IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 2, 2019, accepted January 25, 2019, date of publication February 5, 2019, date of current version February 27, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897597

A Kriging-Assisted Reference Vector Guided Multi-Objective Evolutionary Fuzzy Clustering Algorithm for Image Segmentation

FENG ZHAO[1,2], ZHE ZENG[1,2], HAN QIANG LIU[3], AND JIU LUN FAN[1,2]

[1]Key Laboratory of Electronic Information Application Technology for Scene Investigation, Ministry of Public Security, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
[2]School of Communications and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
[3]School of Computer Science, Shaanxi Normal University, Xi'an 710119, China

Corresponding author: Feng Zhao (fzhao.xupt@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61571361, Grant 61671377, Grant 61202153, and Grant 61102095, and in part by the New Star Team of the Xi'an University of Posts and Telecommunications under Grant xyt2016-01.

ABSTRACT In order to reduce the computational complexity of multi-objective evolutionary optimization-based clustering algorithms, a Kriging-assisted reference vector guided multi-objective robust spatial fuzzy clustering algorithm (KRV-MRSFC) is proposed and then successfully applied to image segmentation. We first construct objective functions with noise robust local spatial information derived from the image to improve the robustness to noise and then use the Kriging model to approximate each objective function to decrease the computational cost. Meanwhile, in order to improve the approximation accuracy of the Kriging model, an angle-penalized distance-based expected improvement sampling criterion is presented in the KRV-MRSFC, which can select individuals with better exploitation and exploration to update the Kriging model. In addition, KRV-MRSFC adopts a clustering validity index with noise robust local image spatial information to select the optimal solution from the final non-dominated solution set to perform image segmentation. The experiments performed on Berkeley and real magnetic resonance images indicate that the proposed method not only achieves satisfactory segmentation performance on noisy images but also requires a low time cost.

INDEX TERMS Image segmentation, multi-objective optimization, fuzzy clustering, Kriging model, reference vector guided evolutionary algorithm.

I. INTRODUCTION

Image segmentation aims to divide an image into multiple heterogeneous regions according to some criteria so that the pixels in the same region have finer similarity while the pixels of different regions have greater differences [1]. It is one of the most challenging tasks in image analysis because segmentation results can directly affect the quality of image processing and recognition [2]. In recent years, researchers have proposed many image segmentation methods, such as histogram threshold methods [3], region growing algorithms [4], graph partitioning methods [5] and clustering algorithms [6], [7]. Fuzzy c-means (FCM) [8] is one of the most popular clustering algorithms. It combines fuzzy set theory with the clustering method so that the given data point can belong to several groups with the degree of belongingness. However, FCM fails to segment images corrupted by noise, outliers and other imaging artifacts because it does not consider any spatial information in the image.

To address the problem of noise sensitivity, many researchers have incorporated the spatial information derived from the image into the objective function of FCM. Ahmed *et al.* [9] proposed a novel FCM algorithm with spatial information (FCM-S) by incorporating a spatial neighborhood term into the objective function of FCM. However, the efficiency of FCM-S is low because the spatial neighborhood term is computed in each iteration. In order to reduce the computational complexity of FCM-S, Chen and Zhang [10]

The associate editor coordinating the review of this manuscript and approving it for publication was Ran Cheng.

proposed two variants, FCM-S1 and FCM-S2, which apply precalculated average filtering and median filtering image to replace the neighborhood term, respectively. Krinidis and Chatzis [11] presented a robust fuzzy local information c-means clustering algorithm (FLICM) which generates a novel local spatial fuzzy factor to make a tradeoff between the noise immunity and preservation of image details. It should be noted that these improved fuzzy clustering algorithms are based on local search and may get into the local optimum due to improper initialization of cluster centers. Furthermore, these approaches always perform image segmentation from one view due to only considering one objective function.

In recent years, multi-objective evolutionary algorithm (MOEA) has become popular in applying to fuzzy clustering due to its heuristic search strategy. Some multi-objective evolutionary fuzzy clustering algorithms have been developed to solve the problems of fuzzy clustering algorithms [12], [13]. Mukhopadhyay *et al.* [14] utilized J_m, Xie-Beni (XB) and PBM as objective functions and obtained the optimal solution from the final non-dominated solution set by using the clustering ensemble strategy. Mukhopadhyay and Maulik [15] proposed a multi-objective variable string length genetic fuzzy clustering algorithm (MOVGA), which uses a variable string length coding strategy to automatically determine the number of clusters and then simultaneously optimizes a global fuzzy compactness function and a fuzzy separation function. Furthermore, MOVGA utilizes I index to obtain the optimal solution from the final non-dominated solution set. Zhao *et al.* [16] incorporated a non-local spatial constrained term into MOVGA and proposed a multi-objective spatial fuzzy clustering algorithm (MSFCA) for image segmentation to overcome the sensitivity to image noise. However, these multi-objective evolutionary fuzzy clustering algorithms require numerous expensive function evaluations, which result in long computation time.

In order to reduce the computational complexity of evolutionary algorithms, researchers used surrogates or metamodels to replace the real function evaluations and proposed surrogate-assisted evolutionary algorithms (SAEAs) [17]–[20]. Pan *et al.* [21] utilized surrogate to estimate the dominance relationship between the candidate solutions and the reference solutions and proposed a classification based surrogate-assisted evolutionary algorithm (CSEA). Sun *et al.* [22] presented a surrogate assisted cooperative swarm optimization algorithm (SA-COSO) which uses a single radial basis function network model to assist the social learning-based particle swarm optimization (SL-PSO) for solving high-dimensional multi-objective problems. Chugh *et al.* [23] used a Kriging model [24] to assist the reference vector guided evolutionary algorithm (RVEA) [25] and proposed a Kriging assisted RVEA called K-RVEA. As the surrogate of K-RVEA, the Kriging model not only approximates the values of objective function, but also gets the uncertainty of approximation which is useful in managing surrogates.

In view of the advantages of K-RVEA, this paper introduces it into fuzzy clustering and proposes a Kriging-assisted reference vector guided multi-objective robust spatial fuzzy clustering algorithm (KRV-MRSFC) for noisy image segmentation. The main contributions of KRV-MRSFC can be summarized as follows: (1) A fuzzy separation function and a global fuzzy compact function with noise robust local spatial information are used as objective functions to improve the robustness to image noise. (2) a Kriging model is employed to approximate objective functions to reduce the running time of multi-objective evolutionary fuzzy clustering algorithm. (3) In order to improve the approximation accuracy of Kriging model, KRV-MRSFC presents an angle penalized distance (APD)-based expected improvement sampling criterion to select individuals for updating the Kriging model. (4) After obtaining the final non-dominated solution set, a clustering validity index with noise robust local spatial information is defined to select the final optimal solution. Experimental results on Berkeley and real magnetic resonance (MR) images show that KRV-MRSFC obtains satisfactory segmentation performance and meanwhile possesses a low time cost.

The rest of this paper is organized as follows. Section II introduces the related techniques used in the proposed algorithm. Section III describes KRV-MRSFC in detail. In Section IV, KRV-MRSFC is verified by segmentation experiments on Berkeley images and MR images. Finally, some concluding remarks and discussions are given in Section V.

II. BACKGROUND

A. REFERENCE VECTOR GUIDED EVOLUTIONARY ALGORITHM

Reference vector guided evolutionary algorithm (RVEA) is a recently proposed efficient multi-objective evolutionary algorithm. Compared with other traditional evolutionary algorithms, e.g. NSGA-II [26], RVEA has two obvious differences: 1) using a set of reference vectors to separate the objective space into several subspaces; 2) proposing a new selection strategy, which is used to balance convergence and diversity. The main framework of RVEA is presented in Algorithm 1.

1) INITIALIZATION OF REFERENCE VECTORS

RVEA uses a set of reference vectors to split the objective space into multiple subspaces and guides the selection process in each subspace. The distribution of the reference vectors in the objective space directly affects the quality of the selected new individuals. Therefore, it is necessary to generate uniformly distributed unit reference vectors in the objective space during initialization. RVEA adopts the approach proposed in [27] to obtain a set of uniformly distributed unit reference vectors. First, a set of uniformly distributed reference points is generated on a unit hyperplane by using

Algorithm 1 Reference Vector Guided Evolutionary Algorithm (RVEA)

Input: a set of unit reference vectors $\mathbf{V}_0$, the maximal number of generations T.

1: Initialize the population P_0 and set the generation counter $t = 0$.
2: **while** $t < T$ **do**
3: Generate the offspring population Q_t
4: Combine the parent and offspring populations: $C_t = P_t \cup Q_t$
5: Select the parent population P_{t+1} from C_t for the next population
6: Update reference vectors $\mathbf{V}_{t+1}$
7: $t = t + 1$
8: **end while**

Output: the final non-dominated solution set.

the simplex-lattice design method [28]:

$$\begin{cases} \mathbf{l}_h = \left(l_h^1, l_h^2, \cdots, l_h^M\right) \\ l_h^j = \left\{\dfrac{0}{G}, \dfrac{1}{G}, \cdots, \dfrac{G}{G}\right\}, \quad \sum_{j=1}^{M} l_h^j = 1 \end{cases} \tag{1}$$

where $h = 1, 2, \cdots, H$ and H is the number of uniformly distributed points. M is the number of objective functions and G is a positive integer. Then, in order to obtain the corresponding reference vectors, these uniformly distributed reference points need to be mapped from a hyperplane to a hypersphere.

$$\mathbf{V}_h = \frac{\mathbf{l}_h}{\|\mathbf{l}_h\|} \tag{2}$$

where $\|\mathbf{l}_h\|$ denotes the L_2-norm of $\mathbf{l}_h$.

2) GENERATION OF OFFSPRING

In RVEA, traditional genetic operations, such as simulating binary crossover [29] and polynomial variations [30], are used to generate the offspring population. At the same time, RVEA adopts the elitism strategy which is the most characteristic part in NSGA-II to preserve the information of individuals among the parents and offspring populations.

3) ANGLE PENALIZED DISTANCE-BASED SELECTION MECHANISM

Since the reference vectors generated by the simplex-lattice design are unit vectors, the objective function values of all individuals in the current generation need to be translated before the reference vectors guiding the selection of new individuals. The j^{th} translated objective function value of the i^{th} individual is obtained as follows:

$$\tilde{f}_i^j = f_i^j - f_{\min}^j \tag{3}$$

where f_i^j is the j^{th} objective function value of the i^{th} individual and $f_{\min}^j$ represents the minimum value of the j^{th} objective function at the current generation.

After the translation, the acute angles between each individual at the current generation and all reference vectors are calculated. Then an individual is assigned to a reference vector if and only if the angle between the individual and the reference vector is minimum among all reference vectors. In this way, the population is divided into multiple subpopulations. One individual is then selected from each subpopulation based on specific selection criteria.

