

ロバストバイラテラルフィルタによる混合雑音除去

井上 光平[†] 原 健二[†] 浦浜 喜一[†]

[†]九州大学大学院芸術工学研究院コミュニケーションデザイン科学部門

〒 815-8540 福岡市南区塩原 4-9-1

E-mail: †{k-inoue,hara,urahama}@design.kyushu-u.ac.jp

あらまし エッジ保存平滑化フィルタの 1 種であるバイラテラルフィルタには、インパルス雑音を除去できないという難点がある。本論文では、バイラテラルフィルタをロバスト化することによって、インパルス雑音も除去可能なエッジ保存平滑化フィルタを提案し、3 種の画質評価指標に基づく比較実験で提案手法の有効性を確認する。

キーワード 雑音除去, インパルス雑音, バイラテラルフィルタ, ロバスト化

1. はじめに

エッジ保存平滑化フィルタの 1 種であるバイラテラルフィルタ [1] (BF と略記する) は、ダイナミックレンジ圧縮 [2], フラッシュ写真強調 [3], メッシュ雑音除去 [4] など、コンピュータビジョンやコンピュータグラフィックスの種々の問題に適用され、その有効性が示されている。また、BF の高速化手法も多数提案されている [2], [5] ~ [16]。BF の拡張例として、Petschnigg ら [17] は、一つのシーンをフラッシュありとフラッシュなしで撮影した 2 枚の画像を入力とし、BF における加重平均の重みをフラッシュありの画像から計算し、その重みを使って雑音を多く含むフラッシュなしの画像を平滑化するジョイント BF (JBF と略記する) を提案した。Eisemann ら [3] も、同時期に同様のフィルタを提案し、それをクロス BF (CBF と略記する) と呼んだ。Kopf ら [18] は、JBF を画像のアップサンプリングに応用した。田中ら [19] は、BF, JBF, NL-means フィルタ [20] を共通の枠組みから導出し、それらフィルタの拡張の可能性を示した。ここでは、ノイズモデルの一般化も行われているが、インパルス雑音については、明示的には言及されていない。しかし、BF では、エッジ保存平滑化とインパルス雑音除去を両立させることは難しく、例えば、ガウス雑音とインパルス雑音が混合した雑音を除去することは困難である。JBF あるいは CBF では、加重平均の重みの計算に雑音を含まない画像を用いることができれば、インパルス雑音もある程度は平滑化できるが、加重平均の計算にインパルス雑音も含まれるため、出力画像にインパルス雑音の影響が残る。また、JBF あるいは CBF では、同一シーンを撮影した 2 枚の画像を必要とするため、BF に比べて適用範囲が限定されるという難点もある。そこで、常ら [21] は、1 枚の画像だけを用いて BF の雑音除去性能を向上させる自己クロス BF を提案し、更にそれをロバスト化することによって、ガウス・インパルス混合雑音の除去性能が向上することを示した。

本論文では、BF をロバスト化したロバスト BF (RBF と略記する) を提案し、その雑音除去性能を、ピーク信号対雑音比などの 3 種の定量的画質評価指標に基づく比較実験で検証する。

以下、2. 章で BF を概説し、3. 章で RBF を提案し、4. 章で実験結果を示し、5. 章でまとめを述べる。

2. バイラテラルフィルタ

n 個の画素からなる画像 f が与えられるとし、第 $i \in \{1, \dots, n\}$ 画素の座標を x_i , 画素値を f_i とすると、バイラテラルフィルタ (BF と略記する) の出力は

$$f_i^B = \frac{\sum_{j \in N_i} u_{ij}^\alpha v_{ij}^\beta f_j}{\sum_{j \in N_i} u_{ij}^\alpha v_{ij}^\beta} \quad (1)$$

