# A convergence proof for local mode filtering

Zhang, Shuoyan
Department of Communication Design Science, Kyushu University

Inoue, Kohei
Department of Communication Design Science, Kyushu University

Hara, Kenji
Department of Communication Design Science, Kyushu University



出版情報：ICT Express. 7 (4), pp.445-448, 2021-12. Elsevier
バージョン：

Available online at www.sciencedirect.com
ScienceDirect
ICT Express 7 (2021) 445–448
www.elsevier.com/locate/icte

# A convergence proof for local mode filtering

Shuoyan Zhang, Kohei Inoue*, Kenji Hara

*Department of Communication Design Science, Kyushu University, Fukuoka, Japan*

Received 30 May 2020; received in revised form 12 February 2021; accepted 22 February 2021
Available online 9 March 2021

**Abstract**

In this paper, we present a convergence proof for an iterative procedure of local mode filtering. We formulate the local mode filtering as a quadratic optimization problem based on the Legendre transform of convex function, from which two closed-form expressions at each iteration step are derived for variables to be optimized. Those analytical solutions ensure that the value of objective function increases monotonically with the progress of the iterative procedure. We also show experimental results using a grayscale image, which support our theoretical results practically.


*Keywords:* Local mode filter; Legendre transform; Convex function

## 1. Introduction

Local mode filter is an edge-preserving smoothing filter, which smooths small variations of a signal, while preserves large ones like sharp edges in an image.

van de Weijer and van den Boomgaard [1] focused on the local mode of a local histogram on an image, and proposed an iterative method for finding the local mode. Kass and Solomon [2] pointed out that the method is equivalent to the restricted mean shift filtering by Barash and Comaniciu [3], and proposed the closest-mode filter, which calculates a local mode by searching for a zero crossing of the derivative of smoothed local histogram. van den Boomgaard and van de Weijer [4] showed the equivalence of local-mode finding, robust estimation and mean-shift analysis by Comaniciu and Meer [5], where the convergence properties of the mean shift are proved. In contrast, Li et al. [6] presented counter examples to them, and presented alternative proofs. Recently, Bao and Yang [7] have used the term "closest-mode filter" to refer to filters whose local histograms are constructed within Gaussian weighted soft spatial windows. Lo [8] has published ImageJ plugins for the mode filter and empirical null filter which are available at GitHub repository (https://github.com/shermanlo77/oxwasp_phd/tree/master/java).

* Corresponding author.
*E-mail address:* k-inoue@design.kyushu-u.ac.jp (K. Inoue).
Peer review under responsibility of The Korean Institute of Communications and Information Sciences (KICS).

In this paper, we present a convergence proof for the iterative procedure of local mode filtering. We formulate the problem of finding local mode in local mode filtering as a quadratic optimization problem based on the Legendre transform of convex function [9]. From this formulation, we derive closed-form expressions for updating variables, and prove the convergence of the sequence of the updated variables.

The rest of this paper is organized as follows. Section 2 introduces the local mode filtering, and Section 3 briefly summarizes the Legendre transform. Then Section 4 presents a Legendre-transformed local mode filter, and shows an algorithm for image processing. Section 5 shows the convergence proof based on the discussion in the preceding sections. Section 6 shows experimental results. Finally, Section 7 concludes this paper.

## 2. Local mode filtering

Let $I$ be an input image. Then the local mode filter [1] can be formulated as follows:

$$J(p) = \arg\max_{J_p} \sum_{q \in N(p)} w_{\sigma_s}(p, q) w_{\sigma_r}\left(I(q), J_p\right), \tag{1}$$

where $J(p)$ denotes the pixel value at the position $p$ on the output image $J$, $N(p)$ denotes a set of neighboring pixels to $p$, and $w_{\sigma_s}(p, q)$ and $w_{\sigma_r}(I(q), J_p)$ denote the spatial domain and range kernels [10] defined by

$$w_{\sigma_s}(p, q) = \exp\left(-\frac{\|p - q\|^2}{2\sigma_s^2}\right) \tag{2}$$

https://doi.org/10.1016/j.icte.2021.02.008
2405-9595/

$$w_{\sigma_r}\left(I(q), J_p\right) = \exp\left(-\frac{\|I(q)-J_p\|^2}{2\sigma_r^2}\right), \tag{3}$$

where $\sigma_s$ and $\sigma_r$ denote their standard deviations.