RVEA proposes a novel selection criterion known as angle penalized distance (APD), which makes a better balance between convergence and diversity. APD is defined as:

$$d_i = (1 + P(\theta_i)) \cdot \left\|\tilde{\mathbf{f}}_i\right\| \tag{4}$$

where $\tilde{\mathbf{f}}_i = \left(\tilde{f}_i^1, \tilde{f}_i^2, \cdots, \tilde{f}_i^M\right)$ represents the translated objective vector and $\left\|\tilde{\mathbf{f}}_i\right\|$ is the distance from the translated objective vector to the origin. θ_i is the angle between the i^{th} individual and the reference vector $\mathbf{v}_i$ to which it is assigned. $P(\theta_i)$ is the penalty function defined as follows:

$$P(\theta_i) = M \cdot \left(\frac{t}{T}\right)^{\chi} \cdot \frac{\theta_i}{\gamma_{\mathbf{v}}} \tag{5}$$

where t and T represent the current and the maximum number of generations, respectively, and χ is a predefined parameter. $\gamma_{\mathbf{v}}$ is the smallest angle between the closest reference vector $\mathbf{v}_i$ of the i^{th} individual and the other reference vectors in the current generation, i.e. $\gamma_{\mathbf{v}} = \min_{h \in \{1,2,\cdots,H\}, h \neq i} \langle \mathbf{v}_i, \mathbf{v}_h \rangle$.

After computing APD of all individuals in each subpopulation, one individual with the minimum APD value is selected from each subpopulation and used for the next generation. Therefore, the number of individuals selected for the next generation is equal to the number of nonempty subpopulations in the current generation.

4) ADAPTATION OF REFERENCE VECTORS

In RVEA, the role of uniformly distributed reference vectors is to guide the selection of non-dominated individuals in the objective space. However, the guidance of the reference vectors becomes invalid when the objective function values of the individuals are scaled to different ranges. One way to solve the above problem is to normalize the values of functions in the same range, which will affect the subsequent selection process. To tackle this issue, an adaptation strategy of reference vectors is proposed in RVEA that dynamically adjusts the distribution of reference vectors in the objective space according to the values of individual objective functions. The adaptation of reference vectors is defined as follows:

$$\mathbf{V}_{t+1,h} = \frac{\mathbf{V}_{0,h} \circ \left(\mathbf{z}_t^{\max} - \mathbf{z}_t^{\min}\right)}{\left\|\mathbf{V}_{0,h} \circ \left(\mathbf{z}_t^{\max} - \mathbf{z}_t^{\min}\right)\right\|} \tag{6}$$

where $\mathbf{V}_{0,h}$ is the uniformly distributed reference vectors generated at the initialized stage, $\mathbf{z}_t^{\max}$ and $\mathbf{z}_t^{\min}$ represent the maximum and minimum values of each objective function in t^{th} generation and $\circ$ is the Hadamard product.

B. KRIGING MODEL

Kriging model, also known as Gaussian process, has been frequently used for surrogate techniques. It is a spatial interpolation model that consists of a regression model and a stochastic process. Kriging model not only predicts the objective function value but also provides the uncertainty information of the approximated values, which is very useful for managing surrogates.

For a new individual $\bar{i}$, the approximated function value is calculated as follows:

$$\bar{y} = \mu + \mathbf{r}^T(\bar{i})\mathbf{R}^{-1}(\mathbf{Y} - \mu\mathbf{1}) \tag{7}$$

and the estimated standard deviation is computed by

$$s = \sqrt{\sigma^2(1 - \mathbf{r}^T(\bar{i})\mathbf{R}^{-1}\mathbf{r}^T(\bar{i}) + \frac{1 - \mathbf{1}^T\mathbf{R}^{-1}\mathbf{r}(\bar{i})}{\mathbf{1}^T\mathbf{R}^{-1}\mathbf{1}}} \tag{8}$$

where

$$\mu = \left(\mathbf{1}^T\mathbf{R}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}^T\mathbf{R}^{-1}\mathbf{Y} \tag{9}$$

$$\sigma^2 = \frac{1}{L}(\mathbf{Y} - \mu\mathbf{1})^T\mathbf{R}^{-1}(\mathbf{Y} - \mu\mathbf{1}) \tag{10}$$

where L represents the number of training data. $\mathbf{Y}$ represents a column vector which contains the function values of training data. $\mathbf{1}$ is a column vector filled with ones. $\mathbf{R}$ and $\mathbf{r}(\bar{i})$ represent the correlation matrix and the correlation vector, respectively. For more details about Kriging model, readers can refer to [31]–[33].

C. KRIGING ASSISTED REFERENCE VECTOR GUIDED EVOLUTION ALGORITHM

Kriging-assisted reference vector guided evolutionary algorithm (K-RVEA) is a novel surrogate-assisted multi-objective evolutionary algorithm that uses Kriging model as surrogate and RVEA as the underlying evolutionary algorithm. The flowchart of K-RVEA is shown in Fig. 1.

As we can see from Fig. 1, K-RVEA uses Kriging model to predict the objective function values of the new individuals instead of calculating the expensive objective functions, which can reduce the consumption of time and enhance optimization efficiency. At the same time, K-RVEA proposes a model management strategy to select individuals to be re-evaluated using the original functions and then used to retrain the surrogate to improve the accuracy of Kriging model. In order to balance the convergence and diversity, this strategy not only considers the information from the RVEA algorithm, but also utilizes the uncertainty information from Kriging model. The individual with the maximum uncertainty is selected to satisfy the diversity while the individual with the minimum APD is selected to ensure the convergence.

It should be noted that the computation time for training surrogate model will dramatically increase with the increasing number of training data. Therefore, the proposed model management strategy should ensure that the number of training data is kept sufficiently small. The details of the management strategy can be found in [23].

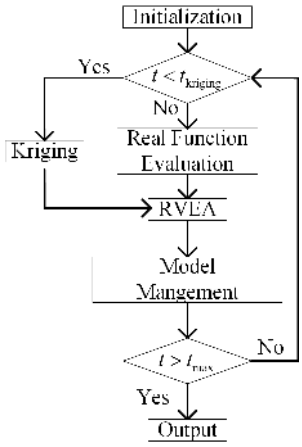


FIGURE 1. The flowchart of K-RVEA.

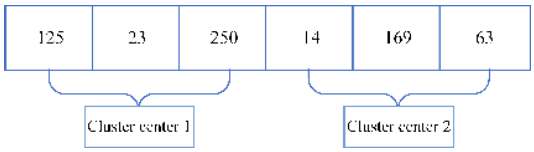


FIGURE 2. Chromosome representation.

III. KRIGING-ASSISTED REFERENCE VECTOR GUIDED MULTI-OBJECTIVE ROBUST SPATIAL FUZZY CLUSTERING ALGORITHM

Taking account of the above advantages of K-RVEA, this paper applies K-RVEA to fuzzy clustering to solve the problem of image segmentation and proposes a Kriging-assisted reference vector guided multi-objective robust spatial fuzzy clustering algorithm (KRV-MRSFC). This algorithm comprises four main parts: initialization, evolutionary operation, model management and optimal solution selection. The details of KRV-MRSFC are shown as follows.

A. POPULATION INITIALIZATION AND CHROMOSOME ENCODING

In KRV-MRSFC, the chromosomes are composed of real numbers which represent the values of cluster centers. If a chromosome encodes the centers of K clusters in M dimensional space, then its length will be $K \times M$. For instance, the chromosome i encoding two cluster centers in three-dimensional space is demonstrated in Fig. 2.

In KRV-MRSFC, cluster centers in each chromosome of the initialized population are randomly selected in the gray level range [0, 255] by using the Latin hypercube sampling [34].

B. OBJECTIVE FUNCTIONS FOR IMAGE SEGMENTATION

Let $X = \{x_1, x_2, \cdots, x_N\}$ denote an image with N pixels. Image segmentation is defined as the union of several non-overlapping meaningful regions with homogeneous characteristics. If this image is separated into K clusters, then these clusters should satisfy $\cup_{k=1}^{K}\Omega_k = X$ and $\cap_{k=1}^{K}\Omega_k = \emptyset$,

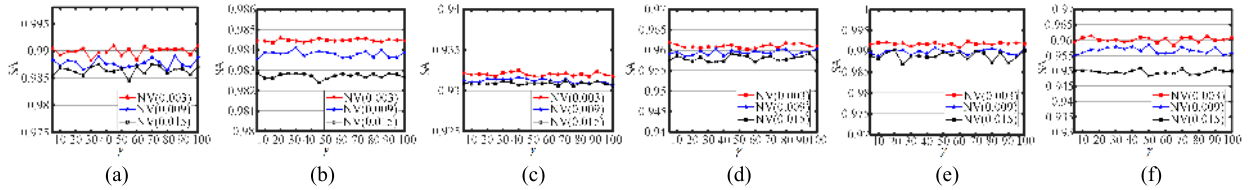


FIGURE 3. Changing trends of SA against the number γ of reference vectors on images corrupted by Gaussian noise: (a) 3063; (b) 3096; (c) 15088; (d) 42049; (e) 135069; (f) 238011.

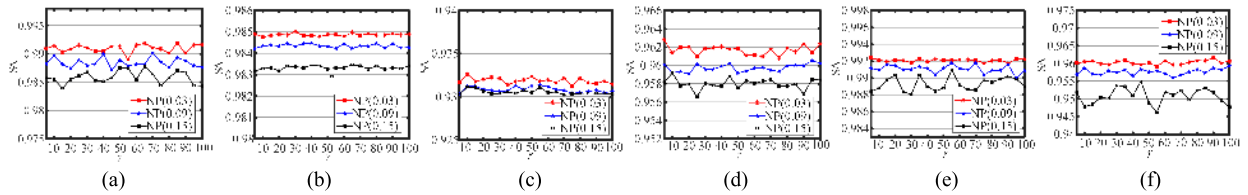


FIGURE 4. Changing trends of SA against the number γ of reference vectors on images corrupted by Salt & pepper noise: (a)3063;(b)3096;(c)15088;(d)42049;(e)135069;(f)238011.

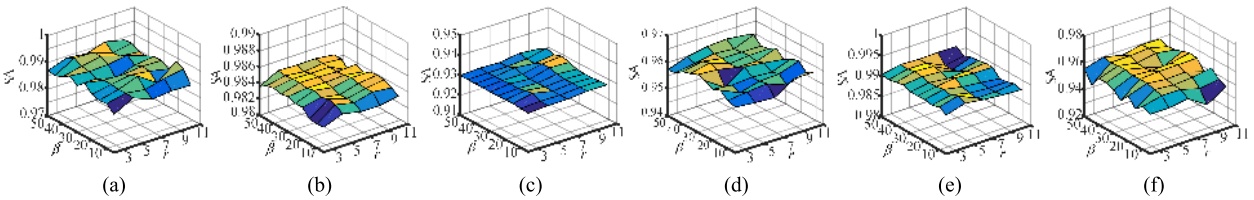


FIGURE 5. SA against the radius r of window and the weighting exponent β on images corrupted by Gaussian noise under NV = 0.009: (a) 3063; (b) 3096; (c) 15088; (d) 42049; (e) 135069; (f) 238011.

where Ω_k represents the k^{th} cluster. It is worth noting that the segmentation criteria will directly affect the segmentation performance. In order to improve the segmentation performance and overcome the influence of noise, a global fuzzy compactness function C_{NRL} with noise robust local spatial information and a fuzzy separation function FS are used as two objective functions in KRV-MRSFC.