と表される。ここで N_i は第 i 画素の近傍画素の集合であり、 u_{ij} , v_{ij} はそれぞれ空間、画素値領域の重みであり、

$$u_{ij} = \exp(-\|x_i - x_j\|^2) \quad (2)$$

$$v_{ij} = \exp(-\|f_i - f_j\|^2) \quad (3)$$

で与えられる。式 (1) の α と β は正定数である。BF とガウスフィルタ (GF と略記する) の違いは重み v_{ij} の有無であり、この v_{ij} がエッジ保存平滑化を可能にしている。しかし、画像にインパルス雑音が含まれる場合には、それも保存されるという欠点がある。例えば、第 i 画素にインパルス雑音があるとすると、 $v_{ii} = 1$, $v_{ij} \approx 0$ ($j \neq i$) より $f_i^B \approx f_i$ となり、インパルス雑音がそのまま出力される。

以上のように、BF では、エッジ保存平滑化とインパルス雑音除去を両立させることが困難である。そこで次章では、BF をロバスト化することによって、インパルス雑音の除去性能を向上させる方法を提案する。

3. ロバストバイラテラルフィルタ

BF をロバスト化するために，各画素 j に対する重み

$$w_j = \sum_{k \in N_j, k \neq j} u_{jk}^{\tilde{\alpha}} v_{jk}^{\tilde{\beta}} \quad (4)$$

$$= \sum_{k \in N_j} u_{jk}^{\tilde{\alpha}} v_{jk}^{\tilde{\beta}} - 1 \quad (5)$$

を考える．ここで $\tilde{\alpha}$ と $\tilde{\beta}$ は正定数である．更に， v_{ij} の代わりに

$$\tilde{v}_{ij} = \max \left\{ v_{ij}, \epsilon^{\frac{1}{\tilde{\beta}}} \right\} \quad (6)$$

を用いて (ϵ は微小な正定数である)

$$f_i^R = \frac{\sum_{j \in N_i} u_{ij}^{\alpha} \tilde{v}_{ij}^{\beta} w_j f_j}{\sum_{j \in N_i} u_{ij}^{\alpha} \tilde{v}_{ij}^{\beta} w_j} \quad (7)$$

とし，これをロバスト BF (RBF と略記する) と呼ぶことにする．

まず， w_j の効果を考える．第 j 画素にインパルス雑音があるとすると， N_j に $f_k \approx f_j$ となるような $k \neq j$ が含まれなければ， $u_{jk}^{\tilde{\alpha}} v_{jk}^{\tilde{\beta}}$ ($k \neq j$) ≈ 0 より $w_j \approx 0$ となる．従って， w_j を導入することによって，式 (7) 中の $j \in N_i$ に関する和からインパルス雑音の項を取り除くことができるかと期待される．

次に， \tilde{v}_{ij} の効果を考える．注目画素 i にインパルス雑音があるとすると， $v_{ij}^{\beta} \approx 0 < \epsilon (j \neq i)$ より $\tilde{v}_{ij}^{\beta} = \epsilon$ となり，式 (7) は

$$f_i^R = \frac{\sum_{j \in N_i} u_{ij}^{\alpha} \epsilon w_j f_j}{\sum_{j \in N_i} u_{ij}^{\alpha} \epsilon w_j} \quad (8)$$

$$= \frac{\sum_{j \in N_i} u_{ij}^{\alpha} w_j f_j}{\sum_{j \in N_i} u_{ij}^{\alpha} w_j} \quad (9)$$

となる．式 (9) は w_j を重みとする重み付き GF である．

以上をまとめると，RBF は，注目画素 i にインパルス雑音が含まれれば重み付き GF になり，そうでなければ重み付き BF になるようなフィルタであるといえる．また，重み w_j の効果によって，インパルス雑音以外のデータを用いた平滑化が可能になる．

4. 実験例

実験に用いた 8 枚のモノクロ画像を図 1 に示す．各画像の画素数は 256×256 である．図 1 の各画像に，平均 0，標準偏差 20 のガウス雑音と 10% のインパルス雑音を付加した画像を図 2 に示す．図 2 の各画像を入力とし，各手法による出力画像が図 1 の元画像にどの程度近くな

表 1 PSNR (dB)

