IEICE TRANS. INF. & SYST., VOL.E104–D, NO.10 OCTOBER 2021

LETTER *Special Section on Picture Coding and Image Media Processing*

Rolling Guidance Filter as a Clustering Algorithm

Takayuki HATTORI[†], ***Nonmember***, **Kohei INOUE**[†a)], ***and*** **Kenji HARA**[†], ***Members***

SUMMARY We propose a generalization of the rolling guidance filter (RGF) to a similarity-based clustering (SBC) algorithm which can handle general vector data. The proposed RGF-based SBC algorithm makes the similarities between data clearer than the original similarity values computed from the original data. On the basis of the similarity values, we assign cluster labels to data by an SBC algorithm. Experimental results show that the proposed algorithm achieves better clustering result than the result by the naive application of the SBC algorithm to the original similarity values. Additionally, we study the convergence of a unimodal vector dataset to its mean vector.

key words: *edge preserving smoothing, rolling guidance filter, clustering*

1. Introduction

Edge-preserving smoothing (EPS) is one of the fundamental techniques used in several research areas such as image processing, computer vision and computer graphics. For example, the bilateral filter proposed by Tomasi and Manduchi [1] has many applications as surveyed by Paris et al. [2]. Recently, Zhang et al. [3] have proposed another EPS filter called the rolling guidance filter (RGF), which can be interpreted as the iterations of joint bilateral filter [4]. The RGF is a scale-aware filter extensible to suit various applications and scenarios, and achieves a realtime performance for large scale images [3]. However, the theoretical convergence in the RGF procedure remains unsolved.

In this paper, we propose a similarity-based clustering (SBC) algorithm based on the rolling guidance framework. The proposed algorithm clusters general vector data on the basis of their similarities. The RGF procedure updates the similarities iteratively, and the resultant similarities become clearer than the original ones. As a result, we can easily cluster given data with the updated similarities. Experimental results show an improved performance of data clustering.

The rest of this paper is organized as follows. Section 2 briefly summarizes the RGF. Section 3 proposes the RGF-based SBC algorithm. Section 4 shows experimental results. Finally, Sect. 5 concludes this paper.

2. Rolling Guidance Filter

Let I be an input image, and let J be a guidance image which has the same size as I. Then the rolling guidance filter (RGF) computes the following equation iteratively:

$$J^{t+1}(p) = \frac{\displaystyle\sum_{q\in N(p)} \exp\left(-\frac{\|p-q\|^2}{2\sigma_s^2} - \frac{\|J^t(p)-J^t(q)\|^2}{2\sigma_r^2}\right) I(q)}{\displaystyle\sum_{q\in N(p)} \exp\left(-\frac{\|p-q\|^2}{2\sigma_s^2} - \frac{\|J^t(p)-J^t(q)\|^2}{2\sigma_r^2}\right)}, \quad (1)$$

where p and q denote the pixel coordinates, $N(p)$ denotes the set of neighboring pixels of p, and σ_s and σ_r are positive constants to control the spatial and range weights, respectively. $J^{t+1}(p)$ denotes the output value at the position p in the $(t+1)$th iteration of the RGF procedure. If we initialize J as $J^0(p) = c$ with a constant c for all pixels p, then the first iteration of (1) becomes the Gaussian filter.

The above RGF updates the guidance image J only, while keeping the input image I fixed. This procedure can be interpreted as the iterations of the joint bilateral filter [4] or cross-bilateral filter [5]. However, as Zhang et al. [3] stated, the proof of the convergence to a meaningful image by the RGF procedure remains an open question.

3. Clustering by RGF

In this section, we would like to generalize the above RGF to the procedure for multidimensional vector data. Instead of two images I and J, we consider two vector datasets $D = \{d_1, d_2, \ldots, d_n\}$ and $F^0 = \{f_1^0, f_2^0, \ldots, f_n^0\}$ where n denotes the number of data, and only the latter set F is updated iteratively. Let $F^t = \{f_1^t, f_2^t, \ldots, f_n^t\}$ be the output at the tth iteration. Then (1) can be rewritten as

$$f_i^{t+1} = \frac{\displaystyle\sum_{j=1}^{n} w_{ij} s\left(f_i^t, f_j^t\right) d_j}{\displaystyle\sum_{j=1}^{n} w_{ij} s\left(f_i^t, f_j^t\right)}, \quad (2)$$

where w_{ij} is the weight between data i and j, e.g., in (1), we have $w_{i=p,j=q} = \exp(-\|p-q\|^2/2\sigma_s^2)$. For the sake of simplicity, we assume that $w_{ij} = 1$ below. The similarity between f_i^t and f_j^t is denoted by $s(f_i^t, f_j^t) = \exp(-\|f_i^t - f_j^t\|^2/2\sigma_r^2)$. We also assume that F^t is initialized as $F^0 = D$.