C_{NRL} is calculated as:

$$C_{NRL} = \sum_{k=1}^{K} \frac{\sum_{n=1}^{N} u_{kn}^{m} \left(\|x_n - v_k\| + \beta \|\tilde{x}_n - v_k\| \right)}{\sum_{n=1}^{N} u_{kn}} \quad (11)$$

where $v_k (1 \le k \le K)$ represents the k^{th} cluster center, and $u_{kn} (1 \le k \le K, 1 \le n \le N)$ denotes the membership degree function value of the n^{th} pixel belonging to the k^{th} cluster. m is the fuzzy factor that determines the amount of fuzziness of the resulting partition, and β is the weighting exponent parameter that controls the penalty effect of the spatial constraint term. u_{kn} is computed as follows:

$$u_{kn} = \frac{1}{\sum_{l=1}^{K} \left(\frac{\|x_n - v_k\|^2 + \beta \|\tilde{x}_n - v_k\|^2}{\|x_n - v_l\|^2 + \beta \|\tilde{x}_n - v_l\|^2} \right)} \quad (12)$$

In (11) and (12), x_n and $\tilde{x}_n$ represent the gray level and noise robust local spatial information values of the n^{th} pixel, respectively. The majority dominated suppressed similarity strategy proposed in our previous work [35] is used to obtain the noise robust local spatial information $\tilde{x}_n$, which utilizes the neighborhood statistics and the competitive learning. For a neighboring window centered at the n^{th} pixel, the pixels which are not or less contaminated by noise should be the majority for the normal noisy image. These pixels should play big roles in obtaining the spatial information of the central pixel. First, all pixels in the neighboring window N_n centered at the n^{th} pixel are sorted in ascending order according to the gray level values. Then the middle $\lceil 0.618^* (|N_n| - 1) \rceil$ pixels with the minimum variance in the sorted set are considered as the majority and called reward subset N_n^R while the other pixels in the sorted set are deemed as the minority and called the punishment subset N_n^P. $\tilde{x}_n$ is defined as follows:

$$\tilde{x}_n = \sum_{j \in N_n} w_{nj} \cdot x_j \quad (13)$$

$|N_n| = r \times r$ is the size of the neighboring window N_n where r is the radius of N_n. x_j represents the gray level value of the j^{th} pixel in the neighboring window N_n. w_{nj} is the weight factor between the n^{th} pixel and the j^{th} pixel. w_{nj} is computed by

$$w_{nj} = \frac{SS_{nj}}{\sum_{l \in N_n} SS_{nl}} \quad (14)$$

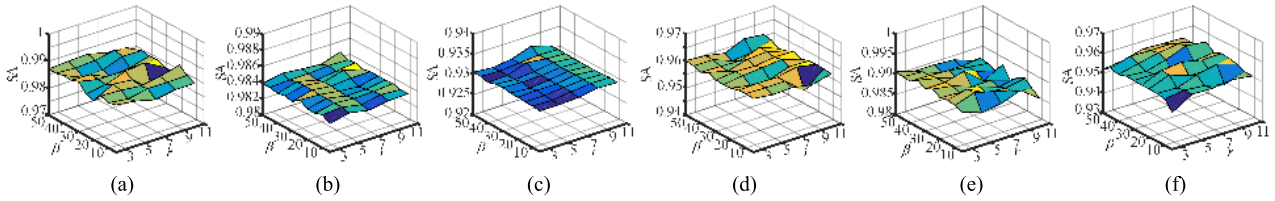


FIGURE 6. SA against the radius r of window and the weighting exponent β on images corrupted by Salt & pepper noise under NP = 0.09: (a) 3063; (b) 3096; (c) 15088; (d) 42049; (e) 135069; (f) 238011.

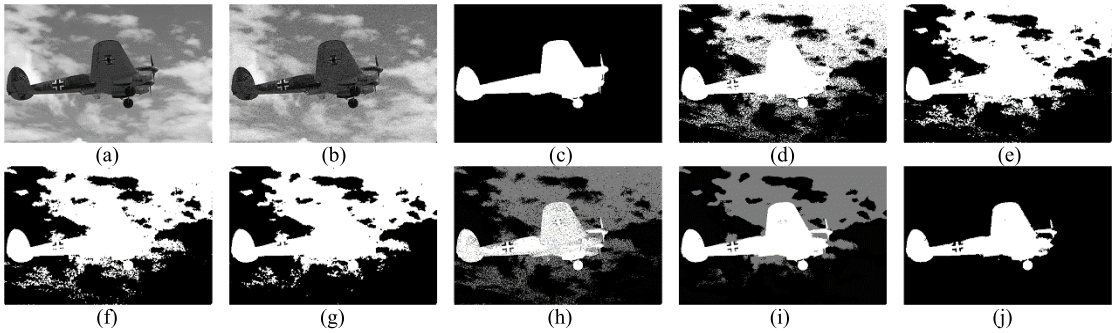


FIGURE 7. Segmentation results on #3063 with Gaussian noise: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

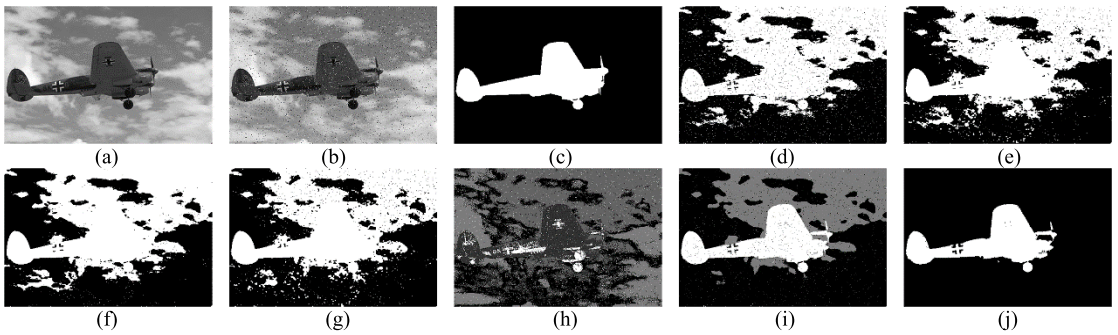


FIGURE 8. Segmentation results on #3063 with Salt & pepper: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

where SS_{nj} represents the majority dominated suppressed similarity and is defined as follows

$$SS_{nj} = \begin{cases} (1+\alpha)\, S_{nj}, & j \in N_n^R \\ (1-\alpha)\, S_{nj}, & j \in N_n^P \end{cases} \tag{15}$$

where α is the suppressed factor. N_n^R and N_n^P represent the reward subset and the punishment subset, respectively. S_{nj} is the local gray level similarity and defined as follows:

$$S_{nj} = \begin{cases} \exp\left(-\left\|x_n - x_j\right\|^2 / \delta_n^2\right), & n \neq j \\ 0, & n = j \end{cases} \tag{16}$$

where $\delta_n = \sqrt{\sum_{j \in N_n} \left\|x_n - x_j\right\|^2 \Big/ |N_n|}$ is a scale factor.

The fuzzy separation function FS is defined as follows:

$$FS = \sum_{p=1}^{K} \sum_{q=1, p \neq q}^{K} \mu_{pq}^m \left\|v_p - v_q\right\| \tag{17}$$

where μ_{pq} denotes the membership degree of the cluster center v_p with respect to the cluster center v_q. It is computed as:

$$\mu_{pq} = \frac{1}{\sum_{l=1, l \neq p}^{K} \left(\frac{\|v_p - v_q\|}{\|v_p - v_l\|}\right)^{\frac{1}{m-1}}}, p \neq q \tag{18}$$

C. EXPECTED IMPROVEMENT SAMPLING STRATEGY BASED ON APD

In this paper, we use the MATLAB Kriging toolbox [36] to train the Kriging model for all initialization data and then use Kriging model to approximate each fitness function to decrease the computational cost.

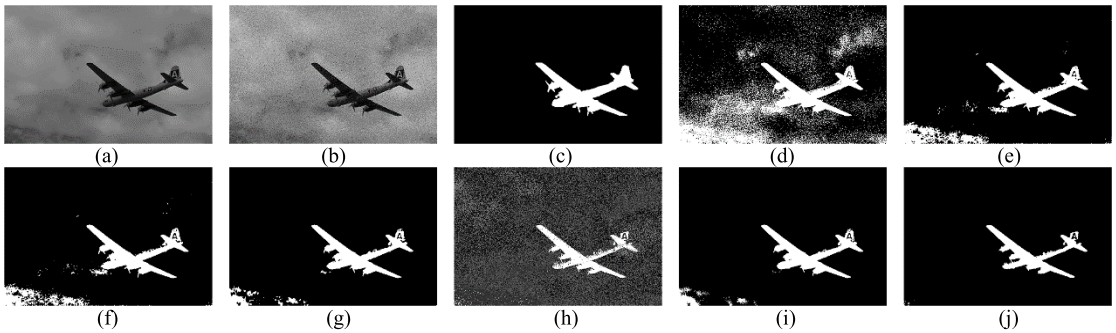


FIGURE 9. Segmentation results on #3096 with Gaussian noise: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

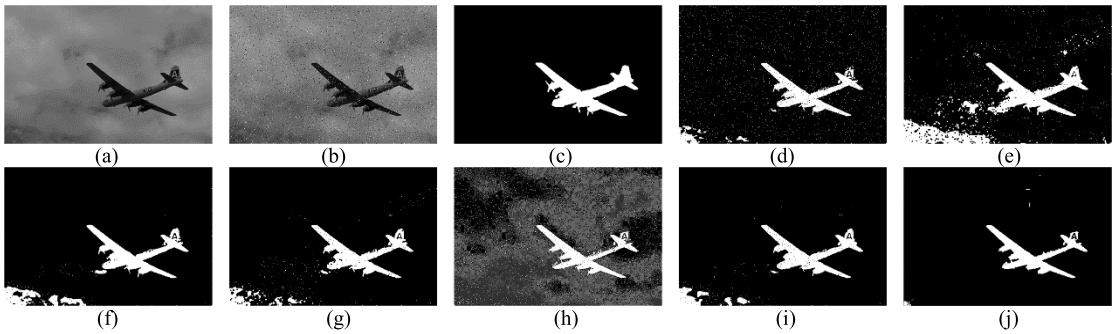


FIGURE 10. Segmentation results on #3096 with Salt & Pepper noise: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

In order to improve the accuracy of the Kriging model approximation, the Kriging model will be updated after finishing a fixed number of generations. The key to solving the above problem is that sample points are selected to update the Kriging model. A number of selection criteria, also known as infill sampling criterion, have been proposed in recent years. Commonly used infill criteria include the lower confidence bound (LCB) [31], the probability of improvement (PI) [37] and the expected improvement (EI) [38]. Due to addressing the trade-off between the exploration and exploitation, EI criterion is one of the most commonly used infilling criteria. In order to well balance the convergence and diversity, this paper introduces the angle penalized distance (APD) into the EI criterion and further proposes a novel expected improvement sampling criterion based on APD. The new criterion is defined as follows:

$$\mathrm{EI}_i^{\mathrm{APD}} = (d_{\min} - d_i) \times \Phi\left(\frac{d_{\min} - d_i}{\tilde{s}_i}\right) + \tilde{s}_i \times \phi\left(\frac{d_{\min} - d_i}{\tilde{s}_i}\right) \quad (19)$$

where d_i and $d_{\min}$ represent the APD value of the i^{th} individual and the current minimum APD value, respectively. $\Phi(\cdot)$ is the normal cumulative distribution function and $\phi(\cdot)$ is the normal probability density function. $\tilde{s}_i$ is the estimated uncertainty of i^{th} individual. It is computed as:

$$\tilde{s}_i = \frac{1}{M}\sum_{j=1}^{M} s_i^j \quad (20)$$

where s_i^j is the estimated standard deviation for the j^{th} objective function value of the i^{th} individual and it is calculated by (8).