	Fig. 2	BF	RSCBF	PSP	RBF
Airplane	19.22	25.17	25.34	26.10	26.14
Barbara	18.92	23.79	22.87	23.05	23.80
Boat	18.80	25.36	27.14	27.63	27.97
Cameraman	19.08	24.99	25.24	25.44	25.50
Girl	19.48	26.50	28.78	29.15	29.55
Lena	18.77	25.05	26.70	26.76	27.14
Mandrill	18.72	23.66	23.65	23.98	24.31
Woman	18.86	25.26	26.33	27.16	27.43

るかを比較した．比較に用いた手法は，BF，ロバスト自己クロス BF (RSCBF と略記する) [21]，Paint Shop Pro ver. 7.04 (PSP と略記する)，RBF である．出力画像を図 3 に示す．図 3 の上の行から順に，BF，RSCBF，PSP，RBF の出力画像を示している．また，図 1,2,3 の拡大図を図 4 に示す．図 4 の 1,2 行目がそれぞれ図 1,2 の拡大図であり，3 行目以降が図 3 の拡大図である．BF (図 4 の 3 行目) ではインパルス雑音が残っているが，その他 (4 行目以降) では除去できている．各フィルタのパラメータは，人手で良好な結果が得られる値に設定した．BF では， $\alpha = 0.1$ ， $\beta = 0.001$ とし，近傍画素の集合 N_i は画素 i を中心とする 15×15 の矩形領域とした．RSCBF については，付録 1. で概説する．PSP では，“ノイズ除去”を行った後に“輪郭以外をぼかす”を行った．後者のパラメータである“スムージングの量”は 6 とした．RBF では， $\alpha = 0.0001$ ， $\beta = 0.5$ ， $\tilde{\alpha} = 0.001$ ， $\tilde{\beta} = 0.5$ ， $\epsilon = 0.61$ とし，式 (5) の N_j のサイズは 7×7 とし，式 (7) の N_i のサイズは 15×15 とした．

次に，出力画像の画質を定量的に評価した結果を示す．画質評価の指標として，ピーク信号対雑音比 (PSNR)，structure similarity (SSIM) [22]，mean SSIM (MSSIM) [22] を用いた．これらの指標の具体的な計算方法は，付録 2. で概説する．

まず，PSNR の値を表 1 に示す．最左列に対象とする画像の名称 (図 1 のキャプションに示している) を示しており，各対象画像における各手法の PSNR 値をその横に並べている．ただし，Fig. 2 と記した列の値は，図 2 の雑音付加画像と図 1 の元画像の PSNR 値である．各行の最大値を太字で示している．どの画像でも，RBF の PSNR 値が最大となった．

次に，SSIM の値を表 2 に示す．Airplane では PSP の SSIM 値が最大となり，Barbara では BF の SSIM 値が最大となったが，その他の画像では RBF の SSIM 値が最大となった．

最後に，MSSIM の値を表 3 に示す．Cameraman で PSP の MSSIM 値が最大となったが，その他の画像では RBF の MSSIM 値が最大となった．

以上のように，3 種の画質評価指標において，ほとんどの画像で，提案手法である RBF が従来手法よりも高い雑音除去性能を示した．

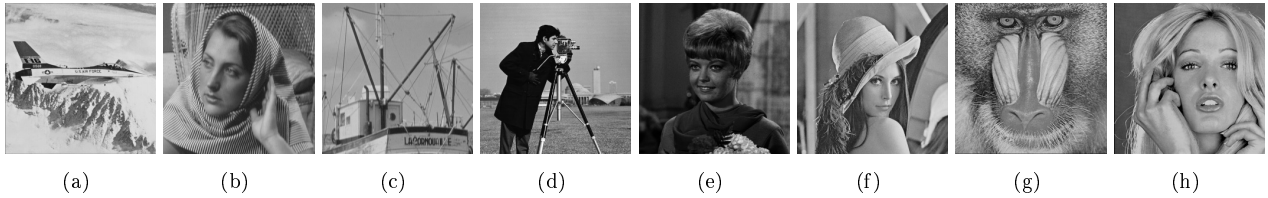


図 1 元画像: (a) Airplane, (b) Barbara, (c) Boat, (d) Cameraman, (e) Girl, (f) Lena, (g) Mandrill, (h) Woman