After the update of every f_i^t by (2), we update the similarities $s(f_i^t, f_j^t)$ for all pairs. Therefore, the RGF procedure can be viewed as an algorithm to update similarities

Manuscript received December 22, 2020.
Manuscript revised March 19, 2021.
Manuscript publicized May 31, 2021.

[†]The authors are with Kyushu University, Fukuoka-shi, 815–8540 Japan.
a) E-mail: k-inoue@design.kyushu-u.ac.jp

DOI: 10.1587/transinf.2021PCL0001

Algorithm 1 Similarity-based clustering

Require: similarity s_{ij} between data i and j
Ensure: a set of labels $\{l_1, l_2, \ldots, l_n\}$
1: $l_i \leftarrow -1$ for $i = 1, 2, \ldots, n$
2: $c \leftarrow 1$
3: **for** $i = 1, 2, \ldots, n$ **do**
4: **if** $l_i < 0$ **then**
5: $l_i \leftarrow c$
6: **for** $j = i, i+1, \ldots, n$ **do**
7: **if** $s_{ij} > \theta$ **then**
8: $l_j \leftarrow c$
9: **end if**
10: **end for**
11: $c \leftarrow c + 1$
12: **end if**
13: **end for**
14: **return** $\{l_1, l_2, \ldots, l_n\}$

between data. The obtained similarities get closer to the minimum value 0 or the maximum value 1. On the basis of this 'crisp' similarities, we can classify the given data into clusters. Here, we use a simple clustering algorithm described in Algorithm 1.

In this algorithm, the first datum gets a label 1, and the following data similar to the first one also get the same label 1. Next, the first datum among the remaining unlabeled data gets the second label 2, and the following unlabeled data similar to the one with label 2 also get the same label 2. Such a procedure is repeated until all data get their labels.

The above RGF-based clustering algorithm shares the open question in the original RGF. Therefore, it is difficult to discuss the convergence of (2). In Appendix A, we study the convergence in a simple case.

4. Experimental Results

We conducted a numerical simulation of data clustering by the proposed algorithm with a two-dimensional dataset S1 [6], which is publicly available on the website "Clustering basic benchmark" [7]. Figure 1 shows 500 point data sampled from the dataset S1 with blue points. This dataset forms 15 Gaussian clusters, whose centers are depicted with yellow points. The similarities between the blue points with $\sigma_r = 50000$ are shown in Fig. 2 (a), where the similarity between ith and jth points is put in the (i, j)th entry of a matrix, which is displayed as a grayscale image in which dark pixels indicate small similarities, and bright pixels indicate large similarities. Here, σ_r is a parameter to be specified by users. In this example, we decided the value as follows: We first find horizontal and vertical ranges in which all data exist as $H = 857681$ and $V = 851014$, respectively. Then we divide the area HV by the number of clusters $K = 15$ as HV/K, which is the mean area per cluster, and take the square root of it $L = \sqrt{HV/K}$, which is the mean length assigned to each cluster. We place that a normal distribution in the range of length L, and assume that the range is covered by a constant multiple of σ_r, e.g., $4\sigma_r \leq L \leq 5\sigma_r$. Typically, if we set $4.5\sigma_r = L$, then we have