The details of the strategy for selecting individuals to update the Kriging model are presented in Algorithm 2. Furthermore, in order to reduce the computational cost of retraining of the Kriging model, this paper adopts the approach proposed in [23] to manage the training data.

D. OPTIMAL SOLUTION SELECTION

In the final generation, KRV-MRSFC produces a set of non-dominant solutions, which are equally important from the algorithm point of view. However, users may only need a single solution in most cases. In order to solve this issue, KRV-MRSFC introduces the noise robust local spatial information into a cluster validity index I [15] and proposes a novel cluster validity index I_{NRL} to select the optimal solution from the final non-dominated solution set. I_{NRL} is computed by

$$I_{NRL} = \frac{1}{K} \times \frac{E_{NRL}(1)}{E_{NRL}(K)} \times D_{NRL}(K) \quad (21)$$

Algorithm 2 Selecting Individuals to Update the Kriging Model

Input: the active adaptive reference vector set $\mathbf{V}_a^a$ and the set J of individuals obtained from RVEA, the number nu of individuals to be used for updating Kriging model.

1: Obtain the size $\left|\mathbf{V}_a^a\right|$ of active adaptive reference.
2: Cluster $\mathbf{V}_a^a$ into min $\left\{nu, \left|\mathbf{V}_a^a\right|\right\}$ groups.
3: Assign J to corresponding active adaptive reference vectors.
4: Calculate the $\mathrm{EI}^{\mathrm{APD}}$ of all individuals in each cluster by using (19).
5: Select one individual from each cluster with the maximum $\mathrm{EI}^{\mathrm{APD}}$.

Output: Individuals for updating the Kriging model.

Algorithm 3 Kriging-Assisted Reference Vector Guided Multi-Objective Robust Spatial Fuzzy Clustering Algorithm (KRV-MRSFC)

Input: the set of unit reference vectors $\mathbf{V}_0$, the size P of the initialized population, the maximum number T of iterations, the prefixed number $\omega_{\max}$ of generations before updating the Kriging model, the number nu of individuals to be used for updating the Kriging model, the image X and the number K of clusters.

1: Compute the noise robust local spatial information for each pixel in the image X.
2: Initialize the population from the gray level range [0, 255].
3: Decode each chromosome in the initialized population.
4: Calculate the corresponding objective functions C_{NRL} and FS for each chromosome in the initialized population.
5: Use individuals in the initialized population to train the Kriging model for each objective function. Set $t = 0$.
6: **while** $t < T$ **do**
7: Set $\omega = 0$
8: **while** $\omega < \omega_{\max}$ **do**
9: Run steps 3-6 of Algorithm 1 and use the Kriging model to estimate the function values.
10: $\omega = \omega + 1$.
11: **end while**
12: Update the Kriging model using Algorithm 2.
13: $t = t + 1$.
14: **end while**
15: Decode each chromosome in the final non-dominated solution set and compute the corresponding index I_{NRL}.

Output: the final optimal solution which is the chromosome with the highest I_{NRL}.

where K is the number of clusters. $E_{NRL}(K)$ represents the fuzzy compactness of the partition and $D_{NRL}(K)$ measures the maximum fuzzy separation between any two clusters among all clusters. They are defined as follows:

$$E_{NRL}(K) = \sum_{k=1}^{K}\sum_{n=1}^{N} u_{kn}\left[\|x_n - v_k\| + \beta\|\tilde{x}_n - v_k\|\right] \quad (22)$$

$$D_{NRL} = \max_{p,q=1}^{K} \|v_p - v_q\| \quad (23)$$

where the membership degree function value u_{kn} is calculated by (12). It is obvious that E_{NRL} (1) is a constant for a given image. The larger value of I_{NRL} implies more compact and well-separated clusters. Therefore, the solution with the highest value of I_{NRL} in the final non-dominated solution set is considered as the optimal solution.

TABLE 1. Parameter settings.

Method	Parameters	Appearance in
FCM	$m=2, T=100, \varepsilon=10^{-5}$	Bezdek *et al.* [8]
FCM-S1	$m=2, T=100, \varepsilon=10^{-5}$, $\beta=6, r=3$	Chen and Zhang [10]
FCM-S2	$m=2, T=100, \varepsilon=10^{-5}$ $\beta=6, r=3$	Chen and Zhang [10]
FLICM	$m=2, T=100, \varepsilon=10^{-5}$ $r=3$	Krinidis and Chatzis [11]
MOVGA	$m=2, K_{\max}=5, T=100$ $P=50, p_c=0.9, p_m=0.1$	Mukhopadhyay and Maulik [15]
MSFCA	$m=2, K_{\max}=5, T=100$ $P=50, p_c=0.9, p_m=0.1$ $\beta=6, r=21, s=7, h=30$	Zhao *et al.* [16]
KRV-MRSFC	$m=2, T=100, P=50$ $p_c=0.9, p_m=0.1, \alpha=0.8$ $nu=5, \omega_{\max}=20, \chi=2$ γ, β, r	Proposed in this paper

E. PROCEDURE OF KRV-MRSFC

The main procedure of KRV-MRSFC is given in Algorithm 3.

IV. EXPERIMENTS AND ANALYSIS

In this paper, we demonstrate the performance of KRV-MRSFC by performing segmentation experiments on Berkeley images [39] and real magnetic resonance (MR) images from IBSR [40]. Six state-of-the-art clustering algorithms, such as FCM [8], FCM-S1 [10], FCM-S2 [10], FLICM [11], MOVGA [15] and MSFCA [16], are adopted as comparative methods. MOVGA and MSFCA can automatically evolve the number of clusters for images while other five methods need users to assign the cluster number in advance. All the methods and their corresponding parameters are presented in Table 1. In the next section, we focus on analyzing the number γ of reference vectors, the radius r of window and the weighting exponent β on the performance of KRV-MRSFC.

A. ANALYSIS OF PARAMETERS OF KRV-MRSFC

We first evaluate the effect of the reference vector number γ. It is tested from 5 to 100 with the increment 5. We respectively

TABLE 2. Segmentation accuracy values of KRV-MRSFC and State-of-Methods on Berkeley images corrupted by Gaussian noise.

Image	NV	FCM	FCM-S1	FCM-S2	FLICM	MOVGA	MSFCA	KRV-MRSFC
3063	0.003	0.6857	0.6728	0.6760	0.6920	0.4533	0.4309	**0.9889**
	0.009	0.6549	0.6771	0.6705	0.6936	0.2361	0.4117	**0.9860**
	0.015	0.6405	0.6784	0.6505	0.6752	0.2011	0.4116	**0.9745**
3096	0.003	0.8123	0.9604	0.9639	0.9782	0.2043	0.9799	**0.9838**
	0.009	0.6744	0.9107	0.9021	0.9725	0.2549	0.9728	**0.9836**
	0.015	0.6283	0.8978	0.8924	0.9585	0.2266	0.9709	**0.9836**
8068	0.003	0.9464,	0.9520	0.9514	0.9541	0.4342	0.7031	**0.9578**
	0.009	0.9395	0.9504	0.9501	0.9520	0.4008	0.7034	**0.9551**
	0.015	0.9370	0.9497	0.9490	0.9515	0.4205	0.7035	**0.9536**
15088	0.003	0.9045	0.9203	0.9193	0.9228	0.4956	0.4091	**0.9307**
	0.009	0.8377	0.9076	0.8957	0.9136	0.3541	0.4033	**0.9311**
	0.015	0.8200	0.8936	0.8907	0.9233	0.3125	0.3845	**0.9311**
24063	0.003	0.9221	0.9490	0.9479	0.9521	0.3694	0.7173	**0.9582**
	0.009	0.8139	0.9416	0.9374	0.9400	0.3174	0.5647	**0.9426**
	0.015	0.7971	0.9325	0.9373	0.9407	0.3072	0.5633	**0.9452**
42049	0.003	0.9562,	0.9609	0.9619	0.9622	0.9565	0.9644	**0.9686**
	0.009	0.9431	0.9584	0.9585	0.9605	0.5127	0.9625	**0.9676**
	0.015	0.9392	0.9583	0.9580	0.9605	0.5018	**0.9625**	0.9585
55067	0.003	0.8802	0.8807	0.8805	0.8809	0.5964	0.8803	**0.9140**
	0.009	0.7362	0.8810	0.8808	0.8811	0.5507	0.8706	**0.9045**
	0.015	0.7099	0.8767	0.8715	0.8603	0.5243	0.6243	**0.9012**
67079	0.003	0.8067	0.8346	0.8351	0.8387	0.3605	0.8361	**0.8457**
	0.009	0.7712	0.8283	0.8269	0.8353	0.2615	0.8217	**0.8369**
	0.015	0.7647	0.8280	0.8259	0.8355	0.2513	0.8214	**0.8360**
86016	0.003	0.8115	0.9427	0.9318	0.9447,	0.1954	0.5635	**0.9662**
	0.009	0.7645	0.9315	0.9163	0.9314	0.1321	0.5560	**0.9466**
	0.015	0.7583	0.9208	0.9142	0.9113	0.1708	0.5260	**0.9266**
100007	0.003	0.8107	0.8152	0.8162	0.8164	0.4865	0.8234	**0.8548**
	0.009	0.7816	0.8138	0.8143	0.8172	0.1107	0.8244	**0.8490**
	0.015	0.7736	0.8141	0.8138	0.8173	0.1176	0.8214	**0.8372**
101027	0.003	0.8107	0.8152	0.8162	0.8164	0.4865	0.8234	**0.8553**
	0.009	0.7816	0.8138	0.8143	0.8172	0.1107	0.8244	**0.8510**
	0.015	0.7735	0.7881	0.7931	0.7980	0.1684	0.4979	**0.8403**
106047	0.003	0.5726	0.5571	0.5577	0.5616	0.2927	0.5324	**0.9184**
	0.009	0.5694	0.5634	0.5637	0.5702	0.3023	0.5263	**0.9168**
	0.015	0.5280	0.5346	0.5538	0.5608	0.2870	0.5008	**0.9082**
108004	0.003	0.6826	0.7172	0.7277	0.8388	0.2992	0.8697	**0.8978**
	0.009	0.6389	0.6983	0.6983	0.8296	0.2870	**0.8645**	0.8511
	0.015	0.6272	0.6966	0.6980	0.8295	0.2330	**0.8620**	0.8475
118035	0.003	0.6341	0.7125	0.6956	0.7315	0.8381	0.8383	**0.8813**
	0.009	0.6007	0.6749	0.6624	0.6675	0.6473	0.8391	**0.8502**
	0.009	0.6268	0.6707	0.6617	0.6540	0.6685	0.8404	**0.8753**
130014	0.003	0.5510	0.5732	0.5749	0.6206	0.3889	0.5730	**0.8447**
	0.009	0.5224	0.5603	0.5602	0.6226	0.3352	0.5186	**0.8288**
	0.015	0.5235	0.5590	0.5569	0.6237	0.3163	0.4986	**0.8116**
135069	0.003	0.5352	0.5399	0.5364	0.5461	0.4680	0.9920	**0.9931**
	0.009	0.5337	0.5223	0.5272	0.5285	0.4349	0.9877	**0.9891**
	0.015	0.5239	0.5145	0.5305	0.5253	0.4255	0.9761	**0.9880**
147091	0.003	0.8993	0.9086	0.9119	0.9095	0.7918	0.7218	**0.9283**
	0.009	0.8746	0.9041	0.9068	0.9061	0.2068	0.7058	**0.9102**
	0.015	0.8683	0.8932	0.8850	0.8953	0.2061	0.6888	**0.9095**
167062	0.003	0.8359	0.9854	0.9121	0.9870	0.9857	0.9859	**0.9922**
	0.009	0.8054	0.9836	0.8606	0.9857	0.9855	0.9852	**0.9906**
	0.015	0.7970	0.9837	0.8582	0.9849	0.9855	**0.9861**	0.9736
238011	0.003	0.5779	0.6306	0.6198	0.6044	0.5685	0.9614	**0.9670**
	0.009	0.5506	0.5984	0.5979	0.9580	0.4464	**0.9607**	0.9506
	0.015	0.5436	0.5958	0.5939	0.9578	0.4306	**0.9606**	0.9558
241004	0.003	0.8579	0.9119	0.9119	0.9052	0.7860	0.9185	**0.9285**
	0.009	0.6668	0.8964	0.8964	0.8860	0.7411	0.9131	**0.9156**
	0.015	0.6326	0.8940	0.8879	0.8741	0.4837	0.9105	**0.9134**

TABLE 3. Segmentation accuracy values of KRV-MRSFC and State-of-Methods on Berkeley images corrupted by salt & pepper noise.

