \sum_{l=1}^n e^{\lambda |s_{ijl}|} \quad (6)$$

である。式 (6) は softmax と呼ばれ¹²⁾、 λ が小さい極限では、 $\lim_{\lambda \rightarrow 0} w_{ijk} = 1/n$ となって平均による融合になり、逆に λ が大きい極限では $\lim_{\lambda \rightarrow \infty} w_{ijk}$ は式 (5) になって最大値選択による融合となる。したがって、 λ を適度な値にすれば両者の中間的な融合になり、それぞれの欠点が緩和された融合ができると期待される。以下の実験では $\lambda = 100$ とした。

なお、式 (6) の w_{ijk} は $\sum_{k=1}^n w_{ijk} = 1$ を満たすので、式 (4) の分母は不要になる。

図4 最大値選択による融合結果
Fusion by maximum selection.

図5 図4の1部分拡大
Partial magnification of Fig.4.

図6 図1の画像での w_{ij1} の値
Values of w_{ij1} in Fig.1.

図3 平均による融合結果
Fusion by averaging.

3.3 重みの平滑化

しかし、式 (6) でも w_{ijk} は画素ごとに計算されるので、図 6 のように乱雑な空間変化をすると思われる。そこで式 (6) の w_{ijk} を平滑化することにし、これにもバイラテラルフィルタを使う。すなわち

$$w'_{ijk} = \frac{\sum_{a=-p}^p \sum_{b=-p}^p e^{-\alpha(a^2+b^2)-\beta(w_{ijk}-w_{i+a,j+b,k})^2} w_{i+a,j+b,k}}{\sum_{a=-p}^p \sum_{b=-p}^p e^{-\alpha(a^2+b^2)-\beta(w_{ijk}-w_{i+a,j+b,k})^2}} \quad (7)$$

とし、この w'_{ijk} を式 (4) の w_{ijk} に代入する。以下の実験では $\alpha = 0.001, \beta = 0.01$ とした。

図 1 の画像での w'_{ij1} の値を図 7 に示す。図 6 が均され重みが連続的に変化している。図 3 の平均による融合では重みは全画素で 0.5 であるから、図 7 は平均による融合の重みと図 6 の中間的な値になっている。

以上の手順をまとめると以下のようになる：

- 1) 各画像 d_{ijk} の対数をとる： $f_{ijk} = \ln(d_{ijk} + \epsilon)$
- 2) f_{ijk} にバイラテラルフィルタをかけて g_{ijk} を求める。
- 4) $g_{ij} = \sum_{k=1}^n g_{ijk} / n$ を計算する。
- 5) $s_{ijk} = f_{ijk} - g_{ijk}$ を式 (6) に代入して w_{ijk} を求める。
- 6) その w_{ijk} を式 (7) に代入して w'_{ijk} を求める。
- 7) その w'_{ijk} を式 (4) の w_{ijk} に代入して s_{ij} を求める。
- 8) $h_{ij} = \delta g_{ij} + s_{ij}$ を計算し、 $e^{h_{ij}}$ を $[0,255]$ に規格化した値を画素値とする画像を出力する。

なお、カラー画像の場合は本節の冒頭に記したように、明度成分 Y に以上の処理を施し、U と V は平均して、YUV を RGB に戻し、RGB それぞれを $[0,255]$ に規格化する。

図 1 の画像の融合結果を図 8 (計算時間 2.87 秒) に示す。平均による融合結果 (図 3) と最大値選択による融合結果 (図 4) の中間的な画像になっている。図 8 の 1 部分 (図 5 と同じ部位) の拡大を図 9 に示す。図 5 に見られたノイズや擬似輪郭がなくなっている。

ただし、平均や最大値選択による融合にかかる時間がほぼバイラテラルフィルタの時間 × 成分画像の枚数であるのに対し、本方法では各成分の重み画像にもフィルタをかけるので倍近くの時間がかかる。また、計算時間は画像サイズと枚数にも比例する。

図 7 図 1 の画像での w'_{ij1} の値
Values of w'_{ij1} in Fig.1.

図 8 Softmax による融合結果
Fusion with softmax.