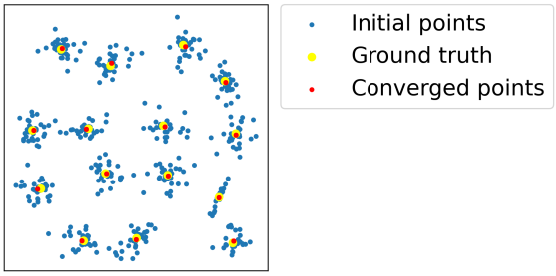
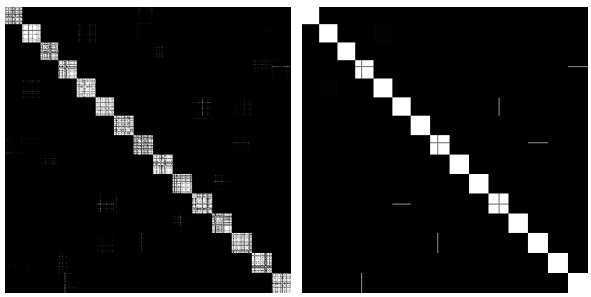
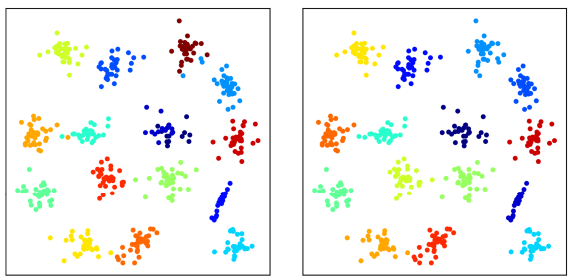


Fig. 1 Two-dimensional data S1 [6]



(a) Initial similarities (b) Converged similarities

Fig. 2 Similarity matrices.

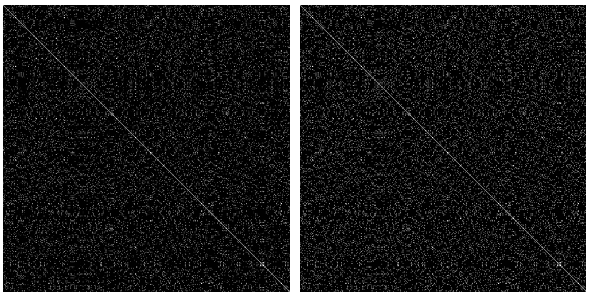


(a) Without RGF (b) With RGF

Fig. 3 Clustering results.

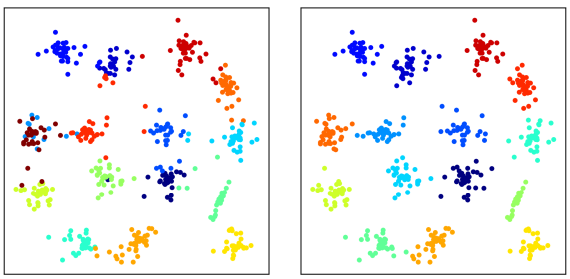
$\sigma_r = L/4.5 = \sqrt{\frac{857681 \times 851014}{15}}/4.5 \approx 49020$, from which we have an approximate number 50000. In Fig. 2, the data are roughly ordered according to their cluster membership, therefore we can see the block diagonal structure in the similarity matrices.

After 10 iterations of the RGF procedure, the blue points moved to the red points in Fig. 1, and the corresponding similarities became crisper than the original ones as shown in Fig. 2 (b) with $\sigma_r = 50000$, which is the same value as Fig. 2 (a). Figure 3 shows the clustering results by Algorithm 1. Figures 3 (a) and (b) show the results without and with the RGF procedure, respectively, i.e., the former result is computed from the similarity matrix in Fig. 2 (a), and the latter from that in Fig. 2 (b). In Fig. 3, the data with the same label have the same color. In Fig. 3 (a), we adjusted the similarity threshold θ in Algorithm 1, line 7, as 0.01 to obtain 15 clusters, where we can see some jumbled data. On the other hand, in Fig. 3 (b), all data are correctly clustered



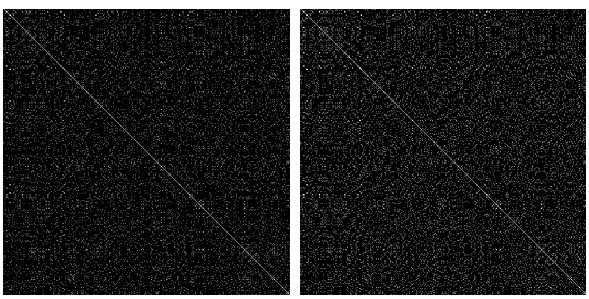
(a) Initial similarities (b) Converged similarities

Fig. 4 Similarity matrices of shuffled data.



(a) Without RGF (b) With RGF

Fig. 5 Clustering results of shuffled data.