Image	NP	FCM	FCM-S1	FCM-S2	FLICM	MOVGA	MSFCA	KRV-MRSFC
3063	0.03	0.6822	0.6722	0.6675	0.6870	0.1286	0.4320	**0.9887**
	0.09	0.7121	0.6797	0.6676	0.6868	0.1492	0.4275	**0.9847**
	0.15	0.6650	0.6657	0.6643	0.6832	0.1449	0.4189	**0.9808**
3096	0.03	0.9650	0.9354	0.9677	0.9738	0.2247	0.9746	**0.9845**
	0.09	0.9370	0.9173	0.9345	0.9536	0.2221	0.9794	**0.9840**
	0.15	0.9283	0.7778	0.9204	0.9585	0.2268	0.9791	**0.9840**
8068	0.03	0.9362	0.9511	0.9507	0.9536	0.6877	0.7003	**0.9564**
	0.09	0.9092	0.9461	0.9497	0.9516	0.6705	0.6866	**0.9552**
	0.15	0.9076	0.9455	0.9493	0.9512	0.6596	0.6826	**0.9562**
15088	0.03	0.9084	0.9173	0.9188	0.9216	0.3721	0.4098	**0.9351**
	0.09	0.8859	0.9193	0.9060	0.9188	0.3729	0.3882	**0.9311**
	0.15	0.8813	0.9109	0.8906	0.9296	0.3610	0.3700	**0.9302**
24063	0.03	0.9317	0.9376	0.9392	0.9373	0.7904	0.9370	**0.9535**
	0.09	0.8878	0.9005	0.9352	0.9021	0.7676	0.8592	**0.9460**
	0.15	0.8812	0.8947	0.9339	0.8960	0.7581	0.5656	**0.9433**
42049	0.03	0.9450	0.9590	0.9616	0.9604	0.9448	0.9430	**0.9686**
	0.09	0.9177	0.9482	0.9600	0.9530	0.8160	0.9092	**0.9606**
	0.15	0.9176	0.9485	0.9516	0.9331	0.7891	0.9091	**0.9585**
55067	0.03	0.8952	0.9054	0.9110	0.9018	0.3449	0.3806	**0.9269**
	0.09	0.8636	0.8580	0.9101	0.8928	0.4067	0.6168	**0.9236**
	0.15	0.8517	0.8573	0.9077	0.8958	0.3883	0.3634	**0.9218**
67079	0.03	0.8144	0.8282	0.8365	0.8307	0.3845	0.8145	**0.8436**
	0.09	0.7943	0.8082	0.8329	0.8107	0.3526	0.8164	**0.8398**
	0.15	0.7855	0.8094	0.8318	0.8132	0.3577	0.8166	**0.8346**
86016	0.03	0.8416	0.9298	0.9332	0.9261	0.1548	0.5395	**0.9365**
	0.09	0.8393	0.8870	0.9198	0.9222	0.1755	0.5182	**0.9296**
	0.15	0.8371	0.8884	0.9185	0.9153	0.1630	0.4929	**0.9261**
100007	0.03	0.8150	0.8112	0.8168	0.8149	0.1368	0.8165	**0.8362**
	0.09	0.8053	0.7927	0.8134	0.8045	0.1234	0.7948	**0.8326**
	0.15	0.8018	0.7941	0.8036	0.8002	0.1217	0.7930	**0.8233**
101027	0.03	0.8667	0.8864	0.8837	0.8957	0.6108	0.4739	**0.9124**
	0.09	0.8439	0.8693	0.8806	0.8852	0.5877	0.4703	**0.8956**
	0.15	0.8399	0.8469	0.8757	0.8774	0.5525	0.4701	**0.8816**
106047	0.03	0.5678	0.5752	0.5603	0.5753	0.4193	0.7606	**0.9169**
	0.09	0.5697	0.5664	0.5567	0.5591	0.3786	0.5592	**0.9163**
	0.15	0.5534	0.5654	0.5484	0.5574	0.3124	0.5586	**0.9074**
108004	0.03	0.7524	0.7112	0.7510	0.8159	0.3250	**0.8730**	0.8607
	0.09	0.7955	0.6811	0.7344	0.7499	0.2995	0.8372	**0.8560**
	0.15	0.7624	0.6812	0.7271	0.7394	0.2710	0.8396	**0.8507**
118035	0.03	0.8436	0.7160	0.7161	0.6977	0.8585	0.8328	**0.8942**
	0.09	0.8137	0.6817	0.7220	0.5933	0.8182	0.8183	**0.8856**
	0.15	0.8033	0.6801	0.7132	0.5978	0.7816	0.8025	**0.8812**
130014	0.03	0.6002	0.5673	0.5834	0.5957	0.4962	0.6730	**0.8421**
	0.09	0.5711	0.5445	0.5761	0.5728	0.4130	0.6191	**0.8346**
	0.15	0.5700	0.5445	0.5699	0.5802	0.4019	0.6126	**0.8322**
135069	0.03	0.9165	0.5512	0.5541	0.5406	0.5141	0.9912	**0.9922**
	0.09	0.9058	0.5476	0.5485	0.5392	0.4177	0.8552	**0.9896**
	0.15	0.9048	0.5485	0.5291	0.5276	0.4347	0.7874	**0.9901**
147091	0.03	0.8946	0.9053	0.9130	0.9072	0.8088	0.7141	**0.9258**
	0.09	0.8672	0.8914	0.9110	0.8950	0.7849	0.6788	**0.9177**
	0.15	0.8629	0.8920	0.9095	0.8966	0.7724	0.6744	**0.9208**
167062	0.03	0.9590	0.9234	0.9810	0.9070	0.9710	0.9765	**0.9900**
	0.09	0.9273	0.8970	0.9539	0.8567	0.9405	0.9464	**0.9788**
	0.15	0.8249	0.8915	0.9548	0.8556	0.9406	0.9194	**0.9616**
238011	0.03	0.9324	0.8602	0.7613	0.6179	0.9249	0.9553	**0.9595**
	0.09	0.8819	0.7168	0.6945	0.6854	0.8745	0.9347	**0.9544**
	0.15	0.8748	0.7193	0.6625	0.6783	0.8746	0.9131	**0.9526**
241004	0.03	0.9008	0.8962	0.9172	0.8949	0.5797	0.9095	**0.9199**
	0.09	0.8601	0.8400	0.9075	0.8297	0.6850	0.8747	**0.9173**
	0.15	0.8327	0.8401	0.9018	0.8281	0.6444	0.8722	**0.9170**

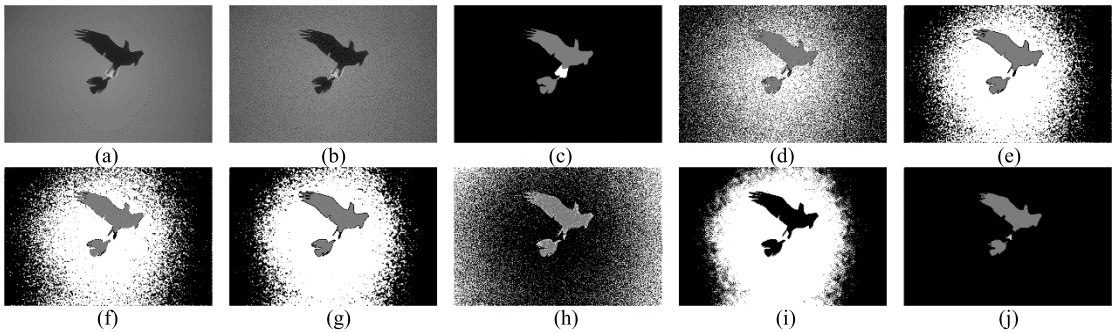


FIGURE 11. Segmentation results on #135069 with Gaussian noise: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

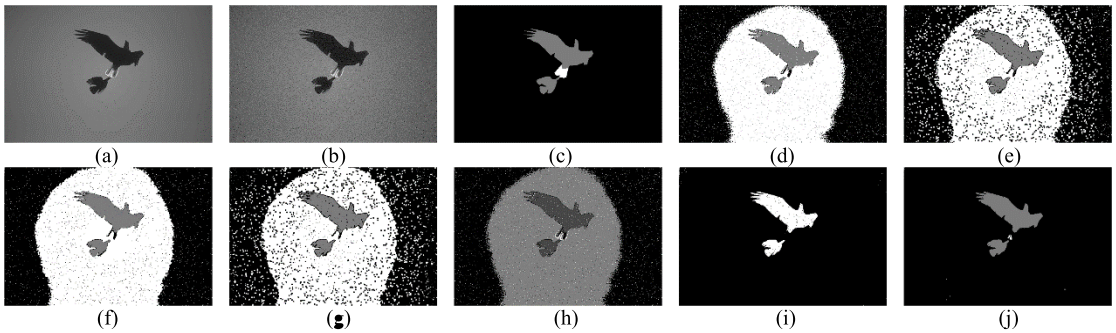


FIGURE 12. Segmentation results on #135069 with Salt & Pepper noise: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

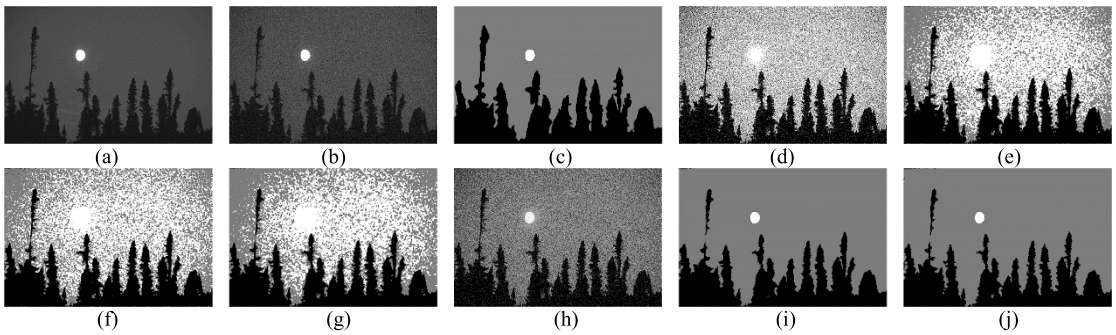


FIGURE 13. Segmentation results on #238011 with Gaussian noise: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

add White Gaussian noise (normalized variance (NV) equal to 0.003, 0.009 and 0.015) and Salt & Pepper noise (noise percentage (NP) equal to 3%, 9% and 15%) to Berkeley images and employ these noisy images to investigate the selection of γ. Figs. 3 and 4 present the curves of segmentation accuracy (SA) [41] with the variation of γ on these noisy images. SA is defined as follows:

$$\mathrm{SA} = \sum_{k=1}^{K} \frac{A_k \cap C_k}{N} \tag{24}$$

where K is the number of clusters and N is the number of pixels. A_k and C_k represent the set of pixels belonging to the k^{th} cluster obtained by the corresponding algorithm and the ground truth image, respectively. It can be found from Figs. 3 and 4 that the values of γ have no obvious effect on the performance of KRV-MRSFC. In this paper, the results of the proposed method under $\gamma = 45$ are provided in the following experiments.