Let $E(J_p)$ be the objective function in (1). Then a recurrence relation for computing the local mode $J(p)$ is derived from the necessary condition for optimality

$$\frac{\partial E}{\partial J_p} = \frac{1}{\sigma_r^2}\sum_{q\in N(p)} w_{\sigma_s}(p,q) w_{\sigma_r}\left(I(q), J_p\right)\left(I(q)-J_p\right) = 0 \tag{4}$$

as follows:

$$J_p^{(t+1)} = \frac{\sum_{q\in N(p)} w_{\sigma_s}(p,q) w_{\sigma_r}\left(I(q), J_p^{(t)}\right) I(q)}{\sum_{q\in N(p)} w_{\sigma_s}(p,q) w_{\sigma_r}\left(I(q), J_p^{(t)}\right)}, \tag{5}$$

where the superscripts $(t+1)$ and $(t)$ denote the numbers of iterations for recurrent computation of (5). For $t=0$, $J_p^{(t)}$ is initialized as $J_p^{(0)} = I(p)$, and (5) is computed iteratively until $J_p^{(t)}$ converges. Finally, the converged value of $J_p^{(t+1)}$ gives the output $J(p)$ in (1).

## 3. Legendre transform

In this section, we briefly summarize the Legendre transform of a convex function [9] as a preparation for the next section.

Let $f(x)$ be a convex function satisfying $f''(x) > 0$. Then the Legendre transform of $f(x)$ is given by

$$g(s) = \max_x\{sx - f(x)\}, \tag{6}$$

which is called the convex conjugate function of $f(x)$. The function $sx - f(x)$ of variable $x$ attains its maximum when $\partial[sx - f(x)]/\partial x = s - f'(x) = 0$ or

$$s = f'(x), \tag{7}$$

where $f'(x)$ is monotonically increasing and invertible. We denote the inverse function as follows:

$$x(s) = (f')^{-1}(s). \tag{8}$$

Substituting this for $x$ in (6), we have

$$g(s) = sx(s) - f(x(s)). \tag{9}$$

The Legendre transform of $g(s)$ in (9) recovers $f(x)$ as follows:

$$\max_s\{xs - g(s)\} = \max_s\{xs - sx(s) + f(x(s))\} \tag{10}$$

$$= xs - sx + f(x) = f(x), \tag{11}$$

where $x(s) = x$ derived from $\partial[xs - sx(s) + f(x(s))]/\partial s = 0$ is substituted for $x(s)$ in (10). Consequently, we have the alternative expression of $f(x)$ as follows:

$$f(x) = \max_s\{xs - s(f')^{-1}(s) + f((f')^{-1}(s))\}, \tag{12}$$

where (8) is substituted for $x(s)$ in (10).

**Example.** Let $f(x) = \exp(-x)$. Then (7) becomes $s = f'(x) = -\exp(-x)$, and therefore (8) is given by $x(s) = (f')^{-1}(s) = -\ln(-s)$, which is substituted for $(f')^{-1}(s)$ in (12), then we have

$$\begin{aligned} f(x) &= \max_s\{xs + s\ln(-s) + \exp(\ln(-s))\} \\ &= \max_s\{s(x + \ln(-s) - 1)\}. \end{aligned} \tag{13}$$

From $\partial[s(x+\ln(-s)-1)]/\partial s = x + \ln(-s) = 0$, we have $s = -\exp(-x)$, which is substituted for $s$ in (13) to confirm that $\max_s\{s(x+\ln(-s)-1)\} = -\exp(-x)(x+\ln(\exp(-x))-1) = \exp(-x)$. ■