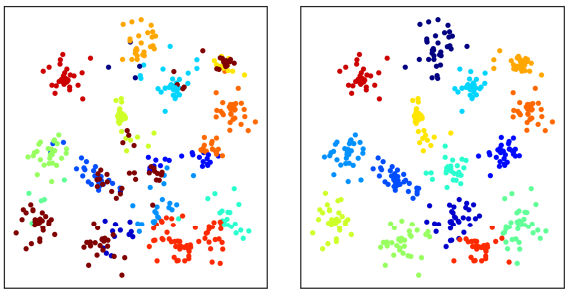
(a) Initial similarities (b) Converged similarities

Fig. 6 Similarity matrices of shuffled data S2.

with $\theta = 0.5$.

Next, we shuffled the order of point data to see the influence of it on the clustering result. Figure 4 shows the similarity matrices with $\sigma_r = 50000$ for both Figs. 4 (a) and (b). We can see that the block diagonal structures in Fig. 2 are untied in Fig. 4. Figures 5 (a) and (b) show the clustering results based on the similarity matrices in Figs. 4 (a) and (b), respectively. In Fig. 5 (a), there are some jumbled data, while Fig. 5 (b) shows the same clustering result as Fig. 3 (b) except for color assignment.

We also clustered another two-dimensional dataset S2 [6]. Figure 6 shows the similarity matrices computed from the shuffled data S2 with $\sigma_r = 50000$. Figures 7 (a) and (b) show the clustering results based on the similarity matrices in Figs. 6 (a) and (b), respectively. A number of jumbled data found in Fig. 7 (a) are correctly clustered in Fig. 7 (b). Figure 8 shows the converged points in red and



(a) Without RGF (b) With RGF

Fig. 7 Clustering results of shuffled data S2.

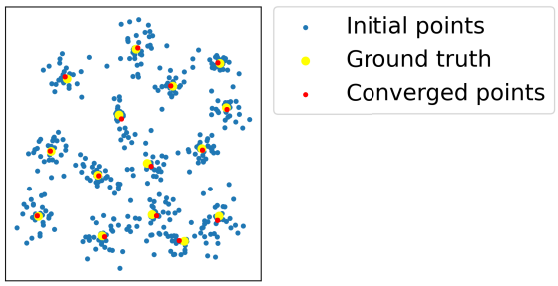


Fig. 8 Two-dimensional data S2 [6]

the ground truth cluster centers in yellow with the original point data in S2 in blue, where the red points have converged near the ground truth yellow points.

5. Conclusion

In this paper, we proposed a generalization of the rolling guidance filter (RGF) to a clustering algorithm for general vector data. The proposed algorithm makes the similarity between data closer to the minimum value 0 or maximum value 1. As a result, we can divide the given data into clusters by a simple algorithm based on a similarity matrix. Experimental results on a benchmark dataset demonstrated the performance of the proposed clustering algorithm, where data points converged to their respective cluster centers or ground truth points without knowing the number of clusters.

Future work will include the development of an RGF-based clustering algorithm which can find clusters of different shapes and non-linearly separable clusters.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP21K11964.

References

[1] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," Sixth International Conference on Computer Vision (IEEE Cat. no.98CH36271), Bombay, India, pp.839–846, 1998, doi: 10.1109/ICCV.1998.710815.

[2] S. Paris, P. Kornprobst, J. Tumblin, F. Durand, "Bilateral Filtering: Theory and Applications," Foundations and Trends® in

Computer Graphics and Vision, vol.4, no.1, pp.1–73, 2009, doi: 10.1561/0600000020.
[3] Q. Zhang, X. Shen, L. Xu, and J. Jia, Rolling Guidance Filter, 2014, In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) Computer Vision – ECCV 2014, ECCV 2014, Lecture Notes in Computer Science, vol.8691, Springer, Cham. doi: 10.1007/978-3-319-10578-9_53
[4] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," ACM Trans. Graphic., vol.23, no.3, pp.664–672, 2004.
[5] E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," ACM Trans. Graph., vol.23, no.3, pp.673–678, Aug. 2004.
[6] P. Fränti and O. Virmajoki: "Iterative shrinking method for clustering problems," Pattern Recognition, vol.39, no.5, pp.761–765, 2006.
[7] P. Fränti and S. Sieranoja: "K-means properties on six clustering benchmark datasets," Applied Intelligence, vol.48, no.12 pp.4743–4759, Dec. 2018. http://cs.joensuu.fi/sipu/datasets/

Appendix A: Convergence to Mean Vector

As Zhang et al. [3] suggested, it is difficult to describe the RGF as an optimization problem from which the RGF procedure will be derived. Therefore, it is also difficult to discuss the convergence of the guidance image to a meaningful image.