Then we analyze the weighting exponent β and the radius of window r. β and r are tested in the sets {5, 10, 15, 20, 25, 30, 35, 40, 45, 50} and {3, 5, 7, 9, 11}, respectively. We also adopt these noisy images to investigate r and β. Figs. 5 and 6 show the curved surfaces of SA with the variations of r and β on these noisy images. In the following experiments, the radius r of window and the weighting exponent β are assigned to 7 and 25, respectively.

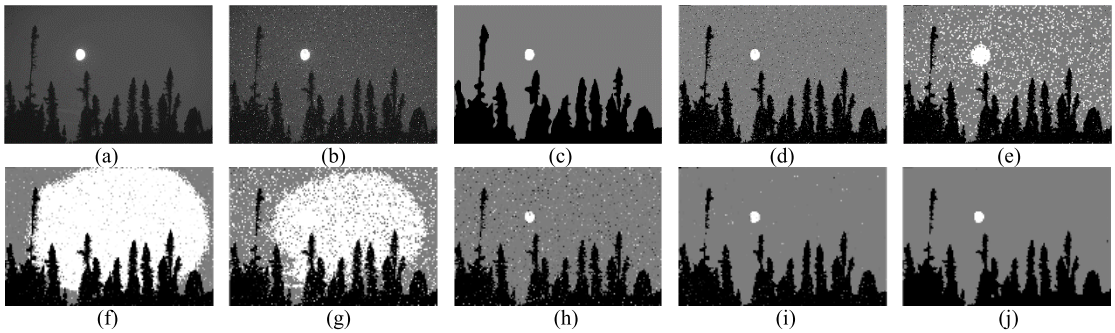


FIGURE 14. Segmentation results on #238011 with Salt & Pepper noise: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

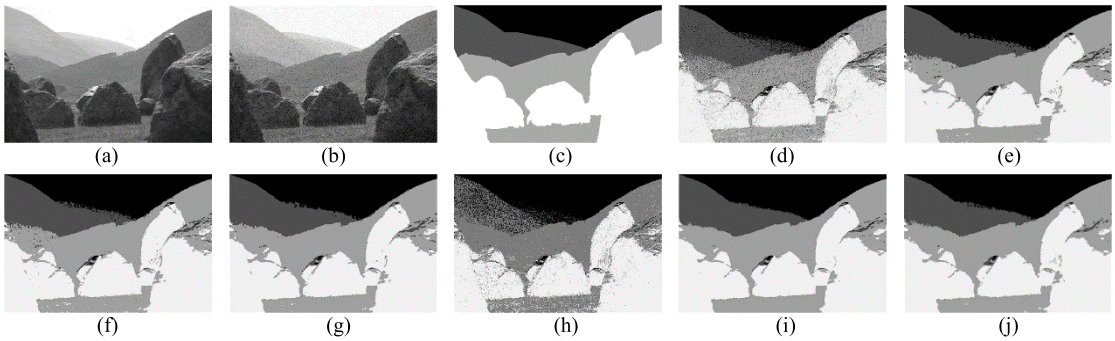


FIGURE 15. Segmentation results on #241004 with Gaussian noise: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

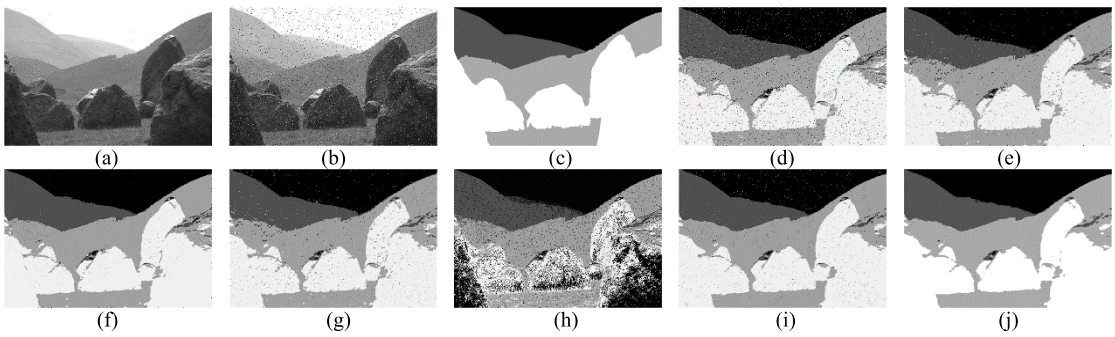


FIGURE 16. Segmentation results on #241004 with Salt & Pepper noise: (a) original image; (b) noisy image; (c) benchmark image; (d) FCM; (e) FCM-S1; (f) FCM-S2; (g) FLICM; (h) MOVGA; (i) MSFCA; (j) KRV-MRSFC.

B. SEGMENTATION EXPERIMENTS ON BERKELEY IMAGES

In this section, we utilize KRV-MRSFC and other comparative methods to segment Berkeley images. White Gaussian noise and Salt & Pepper noise with three different noise levels are added to these Berkeley images. SA is used to quantitatively evaluate the segmentation performance of all the methods. The corresponding results are shown in Tables 2 and 3. It can be found from these tables that KRV-MRSFC outperforms the other methods on most noisy images.

In addition, we select several Berkeley images corrupted by White Gaussian noise with NV $= 0.003$ and Salt & pepper noise with NP $= 3\%$ to visually compare the performance of KRV-MRSFC and other comparative methods. The segmentation results are shown in Figs. 7-16. These results reveal that KRV-MRSFC can overcome the effect of noise and meanwhile obtain satisfactory visual segmentation performance. Compared with MOVGA, KRV-MRSFC introduces the noise-robust local spatial information into two objective functions and the final optimal solution selection criterion to overcome the impact of noise on segmentation results. For instance, Fig. 7 shows that KRV-MRSFC can completely segment the aircraft from the background while MOVGA divide the clouds in the background into the object region. Since the spatial information in KRV-MRSFC is more robust than that in MSFCA, KRV-MRSFC obtains better segmentation result than MSFCA. For example, Fig. 14 reveals

TABLE 4. Segmentation Accuracy Values of KRV-MRSFC and State-of-Methods on MR Images from IBSR.

Image	Slice	FCM	FCM-S1	FCM-S2	FLICM	MOVGA	MSFCA	KRV-MRSFC
1-24	5	0.9725	0.9723	0.9730	0.9758	0.9557	0.9736	**0.9744**
	38	0.9684	0.9724	0.9684	0.9674	0.9688	0.9667	**0.9747**
2-4	20	0.9290	0.9289	0.9295	0.9248	0.9177	0.9256	**0.9466**
	32	0.9485	0.9468	0.9484	0.9449	0.9404	0.9489	**0.9548**
4-8	3	0.9844	0.9844	0.9848	0.9833	0.9753	**0.9929**	0.9903
	32	0.9575	0.9668	0.9554	0.9689	0.9411	0.9534	**0.9736**
5-8	13	0.9589	0.9607	0.9591	0.9558	0.9387	0.9609	**0.9741**
	31	0.9511	0.9516	0.9525	0.9514	0.9390	0.9483	**0.9614**
6-10	10	0.9809	0.9801	0.9810	0.9842	0.9331	0.9820	**0.9874**
	41	0.9657	0.9723	0.9677	0.9683	0.9606	0.9785	**0.9799**
7-8	15	0.9661	0.9652	0.9662	0.9639	0.9660	0.9448	**0.9782**
	22	0.9663	0.9680	0.9652	0.9700	0.9419	0.9356	**0.9779**
8-4	20	0.9520	0.9553	0.9529	0.9541	0.9312	0.9750	**0.9774**
	40	0.9717	0.9727	0.9713	0.9705	0.9811	0.9835	**0.9871**
11-3	15	0.9364	0.9466	0.9373	0.9379	0.9128	0.9575	**0.9746**
	22	0.9342	0.9476	0.9370	0.9488	0.8784	0.9490	**0.9557**
12-3	12	0.9436	0.9447	0.9445	0.9385	0.8927	0.9487	**0.9773**
	35	0.9360	0.9303	0.9350	0.9390	0.9261	0.9313	**0.9609**
13-3	15	0.9110	0.9423	0.9093	0.9078	0.8937	0.9763	**0.9793**
	30	0.9329	0.9418	0.9309	0.9314	0.9227	0.9191	**0.9671**
15-3	15	0.9192	0.9212	0.9185	0.9186	0.9208	0.9471	**0.9525**
	20	0.9154	0.9313	0.9143	0.9123	0.9293	0.9159	**0.9510**
16-3	18	0.9484	0.9508	0.9489	0.9515	0.9466	0.9330	**0.9645**
	20	0.9537	0.9539	0.9554	0.9568	0.9401	0.9312	**0.9604**
17-3	4	0.9897	0.9876	0.9892	0.9892	0.9722	0.9851	**0.9944**
	22	0.9550	0.9584	0.9536	0.9537	0.9545	0.9388	**0.9643**
100-23	20	0.9219	0.9404	0.9192	0.9142	0.8869	0.9667	**0.9674**
	25	0.9291	0.9358	0.9287	0.9317	0.9291	0.9589	**0.9608**
110-3	18	0.9371	0.9535	0.9361	0.9459	0.9283	0.9659	**0.9732**
	25	0.9436	0.9575	0.9452	0.9518	0.8981	0.9660	**0.9670**
111-2	5	0.9843	0.9825	0.9832	0.9851	0.9575	0.9788	**0.9914**
	35	0.9387	0.9476	0.9398	0.9425	0.9569	0.9617	**0.9647**
112-2	7	0.9737	0.9700	0.9730	0.9775	0.9461	0.9713	**0.9833**
	10	0.9588	0.9655	0.9590	0.9671	0.9229	0.9639	**0.9750**
191-3	7	0.9765	0.9759	0.9770	0.9791	0.9435	0.9730	**0.9899**
	42	0.9887	0.9871	0.9883	0.9880	0.9366	0.9579	**0.9898**
202-3	10	0.9630	0.9651	0.9627	0.9627	0.9255	0.9557	**0.9770**
	20	0.9460	0.9532	0.9463	0.9488	0.9376	**0.9769**	0.9698

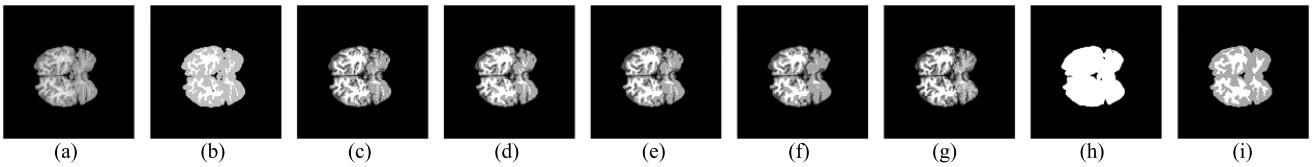


FIGURE 17. Segmentation results on the slice 12 of 12-3: (a) original image; (b) benchmark image; (c) FCM; (d) FCM-S1; (e) FCM-S2; (f) FLICM; (g) MOVGA; (h) MSFCA; (i) KRV-MRSFC.