## 4. Legendre-transformed local mode filter

Applying the above alternative expression in (12) or (13) to the exponential function in (3), we can rewrite the maximization problem in (1) as follows:

$$\max_{J_p,\{s_q\}} \sum_{q\in N(p)} w_{\sigma_s}(p,q)\left[s_q\left(\frac{\|I(q)-J_p\|^2}{2\sigma_r^2} + \ln(-s_q) - 1\right)\right], \tag{14}$$

where the variable $s_q$ is prepared for each $q$ in $N(p)$.

Let $\tilde{E}(J_p, \{s_q\})$ be the objective function in (14). Then $\tilde{E}(J_p, \{s_q\})$ is a quadratic function of $J_p$ for a fixed $\{s_q\}$. On the other hand, it is also a function of $s_q$ including $s_q\ln(-s_q)$ whose degree is less than 2. Therefore, we have two analytically solvable necessary conditions for optimality as follows:

$$\frac{\partial \tilde{E}}{\partial s_q} = w_{\sigma_s}(p,q)\left(\frac{\|I(q)-J_p\|^2}{2\sigma_r^2} + \ln(-s_q)\right) = 0, \tag{15}$$

$$\frac{\partial \tilde{E}}{\partial J_p} = -\frac{1}{\sigma_r^2}\sum_{q\in N(p)} w_{\sigma_s}(p,q) s_q\left(I(q)-J_p\right) = 0. \tag{16}$$

From (15), we have a closed-form expression for $s_q$ as follows:

$$\begin{aligned} s_q(J_p) &= -\exp\left(\frac{\|I(q)-J_p\|^2}{2\sigma_r^2}\right) \\ &= -w_{\sigma_r}\left(I(q), J_p\right) \end{aligned} \tag{17}$$

for $q$ in $N(p)$. Therefore, for a fixed $J_p = J_p^{(t)}$, we have

$$\tilde{E}\left(J_p^{(t)}, \{s_q^{(t+1)}\}\right) \geq \tilde{E}\left(J_p^{(t)}, \{s_q^{(t)}\}\right), \tag{18}$$

where $s_q^{(t+1)} = s_q(J_p^{(t)})$.

On the other hand, from (16), we have another closed-form expression for $J_p$ as follows:

$$J_p(\{s_q\}) = \frac{\sum_{q\in N(p)} w_{\sigma_s}(p,q) s_q I(q)}{\sum_{q\in N(p)} w_{\sigma_s}(p,q) s_q}. \tag{19}$$

Therefore, for a fixed set $\{s_q\} = \{s_q^{(t+1)}\}$, we have

$$\tilde{E}\left(J_p^{(t+1)}, \{s_q^{(t+1)}\}\right) \geq \tilde{E}\left(J_p^{(t)}, \{s_q^{(t+1)}\}\right), \tag{20}$$

where $J_p^{(t+1)} = J_p(\{s_q^{(t+1)}\})$.

The procedure of the above Legendre-transformed local mode filtering is summarized in Algorithm 1.

**Algorithm 1** Legendre-transformed local mode filtering

**Require:** Input image $I$
**Ensure:** Output image $J$

```
for each pixel p in I do
    Initialize t as t = 0
    Initialize J_p as J_p^(t) = I(p)
    while J_p^(t) is not convergent do
        for each pixel q in N(p) do
            Compute s_q^(t+1) = s_q(J_p^(t)) by (17)
        end for
        Compute J_p^(t+1) = J_p({s_q^(t+1)}) by (19)
        t ← t + 1
    end while
    J(p) ← J_p^(t)
end for
```

## 5. Convergence proof

In this section, we prove that the sequence $\{J_p^{(t)}\}_{t=0,1,\ldots}$ given by (5) converges to a solution to the problem (1).