図 2 図 1 の画像に平均 0, 標準偏差 20 のガウス雑音と 10% のインパルス雑音を付加した画像

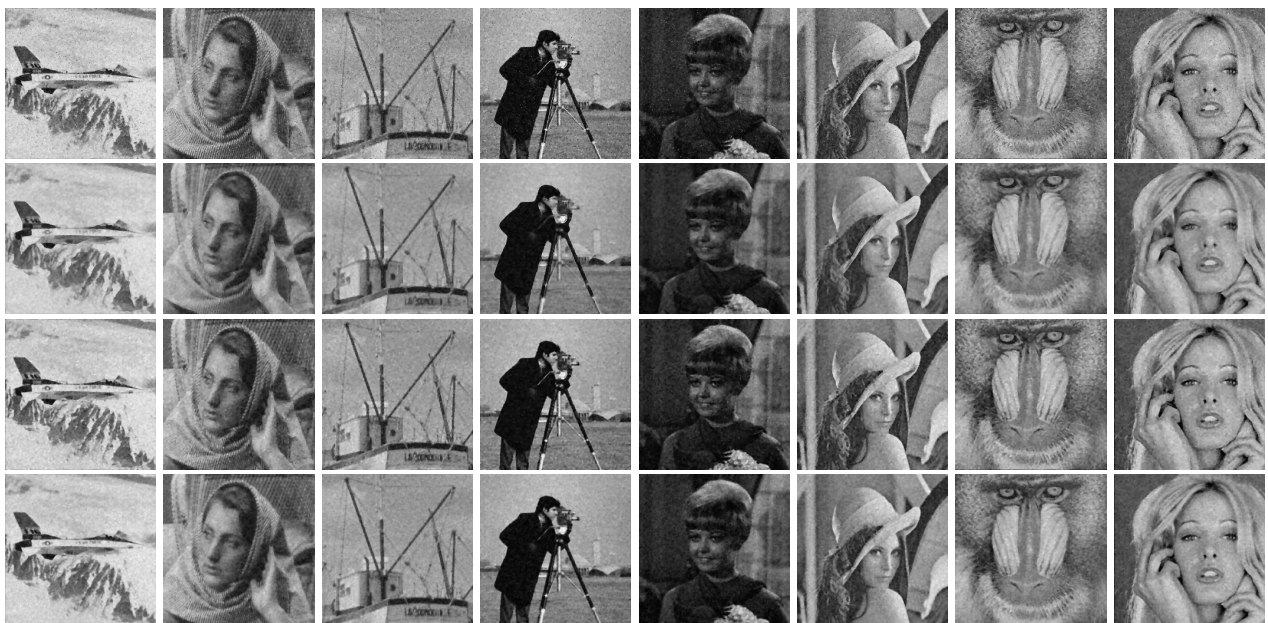


図 3 出力画像: 上の行から順に, BF, RSCBF, PSP, RBF の出力画像

表 2 SSIM ($\times 10^{-2}$)

	Fig. 2	BF	RSCBF	PSP	RBF
Airplane	84.45	95.28	95.03	96.18	96.16
Barbara	84.43	93.82	91.87	92.55	93.68
Boat	72.64	91.77	94.56	95.05	95.26
Cameraman	90.37	97.27	96.88	97.54	97.58
Girl	80.84	95.43	97.53	97.50	97.73
Lena	84.15	95.60	97.11	97.02	97.24
Mandrill	77.45	90.36	89.76	90.80	91.32
Woman	80.72	94.56	96.06	96.45	96.60

表 3 MSSIM ($\times 10^{-2}$)