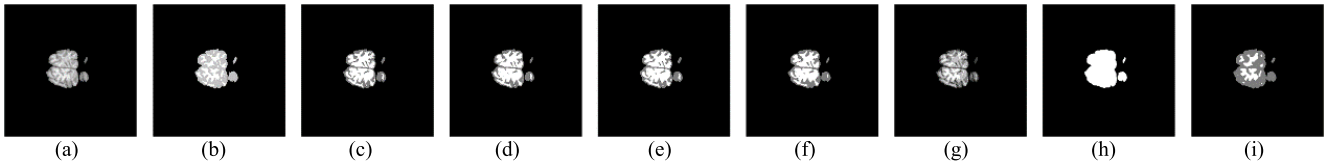


FIGURE 18. Segmentation results on the slice 5 of 111-2: (a) original image; (b) benchmark image; (c) FCM; (d) FCM-S1; (e) FCM-S2; (f) FLICM; (g) MOVGA; (h) MSFCA; (i) KRV-MRSFC.

that KRV-MRSFC can not only well segment the moon and the trees from the background, but also suppress the effects of Salt & Pepper noise in the image compared to MSFCA.

C. SEGMENTATION EXPERIMENTS ON MR IMAGES

Real MR images from IBSR are used to verify the performance of KRV-MRSFC and comparative methods in this section. We do not artificially add noise to MR images

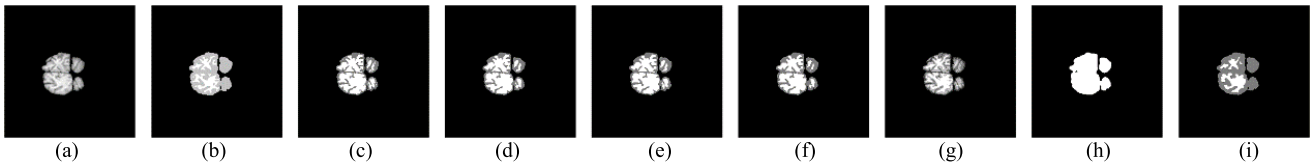


FIGURE 19. Segmentation results on the slice 7 of 191-3: (a) original image; (b) benchmark image; (c) FCM; (d) FCM-S1; (e) FCM-S2; (f) FLICM; (g) MOVGA; (h) MSFCA; (i) KRV-MRSFC.

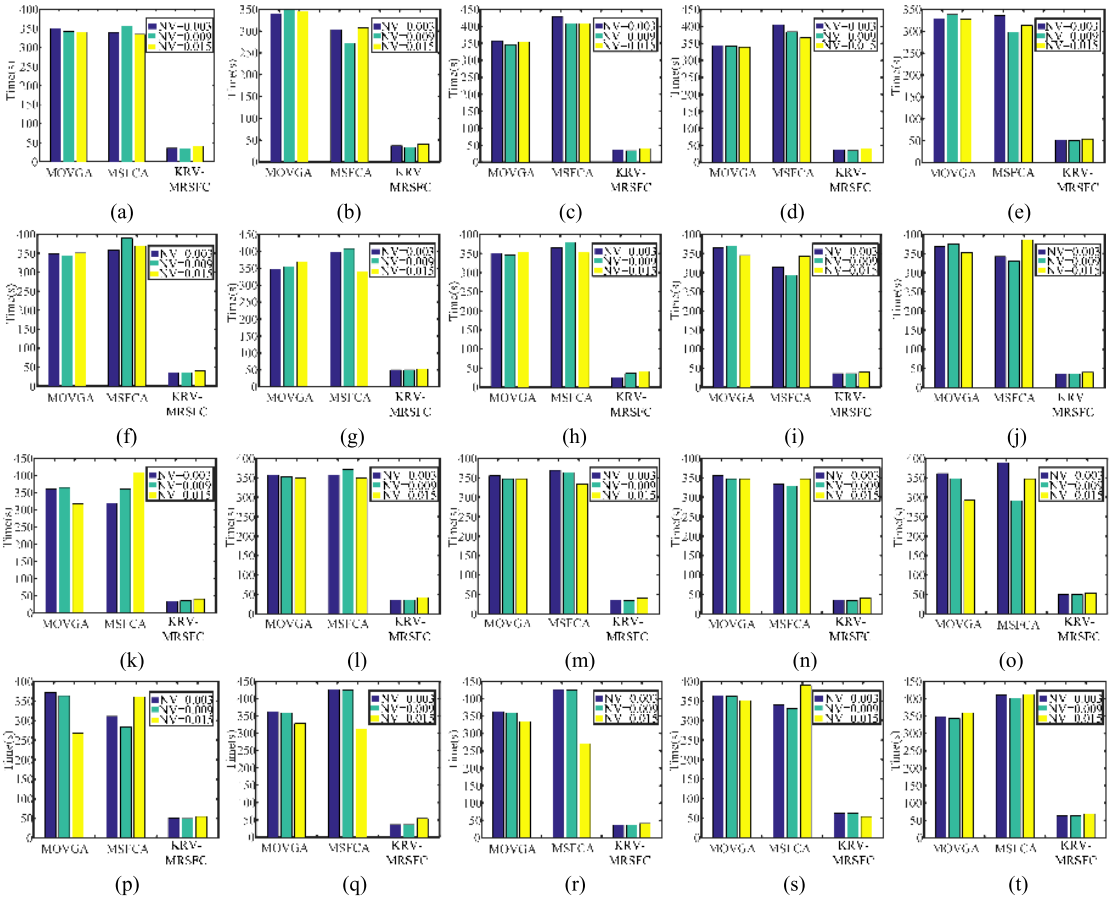


FIGURE 20. Running time comparison of MOVGA, MSFCA and KRV-MRSFC on Berkeley images corrupted by Gaussian noise: (a) 3063; (b) 3096; (c) 8068; (d) 15088; (e) 24063; (f) 42049; (g) 55067; (h) 67079; (i) 86016; (j) 100007; (k) 101027; (l) 106047; (m) 108004; (n) 118035; (o) 130014; (p) 135069; (q) 147091; (r) 167062; (s) 238011; (t) 241004.

because MR images are corrupted by Rician noise. The SA values of all the methods are shown in Table 4. It demonstrates that KRV-MRSFC behaves much better than the other methods on most MR images.

Some MR images are selected to show the visual segmentation performance of KRV-MRSFC and comparative methods. The segmentation results are presented in Figs. 17-19. It can be found from these figures that KRV-MRSFC can well segment the white and gray matter regions in the image. In contrast, the white matter regions in the results of FCM, FCM-S1, FCM-S2 and FLICM are expanded too much. MOVGA and MSFCA cannot obtain satisfactory segmentation results due to failing to evolve a correct cluster number on these images.

D. RUNNING TIME ANALYSIS

In this section, we compare the time efficiency of three multi-objective evolutionary clustering algorithms, MOVGA, MSFCA and KRV-MRSFC. All the methods are implemented with MATLAB and performed on a computer with Inter Core i5-6500 CPU, 8G RAM and Windows 7. Figs. 20 and 21 show the running time of these three multi-objective evolutionary clustering algorithms on several Berkeley images corrupted by White Gaussian noise and Salt & pepper noise with three different noise levels, respectively. Due to adopting Kriging model to predict the values of objective functions of new individuals instead of directly calculating the expensive objective functions, KRV-MRSFC consumes the fewest running time among all the multi-objective evolutionary

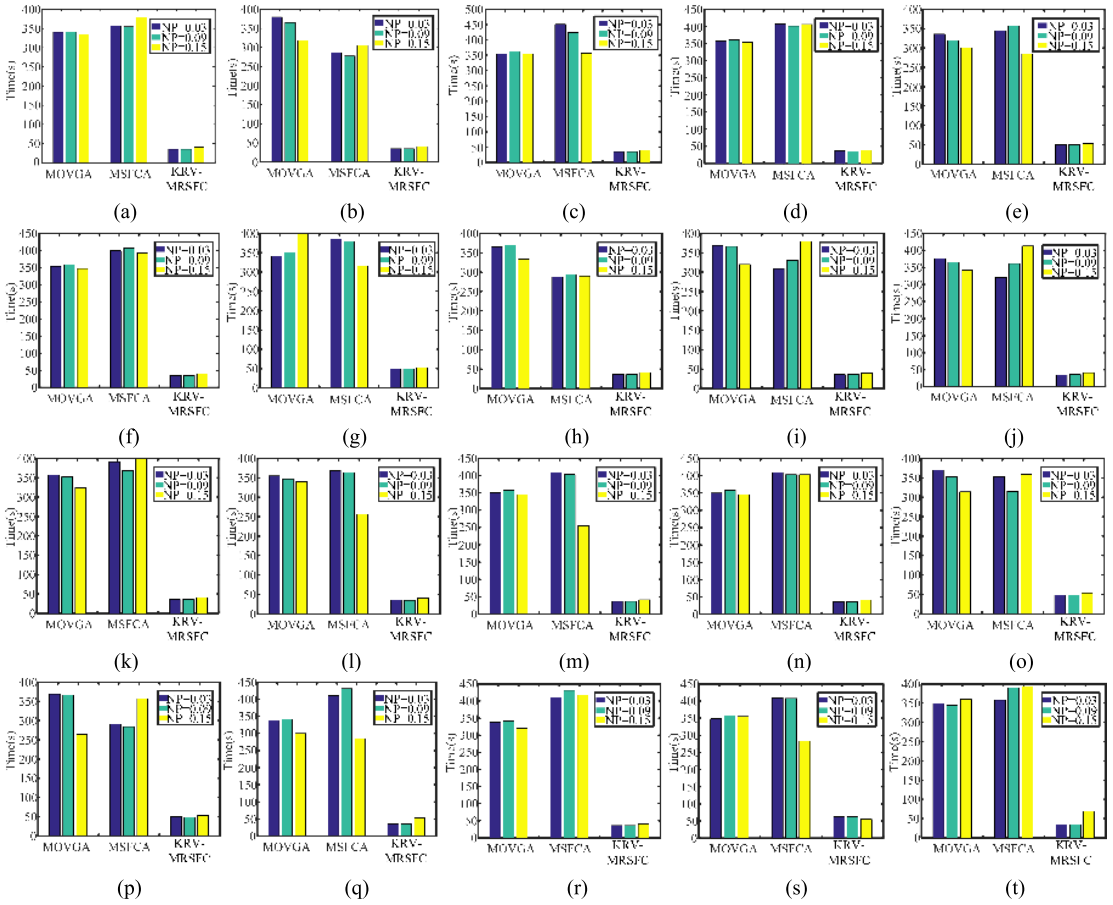


FIGURE 21. Running time comparison of MOVGA, MSFCA and KRV-MRSFC on Berkeley images corrupted by Salt & Pepper noise: (a) 3063; (b) 3096; (c) 8068; (d) 15088; (e) 24063; (f) 42049; (g) 55067; (h) 67079; (i) 86016; (j) 100007; (k) 101027; (l) 106047; (m) 108004; (n) 118035; (o) 130014; (p) 135069; (q) 147091; (r) 167062; (s) 238011; (t) 241004.

clustering algorithms. Therefore, the optimization efficiency of KRV-MRSFC is significantly better than MOVGA and MSFCA.