図 9 図 8 の 1 部分拡大
Partial magnification of Fig.8.

4. 他 の 実 験 例

本提案法は撮影条件に特化しない汎用的な画像融合法であり、いろいろな多重画像に適用できる。以下では、多重合焦画像、多方向照明画像、多重露光画像、多モード医用画像などでの実験結果を示す。

4.1 多重合焦画像

図 10¹³⁾ は焦点距離の異なる 2 枚のカラー写真 (サイズ 415×315) である。左の画像では手前左の女性に合焦しており、右後方はぼけている。逆に右の画像では右奥の男性が合焦しており、左手前はぼけている。図 11 はこれら 2 枚を融合した結果 (計算時間 5.54 秒) であり、両端の距離に合焦した画像が得られている。

図 10 焦点距離の異なる 2 枚の写真
Two photographs with different focal length.

図 11 融合結果
Result of fusion.

4.2 多方向照明画像

図 12(a),(b),(c)¹⁴⁾ は照明の向きを変えて撮影した 3 枚のカラー写真 (サイズ 154 × 300) である。陰影によって葉の表面の凹凸が見やすくなっているところもあるが、影になって葉脈が見えなくなっているところもある。図 12(d) はこれら 3 枚を融合した結果 (計算時間 2.91 秒) であり、影で暗くなっていた部分でも葉脈が見えるようになっている。

4.3 多重露光画像

図 13 は、露出を変えて撮影した 2 枚のモノクロ写真 (サイズ 586 × 393)¹⁵⁾ である。上の画像では中央の人物群は見えるが、周辺の室内壁は黒くつぶれてしまっている。逆に下の画像では室内壁のレンガが見えるようになっているが、中央の室外景色が白くとんでいいる。図 14 左は融合結果 (計算時間 9.75 秒) であり、室内、室外ともに細部が見えるようになっている。なお、図 14 右は同左の画像の右下

図 12 照明の向きの異なる 3 枚の写真 (a),(b),(c) の融合結果 (d)

Fusion (d) of three photographs with different lighting (a),(b),(c).

図 13 露光の異なる 2 枚の写真
Two photographs with different exposure.

図 14 融合結果 (左), 1 部拡大 (右)
Result of fusion (left), partial magnification (right).

部分であり、後で図 21 と比較するために拡大している。

このような多重露光撮影は HDR(high dynamic range) 画像を作るのに利用される。8 段階の多重露光撮影のなかの図 15 の 4 枚のカラー写真 (サイズ 222 × 142)¹⁶⁾ を融合した結果 (計算時間 2.65 秒) を図 16 に示す。

4.4 多モード医用画像

医療撮影ではマルチモーダルな写真が複数撮られることが多い。図 17 は、撮影条件の異なる 2 枚の MR 画像 (サイズ 256 × 256)²⁾ である。左の画像では脳の右下にある腫瘍が見えない。右の画像では腫瘍は明瞭であるが左のほうにある眼が判別できない。図 18 はこれら 2 枚の融合結果 (計算時間 2.77 秒) であり、腫瘍も眼も見やすくなっている。

図 15 露光の異なる 4 枚の写真
Four photographs with different exposure.

図 16 融合結果
Result of fusion.

図 17 脳の MR 画像
MR images of brain.

図 18 融合結果
Result of fusion.

図 19 非線形射影追跡による融合
Fusion with nonlinear projection pursuit.

図 20 CT 画像 (a) と PET 画像 (b) の融合結果 (c)
Fusion (c) of CT image (a) and PET image (b).

図 21 Retinex による融合結果 (左), 1 部拡大 (右)
Fusion with Retinex (left), partial magnification (right).