	Fig. 2	BF	RSCBF	PSP	RBF
Airplane	97.72	99.43	99.37	99.50	99.51
Barbara	94.25	98.57	98.38	98.41	98.62
Boat	96.84	99.32	99.48	99.53	99.56
Cameraman	88.12	97.26	97.44	97.83	97.71
Girl	81.78	96.35	97.23	97.05	97.24
Lena	94.96	98.96	99.21	99.15	99.25
Mandrill	95.99	98.83	98.78	98.87	98.94
Woman	95.88	98.26	98.44	98.53	98.52

5. ま と め

バイラテラルフィルタでは除去するのが困難なインパルス雑音が混合した雑音を除去することを目的としたロバストバイラテラルフィルタを提案し, 3 種の画質評価指標に基づく比較実験でその有効性を確認した.

カラー画像に対する有効性の検証と, 雑音除去性能の更なる向上が今後の課題である.

文 献

- [1] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," Proc. ICCV '98, 1998.

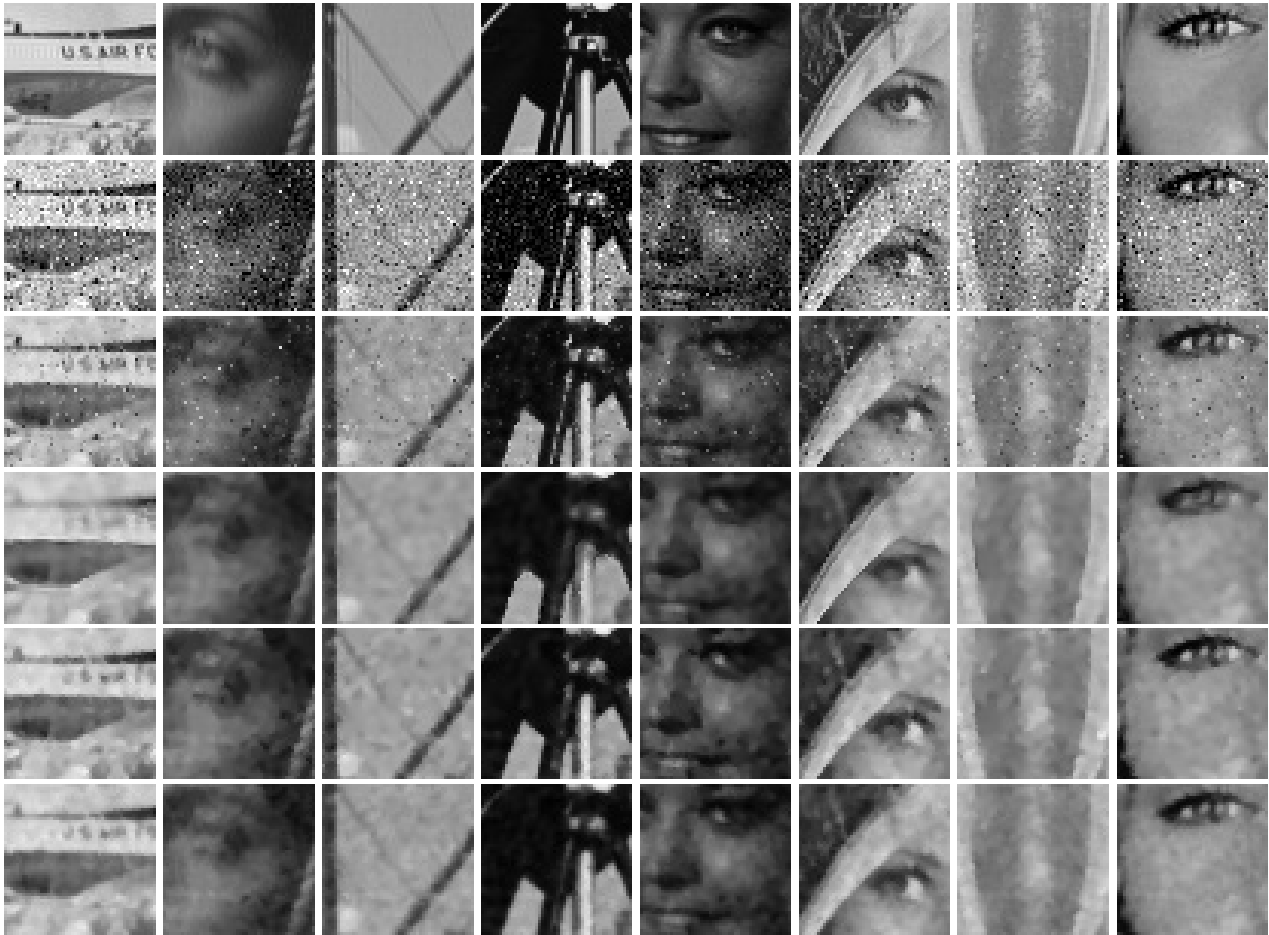


図4 図1,2,3の拡大図: 上の行から順に, 元画像, 雑音付加画像, BF, RSCBF, PSP, RBFの出力画像

- [2] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," Proc. SIGGRAPH '02, pp. 257-266, 2002.
- [3] E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," ACM Trans. Graph., vol. 23, no. 3, pp. 673-678, Aug. 2004.
- [4] S. Fleishman, I. Drori, and D. Cohen-Or, "Bilateral mesh denoising," ACM Trans. Graphics, vol. 22, no. 3, pp. 950-953, Jul. 2003.
- [5] M. Elad, "On the origin of the bilateral filter and ways to improve it," IEEE Trans. Image Processing, vol. 11, no. 10, pp. 1141 - 1151, Oct. 2002.
- [6] T. Q. Pham and L. J. van Vliet, "Separable bilateral filtering for fast video preprocessing," Proc. ICME, Jul. 2005.
- [7] B. Weiss, "Fast median and bilateral filtering," ACM Trans. Graph., vol. 25, no. 3, pp. 519-526, Jul. 2006.
- [8] 景琳琳, 浦浜喜一, "加重係数のテイラー展開に基づく空間分解バイラテラルフィルタの高精度化," 信学論, vol. J89-A, no. 7, pp. 643-645, Feb. 2006.
- [9] F. Porikli, "Constant time $O(1)$ bilateral filtering," Proc. CVPR, pp. 1-8, 2008.