V. CONCLUSION

A Kriging-assisted reference vector guided multi-objective robust spatial fuzzy clustering algorithm (KRV-MRSFC) for image segmentation is proposed in this paper. To acquire satisfactory segmentation results on noisy images, KRV-MRSFC introduces the noise robust local spatial information into objective functions to be optimized.

Meanwhile, the noise robust local spatial information is also introduced into a cluster validity index for selecting the optimal solution from the final non-dominated solution set. In addition, KRV-MRSFC uses the Kriging model to approximate the original objective functions to reduce the performing time. Moreover, an expected improvement sampling criterion based on APD is proposed in KRV-MRSFC to select individuals to update the Kriging model. This strategy improves the approximation accuracy of the Kriging model. Experimental results on Berkeley and real MR images indicate that KRV-MRSFC outperforms state-of-the-art methods in segmentation performance and meanwhile possesses a low time cost.

Our future research includes automatically determining the cluster number and designing more effective infilling criteria to improve the accuracy of surrogate model. In addition, how to obtain more effective image spatial information to construct fitness functions deserves further research.

REFERENCES

[1] S. Shirakawa and T. Nagao, "Evolutionary image segmentation based on multiobjective clustering," in *Proc. CEC*, Trondheim, Norway, May 2009, pp. 2466–2473.

[2] B. Wang and Z. Tu, "Affinity learning via self-diffusion for image segmentation and clustering," in *Proc. CVPR*, Providence, RI, USA, Jun. 2012, pp. 2312–2319.

[3] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.

[4] R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 6, pp. 641–647, Jun. 1994.

[5] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[6] T. N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 901–914, Apr. 1992.

[7] L. Cinque, G. Foresti, and L. Lombardi, "A clustering fuzzy approach for image segmentation," *Pattern Recognit*, vol. 37, no. 9, pp. 1797–1807, 2004.

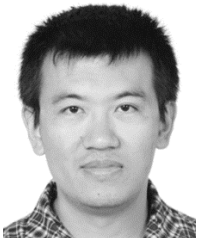
[8] J. C. Bezdek, W. Full, and R. Ehrlich, "FCM: The fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, nos. 2–3, pp. 191–203, 1984.
[9] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. Farag, and T. Moriarty, "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 21, no. 3, pp. 193–199, Mar. 2002.
[10] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 4, pp. 1907–1916, Aug. 2004.
[11] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1328–1337, May 2010.
[12] J. Handl and J. Knowles, "An evolutionary approach to multiobjective clustering," *IEEE Trans. Evol. Comput.*, vol. 11, no. 1, pp. 56–76, Feb. 2007.
[13] S. Bandyopadhyay and U. Maulik, "An evolutionary technique based on K-means algorithm for optimal clustering in RN," *Inf. Sci.*, vol. 146, nos. 1–4, pp. 221–237, Oct. 2002.
[14] A. Mukhopadhyay, U. Maulik, and S. Bandyopadhyay, "Multiobjective genetic clustering with ensemble among Pareto front solutions: Application to MRI brain image segmentation," in *Proc. ICAPR*, Kolkata, India, Feb. 2009, pp. 236–239.
[15] A. Mukhopadhyay and U. Maulik, "A multiobjective approach to MR brain image segmentation," *Appl. Soft Comput.*, vol. 11, no. 1, pp. 872–880, Jan. 2011.
[16] F. Zhao, H. Liu, and J. Fan, "A multiobjective spatial fuzzy clustering algorithm for image segmentation," *Appl. Soft Comput.*, vol. 30, pp. 48–57, May 2015.
[17] Y. Jin, "Surrogate-assisted evolutionary computation: Recent advances and future challenges," *Swarm Evol. Comput.*, vol. 1, no. 2, pp. 61–70, Jun. 2011.
[18] A. Bhosekar and M. Ierapetritou, "Advances in surrogate based modeling, feasibility analysis, and optimization: A review," *Comput. Chem. Eng.*, vol. 108, no. 4, pp. 250–267, Jan. 2018.
[19] H. Wang, Y. Jin, and J. Doherty, "Committee-based active learning for surrogate-assisted particle swarm optimization of expensive problems," *IEEE Trans. Cyber*, vol. 47, no. 9, pp. 2664–2677, Sep. 2017.
[20] H. Wang and Y. Jin, "A random forest-assisted evolutionary algorithm for data-driven constrained multiobjective combinatorial optimization of trauma systems," *IEEE Trans. Cybern.*, to be published. doi: 10.1109/TCYB.2018.2869674.
[21] L. Pan, C. He, Y. Tian, H. Wang, X. Zhang, and Y. Jin, "A classification based surrogate-assisted evolutionary algorithm for expensive many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 23, no. 1, pp. 74–88, Feb. 2018. doi: 10.1109/TEVC.2018.2802784.
[22] C. Sun, Y. Jin, R. Cheng, J. Ding, and J. Zeng, "Surrogate-assisted cooperative swarm optimization of high-dimensional expensive problems," *IEEE Trans. Evol. Comput.*, vol. 21, no. 4, pp. 644–660, Aug. 2017.
[23] T. Chugh, Y. Jin, K. Miettinen, J. Hakanen, and K. Sindhya, "A surrogate-assisted reference vector guided evolutionary algorithm for computationally expensive many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 129–142, Feb. 2018.
[24] D. Buche, N. N. Schraudolph, and P. Koumoutsakos, "Accelerating evolutionary algorithms with Gaussian process fitness function models," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 35, no. 2, pp. 183–194, May 2005.
[25] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff, "A reference vector guided evolutionary algorithm for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 20, no. 5, pp. 773–791, Oct. 2016.
[26] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, "A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II," in *Parallel Problem Solving from Nature PPSN VI*, M. Schoenauer *et al.*, Eds. Berlin, Germany: Springer, Sep. 2000, pp. 849–858.
[27] R. Cheng, Y. Jin, K. Narukawa, and B. Sendhoff, "A multiobjective evolutionary algorithm using Gaussian process-based inverse modeling," *IEEE Trans. Evol. Comput.*, vol. 19, no. 6, pp. 838–856, Dec. 2015.
[28] J. A. Cornell, *Experiments With Mixtures: Designs, Models, and the Analysis of Mixture Data*. Hoboken, NJ, USA: Wiley, 2011, pp. 30–50.
[29] K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space," *Complex Syst.*, vol. 9, no. 2, pp. 115–148, 1995.
[30] K. Deb and M. Goyal, "A combined genetic adaptive search (GeneAS) for engineering design," *Comput. Sci. Inf.*, vol. 26, pp. 30–45, Aug. 1996.
[31] B. Liu, Q. Zhang, and G. G. E. Gielen, "A Gaussian process surrogate model assisted evolutionary algorithm for medium scale expensive optimization problems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 2, pp. 180–192, Apr. 2014.
[32] J. Tian, Y. Tan, J. Zeng, C. Sun, and Y. Jin, "Multi-objective infill criterion driven Gaussian process assisted particle swarm optimization of high-dimensional expensive problems," *IEEE Trans. Evol. Comput.*, to be published. doi: 10.1109/TEVC.2018.2869247.
[33] D. Guo, Y. Jin, J. Ding, and T. Chai, "Heterogeneous ensemble-based infill criterion for evolutionary multiobjective optimization of expensive problems," *IEEE Trans. Cybern.*, to be published. doi: 10.1109/TCYB.2018.2794503.
[34] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979.
[35] F. Zhao, J. Fan, H. Liu, R. Lan, and C. W. Chen, "Noise robust multiobjective evolutionary clustering image segmentation motivated by the intuitionistic fuzzy information," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 2, pp. 387–401, Feb. 2019. doi: 10.1109/TFUZZ.2018.2852289.
[36] S. N. Lophaven, H. B. Nielsen, and J. Søndergaard, "DACE: A MATLAB Kriging toolbox," Ph.D. dissertation, Dept. Inform. Math. Model., Tech. Univ. Denmark, Kongens Lyngby, Denmark, 2002, pp. 2–12.
[37] Z. Zhou, Y. S. Ong, P. B. Nair, A. J. Keane, and K. Y. Lum, "Combining global and local surrogate models to accelerate evolutionary optimization," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 37, no. 1, pp. 66–76, Jan. 2007.
[38] J. Luo, A. Gupta, Y.-S. Ong, and Z. Wang, "Evolutionary optimization of expensive multiobjective problems with co-sub-Pareto front Gaussian process surrogates," *IEEE Trans. Cybern.*, to be published. doi: 10.1109/TCYB.2018.2811761.
[39] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.
[40] A. J. Worth. *The Internet Brain Segmentation Repository (IBSR)*. Accessed: 2009. [Online]. Available: http://www.nitrc.org/projects/ibsr
[41] T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, and A. K. Nandi, "Significantly fast and robust fuzzy C-means clustering algorithm based on morphological reconstruction and membership filtering," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 5, pp. 3027–3041, Oct. 2018.

FENG ZHAO received the B.S. degree in computer science and technology from Heilongjiang University, in 2004, the M.S. degree in signal and information processing from the Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi, China, in 2007, and the Ph.D. degree in pattern recognition and intelligent system from Xidian University, in 2010.

She has been a Professor with the School of Telecommunication and Information Engineering, Xi'an University of Posts and Telecommunications, since 2015. She has authored two books and more than 30 articles. Her research interests include computational intelligence, pattern recognition, and image processing.

She was a recipient of New-Star of Young Science and Technology supported by Shaanxi, in 2014, and the IET International Conference on Ubi-media Computing Best Paper Award, in 2012.

ZHE ZENG was received the B.S. degree in electronic and information engineering from the Xi'an University of Posts and Telecommunications, in 2017, where he is currently pursuing the M.S. degree. His research interests include computational intelligence, image processing, and machine learning.



HAN QIANG LIU received the B.S. degree in computer science and technology from Heilongjiang University, in 2004, and the M.S. and Ph.D. degrees in pattern recognition and intelligent system from Xidian University, in 2007 and 2011, respectively.

He is currently an Associate Professor with the School of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi, China. His research interests include pattern recognition and image processing.

JIU LUN FAN received the B.S. and M.S. degrees in fundamental mathematics from Shaanxi Normal University, in 1984 and 1988, respectively, and the Ph.D. degree in signal and information processing from Xidian University, in 1998.

He is currently the President and a Professor with the Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi, China. He has authored five books and more than 200 articles. His research interests include fuzzy set theory, pattern recognition, and image processing.