In this section, we consider a simplified situation where data form a single cluster, i.e., the kernel density estimate

$$\hat{p}(f) \propto \sum_{j=1}^{n} s\left(f, d_j\right) \tag{A·1}$$

is unimodal. The first iteration of (2) with $w_{ij} = 1$ is given by

$$f_i^1 = \frac{\sum_{j=1}^{n} s\left(f_i^0, f_j^0\right) d_j}{\sum_{j=1}^{n} s\left(f_i^0, f_j^0\right)} = \frac{\sum_{j=1}^{n} s\left(d_i, d_j\right) d_j}{\sum_{j=1}^{n} s\left(d_i, d_j\right)}, \tag{A·2}$$

which satisfies the following inequality:

$$\sum_{j=1}^{n} s\left(f_i^1, d_j\right) \geq \sum_{j=1}^{n} s\left(d_i, d_j\right), \tag{A·3}$$

since (A·2) can be derived from the necessary condition for optimality of $\max_f \hat{p}(f)$:

$$\frac{\partial}{\partial f} \sum_{j=1}^{n} s\left(f, d_j\right) = 0 \tag{A·4}$$

with the initialization $f_i^0 = d_i$ for $i = 1, 2, \ldots, n$. The unimodality of $\hat{p}(f)$ and (A·3) mean that f_i^1 is closer to the only mode of $\hat{p}(f)$ than d_i. On the basis of this observation, we assume that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} s\left(d_i, d_j\right) \left\|d_i - d_j\right\|^2 \geq \sum_{i=1}^{n} \sum_{j=1}^{n} s\left(d_i, d_j\right) \left\|f_i^1 - f_j^1\right\|^2, \tag{A·5}$$

and introduce a convex function $\phi(x) = \exp(-x/2\sigma_r^2)$ which satisfies $\phi(x_2) - \phi(x_1) \geq \phi'(x_1)(x_2 - x_1)$ for all x_1 and x_2, where $\phi'(x_1) = -\exp(-x_1/2\sigma_r^2)/2\sigma_r^2$. Then we have $s(d_i, d_j) = \phi(\|d_i - d_j\|^2)$, and

$$\begin{aligned}
&\sum_{i=1}^{n} \sum_{j=1}^{n} \left[s\left(f_i^1 - f_j^1\right) - s\left(d_i, d_j\right) \right] \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\phi\left(\left\|f_i^1 - f_j^1\right\|^2\right) - \phi\left(\left\|d_i - d_j\right\|^2\right) \right] \\
&\geq \sum_{i=1}^{n} \sum_{j=1}^{n} \phi'\left(\left\|d_i - d_j\right\|^2\right) \left(\left\|f_i^1 - f_j^1\right\|^2 - \left\|d_i - d_j\right\|^2 \right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{2\sigma_r^2} \exp\left(-\frac{\left\|d_i - d_j\right\|^2}{2\sigma_r^2} \right) \left(\left\|d_i - d_j\right\|^2 - \left\|f_i^1 - f_j^1\right\|^2 \right) \\
&= \frac{1}{2\sigma_r^2} \sum_{i=1}^{n} \sum_{j=1}^{n} s\left(d_i, d_j\right) \left(\left\|d_i - d_j\right\|^2 - \left\|f_i^1 - f_j^1\right\|^2 \right) \geq 0,
\end{aligned} \tag{A·6}$$

where the final inequality follows from (A·5). That is, the total sum of similarity $s(f_i, f_j)$ increases with the update of f_i by (A·2). On the other hand, the similarity $s(f_i, f_j)$ is bounded by 1 from above. Therefore, when we iterate (2), each $s(f_i, f_j)$ will approach 1. Then each f_i^t approaches the unweighted average of D or the mean vector $\bar{d} = \sum_{i=1}^{n} d_i/n$.

In the case that there exists more than one cluster in the given data D, the above discussion will be valid for the partial data belonging to the same cluster as shown in Fig. 1, where the converged red points almost overlap with the ground truth yellow points.