- [10] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," Int. J. Comput. Vision, vol. 81, no. 1, pp. 24-52, Jan. 2009.
- [11] Q. Yang, K.-H. Tan, and N. Ahuja, "Real-time $O(1)$ bilateral filtering," Proc. CVPR, Jun. 2009.
- [12] A. Adams, N. Gelfand, J. Dolson, and M. Levoy, "Gaussian kd-trees for fast high-dimensional filtering," ACM Trans. Graphics, vol. 28, no. 3, pp. 1-12, Jul. 2009.
- [13] 五十嵐正樹, 池辺将之, 下山荘介, 山野健太, 本久順一, "フィルタカーネルサイズの依存性を持たない高速2-Dバイラテラルフィルタ演算手法の提案," 映情学誌, vol. 64, no. 3, pp. 389-398, Mar. 2010.
- [14] W. Yu, F. Franchetti, J. C. Hoe, Y.-J. Chang, and T. Chen, "Fast bilateral filtering by adapting block size," Proc. ICIP, pp. 3281-3284, Sept. 2010.
- [15] B. Gunturk, "Fast bilateral filter with arbitrary range and domain kernels," Proc. ICIP, pp. 3289-3292, Sept. 2010.
- [16] M. Igarashi, M. Ikebe, S. Shimoyama, K. Yamano, and J. Motohisa, " $O(1)$ bilateral filtering with low memory usage," Proc. ICIP, pp. 3301-3304, Sept. 2010.
- [17] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," ACM Trans. Graph., vol. 23, no. 3, pp. 664-672, 2004.
- [18] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," ACM Trans. Graphics, vol. 26, no. 3, 96, Jul. 2007.
- [19] 田中正行, 奥富正敏, "バイラテラルフィルタとノンローカルミーンフィルタの統一的解釈とその発展へ向けて," 信学論, vol. J91-D, no. 8, pp. 1470-1479, Aug. 2010.
- [20] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," Proc. CVPR, vol. 2, pp. 60-65, 2005.
- [21] 常セン, 井上光平, 浦浜喜一, "自己クロスバイラテラル

フィルタによるブートストラップ雑音除去,” 信学技報, IE2009-111, vol. 109, no. 292, pp. 73-78, 2009.