る。ちなみに図 19 は、非線形射影追跡による融合結果³⁾であり、本提案法は、計算量が多い非線形射影追跡よりも簡単な処理で良好な融合画像が得られている。

図 20(a) は CT 画像 (サイズ 110×300) で、同 (b) は PET 画像である。図 20(c) はこれらの融合結果 (計算時間 1.40 秒) である。

以上のように、本提案法は種々の多重画像の融合に適用できる。このような汎用性が本提案法の長所である。本提案法は、バイラテラルフィルタの使用が特徴であり、線形フィルタではエッジ付近にハローが生じる。Jobson ら⁹⁾の手法は単一画像のコントラスト強調法であるが、多重平滑化によりハローを軽減している。しかしパラメータ調整が難しく、文献 9) のいくつかの結果にはハローが生じている。この Retinex 法を応用して図 13 の 2 枚の画像を融合した結果を図 21 左に示す。同右は左の画像の右下部分 (図 14 右と同じ部位) の拡大である。暗い影に接した石の縁にハローが生じている。一方、提案法による結果 (図 14 右) にはそのようなハローは見られない。

このように、細かなパラメータ調整を要しないので使い易く、また人工的なハローも生じない安定性が本提案法の長所である。

5. む す び

単一画像のコントラスト強調法を多重画像に拡張した画像融合アルゴリズムを提案し、いろいろな多重画像で性能を検証した。本方法による融合画像のエッジ付近にハローが生じていないのは、バイラテラルフィルタを使っているからであり、Retinex 等の線形フィルタを用いる融合法では人工的なハローが生じる。

本提案法は汎用的な手法であるが、それぞれの多重画像の特徴に適合させて、多重画像ごとの改変を付け加えていくのが今後の課題である。

〔文 献〕

- 1) U. Schmiedl: "The utility of principal component analysis for the image display of brain lesions: a preliminary comparative study", Mag. Reso. Med., 4, pp.471-476 (1987)
- 2) G. Harikumar and Y. Bresler: "Feature extraction techniques for exploratory visualization of vector-valued imagery", IEEE Trans. Image Proc., 5, 9, pp.1324-1334 (1996)
- 3) 小野直樹, 浦浜喜一: "射影追跡による多重画像のモノクロ表示", 映像学誌, 51, 4, pp.592-597 (1997)
- 4) A. Manduka: "Multispectral image visualization with nonlinear projections", IEEE Trans. Image Proc., 5, 10, pp.1486-1490 (1996)
- 5) P. Hill, C. N. Canagarajah and D. R. Bull: "Image fusion using complex wavelets", Proc. BMVC (2002)
- 6) M. Grundland, R. Vohra, G. P. Williams and N. A. Dodgson: "Nonlinear multiresolution image blending", Proc. ICCVG (2006)
- 7) Q. Wang and Y. Shen: "Performance assessment of image fusion", Proc. PSIVT, pp.373-382 (2006)
- 8) P. Scheunders: "Multispectral image fusion using local mapping techniques", Proc. ICPR, pp.311-314 (2000)
- 9) D. J. Jobson, Z. Rahman and G. A. Woodell: "A multiscale retinex for bridging the gap between color images and the human observations of scenes", IEEE Trans. Image Proc., 6, 7, pp.965-976 (1997)
- 10) R. Fattal, M. Agrawala and S. Rusinkiewicz: "Multiscale shape and detail enhancement from multi-light image collections", Proc. SIGGRAPH, 51 (2007)
- 11) 井上光平, 浦浜喜一: "多重スケール非線形平滑化によるコントラスト強調", 信学論, J83-A, 6, pp.831-833 (2000)
- 12) A. Tuerk and S. J. Young: "Polynomial softmax functions for pattern classification", Tech. Report, CUED/F-INFENG/TR 402 (2001)
- 13) <http://www.pixeltopixel.com/ToPutSelectiveItemInFocus.html>
- 14) <http://vis.berkeley.edu/papers/mlc/supplemental/>
- 15) <http://www.eyetap.org/wearcomp/wyckoff/wyckoff.html>
- 16) <http://research.edm.uhasselt.be/tmtenens/>

よし たけ けん 2007 年、北九州市立大学国際環境工学部情報メディア工学科卒業。現在、九州大学大学院芸術工学府博士前期課程在学中。画像処理、コンピュータビジョンに関する研究に従事。学生会員。

う ら は ま き い ち 1980 年、九州大学大学院工学研究科博士後期課程修了。同年、同大学助手。1989 年、九州工業大学助教授。現在、九州大学大学院芸術工学研究院教授。パターン認識、画像情報処理の研究に従事。