Combining (18) and (20), we have

$$\tilde{E}\left(J_p^{(t+1)}, \{s_q^{(t+1)}\}\right) \geq \tilde{E}\left(J_p^{(t)}, \{s_q^{(t+1)}\}\right) \geq \tilde{E}\left(J_p^{(t)}, \{s_q^{(t)}\}\right), \tag{21}$$

that is, the sequence $\{\tilde{E}(J_p^{(t)}, \{s_q^{(t)}\})\}_{t=0,1,\ldots}$ is nondecreasing and bounded above since $w_{\sigma_s}(p,q) \leq 1$ and

$$\max_{s_q}\left\{s_q\left(\frac{\|I(q)-J_p\|^2}{2\sigma_r^2} + \ln(-s_q) - 1\right)\right\} = \exp\left(-\frac{\|I(q)-J_p\|^2}{2\sigma_r^2}\right) \leq 1. \tag{22}$$

Substituting $s_q^{(t+1)} = s_q(J_p^{(t)})$ given by (17) for $s_q$ in (19), we have (5). Hence, the procedure of the Legendre-transformed local mode filtering is equivalent to that of the original local mode filtering, and since $\tilde{E}(J_p^{(t)}, \{s_q^{(t)}\}) = E(J_p^{(t)})$, the sequence $\{E(J_p^{(t)})\}_{t=0,1,\ldots}$ is also nondecreasing and bounded above. Therefore, it converges to the local maximum.

Here we would like to rewrite (5) as a gradient method as follows:

$$J_p^{(t+1)} = J_p^{(t)} + \alpha^{(t)} \left.\frac{\partial E}{\partial J_p}\right|_{J_p=J_p^{(t)}}, \tag{23}$$

where $\alpha^{(t)}$ is given by

$$\alpha^{(t)} = \frac{\sigma_r^2}{E\left(J_p^{(t)}\right)}. \tag{24}$$

After the convergence of $\{E(J_p^{(t)})\}_{t=0,1,\ldots}$, the necessary condition for optimality in (4) is satisfied. Substituting (4) into (23), we have

$$J_p^{(t+1)} = J_p^{(t)}, \tag{25}$$

which means the convergence of $\{J_p^{(t)}\}_{t=0,1,\ldots}$.

**Fig. 1.** Input image of 256 by 256 pixels.

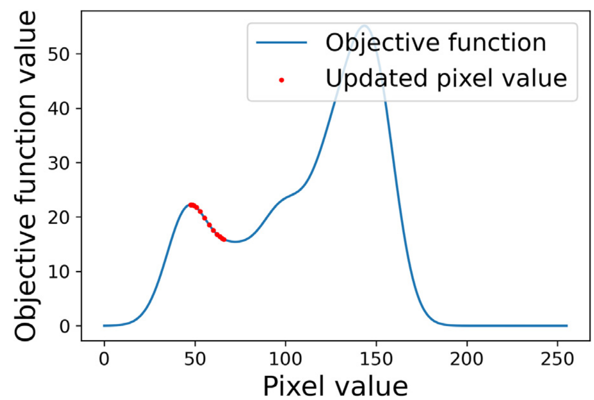


**Fig. 2.** Objective function at the center of white square in Fig. 1. The red points illustrate the convergence to the local mode.

## 6. Experimental results

We have conducted experiments to confirm the above theoretical results. Fig. 1 shows an input grayscale image of 8-bit per pixel; each pixel has an integer in $\{0, 1\ldots, 255\}$. The objective function in (1) at the center of a square region marked out with white line on the upper right side of Fig. 1 is shown in Fig. 2, where the vertical and horizontal axes denote the function value and the pixel value, respectively. In this example, we set $\sigma_s = 5$ and $\sigma_r = 10$, and $N(p)$ is the set of pixels in the square region with a side length of $6\sigma_s$ as shown in Fig. 1. The value of the center pixel in the square region is 66, which is near the valley floor in Fig. 2 as denoted by the rightmost red point, from which, it climbs to the top of the left peak or the local mode of this distribution. The loci of updated pixel values and the corresponding values of the objective function are shown with red points as well, which demonstrates the convergence of the sequence to a local mode. We judge the convergence when the condition $(J_p^{(t+1)} - J_p^{(t)})^2 < 10^{-3}$ is satisfied.