- [22] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE Trans. Image Processing., vol. 13, no. 4, pp. 600-612, 2004.

付 録

1. ロバスト自己クロスバイラテラルフィルタ

常ら [21] のロバスト自己クロスバイラテラルフィルタ (RSCBF) を概説する。まず, 入力画像 f にウィンドウサイズ 11×11 のメジアンフィルタをかけた結果を f^{med} とし, これをモードフィルタの反復法の初期値とする。すなわち, 反復式

$$f_i^{(t+1)} = \frac{\sum_{j \in N_i} \tilde{u}_{ij} \exp(-\beta \|f_i^{(t)} - f_j\|^2) f_j}{\sum_{j \in N_i} \tilde{u}_{ij} \exp(-\beta \|f_i^{(t)} - f_j\|^2)} \quad (\text{A.1})$$

において, $f_i^{(0)} = f_i^{\text{med}}$ とする。ここで,

$$\tilde{u}_{ij} = \begin{cases} (1 - \|x_i - x_j\|/p)^2 & \|x_i - x_j\| \leq p \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.2})$$

であり, $p = 5$ とする。モードフィルタの出力を f^M とすると, RSCBF の反復式は

$$\tilde{f}_i^{(t+1)} = \frac{\sum_{j \in N_i} \tilde{u}_{ij} \tilde{w}_{ij}^{(t)} \exp(-\beta \|f_i^M - f_j^M\|^2) f_j}{\sum_{j \in N_i} \tilde{u}_{ij} \tilde{w}_{ij}^{(t)} \exp(-\beta \|f_i^M - f_j^M\|^2)} \quad (\text{A.3})$$

と表される。ここで,

$$\tilde{w}_{ij}^{(t)} = \exp(-\gamma \|\tilde{f}_j^{(t)} - f_j\|^2) \quad (\text{A.4})$$

であり, $\tilde{f}_j^{(t)}$ の初期値は $\tilde{f}_j^{(0)} = f_j$ とする。文献 [21] と同様に, 式 (A.1), (A.3) の β は 0.01 とし, 式 (A.4) の γ は 0.0001 とする。

2. 画質評価指標

実験で用いた画質評価指標 PSNR, SSIM, MSSIM について概説する。

2.1 ピーク信号対雑音比 (PSNR)

雑音を含まないモノクロ画像を $f = [f_i]$, $i = 1, \dots, n$ とし, f に雑音を付加した画像を $g = [g_i]$ とすると, PSNR は

$$PSNR = 10 \log_{10} \left(\frac{P^2}{\frac{1}{n} \sum_{i=1}^n (f_i - g_i)^2} \right) \quad (\text{A.5})$$

で定義される。ここで P は画素値の最大値であり, 1 画素が 8 ビットで表現される場合には $P = 255$ である。

2.2 Structural similarity (SSIM)

画像 f, g の SSIM は

$$SSIM(f, g) = \frac{(2\mu_f \mu_g + C_1)(2\sigma_{fg} + C_2)}{(\mu_f^2 + \mu_g^2 + C_1)(\sigma_f^2 + \sigma_g^2 + C_2)} \quad (\text{A.6})$$

で定義される。ここで μ_f, μ_g は各画像の画素値の平均値

$$\mu_f = \frac{1}{n} \sum_{i=1}^n f_i, \quad \mu_g = \frac{1}{n} \sum_{i=1}^n g_i \quad (\text{A.7})$$

であり, σ_f, σ_g は標準偏差

$$\sigma_f = \left(\frac{1}{n-1} \sum_{i=1}^n (f_i - \mu_f)^2 \right)^{\frac{1}{2}} \quad (\text{A.8})$$

$$\sigma_g = \left(\frac{1}{n-1} \sum_{i=1}^n (g_i - \mu_g)^2 \right)^{\frac{1}{2}} \quad (\text{A.9})$$

であり, σ_{fg} は共分散

$$\sigma_{fg} = \frac{1}{n-1} \sum_{i=1}^n (f_i - \mu_f)(g_i - \mu_g) \quad (\text{A.10})$$

である。 C_1, C_2 は 0 による除算を避けるためのものであり,

$$C_1 = (K_1 L)^2, \quad C_2 = (K_2 L)^2 \quad (\text{A.11})$$

で定義される。ここで L は画素値のダイナミックレンジであり, 8 ビットのモノクロ画像の場合, $L = 255$ である。 K_1, K_2 は微小な定数であり, 文献 [22] と同様に $K_1 = 0.01, K_2 = 0.03$ とする。

2.3 Mean SSIM (MSSIM)

画像 f, g の MSSIM は

$$MSSIM(f, g) = \frac{1}{n} \sum_{i=1}^n SSIM(f_{N_i}, g_{N_i}) \quad (\text{A.12})$$

で定義される。ここで f_{N_i}, g_{N_i} はそれぞれ f, g の画素 i を中心とする矩形領域を切り出した部分画像であり, 矩形のサイズは 11×11 とする。部分画像 $f_{N_i} = [f_{N_i,j}]$, $g_{N_i} = [g_{N_i,j}]$ の SSIM は

$$SSIM(f_{N_i}, g_{N_i}) = \frac{(2\tilde{\mu}_{f_{N_i}} \tilde{\mu}_{g_{N_i}} + C_1)(2\tilde{\sigma}_{f_{N_i} g_{N_i}} + C_2)}{(\tilde{\mu}_{f_{N_i}}^2 + \tilde{\mu}_{g_{N_i}}^2 + C_1)(\tilde{\sigma}_{f_{N_i}}^2 + \tilde{\sigma}_{g_{N_i}}^2 + C_2)} \quad (\text{A.13})$$

とする。ここで

$$\tilde{\mu}_{f_{N_i}} = \sum_{j \in N_i} \hat{w}_{ij} f_{N_i,j}, \quad \tilde{\mu}_{g_{N_i}} = \sum_{j \in N_i} \hat{w}_{ij} g_{N_i,j} \quad (\text{A.14})$$

$$\tilde{\sigma}_{f_{N_i}} = \left(\sum_{j \in N_i} \hat{w}_{ij} (f_{N_i,j} - \tilde{\mu}_{f_{N_i}})^2 \right)^{\frac{1}{2}} \quad (\text{A.15})$$

$$\tilde{\sigma}_{g_{N_i}} = \left(\sum_{j \in N_i} \hat{w}_{ij} (g_{N_i,j} - \tilde{\mu}_{g_{N_i}})^2 \right)^{\frac{1}{2}} \quad (\text{A}\cdot 16)$$

$$\tilde{\sigma}_{f_{N_i} g_{N_i}} = \sum_{j \in N_i} \hat{w}_{ij} (f_{N_i,j} - \tilde{\mu}_{f_{N_i}}) (g_{N_i,j} - \tilde{\mu}_{g_{N_i}}) \quad (\text{A}\cdot 17)$$

であり, \hat{w}_{ij} は $\sum_{j \in N_i} \hat{w}_{ij} = 1$ を満たす重み

$$\hat{w}_{ij} = \frac{\exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma_w^2}\right)}{\sum_{j' \in N_i} \exp\left(-\frac{\|x_i - x_{j'}\|^2}{2\sigma_w^2}\right)} \quad (\text{A}\cdot 18)$$

であり, σ_w は文献[22]と同様に $\sigma_w = 1.5$ とする.