Fig. 3 shows the values of the objective function and its derivative, where the horizontal axis denotes the number of iterations, and the left and right vertical axes denote the

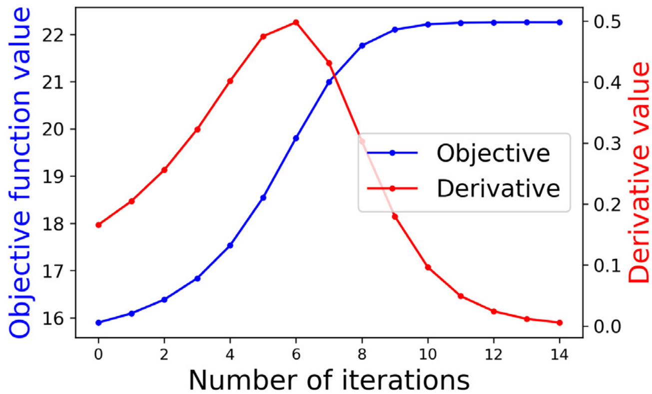


**Fig. 3.** The values of objective function and its derivative.

**Fig. 4.** Output image of local mode filtering.

objective function value and its derivative value, respectively. These two vertical axes are related to two curves in this graph by their colors, blue and red. The blue curve shows that the objective function is monotonically increasing and bounded above, and the red curve shows that the necessary condition for optimality in (4) is eventually satisfied. To satisfy the exact convergence condition $(J_p^{(t+1)} - J_p^{(t)})^2 = 0$, it took 59 iterations in 0.5 s measured on a laptop computer.

Fig. 4 shows the output image of local mode filtering for the input image in Fig. 1. All pixel values converged to the corresponding local modes in a dozen times of iterations.

## 7. Conclusion

In this paper, we proved that the iterative procedure of local mode filtering takes any initial values to their corresponding local modes by using the Legendre transform of convex function. Experimental results demonstrated the convergence of pixel values of a grayscale image to their local optima by the local mode filtering procedure.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] J. van de Weijer, R. van den Boomgaard, Local mode filtering, in: Computer Vision and Pattern Recognition (CVPR), Vol. 2, 2001, pp. 428–433.
[2] M. Kass, J. Solomon, Smoothed local histogram filters, ACM Trans. Graph. 29 (4) (2010) 100:1–10.
[3] D. Barash, D. Comaniciu, A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering and mean shift, Image Vis. Comput. 22 (1) (2004) 73–81.
[4] R. van den Boomgaard, J. van de Weijer, On the equivalence of local-mode finding, robust estimation and mean-shift analysis as used in early vision tasks, in: International Conference on Pattern Recognition (ICPR'02), Vol. 3, 2002, pp. 927–930.
[5] D. Comaniciu, P. Meer, Mean shift: a robust approach toward feature space analysis, IEEE Trans. Pattern Anal. Mach. Intell. 24 (5) (2002) 603–619.
[6] X. Li, Z. Hu, F. Wu, A note on the convergence of the mean shift, Pattern Recognit. 40 (6) (2007) 1756–1762.
[7] L. Bao, Q. Yang, Robust Piecewise-Constant Smoothing: M-smoother revisited, 2014, CoRR, abs/1410.7580.
[8] S.E. Lo, Characterisation of Computed Tomography Noise in Projection Space with Applications to Additive Manufacturing (Ph.D. thesis), University of Warwick, Department of Statistics, 2020.
[9] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.
[10] B.K. Gunturk, Fast bilateral filter with arbitrary range and domain kernels, IEEE Trans. Image Process. 20 (9) (2011) 2690–2696.